FACTEURS DES SUITES DE RUDIN-SHAPIRO GÉNÉRALISÉES

Jean-Paul Allouche Mireille Bousquet-Mélou

Résumé

Par une sorte d'intégration des suites de pliage de papier on peut obtenir des suites de Rudin-Shapiro généralisées. Nous étudions les facteurs de ces suites, donnant en particulier une propriété (optimale) de "semi-synchronisation" et une majoration linéaire uniforme de la fonction de récurrence. Nous étudions aussi les puissances qui apparaissent dans ces suites, et nous montrons que le langage de tous leurs facteurs n'est pas algébrique.

Abstract

"Integrating" paperfolding sequences yields generalized Rudin-Shapiro sequences. We study the factors (subwords) of these sequences, giving an (optimal) property of "half-synchronization" and a uniform linear bound for the recurrence function. We also study the powers occurring in these sequences, and we show that the language consisting of all their factors is not context-free.

1 Introduction

L'étude des facteurs d'une suite infinie à valeurs dans un alphabet fini $\mathcal{A}$ (le plus souvent $\mathcal{A} = \{0, 1\} = \mathbf{Z}/2\mathbf{Z}$) est une manière de décrire si ces suites sont "compliquées" ou pas. Par exemple le nombre de facteurs de longueur n d'une suite binaire est au plus 2^n et on peut concevoir qu'une suite pour laquelle ce nombre est effectivement 2^n quel que soit l'entier n, est une suite "complexe". De même, et sans confondre les notions intuitives de complexité et de hasard, le nombre de facteurs de longueur n d'une suite "aléatoire" devrait être aussi égal à 2^n.

Une autre approche possible du caractère aléatoire d'une suite est due à Shapiro ([21]) puis Rudin ([20]), qui ont étudié, pour une suite binaire $(a_n)_{n\geq 0}$, le comportement asymptotique des moyennes M_N définies par :

$$M_N(a) = \sup_{\theta\in[0,1[} \left| \sum_{0\leq n\leq N-1} (-1)^{a_n} e^{2i\pi n\theta} \right|.$$

Received by the editors April 1993, revised December 1993.
Communicated by M. Boffa.
AMS Mathematics Subject Classification : 68R15, 68Q45, 11B85.
Keywords : Rudin-Shapiro sequences, complexity, recurrence function, algebraic languages, k-powerfree words, paperfolding sequences.

On vérifie facilement que l'on a, quelle que soit la suite a, l'encadrement suivant pour $M_N(a)$:

$$\sqrt{N} = \left\| \sum_{0 \le n \le N-1} (-1)^{a_n} e^{2i\pi n\theta} \right\|_{\mathbf{L}^2} \le M_N(a) \le N.$$

Pour une suite périodique, ou même seulement presque-périodique, le lecteur se convaincra aisément que $M_N(a)$ est de l'ordre de N. Par ailleurs, pour presque toute suite a, au sens de la mesure de Lebesgue, on sait que (voir par exemple [13]) :

$$M_N(a) \le \sqrt{N \log N}.$$

En d'autres termes cette dernière inégalité jointe à la minoration précédente indique qu'une suite "prise au hasard" a une moyenne M_N qui varie à peu près comme $\sqrt{N}$.

Shapiro ([21]) puis Rudin ([20]) ont construit une suite $b = (b_n)$ telle que

$$M_N(b) \le C\sqrt{N}.$$

Cette suite $(b_n)_n$ a donc un comportement "aléatoire" au sens de la moyenne M_N. Pourtant, le nombre de ses facteurs de longueur n est beaucoup plus petit que 2^n (il vaut précisément $8n - 8$ pour $n \ge 8$). Cette suite est par ailleurs engendrée par un algorithme très simple : b_n est la parité du nombre d'occurrences du facteur 11 dans le développement binaire de l'entier n (voir [9]).

On peut rechercher d'autres suites à la fois "aléatoires" au sens de la moyenne M_N et ayant "peu" de facteurs de longueur donnée. Il se trouve que les suites dites de pliage de papier permettent, par une sorte d'intégration modulo 2, de construire une infinité non-dénombrable de suites v ayant peu de facteurs (voir [2]) et vérifiant la propriété "du $\sqrt{N}$" :

$$\exists C_v,\ \forall N \ge 1,\ M_N(v) \le C_v\sqrt{N}.$$

Ces suites s'appellent suites de Rudin-Shapiro généralisées, et ont été introduites par Mendès France et Tenenbaum dans [15].

Nous avons étudié dans [2] et [3] les facteurs des suites de pliage (complexité, synchronisation des facteurs, étude du langage formé de tous les facteurs des suites de pliage), et indiqué que des résultats analogues pouvaient être obtenus pour les suites de Rudin-Shapiro généralisées au sens de [15]. C'est cette question que nous abordons plus en détail ici, montrant en fait quels résultats peuvent être transposés (complexité, semi-synchronisation, langage de tous les facteurs), mais aussi quelles différences apparaissent (impossibilité d'une synchronisation complète). Plus précisément nous nous proposons de démontrer ici les résultats suivants :

- le nombre de facteurs d'une suite de Rudin-Shapiro généralisée au sens de [15] est $8n - 8$, *(résultat déjà obtenu dans [2]), et on peut obtenir la moitié du nombre total de facteurs de longueur donnée de chacune de ces suites en regardant des ensembles de positions indépendants de la suite considérée, ("demi-synchronisation"). L'autre moitié est obtenue en remplaçant dans ces facteurs les* 0 *par des* 1 *et les* 1 *par des* 0*;*

- on ne peut "synchroniser" tous les facteurs de longueur donnée de ces suites;

- il existe une fonction linéaire qui majore uniformément toutes les fonctions de récurrence de ces suites. Plus précisément, elles vérifient toutes l'inégalité $R_v(k+1) < 172k$;

- nous calculons le nombre total de facteurs de longueur donnée qui apparaissent dans au moins une de ces suites;

- nous étudions le langage $\mathcal{L}$ *de tous les facteurs des suites de Rudin-Shapiro (toujours au sens de [15]) : nous recherchons les puissances et les chevauchements (la même question est étudiée pour les facteurs de toutes les suites de pliage), nous calculons le nombre de mots de* $\mathcal{L}$ *de longueur donnée, nous montrons que* $\mathcal{L}$ *n'est pas algébrique et que sa série génératrice est transcendante (donc si le complémentaire de* $\mathcal{L}$ *est algébrique, il est ambigu).*

Notons pour finir cette introduction qu'il existe d'autres généralisations de la suite classique de Rudin-Shapiro, par exemple les suites proposées par Mendès France et étudiées dans [6] et [4], ou encore les suites de Rudin-Shapiro sur plusieurs lettres introduites par Rider (voir [19] et [18]).

2 Suites de pliage et suites de Rudin-Shapiro généralisées

Les suites de pliage sont obtenues en répétant une infinité de fois l'opération consistant à plier en deux une feuille de papier, dans un sens ou dans l'autre. Il existe fort heureusement des façons plus formelles de les décrire, (voir par exemple [10], ou la bibliographie de [2]). On peut tout d'abord donner une définition récursive. La suite $u = (u_n)_{n\geq 1}$ est une suite de pliage si et seulement si elle vérifie les deux conditions suivantes :

- il existe $f_0 \in \{0,1\}$ tel que $\forall\, n \geq 0,\ u_{2n+1} = f_0 + n \mod 2$,
- la suite $(u_{2n})_{n\geq 1}$ est elle-même une suite de pliage.

La suite u est alors caractérisée par la donnée des valeurs $f_m = u_{2^m}$ pour $m \geq 0$. Inversement, étant donnée une suite $(f_m)_{m\geq 0}$, la suite u définie par

$$\forall\, m \geq 0, \forall\, i \geq 0,\ u_{2^m(2i+1)} = f_m + i \mod 2$$

est une suite de pliage. La suite $(f_m)_{m\geq 0}$ est la *suite d'instructions de pliage* associée à u.

