Chapitre 7

Une autre application des morphismes circulaires : calculs de complexités de suites

Dans ce chapitre, nous étudions un problème assez différent de celui de l'évitabilité des motifs, mais auquel la notion de morphisme circulaire utilisée au chapitre précédent apporte un éclairage nouveau.

7.1 Introduction

Soit $\mathbf{u}$ un suite infinie sur un alphabet (fini) Σ. On appelle *complexité* de $\mathbf{u}$ la fonction $p\colon \mathbb{N} \to \mathbb{N}$ où, pour tout $n \geqslant 0$, $p(n)$ est le nombre de facteurs de $\mathbf{u}$ dont la longueur est n.

J.-P. Allouche et J. Berstel ont posé le problème suivant : pour quels entiers α et β existe-t-il des suites dont la complexité est la fonction affine $p(n) = \alpha n + \beta$? Si on veut que ce soit vrai pour tout $n \geqslant 0$, la seule possibilité est $\beta = 1$ et $\alpha = \#\,\mathrm{Alph}\,(\mathbf{u}) - 1$ (sur un alphabet binaire, cela correspond aux suites sturmiennes, de complexité $p(n) = n + 1$). On se contente donc de fonctions de complexité ultimement affines, c'est-à-dire qu'il existe un rang n_0 tel que pour tout $n \geqslant n_0$ on a $p(n) = \alpha n + \beta$.

Allouche a développé plusieurs techniques pour construire, à partir d'une suite de complexité ultimement $\alpha n + \beta$, des suites de complexités différentes, par exemple $\alpha n + \beta + 1$. Cependant il reste des couples (α, β) qui ne peuvent être obtenus par ces techniques en partant des quelques suites de complexité ultimement affine déjà connues. Il est donc nécessaire de construire de nouveaux points de départ. C'est dans cet objectif que cette étude a été faite, mais celui-ci a été dépassé puisqu'on obtient directement toutes les complexités ultimement affines possibles.

7.2 Suites ou langages ?

À toute suite $\mathbf{u}$ on associe le langage de ses facteurs $F(\mathbf{u})$. C'est un langage factoriel et prolongeable à droite.

Réciproquement, soit $L \subset \Sigma^*$ un langage factoriel prolongeable à droite et *transitif*, c'est-à-dire tel que

$$\forall x, y \in L, \exists z \in L, xzy \in L \ .$$

Alors il existe une suite $\mathbf{u}$ telle que $L = F(\mathbf{u})$. Mais en fait on n'a pas besoin d'une telle suite pour étudier la complexité. Par contre, il est utile de supposer le langage L *prolongeable* :

$$\forall x \in L, \exists a, b \in \Sigma, axb \in L$$

ce qui interdit les facteurs apparaissant uniquement en début de suite.

On s'intéresse donc à la classe des langages factoriels prolongeables, qui est stable par union et par application de morphisme suivie d'une clôture factorielle (partant du langage $L \subset \Sigma^*$, on construit $F(f(L)) \subset \Sigma'^*$ pour tout morphisme $f\colon \Sigma^* \to \Sigma'^*$). Les langages résultant de nos constructions étant tous transitifs, on pourra donc les voir indifféremment comme des constructions de langages ou des constructions de suites.

À tout morphisme $f\colon \Sigma^* \to \Sigma^*$, on associe le langage

$$FL(f) = F\left(\{\ f^n(x)\ \mid n \in \mathbb{N} \text{ et } x \in \Sigma\ \}\right)$$

(c'est la clôture factorielle d'un *DOL-langage sans axiome*, c'est-à-dire un DOL-langage où toutes les lettres de l'alphabet servent d'axiome), qui est prolongeable pourvu que toutes les lettres apparaissent comme facteurs internes, c'est-à-dire

$$\forall x \in \Sigma, \exists a, b \in \Sigma, axb \in FL(f)$$

ce qui exclut les morphismes du type $(a \mapsto ab, b \mapsto bb)$. Si de plus il existe une lettre $a \in \Sigma$ telle que $f(a) \in a\Sigma^*$ et que $\bigcup_{n\in\mathbb{N}} \mathrm{Alph}\,(f^n(a)) = \Sigma$, la suite de mots $(f^n(a))$ a pour limite un mot infini $\mathbf{u} = f^\omega(a)$ dont l'ensemble des facteurs est exactement $FL(f)$.

7.3 Facteurs spéciaux et bispéciaux

7.3.1 Facteurs spéciaux

On suppose pour le moment que tous les mots sont sur l'alphabet $\Sigma = \{a, b\}$. Soit L un langage factoriel prolongeable sur Σ.