Il est aussi possible de définir les suites de pliage à l'aide des opérateurs de *symétrie perturbée* [8]. Si $a \in \{0,1\}$, on note T_a la transformation de $\{0,1\}^*$ définie par :

$$\forall\, w \in \{0,1\}^*,\ T_a(w) = wa\tilde{w},$$

où $\tilde{w}$ est le mot obtenu en lisant w de droite à gauche et en remplaçant toute occurrence de 0 (resp. 1) par 1 (resp. 0). Alors la suite de pliage associée à la suite d'instructions $(f_m)_{m\geq 0}$ est

$$\lim_{m\to\infty} T_{f_m} T_{f_{m-1}} \ldots T_{f_0}(\emptyset).$$

Les suites de Rudin-Shapiro généralisées – que nous nommerons simplement suites de Rudin-Shapiro – se construisent à partir des suites de pliage en utilisant un opérateur d'*intégration* I, défini comme suit :

$$\begin{array}{rccc} I: & \{0,1\}^{\mathbf{N}^*} & \rightarrow & \{0,1\}^{\mathbf{N}} \\ & (u_n)_{n\geq 1} & \mapsto & (v_n)_{n\geq 0} \end{array}$$

où

$$v_n = \sum_{i=1}^{n} u_i \mod 2.$$

Bien entendu, $v_0 = 0$.

Les suites de Rudin-Shapiro sont les images par I des suites de pliage. Remarquons que l'opérateur I, étant injectif, induit une bijection entre les suites de pliage et les suites de Rudin-Shapiro.

Rappelons maintenant les résultats que nous avons obtenus dans [3] et [2] sur les facteurs des suites de pliage, en traitant dans un premier temps des questions liées au nombre et à la position de ces facteurs.

Notation 2.1 *Soit $u = (u_n)_{n\geq n_0}$ une suite, et $k \geq 1$. Nous noterons :*

- $\mathcal{F}_u(k) = \{u_n u_{n+1} \dots u_{n+k-1}, n \geq n_0\}$ *l'ensemble des facteurs de u de longueur k, et $p_u(k) = \#\mathcal{F}_u(k)$ la complexité de u,*
- $\mathcal{F}_u^p(k) = \{u_{2n} u_{2n+1} \dots u_{2n+k-1}, 2n \geq n_0\}$ *l'ensemble des facteurs de u de longueur k placés en position paire, et $p_u^p(k) = \#\mathcal{F}_u^p(k)$ la complexité paire de u,*
- $\mathcal{F}_u^i(k) = \{u_{2n+1} u_{2n+2} \dots u_{2n+k}, 2n+1 \geq n_0\}$ *l'ensemble des facteurs de u de longueur k placés en position impaire, et $p_u^i(k) = \#\mathcal{F}_u^i(k)$ la complexité impaire de u.*

Par ailleurs, si $a \in \{0,1\}$, on note $\bar{a}$ le nombre $1-a$. De même, on pose $\bar{p} = i$ et $\bar{i} = p$.

Résultat 2.2 *Toutes les suites de pliage ont même complexité paire et impaire : pour toute suite de pliage u, pour tout $k \geq 1$,*

$$p_u^p(k) = P^p(k) = 4\left\lfloor \frac{k+1}{2} \right\rfloor \quad \text{et} \quad p_u^i(k) = P^i(k) = 4\left\lfloor \frac{k}{2} \right\rfloor.$$

De plus, pour tout $k \geq 1$, il existe des ensembles $\mathcal{P}_k^p \subset 2\mathbf{N}^*$ et $\mathcal{P}_k^i \subset 2\mathbf{N}+\mathbf{1}$ tels que, pour α dans $\{p,i\}$,

· $\#\mathcal{P}_k^\alpha = P^\alpha(k)$,

· *pour toute suite de pliage u, $\#\{u_n u_{n+1} \dots u_{n+k-1}, n \in \mathcal{P}_k^\alpha\} = P^\alpha(k)$.*

Pour passer ensuite des complexités paire et impaire à la complexité globale, nous utilisons le résultat suivant.

Résultat 2.3 *Soient u et v deux suites de pliage, et $k \geq 7$. Alors :*

$$\mathcal{F}_u^p(k) \cap \mathcal{F}_v^i(k) = \emptyset.$$

Ceci permet d'énoncer un résultat global.

Résultat 2.4 *Toutes les suites de pliage ont la même fonction de complexité $P(k)$ définie par ([2]) :*
· $P(0) = 1, P(1) = 2, P(2) = 4, P(3) = 8, P(4) = 12, P(5) = 18, P(6) = 23$,
· pour $k \geq 7$, $P(k) = 4k$.
De plus, ([3]), pour tout $k \geq 1$, il existe un ensemble $\mathcal{P}_k$ de cardinal $4k$ tel que, pour toute suite de pliage u, et pour $k \geq 7$:

$$\#\{u_n u_{n+1} \dots u_{n+k-1}, n \in \mathcal{P}_k\} = 4k.$$

Il suffit en effet de choisir $\mathcal{P}_k = \mathcal{P}_k^p \cup \mathcal{P}_k^i$. Dans [3], nous démontrons que les ensembles $\mathcal{P}_k$ peuvent en fait être construits à partir de $\mathcal{P}_1$ par la récurrence suivante :

$$\forall k \geq 1, \quad \mathcal{P}_{2k} = (2\mathcal{P}_k - 1) \cup 2\mathcal{P}_k \quad \text{et} \quad \mathcal{P}_{2k+1} = (2\mathcal{P}_k - 1) \cup 2\mathcal{P}_{k+1}.$$

En particulier, lorsque $\mathcal{P}_1 = \{2, 3, 5, 6\}$, l'ensemble $\mathcal{P}_k$ est de diamètre minimal, a pour élément minimal $1 + 2^{H(k+1)}$ et pour élément maximal $12.2^{H(k)}$, l'entier $H(k)$ étant caractérisé par l'inégalité $2^{H(k)} < k \leq 2^{H(k)+1}$. Nous utiliserons ce résultat lors de l'étude de la fonction de récurrence des suites de Rudin-Shapiro.

Dans un deuxième temps, nous étudions dans [3] l'ensemble des *mots de pliage*, c'est-à-dire des mots qui apparaissent comme facteurs d'au moins une suite de pliage. Nous énumérons ces mots, et montrons qu'ils forment un langage non-algébrique.

L'objet de cet article est de voir de quelle façon ces résultats peuvent être transposés aux suites de Rudin-Shapiro.

Nous commençons par établir deux lemmes qui seront utiles par la suite. Nous notons ici $\Phi(n)$ la valuation 2-adique de n, c'est-à-dire l'entier m défini par $n = 2^m(2i+1)$.

Lemme 2.5 1. *Soient m et n deux entiers tels que $m < n$. Il existe un entier $j \in [m, n]$ tel que, pour tout $i \in [m, n]$, $i \neq j$, on ait $\Phi(i) \neq \Phi(j)$.*

2. *Soit $n = 2^j(2k+1)$ et $i > j$. Alors, pour tout $r \geq 0$ et pour toute suite de pliage u, $u_n = u_{n+r2^{i+1}}$.*

Preuve. 1. Soit $k = \max\{\Phi(i), i \in [m, n]\}$. Soit $j \in [m, n]$ tel que $\Phi(j) = k$. Supposons qu'il existe $j' \in [m, n]$, $j' \neq j$, tel que $\Phi(j') = k$. Alors $j = 2^k(2r+1)$ et $j' = 2^k(2s+1)$ avec $r < s$ par exemple. Mais alors $l = 2^k(2r+2) = 2^{k+1}(r+1)$ est un élément de $[m, n]$ qui vérifie $\Phi(l) > k$, ce qui contredit le choix de k.
2. Il suffit d'utiliser la relation $u_{2^m(2i+1)} = f_m + i \bmod 2$.