Un mot $u \in L$ est dit *spécial à droite* si les mots ua et ub sont tous deux dans L, *spécial à gauche* si au et bu sont dans L. Un facteur spécial correspond à un choix de prolongement, ce qui se voit aisément dans les *arbres de facteurs* et les *graphes de Rauzy* de L, dont nous allons détailler la construction.

L'*arbre des facteurs à droite* (resp. *à gauche*) de L est un arbre infini dont les sommets sont étiquetés par les éléments de L et les arêtes par les lettres, de sorte qu'il y a une arête d'étiquette x entre les sommets u et v si et seulement si $v = ux$ (resp. $v = xu$). On le représente conventionnellement en plaçant la racine, étiquetée par le mot vide, à gauche (resp. à droite), si bien que quand on lit de la gauche vers la droite les étiquettes du chemin entre la racine et le sommet étiqueté u on retrouve le mot u; on peut alors se contenter d'indiquer les étiquettes des arêtes, ou la dernière lettre des étiquettes de sommet (cf. figure 7.1). Dans cet arbre, chaque facteur spécial à droite (resp. à gauche) correspond à une bifurcation.

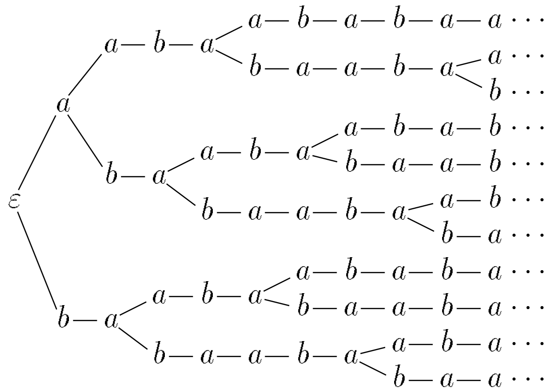

FIG. 7.1 - *Arbre à droite des facteurs de la suite de Fibonacci*

Pour tout entier n, le *graphe de Rauzy d'ordre n* de L est le graphe fini dont les sommets sont étiquetés par les mots de L de longueur n et tel qu'il y a une arête entre les sommets u et v si et seulement s'il existe deux lettres x et y telles que $ux = yv$. Un mot spécial à droite correspond à un sommet d'où partent deux arêtes, et un mot spécial à gauche à un sommet où arrivent deux arêtes (cf. figure 7.2).

Le nombre de facteurs spéciaux de longueur n à gauche ou à droite est le même : c'est $s(n) = sg(n) = sd(n) = p(n+1) - p(n)$.

7.3.2 Facteurs bispéciaux

Lorsqu'un facteur u est spécial à la fois à gauche et à droite, on dit qu'il est *bispécial*. Trois cas peuvent se présenter, selon le nombre d'éléments de $L \cap \Sigma u \Sigma$: si cet ensemble a quatre éléments (le maximum), on dit que u est *bispécial strict*. S'il n'en a que trois, u est *bispécial ordinaire*. Enfin, il est