■

Lemme 2.6 *Soient $m \geq 0$ et $i \geq 0$ tels que $m < 2^i$. Soit u une suite de pliage, et $v = I(u)$ la suite de Rudin-Shapiro associée. Pour $r \geq 0$, notons $w_r = v_{m+r2^{i+1}}$. Cinq termes consécutifs de la suite $(w_r)_{r \geq 0}$ ne sont alors jamais identiques.*

Preuve. Soit $(f_j)_{j\geq 0}$ la suite d'instructions de pliage à partir de laquelle est construite u. Soit $(u'_r)_{r\geq 1}$ la suite de pliage associée à la suite d'instructions $(f_j)_{j\geq 1+i}$. Alors pour $r \geq 1$, $u'_r = u_{r2^{i+1}}$. Étant donné un mot $a = a_1 a_2 \dots a_k$ de $\{0,1\}^*$, on notera ici $s(a)$ la somme $a_1 + \cdots + a_k$, prise modulo 2.

La construction de u par symétrie perturbée permet d'écrire :

$$u = (F f_i \tilde{F}) u'_1 (F \bar{f}_i \tilde{F}) u'_2 (F f_i \tilde{F}) u'_3 (F \bar{f}_i \tilde{F}) u'_4 \dots$$

où $F = u_1 \dots u_{2^i - 1}$.
Notons $F = PS$, avec $P = u_1 \dots u_m$; on a, pour $r \geq 1$,

$$\begin{aligned} w_r - w_{r-1} &= s(S) + s(f_i + r - 1) + s(\tilde{S}) + s(\tilde{P}) + s(u'_r) + s(P) \\ &= \mathrm{long}(SP) + f_i + r - 1 + u'_r \mod 2 \\ &= 2^i - 1 + f_i + r - 1 + u'_r \mod 2 \\ &= f_i + r + u'_r \mod 2. \end{aligned}$$

Donc $w_r = w_{r-1}$ si et seulement si $u'_r = f_i + r \mod 2$. En particulier, si cinq termes consécutifs de la suite $(w_r)_{r\geq 0}$ sont égaux, on a, pour une certaine valeur de r,

$$u'_r u'_{r+1} u'_{r+2} u'_{r+3} = a\bar{a}a\bar{a}$$

avec $a = f_i + r \mod 2$. Mais le facteur $a\bar{a}a\bar{a}$ n'est pas un facteur de pliage, et on aboutit donc à une contradiction.

■

La proposition suivante (plus précise que celle donnée dans [2] qui donne seulement la dernière égalité) décrit le passage des facteurs de pliage aux facteurs de Rudin-Shapiro.

Proposition 2.7 *Soit Ψ l'application suivante :*

$$\begin{array}{llll} \Psi : & \{0,1\}^+ & \to & \{0,1\} \times \{0,1\}^* \\ & a_1 a_2 \dots a_k & \mapsto & (a_1, (a_2 - a_1)(a_3 - a_2) \dots (a_k - a_{k-1})), \end{array}$$

tous les termes étant évalués modulo 2. Soit u une suite de pliage, et $v = I(u)$ la suite de Rudin-Shapiro associée. Soit $k \geq 1$. Par restriction des ensembles de départ et d'arrivée, Ψ induit :
· une bijection Ψ_1 de $\mathcal{F}_v^p(k)$ dans $\{0,1\} \times \mathcal{F}_u^i(k-1)$,
· une bijection Ψ_2 de $\mathcal{F}_v^i(k)$ dans $\{0,1\} \times \mathcal{F}_u^p(k-1)$,
· une bijection Ψ_3 de $\mathcal{F}_v(k)$ dans $\{0,1\} \times \mathcal{F}_u(k-1)$.
Par conséquent, toutes les suites de Rudin-Shapiro ont même fonction de complexité paire, impaire et globale : pour tout $k \geq 1$,
· $p_v^p(k) = 2P^i(k-1)$,
· $p_v^i(k) = 2P^p(k-1)$,
· $p_v(k) = 2P(k-1)$.
(Comme d'habitude le mot vide est facteur de toute suite, donc les égalités ci-dessus ont un sens pour $k = 1$).

Preuve. Les transformations Ψ_1, Ψ_2 et Ψ_3 sont injectives, puisque Ψ l'est. D'autre part, si v est une suite de Rudin-Shapiro et $v_n v_{n+1} \dots v_{n+k-1}$ l'un de ses facteurs, alors $\Psi(v_n \dots v_{n+k-1}) = (v_n, u_{n+1} \dots u_{n+k-1})$ appartient visiblement à $\{0,1\} \times \mathcal{F}_u^i(k-1)$ (resp. $\{0,1\} \times \mathcal{F}_u^p(k-1)$, $\{0,1\} \times \mathcal{F}_u(k-1)$) dès que $v_n v_{n+1} \dots v_{n+k-1}$ appartient à $\mathcal{F}_v^p(k)$ (resp. $\mathcal{F}_v^i(k)$, $\mathcal{F}_v(k)$). Le seul point restant à prouver est donc la surjectivité de Ψ_1 et Ψ_2, qui implique celle de Ψ_3 puisque, pour toute suite w et tout entier l, $\mathcal{F}_w(l) = \mathcal{F}_w^p(l) \cup \mathcal{F}_w^i(l)$.

Soit donc $\alpha \in \{i, p\}$ et w un élément de $\mathcal{F}_u^\alpha(k-1)$. Il existe un entier $m \geq 0$, de parité $\bar{\alpha}$, tel que $w = u_{m+1} \dots u_{m+k-1}$. Alors $(v_m, w) = \Psi(v_m \dots v_{m+k-1})$ appartient à $\Psi(\mathcal{F}_v^{\bar{\alpha}}(k))$. Nous cherchons à prouver que $(1 - v_m, w)$ est aussi dans $\Psi(\mathcal{F}_v^{\bar{\alpha}}(k))$.

Soit $i \geq 1$ tel que $m + k - 1 < 2^i$.

- D'après le lemme 2.5, on a

$$\forall r \geq 0, \; u_{m+1+r2^{i+1}} \dots u_{m+k-1+r2^{i+1}} = u_{m+1} \dots u_{m+k-1}.$$

- Par ailleurs, le lemme 2.6 implique que, pour une certaine valeur r_0 de r, $v_{m+r2^{i+1}}$ est différent de v_m, et vaut donc $1 - v_m$.

Et finalement

$$(1 - v_m, w) = \Psi(v_{m+r_0 2^{i+1}} \dots v_{m+k-1+r_0 2^{i+1}}) \in \Psi(\mathcal{F}_v^{\bar{\alpha}}(k)).$$

■

Remarque 2.8
- Il résulte de la proposition 2.7 que, si w est un facteur de v, $\bar{w}$ en est aussi un.
- Si $u_{m+1} \dots u_{m+k-1}$ et $u_{n+1} \dots u_{n+k-1}$ sont deux facteurs différents de u, on a alors bien sûr

$$v_m \dots v_{m+k-1} \neq v_n \dots v_{n+k-1},$$

mais aussi

$$v_m \dots v_{m+k-1} \neq \bar{v}_n \dots \bar{v}_{n+k-1}.$$

Ceci est d'ailleurs vrai pour tout couple (u, v) de suites à valeurs dans $\{0,1\}$ telles que $v = I(u)$.

3 Possibilités de synchronisation

La deuxième partie du résultat 2.4, c'est-à-dire l'existence de ce que nous nommons “positions canoniques” dans [3], peut-être vue comme une “synchronisation” d'une occurrence des facteurs de toutes les suites de pliage de papier. Cette synchronisation peut-elle aussi être réalisée pour les facteurs de toutes les suites de Rudin-Shapiro ?

Des résultats 2.2 et 2.4, combinés à la proposition 2.7, on déduit immédiatement la proposition suivante, qui garantit l'existence d'une “semi-synchronisation” pour les facteurs des suites de Rudin-Shapiro.