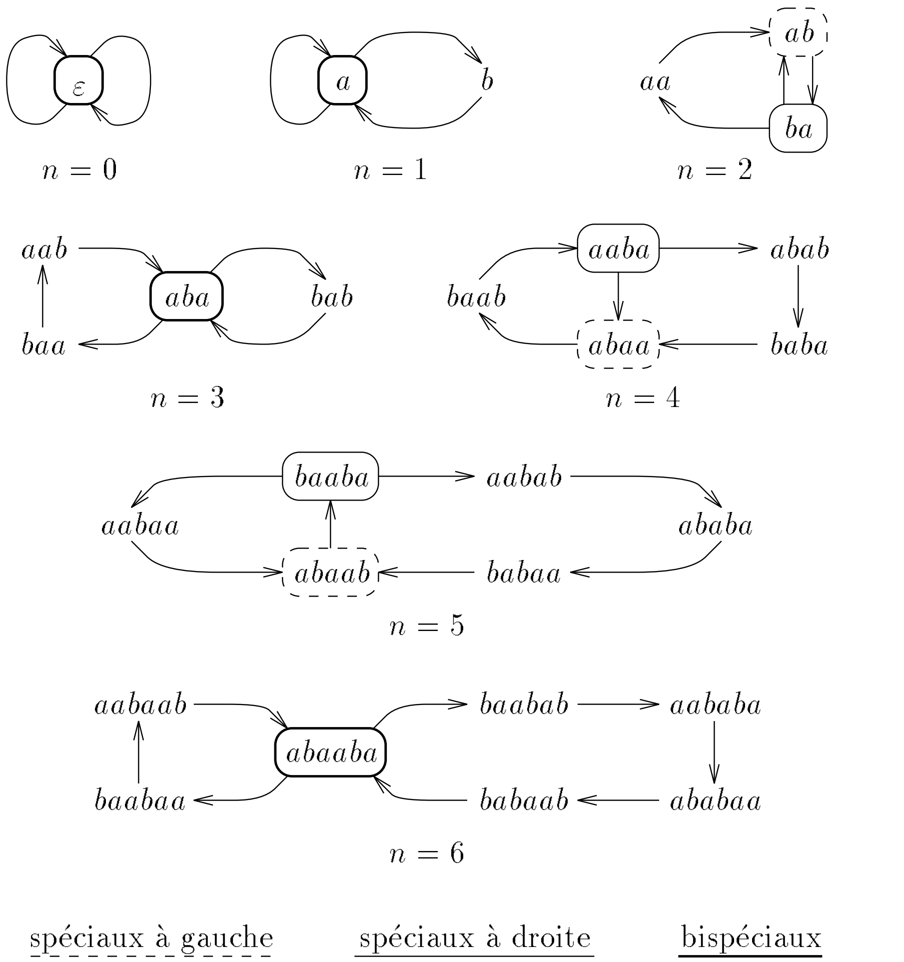

FIG. 7.2 - *Graphes de Rauzy d'ordre 0 à 6 de la suite de Fibonacci*

possible que cet ensemble ait deux éléments ($\{aua, bub\}$ ou $\{aub, bua\}$) et on dit dans ce cas que u est *bispécial faible*.

Quand on prolonge à droite un facteur spécial à gauche u, on obtient un ou deux facteurs qui peuvent être spéciaux à gauche ou non, en fonction de la nature de u :

- aucun n'est spécial à gauche si u est bispécial faible;

- un seul est spécial à gauche si u est non bispécial ou bispécial ordinaire;

- les deux sont spéciaux à gauche si u est bispécial strict.

Réciproquement, un préfixe de facteur spécial à gauche est spécial à gauche, donc on peut construire l'*arbre des facteurs spéciaux à gauche*, qui est un sous-arbre de l'arbre des facteurs à droite, dans lequel on ne garde que les sommets dont les étiquettes sont spéciales à gauche. Dans cet arbre, un bispécial strict indique un embranchement et un bispécial faible indique une fin de branche. Cet arbre est souvent plus facile à décrire que l'arbre de tous les facteurs, le cas extrême étant celui des suites sturmiennes pour lesquelles il est filiforme. On peut bien entendu aussi construire l'arbre des facteurs spéciaux à droite, qui est orienté vers la gauche. Pour bien indiquer qu'il s'agit de facteurs spéciaux, on indique à la racine les deux lettres a et b par lesquelles les facteurs spéciaux peuvent se prolonger. (figures 7.3 et 7.4).

$$\cdots\ b - a - b - a - a - b - a - a - b - a - b - a - a - b - a - \varepsilon\ \left|\ \begin{array}{l} a \\ b \end{array}\right.$$

FIG. 7.3 - *Arbre des facteurs spéciaux à droite pour la suite de Fibonacci*

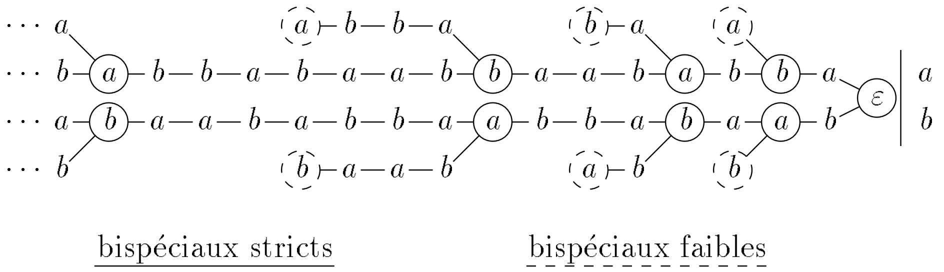

FIG. 7.4 - *Arbre des facteurs spéciaux à droite pour la suite de Thue-Morse*

Les différents types de facteurs bispéciaux interviennent également lors de l'expansion des graphes de Rauzy : pour passer du graphe des facteurs de longueur n à celui des facteurs de longueur $n+1$, on remplace chaque arête par un sommet et chaque sommet par une, deux, trois ou quatre arêtes selon que le mot qui étiquette ce sommet est non spécial, spécial non bispécial ou bispécial faible, bispécial ordinaire, bispécial strict (cf. figure 7.5). En particulier, si aucun facteur de longueur n n'est bispécial, il n'y a qu'une expansion possible.

Si on note $bs(n)$, $bo(n)$ et $bf(n)$ le nombre de facteurs bispéciaux de longueur n de chaque type, on établit aisément la relation $bs(n)-bf(n) = s(n+1)-s(n)$.

Pour calculer $p(n)$, il est donc suffisant de connaître $bs(n)$ et $bf(n)$.

On peut également utiliser des séries génératrices : à la fonction p on associe la série $P(X) = \sum_{n=0}^{\infty} p(n)X^n$, et de même $S(X)$, $BS(X)$ et $BF(X)$ pour s, bs et bf. En supposant que $s(0) = 1$, c'est-à-dire que les lettres a et b sont toutes

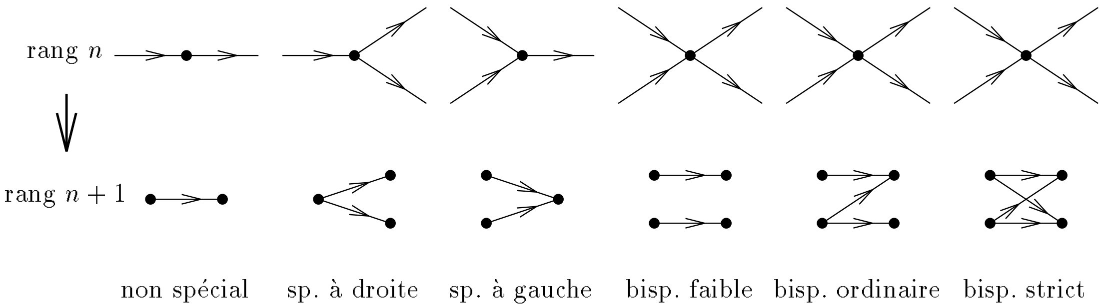

FIG. 7.5 - *Expansion des graphes de Rauzy*

deux utilisées (sinon le mot vide ne serait pas spécial), on a alors les relations suivantes :

$$P(X) = \frac{1}{1-X}\,(1 + XS(X))$$

$$S(X) = \frac{1}{1-X}\,(1 + X(BS(X) - BF(X)))$$

soit

$$S(X) = \frac{1}{(1-X)^2}\Big(1 + X^2(BS(X) - BF(X))\Big) \ .$$

7.3.3 Le cas des alphabets plus grands

Quand l'alphabet Σ a plus de deux éléments, la situation se complique quelque peu, surtout en ce qui concerne les facteurs bispéciaux. Supposons donc que Σ a k éléments, et soit comme ci-dessus L un langage factoriel prolongeable sur Σ.

Nous dirons qu'un mot $u \in L$ est *spécial à droite d'ordre* m, où m est un entier strictement positif, si le mot u se prolonge à droite de $m+1$ manières différentes, c'est-à-dire si l'ensemble $\{\ x \in \Sigma \mid ux \in L\ \}$ a exactement $m+1$ éléments. Sur l'alphabet binaire, tous les facteurs spéciaux à droite sont d'ordre 1, mais si l'alphabet est plus grand ils n'ont pas nécessairement tous le même ordre. Les facteurs spéciaux à gauche d'ordre m sont définis de la même manière.

Comme avec un alphabet binaire, les facteurs spéciaux correspondent à des embranchements dans les arbres de facteurs et les graphes de Rauzy. Le nombre de branches est égal à l'ordre du facteur spécial, augmenté d'une unité.

Pour calculer la complexité, ce n'est plus le nombre de facteurs spéciaux qui est important, car il faut les compter avec multiplicité : si $sd_m(n)$ est le nombre de facteurs spéciaux à droite de longueur n et d'ordre m, et de même $sg_m(n)$

à gauche, alors on a

$$s(n) = \sum_{m=1}^{k-1} m.sg_m(n) = \sum_{m=1}^{k-1} m.sd_m(n) = p(n+1) - p(n) \ .$$