Proposition 3.1 *Soit $k \geq 1$. Il existe des ensembles $\mathcal{Q}_k^p \subset 2\mathbf{N}$ et $\mathcal{Q}_k^i \subset 2\mathbf{N}+\mathbf{1}$ tels que, pour $\alpha \in \{p, i\}$:*
$\cdot$ $\#\mathcal{Q}_k^\alpha = P^\alpha(k-1)$,
$\cdot$ *pour toute suite de Rudin-Shapiro v,*

$$\#\left(\{v_n \ldots v_{n+k-1}, n \in \mathcal{Q}_k^\alpha\} \cup \{\bar{v}_n \ldots \bar{v}_{n+k-1}, n \in \mathcal{Q}_k^\alpha\}\right) = 2P^\alpha(k-1) = p_v^{\bar{\alpha}}(k).$$

Il existe $\mathcal{Q}_k$, de cardinal $4(k-1)$, tel que, pour $k \geq 8$, et pour toute suite v de Rudin-Shapiro,

$$\#\left(\{v_n \ldots v_{n+k-1}, n \in \mathcal{Q}_k\} \cup \{\bar{v}_n \ldots \bar{v}_{n+k-1}, n \in \mathcal{Q}_k\}\right) = 8(k-1) = p_v(k).$$

Notons que ceci nous permet déjà de connaître tous les facteurs d'une suite de Rudin-Shapiro de longueur k par le seul examen de ses facteurs de longueur k commençant aux indices de $\mathcal{Q}_k$.

Nous allons maintenant démontrer que l'on ne peut *pas* obtenir mieux, ce qui reviendrait à synchroniser complètement les facteurs des suites de Rudin-Shapiro. Intuitivement, ce résultat provient du fait que, dans l'énoncé du lemme 2.6, la localisation des changements de valeur de la suite $(w_r)_{r\geq 0}$ dépend de la suite de pliage considérée.

Le résultat que nous allons prouver est en fait un peu plus général. Le seul objectif de cette généralité est de pouvoir mener commodément un raisonnement par récurrence.

Définition 3.2 *Soit g une application de $\mathbf{N}$ dans $\{0,1\}$ et $k \in \mathbf{N}$. On appelle g-synchronisation (resp. g-synchronisation paire, g-synchronisation impaire) des facteurs des suites de pliage de longueur k tout ensemble $\mathcal{P}$ inclus dans $\mathbf{N}^*$ (resp. $2\mathbf{N}^*$, $2\mathbf{N}+\mathbf{1}$), tel que :*
$\cdot$ $\#\mathcal{P} = 2P(k)$ *(resp. $2P^p(k)$, $2P^i(k)$),*
$\cdot$ $\forall u$ *suite de pliage,*

$$\#\{(u_1 + \cdots + u_{n-1} + g(n-1), u_n \ldots u_{n+k-1}), n \in \mathcal{P}\} = 2P(k),$$

(resp. $2P^p(k)$, $2P^i(k)$).

Proposition 3.3 *Soit g une application de $\mathbf{N}$ dans $\{0,1\}$.*
1. *Quel que soit l'entier $k \geq 0$, il n'existe pas de g-synchronisation impaire pour les facteurs de longueur k des suites de pliage.*
2. *Quel que soit l'entier $k \geq 7$, il n'existe pas de g-synchronisation pour les facteurs de longueur k des suites de pliage.*

Preuve. 1. Nous allons prouver ce résultat par récurrence sur k.

Pour $k = 0$, l'existence d'une g-synchronisation impaire des facteurs de longueur 0 équivaut à celle de deux entiers m et n tels que, pour toute suite de pliage u, on ait $u_1 + \cdots + u_{2m} + g(2m) \neq u_1 + \cdots + u_{2n} + g(2n)$. Soient donc m et n deux entiers, avec $m < n$, et u une suite de pliage. Si $u_{2m+1} + \cdots + u_{2n} = g(2m) + g(2n)$, alors le choix $\{m,n\}$ ne constitue évidemment pas une g-synchronisation. Sinon, soit $j \in \{2m+1,\dots,2n\}$ tel que $\Phi(j) \neq \Phi(i)$ pour tout $i \in \{2m+1,\dots,2n\}$, $i \neq j$ (cf. lemme 2.5). Soit $(u'_k)_{k\geq 1}$ la suite de pliage construite à l'aide des instructions $(f'_i)_{i\geq 0}$, où $f'_i = f_i$ si $i \neq j$ et $f'_j = 1 - f_j$. Alors $u'_{2m+1} + \cdots + u'_{2n} = g(2m) + g(2n)$, et le choix $\{m,n\}$ ne constitue pas une synchronisation.

Soit maintenant $k \geq 1$, et supposons qu'on ait trouvé une g-synchronisation impaire pour les facteurs de longueur k. Il existe un ensemble $\mathcal{P} \subset \mathbf{N}$, tel que :

$\cdot$ $\#\mathcal{P} = 2P^i(k)$,

$\cdot$ pour toute suite de pliage u,

$$\#\{(u_1 + \cdots + u_{2m} + g(2m), u_{2m+1} \dots u_{2m+k}), m \in \mathcal{P}\} = 2P^i(k).$$

Cette équation signifie notamment qu'en regardant tous les facteurs de longueur k débutant en position $2m+1$, pour $m \in \mathcal{P}$, on trouvera deux fois chaque facteur de $\mathcal{F}^i_u(k)$.

Soit $\mathcal{P}^p$ l'ensemble des éléments pairs de $\mathcal{P}$. Nous allons montrer qu'il existe une fonction h telle que $1 + \mathcal{P}^p$ soit une h-synchronisation impaire des facteurs de longueur $\lfloor k/2 \rfloor$. Soit $(u'_n)_{n\geq 1}$ la suite de pliage construite à partir des instructions $(f_i)_{i\geq 1}$. Alors, pour tout entier m,

$$u_{2m+1} \dots u_{2m+k} = (f_0 + m)u'_{m+1}(\overline{f_0 + m})u'_{m+2} \dots,$$

ce que l'on note :

$$u_{2m+1} \dots u_{2m+k} = (f_0 + m)(\overline{f_0 + m})(f_0 + m)(\overline{f_0 + m}) \dots \ \odot u'_{m+1} \dots u'_{m+\lfloor \frac{k}{2} \rfloor}. \quad (1)$$

En particulier, si m et m' sont de parités différentes,

$$u_{2m+1} \dots u_{2m+k} \neq u_{2m'+1} \dots u_{2m'+k}.$$

Par conséquent, en regardant tous les facteurs de longueur k débutant en position $2m+1$, pour $m \in \mathcal{P}^p$, on trouvera deux fois chaque facteur de longueur k commençant en position $4n+1$, pour une certaine valeur de n. D'après l'identité (1), les facteurs distincts ainsi obtenus sont en bijection avec les facteurs de longueur $\lfloor k/2 \rfloor$ de u' commençant en position impaire. Ainsi, leur nombre est $P^i(\lfloor k/2 \rfloor)$. Donc :

$$\#\{(u_1 + \cdots + u_{2m} + g(2m), u_{2m+1} \dots u_{2m+k}), m \in \mathcal{P}^p\} = 2P^i(\lfloor k/2 \rfloor).$$

Soit maintenant m un élément de $\mathcal{P}^p$:

$\cdot$ $u_{2m+1} \dots u_{2m+k} = f_0\bar{f}_0 f_0 \bar{f}_0 \dots \ \odot u'_{m+1} \dots u'_{m+\lfloor \frac{k}{2} \rfloor}$,

$$\begin{aligned} \cdot\ u_1 + \cdots + u_{2m} + g(2m) &= u_1 + u_3 + \cdots + u_{2m-1} + u'_1 + \cdots + u'_m + g(2m) \\ &= \tfrac{m}{2}(f_0 + \bar{f}_0) + u'_1 + \cdots + u'_m + g(2m) \\ &= u'_1 + \cdots + u'_m + g(2m) + \tfrac{m}{2}. \end{aligned}$$

Nous avons donc, pour toute suite de pliage u',

$$\#\left\{\left(u'_1 + \cdots + u'_m + g(2m) + m/2, u'_{m+1} \dots u'_{m+\lfloor k/2 \rfloor}\right), m \in \mathcal{P}^p\right\} = 2P^i(\lfloor k/2 \rfloor),$$

ce qui signifie que $1+\mathcal{P}^p$ est une h-synchronisation impaire des facteurs de longueur $\lfloor k/2 \rfloor$, avec $h(m) = g(2m) + m/2$.

2. Soit $k \geq 7$, et $\mathcal{P}$ une g-synchronisation des facteurs de longueur k. Soit u une suite de pliage. D'après le résultat 2.3, $\mathcal{F}_u^p(k) \cap \mathcal{F}_u^i(k) = \emptyset$, et l'ensemble des éléments impairs de $\mathcal{P}$ forme une g-synchronisation impaire pour k, ce qui est impossible d'après le premier point.

■

Remarque 3.4 On peut montrer de façon analogue que, quels que soient la fonction g et l'entier k, il n'existe pas de g-synchronisation paire pour k. Le corps de la récurrence consiste à construire une h-synchronisation paire pour $\lfloor (k+1)/2 \rfloor$ à partir de la g-synchronisation paire d'origine. Il reste alors à traiter "à la main" les cas $k = 0$ et $k = 1$ pour initialiser la récurrence.

Corollaire 3.5 1. *Soit $k \geq 1$. Il n'existe pas d'ensemble $\mathcal{P} \subset 2\mathbf{N} + \mathbf{1}$ tel que l'on ait simultanément :*
- $\#\mathcal{P} = 2P^i(k-1)$,
- *$\forall v$ suite de Rudin-Shapiro,* $\#\{v_n \ldots v_{n+k-1}, n \in \mathcal{P}\} = 2P^i(k-1) = p_v^p(k)$.

2. *Soit $k \geq 8$. Il n'existe pas d'ensemble $\mathcal{P} \subset \mathbf{N}$ tel que l'on ait simultanément :*
- $\#\mathcal{P} = 2P(k-1)$,
- *$\forall v$ suite de Rudin-Shapiro,* $\#\{v_n \ldots v_{n+k-1}, n \in \mathcal{P}\} = 2P(k-1) = p_v(k)$.

Preuve. 1. La proposition 2.7 ramène l'existence de l'ensemble $\mathcal{P}$ à celle d'une g-synchronisation paire des facteurs de longueur $k-1$ des suites de pliage, où g est la fonction nulle.
2. La proposition 2.7 ramène l'existence de l'ensemble $\mathcal{P}$ à celle d'une g-synchronisation des facteurs de longueur $k-1$ des suites de pliage, où g est la fonction nulle.

Remarque 3.6
1. Il est vraisemblable que ces résultats s'étendent lorsque $k \leq 7$, c'est-à-dire qu'il est possible de "semi-synchroniser" les facteurs, mais pas de les synchroniser complètement. Toutefois, ces résultats doivent être démontrés cas par cas, ce qui constitue un travail un peu fastidieux ...
2. Pourquoi introduire la notion de g-synchronisation? Nous l'avons dit, synchroniser tous les facteurs des suites de Rudin-Shapiro équivaudrait à trouver une g-synchronisation des facteurs de pliage pour $g = 0$. L'idée de procéder par récurrence pour démontrer l'impossibilité d'une telle g-synchronisation est naturelle. En suivant le schéma de la preuve de la proposition 3.3, on est alors naturellement amené à étudier les g-synchronisations pour la fonction définie par $g(m) = \lfloor m/2 \rfloor$. Ceci prouve que l'on ne peut pas se contenter d'examiner le cas $g = 0$.

Ainsi, c'est l'impossibilité de synchroniser les changements de valeur de la suite $(w_r)_{r \geq 0}$ du lemme 2.6 qui empêche de synchroniser complètement les facteurs des suites de Rudin-Shapiro. Toutefois, ce lemme affirme que ces changements de valeur ont lieu avec une certaine régularité, qui, elle, est *indépendante* de la suite considérée, et ce résultat va nous permettre de donner une majoration uniforme des fonctions de récurrence des suites de Rudin-Shapiro.

4 Fonction de récurrence

Rappelons qu'une suite u est dite *minimale* si, pour tout entier k, il existe un entier r tel que tout facteur de u de longueur r contienne comme sous-facteurs tous les facteurs de u de longueur k. La plus petite valeur de r vérifiant cette propriété est alors notée $R_u(k)$, et la fonction R_u est appelée *fonction de récurrence* de la suite u.

Nous démontrons ici que toutes les suites de Rudin-Shapiro sont minimales, et donnons une majoration linéaire de leur fonction de récurrence, indépendante de la suite étudiée.

Proposition 4.1 *Soit v une suite de Rudin-Shapiro. Alors v est minimale, et l'on a les inégalités suivantes :*

$$R_v(1) \leq 10 \quad et \quad \forall k \geq 2,\ R_v(k) \leq 172.2^{H(k-1)} - 2^{H(k)} + k - 3,$$

où, pour $k \geq 1$, $H(k)$ est l'entier (≥ -1) caractérisé par $2^{H(k)} < k \leq 2^{H(k)+1}$.
En particulier, pour toute suite de Rudin-Shapiro v et tout entier $k \geq 1$:

$$R_v(k+1) < 172k.$$

Preuve. Soit v une suite de Rudin-Shapiro, et u la suite de pliage associée. Soit $k \geq 1$. Si $k = 1$, le lemme 2.6 , appliqué à $m = 0$ et $i = 0$ permet d'affirmer que tout facteur de v de longueur 10 contient les lettres 0 et 1. Supposons maintenant que $k \geq 2$. D'après la proposition 2.7, dire que tout facteur de v de longueur m contient tous les facteurs de v de longueur k équivaut à l'identité suivante :

$$\forall n \geq 0, \#\{(u_1 + \cdots + u_i, u_{i+1} \ldots u_{i+k-1}), n \leq i \leq n + m - k\} = 2P(k-1).$$

Notons a l'entier $1 + 2^{H(k)}$, et b l'entier $12.2^{H(k-1)} + k - 2$. D'après les résultats de [3] rappelés à la suite du résultat 2.4, le mot $u_a u_{a+1} \ldots u_b$ contient tous les facteurs de u de longueur $k - 1$.
Soit i un entier tel que $b < 2^i$. Pour $r \geq 0$, considérons l'intervalle d'entiers $I_r = [a + r2^{i+1}, b + r2^{i+1}]$. Soit w un facteur de u de longueur $k-1$, et $n \leq b-k+2$ un entier tel que $w = u_n \ldots u_{n+k-2}$.

· Pour $r \geq 0$, notons w_r le mot $u_{n+r2^{i+1}} \ldots u_{n+k-2+r2^{i+1}}$. D'après le lemme 2.5, pour tout $r \geq 0$, $w = w_r$.

· D'après le lemme 2.6, pour cinq valeurs consécutives de r, la "somme préfixe" $u_1 + \cdots + u_{n-1+r2^{i+1}}$ prendra les valeurs 0 et 1.

Ainsi, tout intervalle d'entiers contenant cinq des intervalles I_r contiendra tous les facteurs de u de longueur $k - 1$, chacun d'eux apparaissant avec les deux "sommes préfixes" possibles. Or, tout intervalle de longueur supérieure ou égale à $172.2^{H(k-1)} - 2^{H(k)} + k - 3$ contient cinq des intervalles I_r, ce qui mène finalement au résultat annoncé, puis à la majoration finale grâce à la croissance de H.

■

5 Puissances dans les suites de pliage et dans les suites de Rudin-Shapiro généralisées

Dans [3] nous avons montré que le langage de tous les facteurs des suites de pliage est non-algébrique (voir aussi [14]), en étudiant les puissances dans ces suites. Pour ce faire nous avons utilisé un résultat évoqué dans [1] qui généralise un résultat de [17] : *une suite de pliage ne contient aucun carré ww tel que la longueur de w soit supérieure ou égale à* 6.

Nous nous proposons ici de raffiner ce résultat en étudiant les facteurs des suites de pliage du type wxw où w est un mot non vide et x une lettre, puis d'en déduire l'inexistence de carrés trop longs pour les suites de pliage de papier, (par une preuve autonome légèrement différente de celle évoquée dans [1]), et pour les suites de Rudin-Shapiro. Ceci nous permettra de montrer que les premières n'admettent pas de puissance quatrième et que les secondes n'admettent pas de puissance cinquième. Dans le paragraphe suivant, nous appliquerons ces résultats à l'étude de la non-algébricité du langage de tous les facteurs de Rudin-Shapiro.

Proposition 5.1 *Si une suite de pliage contient un facteur wxw, où w est un mot non vide dans $\{0,1\}^*$ et x une lettre dans $\{0,1\}$, alors la longueur de w, notée $|w|$, vérifie : $|w| \in \{1,2,3,4,7\}$.*

Preuve. La démonstration se fait en plusieurs étapes.
1. *Une suite de pliage ne peut contenir un facteur wxw avec $|w|$ pair et $|w| \geq 8$.*
En effet on aurait alors deux occurrences de w qui commenceraient à des indices de parités différentes ce qui contredirait le résultat 2.3.
2. *Une suite de pliage ne peut contenir un facteur wxw avec $|w| = 4k+1$ et $k \geq 1$.*
Supposons en effet que l'on ait

$$u_n u_{n+1} \cdots u_{n+|w|-1} u_{n+|w|} u_{n+|w|+1} \cdots u_{n+2|w|} = wxw,$$

et soit $j \in \{n, n+1\}$ tel que j soit impair. Alors

$$u_j u_{j+2} u_{j+4} u_{j+6} \cdots = a\bar{a}a\bar{a} \cdots,$$

pour un a dans $\{0,1\}$. Mais $u_{j+|w|+1}$ devrait être égal à la fois à $\bar{a}$ (puisque $j+2$ et $j+|w|+1$ sont congrus modulo 4) et à a (puisque $u_{j+|w|+1} = u_j$), ce qui est impossible.
3. *Si wxw est un facteur d'une suite de pliage avec $|w| = 4k+3$, et $k \geq 2$, alors $k = 2j+1$, (donc $j \geq 1$ et $|w| = 8j+7$).*
Si wxw commence à un indice pair dans la suite de pliage, le sous-mot formé des lettres en position paire de cette occurrence de wxw est un facteur de la suite de pliage $(u_{2n})_{n\geq 1}$, de la forme zz avec $|z| = 2k+2$.
Si wxw commence à un indice impair dans la suite de pliage, le sous-mot formé des lettres en position paire de cette occurrence de wxw est un facteur de la suite de pliage $(u_{2n})_{n\geq 1}$, de la forme txt avec $|t| = 2k+1$.
Dans les deux cas on obtient donc nécessairement un facteur d'une suite de pliage, de la forme mym, avec $|m| = 2k+1$ et $y \in \{0,1\}$. Comme $k \geq 2$ par hypothèse, k ne peut être pair, d'après le 2, donc k est impair (et donc ≥ 3).
4. *Une suite de pliage ne peut contenir de facteur zz avec $|z| = 2^k$, $k \geq 1$.*
Ceci se montre par récurrence sur k :

- pour $k = 1$, soit $abcd$ un facteur de longueur 4 d'une suite de pliage. Si ce facteur commence à une position impaire, on a $c = \bar{a}$, s'il commence à une position paire, alors $d = \bar{b}$, et donc on ne peut jamais avoir $ab = cd$.

- si une suite de pliage admet un facteur zz avec $|z| = 2^{k+1}$, $k \geq 1$, le sous-mot d'une occurrence de zz formé des lettres apparaissant aux indices pairs est de la forme tt, avec $|t| = 2^k$, et c'est un facteur de la suite de pliage $(u_{2n})_{n\geq 1}$.

5. *Une suite de pliage ne peut avoir de facteur de la forme wxw avec $|w| = 8j + 7$, $j \geq 1$.*
Ceci se montre par récurrence sur j :

- Établissons d'abord le résultat pour $j = 1$, c'est-à-dire $|w| = 15$. On a donc, pour une certaine suite de pliage u,

$$u_n u_{n+1} \cdots u_{n+14} u_{n+15} u_{n+16} \cdots u_{n+30} = wxw.$$

On peut affirmer que $u_{n+31} = \bar{x}$, sinon

$$u_n u_{n+1} \cdots u_{n+14} u_{n+15} u_{n+16} \cdots u_{n+30} u_{n+31} = wxwx,$$

et ceci contredirait le 4.
On remarque que n est nécessairement impair : dans le cas contraire le mot $u_n u_{n+2} \cdots u_{n+30}$ serait un facteur de la suite de pliage $(u_{2n})_{n\geq 1}$, de la forme zz, avec $|z| = 8$, ce qui contredirait le 4.
On a $n \equiv 1 \bmod 4$. Sinon $n + 3 \equiv 2 \bmod 4$ et

$$u_{n+3} u_{n+7} u_{n+11} u_{n+15} u_{n+19} u_{n+23} u_{n+27} u_{n+31} = (b\bar{b})^4$$

pour un $b \in \{0, 1\}$, ce qui contredit le fait que $u_{n+15} = x$ et $u_{n+31} = \bar{x}$.
Le mot

$$u_{n+3} u_{n+7} u_{n+11} u_{n+15} u_{n+19} u_{n+23} u_{n+27} u_{n+31}$$

est donc un facteur de la suite de pliage $(u_{4n})_{n\geq 1}$. Or ce mot s'écrit

$$u_{n+3} u_{n+11} u_{n+19} u_{n+27} \odot u_{n+7} u_{n+15} u_{n+23} u_{n+31}.$$

Par conséquent l'un des mots écrits de part et d'autre du signe $\odot$ est de la forme $c\bar{c}c\bar{c}$ pour un $c \in \{0, 1\}$, et l'autre est un facteur de pliage.
Comme $u_{n+15} = x$ et $u_{n+31} = \bar{x}$, c'est donc que le facteur

$$u_{n+7} u_{n+15} u_{n+23} u_{n+31}$$

est un facteur de pliage (de la suite $(u_{8n})_{n\geq 1}$ donc), mais il est de la forme zz, avec $|z| = 2$, ce qui contredit 4.

- Supposons maintenant que l'on ait un facteur d'une suite de pliage de la forme wxw avec $|w| = 8j + 7$, $j \geq 2$. En ne conservant que les lettres de cette occurrence de wxw qui apparaissent aux indices pairs, et suivant que cette occurrence de wxw commence à un indice pair ou impair, on déduit finalement, de la même façon qu'au 3, l'existence d'un facteur de la forme mym avec $|m| = 4j + 3$, dans la suite de pliage $(u_{2n})_{n\geq 1}$. Comme $j \geq 2$, on applique le 3 : j est impair, $j = 2i + 1$, avec $i \geq 1$, et on a donc un facteur de pliage mym, avec $|m| = 8i + 7$ et $1 \leq i \leq j - 1$, ce qui permet de faire fonctionner la récurrence.

6. *Une suite de pliage ne peut avoir de facteur de la forme wxw avec $|w| = 6$.*
Supposons en effet que pour une suite de pliage u on ait :

$$(u_n u_{n+1} u_{n+2} u_{n+3} u_{n+4} u_{n+5}) u_{n+6} (u_{n+7} u_{n+8} u_{n+9} u_{n+10} u_{n+11} u_{n+12}) = wxw.$$

Si n est impair, alors

$$u_n u_{n+2} u_{n+4} u_{n+6} u_{n+8} u_{n+10} u_{n+12} = a\bar{a}a\bar{a}a\bar{a}a$$

pour un a dans $\{0, 1\}$, d'où nécessairement :

$$wxw = (aa\bar{a}\bar{a}aa)\bar{a}(aa\bar{a}\bar{a}aa).$$

Le sous-mot des lettres en position paire est un facteur de pliage (de la suite $(u_{2n})_{n \geq 1}$). Or il vaut $a\bar{a}aa\bar{a}a = aa\bar{a} \odot \bar{a}aa$, ce qui est impossible car aucun des deux mots de part et d'autre du symbole $\odot$ ne respecte l'alternance des lettres en position impaire d'un facteur de pliage.
De même, si n est pair, alors

$$u_{n+1} u_{n+3} u_{n+5} u_{n+7} u_{n+9} u_{n+11} = c\bar{c}c\bar{c}c\bar{c}c$$

pour un c dans $\{0, 1\}$, d'où nécessairement :

$$wxw = (\bar{c}cc\bar{c}\bar{c}c)x(\bar{c}cc\bar{c}\bar{c}c).$$

Le sous-mot des lettres en position paire est un facteur de pliage (de la suite $(u_{2n})_{n \geq 1}$). Or il vaut $\bar{c}c\bar{c}xc\bar{c}c = \bar{c}\bar{c}cc \odot cx\bar{c}$, ce qui est impossible car aucun des deux mots de part et d'autre du symbole $\odot$ ne respecte l'alternance des lettres en position impaire d'un facteur de pliage.

En regroupant tous ces résultats on arrive finalement à la conclusion annoncée.

■

Corollaire 5.2

- *Si ww est un facteur de pliage non vide, alors la longueur de w vérifie $|w| \in \{1, 3, 5\}$.*

- *Si zz est un facteur de Rudin-Shapiro non vide, alors la longueur de z vérifie $|z| \in \{1, 2, 3, 4, 5, 8\}$.*

Preuve.

- Si ww est un facteur de pliage non vide, posons $w = mx$, où m est un mot, et x une lettre. Alors mxm est facteur de pliage, donc, en appliquant la proposition 5.1, $m = \emptyset$ ou $|m| \in \{1, 2, 3, 4, 7\}$. Donc : $|w| \in \{1, 2, 3, 4, 5, 8\}$. Mais les cas $|w| = 2$, $|w| = 4$ et $|w| = 8$ sont exclus d'après le point 4 de la preuve de la proposition 5.1.

- Soit zz un facteur non vide d'une suite de Rudin-Shapiro généralisée v, et soit u la suite de pliage telle que $v = I(u)$. En d'autres termes,

$$\forall n \geq 1, \quad u_n = v_n + v_{n-1} \bmod 2.$$

  Soit k la longueur de z. Il existe un entier $n \geq 0$ tel que :

$$v_n v_{n+1} \cdots v_{n+k-1} v_{n+k} \cdots v_{n+2k-1} = zz.$$

  On a alors :

$$\forall j \in [1, k-1], \quad u_{n+j} = v_{n+j} + v_{n+j-1} = v_{n+j+k} + v_{n+j-1+k} = u_{n+j+k}.$$

  La suite u admet donc le facteur

$$(u_{n+1} \cdots u_{n+k-1}) u_{n+k} (u_{n+k+1} \cdots u_{n+2k-1}) = wxw,$$

  avec $|w| = k-1$. Donc, d'après la proposition 5.1, $k \in \{2, 3, 4, 5, 8\}$ ou $w = \emptyset$.

■

Corollaire 5.3

- *Dans une suite de pliage quelconque, il n'y a pas de puissance quatrième, et les seuls cubes sont* 000 *et* 111, *(ces cubes apparaissent effectivement dans toute suite de pliage).*
- *Dans une suite de Rudin-Shapiro généralisée, il n'y a pas de puissance cinquième, et les seules puissances quatrièmes sont* 0000 *et* 1111, *(ces puissances quatrièmes apparaissent effectivement dans toute suite de Rudin-Shapiro généralisée).*

Preuve.

- Soit en effet w^k un facteur de pliage tel que $w \neq \emptyset$, avec $k \geq 4$. Alors w^4 est facteur. Comme $w^4 = (w^2)^2$, on déduit du corollaire précédent que $w^2 \in \{1, 3, 5\}$ ce qui n'est pas possible parce que $|w^2|$ est un nombre pair.
  Supposons que w^3 soit un facteur de pliage avec $w \neq \emptyset$. Alors w^2 est aussi facteur, donc, d'après le corollaire précédent, $|w| \in \{1, 3, 5\}$.
  Si $|w| = 5$, alors suivant que w^3 apparaît à une position paire ou impaire, il est de la forme :

$$(*\, a\, *\, \bar{a}\, *)(a\, *\, \bar{a}\, *\, a)(*\, \bar{a}\, *\, a\, *),$$

  pour un a dans $\{0, 1\}$, ou

$$(a\, *\, \bar{a}\, *\, a)(*\, \bar{a}\, *\, a\, *)(\bar{a}\, *\, a\, *\, \bar{a}),$$

  pour un a dans $\{0, 1\}$, ce qui est impossible (comparer le premier et le troisième facteur dans chaque cas).
  Si $|w| = 3$, alors suivant que w^3 apparaît à une position paire ou impaire, il est de la forme :

$$(*\, a\, *)(\bar{a}\, *\, a)(*\, \bar{a}\, *),$$

ou

$$(a \, * \, \bar{a})(* \, a \, *)(\bar{a} \, * \, a),$$

ce qui est impossible pour la même raison.
Donc $|w| = 1$. Pour montrer que les deux facteurs 000 et 111 apparaissent effectivement dans toute suite de pliage, on remarque d'abord que la construction de ces suites par symétrie perturbée permet de se limiter à montrer l'apparition de l'un seulement de ces deux facteurs dans toute suite de pliage. Puis, si l'on appelle a, b, c, d les quatre premières instructions de pliage de la suite considérée, cette suite commence par le facteur $a\ b\ \overline{a}\ c\ a\ \overline{b}\ \overline{a}\ d\ a\ b\ \overline{a}\ \overline{c}$. Si $a = b$, on regarde les deux facteurs $\overline{b}\ \overline{a}\ d$ et $d\ a\ b$. Si $a = \overline{b}$, on regarde les deux facteurs $b\ \overline{a}\ c$ et $b\ \overline{a}\ \overline{c}$.

- Cherchons maintenant les facteurs des suites de Rudin-Shapiro de la forme z^4. Comme $z^4 = (z^2)^2$, alors, d'après le corollaire précédent, $|z^2|$ appartient à $\{1, 2, 3, 4, 5, 8\}$, et donc nécessairement $|z| \in \{1, 2, 4\}$.
  Si $|z| = 4$, alors on a en particulier un facteur de la forme

  $$(abcd)(abcd)(a \cdots,$$

  dans une suite de Rudin-Shapiro v, d'où, dans la suite de pliage u telle que $I(u) = v$ un facteur :

  $$b + a \quad c + b \quad d + c \quad a + d \quad b + a \quad c + b \quad d + c \quad a + d,$$

  ce qui contredit le point 4 de la preuve de la proposition 5.1.
  Si $|z| = 2$, on a un facteur de Rudin-Shapiro de la forme $xyxyxyxy$, d'où un facteur de pliage de la forme

  $$y + x \quad x + y \quad y + x \quad x + y \cdots,$$

  ce qui contredit aussi le point 4 de la preuve de la proposition 5.1.
  Donc les seules puissances quatrièmes possibles sont 0000 et 1111, et ces deux facteurs apparaissent effectivement dans toutes les suites de Rudin-Shapiro : il suffit en effet de remarquer qu'un facteur 000 dans une suite de pliage donne un facteur $aaaa$ dans la suite de Rudin-Shapiro associée et d'utiliser la première partie de la remarque qui suit la proposition 2.7.
  Enfin une puissance cinquième dans une suite de Rudin-Shapiro serait nécessairement le "prolongement" d'une puissance quatrième, donc soit 00000, soit 11111. Mais chacun de ces deux facteurs donnerait un facteur de pliage 0000 ce qui est exclu.

■

Remarque 5.4
- Les valeurs 5 pour les suites de pliage et 8 pour les suites de Rudin-Shapiro sont atteintes. Par exemple la suite de pliage régulier où toutes les instructions de pliage sont égales à zéro commence par

$$001(00110)(00110)110001001 \cdots,$$

et la suite de Rudin-Shapiro généralisée qui s'en déduit commence par :

$$0001(11011110)(11011110)\cdots.$$

- On peut rechercher plus précisément tous les carrés des suites de pliage et ceux des suites de Rudin-Shapiro, ainsi que les chevauchements dans ces deux familles de suites, puis les cubes des suites de Rudin-Shapiro, ce qui donnera finalement avec le corollaire 5.3 toutes les puissances et chevauchements dans ces deux familles de suites. Voici le résultat :

- *les carrés qui apparaissent dans au moins une suite de pliage sont les mots zz où z est dans l'ensemble*

$$\{0, 1, 011, 001, 110, 100, 00110, 11001, 10011, 01100\};$$

- *les chevauchements qui apparaissent dans au moins une suite de pliage sont les mots de l'ensemble*

$$\{000, 111, 0110110, 1001001\};$$

- *les carrés qui apparaissent dans au moins une suite de Rudin-Shapiro sont les mots zz où z est l'un des mots* 0, 1, 00, 11, 01, 10, 001, 110, 010, 101, 011, 100, 0001, 1110, 0010, 1101, 0100, 1011, 0111, 1000, 00010, 11101, 00100, 11011, 01000, 10111, 00011101, 11100010, 00010010, 11101101, 00100001, 11011110, 00101110, 11010001, 01001000, 10110111, 01000111, 10111000, 01110100, 10001011, 01111011, 10000100*;*

- *les chevauchements qui apparaissent dans au moins une suite de Rudin-Shapiro sont les mots* 000, 111, 0010010, 1101101, 0100100, 1011011, 00010000100, 11101111011, 00100001000, 11011110111*;*

- *les cubes qui apparaissent dans au moins une suite de Rudin-Shapiro sont les mots* 000 *et* 111*.*

La démonstration de ce résultat sera épargnée au lecteur. Elle utilise la proposition 5.1, le lien entre les facteurs de pliage et ceux des suites de Rudin-Shapiro déjà évoqué à maintes reprises, le fait que tout facteur de pliage s'obtient en intercalant dans un facteur de pliage l'un des mots $0101\cdots$ ou $1010\cdots$, la remarque que si w est un facteur de pliage, alors $\overline{w}$ aussi, la remarque analogue pour les facteurs de Rudin-Shapiro, et ... la construction effective de préfixes de suites de pliage et de suites de Rudin-Shapiro pour exhiber effectivement les facteurs annoncés ci-dessus. L'étape initiale (et décisive) est la recherche des facteurs de pliage de la forme wxw avec $w \neq \emptyset$ (il y en a exactement 40).

6 Le langage des facteurs de Rudin-Shapiro : fonction génératrice et non-algébricité

Notons $\mathcal{L}$ le langage de tous les facteurs de toutes les suites de Rudin-Shapiro généralisées étudiées ici. Quelles sont les propriétés de ce langage ? Rappelons qu'une question analogue pour les suites sturmiennes est traitée dans [11], [16] et [7], et que le cas des suites de pliage de papier a été étudié dans [3] (voir aussi [14] pour la non-algébricité).

Nous nous proposons ici de donner trois résultats semblables à ceux prouvés dans [3] pour les facteurs de toutes les suites de pliage.

Théorème 6.1 *Le langage $\mathcal{L}$ de tous les facteurs de toute les suites de Rudin-Shapiro généralisées étudiées ici est non-algébrique.*

Théorème 6.2 *Le nombre de facteurs de longueur n qui apparaissent dans une des suites de Rudin-Shapiro, soit $h'(n)$, vérifie :*
$\forall k \geq 4 \;\; h'(2k) = 2h'(k) + 2h'(k+1)$,
$\forall k \geq 4 \;\; h'(2k+1) = 4h'(k+1)$.
On a les valeurs suivantes :
$h'(1) = 2, \; h'(2) = 4, \; h'(3) = 8, \; h'(4) = 16$,
et, pour $k \geq 5$:

$$\begin{array}{lll} h'(k) = 2^{m+2}k - 2^{m+2} - 10.4^{m-1} & \textit{si} & k \in [2^m + 1, \frac{3}{2}.2^m + 1], \\ h'(k) = 3.2^{m+1}k - 3.2^{m+1} - 22.4^{m-1} & \textit{si} & k \in [\frac{3}{2}.2^m + 1, \frac{7}{4}.2^m + 1], \\ h'(k) = 2^{m+2}k - 2^{m+2} - 2.4^{m} & \textit{si} & k \in [\frac{7}{4}.2^m + 1, 2^{m+1}]. \end{array}$$

Théorème 6.3 *La série génératrice du langage $\mathcal{L}$ est transcendante.*

Preuves. Le théorème 6.1 se démontre comme dans [3] : le lemme de l'étoile pour les langages algébriques (voir [12] par exemple) implique facilement que dans tout langage algébrique infini il y a des mots qui admettent comme facteurs des puissances arbitrairement grandes. Ce n'est pas le cas pour le langage $\mathcal{L}$ d'après le corollaire 5.3 ci-dessus.

$\blacksquare$

Le théorème 6.2 se déduit immédiatement du théorème d'énumération de tous les facteurs de longueur donnée de toutes les suites de pliage de papier, ([3]). On remarque en effet que l'application Ψ de la proposition 2.7 induit une bijection

$$\Psi : \bigcup_{v \text{ Rudin-Shapiro}} \mathcal{F}_v(k) \to \{0,1\} \times \bigcup_{u \text{ pliage}} \mathcal{F}_u(k-1),$$

Donc $h'(k) = 2h(k-1)$, où $h(k)$ est le nombre de tous les facteurs de pliage de longueur k.

$\blacksquare$

Quant au théorème 6.3, il résulte facilement du théorème donné dans [3] qui stipule que la série $\sum h(n)X^n$ est transcendante.

$\blacksquare$

Remarque 6.4

- On peut montrer facilement comme dans [3] que $\frac{h'(k)}{k^2}$ est compris entre deux constantes strictement positives.
- Il est presque immédiat que la suite $(h'(k))_{k\geq 1}$ est 2-régulière au sens de [5].
- Comme dans [14], le théorème 6.1 montre que tout sous-langage infini de $\mathcal{L}$ est non-algébrique.
- Comme la série génératrice de $\mathcal{L}$ est transcendante, il en est de même de celle du langage complémentaire $\{0,1\}^* \setminus \mathcal{L}$. Nous ne savons pas si ce langage est algébrique (comme pour le complémentaire du langage des mots de Sturm, voir [11]), mais s'il est algébrique, il est donc nécessairement ambigu.

Références

[1] J.-P. Allouche, Suites infinies à répétitions bornées, *Séminaire de Théorie des Nombres de Bordeaux* (1983–1984), Exposé 20, 20-01–20-11.

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

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

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

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

[6] J.-P. Allouche et J. Shallit, Suites à complexité ultimement affine, *Congrès "Thémate"*, Luminy (1991), à paraître.

[7] J. Berstel et M. Pocchiola, A geometric proof of the enumeration formula for Sturmian words, *Intern. J. Alg. Comput.* **3** (1993) 349–355.

[8] A. Blanchard et M. Mendès France, Symétrie et transcendance, *Bull. Sci. Math.* **106** (1982) 325–335.

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

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

[11] S. Dulucq et D. Gouyou-Beauchamp, Sur les facteurs des suites de Sturm, *Theoret. Comp. Sci.* **71** (1990) 381–400.

[12] J.E. Hopcroft et J.D. Ullman, *Introduction to automata theory, languages and computation*, Addison-Wesley (1979).

[13] J.-P. Kahane, *Some random series of functions*, Heath Mathematical Monographs (1968).

[14] S. Lehr, A result about languages concerning paperfolding sequences, *Math. Systems Theory* **25** (1992) 309–313.

[15] 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.

[16] F. Mignosi, Sturmian words and ambiguous context-free languages, *Internat. J. Found. Comp. Sci.* **1** (1990) 309–323.

[17] H. Prodinger et F.J. Urbanek, Infinite $0-1$ sequences without long adjacent identical blocks, *Discrete Math.* **28** (1979) 277–289.

[18] M. Queffélec, Une nouvelle propriété des suites de Rudin-Shapiro, *Ann. Inst. Fourier, Grenoble* **37** (1987) 115–138.

[19] D. Rider, Transformations of Fourier coefficients, *Pacific J. Math.* **19** (1966) 347.

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

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

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

Mireille Bousquet-Mélou
C.N.R.S., U.R.A. 1304
LaBRI
351, cours de la Libération
F-33405 Talence Cedex
France