Un facteur u spécial à droite et à gauche est encore dit *bispécial.* A tout facteur bispécial sont associés deux entiers m_d et m_g, que nous appellerons son *ordre à droite* et son *ordre à gauche*, et qui correspondent à son ordre en tant que facteur à droite et à gauche respectivement. Ces deux nombres n'ont aucune raison d'être égaux. Nous allons encore classer les facteurs bispéciaux en fonction du nombre d'éléments de l'ensemble $L \cap \Sigma u \Sigma$. Celui-ci a au plus $(m_d+1)(m_g+1)$ éléments, et si c'est le cas on dit que u est *bispécial strict*; il a au moins $\max(m_d, m_g)+1$ éléments, car chaque prolongement de u à droite ou à gauche doit correspondre à au moins un prolongement bilatéral : c'est le cas que nous avons appelé *bispécial faible* pour un alphabet binaire. Mais cette fois tous les intermédiaires sont possibles entre ces deux extrêmes : on peut quantifier cela au moyen d'un troisième entier m, l'*ordre* du facteur bispécial, défini ainsi :

$$m = \#\,(L \cap \Sigma u \Sigma) - m_d - m_g - 1 \ .$$

L'ordre m est un entier relatif, compris entre $-\min(m_d, m_g)$ et $m_d m_g$, donc entre $-(k-1)$ et $(k-1)^2$; si $k=2$, on retrouve les trois catégories de facteurs bispéciaux, qui correspondent aux trois ordres possibles, -1, 0 et 1.

Quand on prolonge à droite un facteur spécial à gauche u, la nature de u permet de déterminer, non le nombre de facteurs spéciaux à gauche obtenus, mais leur ordre total :

- si u n'est pas bispécial, on obtient un facteur spécial à gauche de même ordre que u;
- si u est bispécial d'ordre m, on obtient un certain nombre de facteurs spéciaux à gauche dont l'ordre total est l'ordre à gauche de u augmenté de l'ordre de u.

Cela donne la relation suivante :

$$b(n) = \sum_{m=-(k-1)}^{(k-1)^2} m.b_m(n) = s(n+1) - s(n)$$

où $b_m(n)$ est le nombre de facteurs bispéciaux d'ordre m et de longueur n. Si $P(X)$, $S(X)$ et $B(X)$ sont les séries génératrices associées à $p(n)$, $s(n)$ et $b(n)$, on a les formules suivantes :

$$P(X) = \frac{1}{1-X}\,(1 + XS(X))$$

$$S(X) = \frac{1}{1-X}\,(s(0) + XB(X))$$

(où $s(0) = p(1) - 1$ vaut $k - 1$ si toutes les lettres sont utilisées), soit

$$S(X) = \frac{1}{(1-X)^2}\left(1 + (p(1) - 2)X + X^2 B(X)\right) \ .$$

Pour pouvoir parler d'arbre de facteurs spéciaux, il faut un critère plus fin que l'ordre : étant donné une partie Σ' de Σ à au moins deux éléments, on construit l'*arbre des facteurs spéciaux à gauche de type* Σ' en ne conservant dans l'arbre des facteurs à droite que ceux qui se prolongent à gauche par chacune des lettres de Σ'. L'ensemble des facteurs spéciaux à gauche est donc décrit par une famille d'arbres (un pour chaque partie Σ'), mais en pratique on peut se limiter aux arbres non vides avec Σ' maximal (i.e. aucune partie plus grande ne donne le même arbre), ce qui en général donne un petit nombre d'arbres, qu'on pourra appeler *base* de la famille des arbres des facteurs spéciaux à gauche. Dans ces arbres, tous les embranchements et toutes les extrémités de branches sont étiquetés par des facteurs bispéciaux, et réciproquement tous les facteurs bispéciaux d'ordre non nul donnent un embranchement ou une extrémité sur au moins l'un des arbres de base : on peut donc déterminer les facteurs bispéciaux qui contribuent à la complexité par simple observation de la base, et il est alors facile de déterminer leur ordre.

7.4 Action d'un morphisme circulaire

Soit $f\colon \Sigma^* \to \Sigma'^*$ un morphisme circulaire de délai de synchronisation l (voir page 89), qu'on supposera de plus préfixe (si $f(u)$ est préfixe de $f(v)$, alors u est préfixe de v) pour simplifier. En fait, ce qui compte est que f est à délai de déchiffrage borné, et c'est le cas de tout morphisme circulaire.

Soit $L \subset \Sigma^*$ un langage factoriel prolongeable, et $L' = F(f(L))$ la clôture factorielle de son image par f. Nous allons montrer qu'il est possible de construire l'arbre (ou les arbres) des facteurs spéciaux à droite de L' en fonction de celui (ceux) de L.

Soit $u \in \Sigma'^*$ spécial à droite dans L' de longueur supérieure ou égale à l. Le mot u peut s'écrire $u = u_1|f(u_2)|u_3$ (rappelons que les barres de mesure indiquent les points de synchronisation) où u_1 et u_3 sont des mots de L' de longueur au plus l, et u_2 est un mot de L. En fait, puisqu'on a supposé f préfixe, u_3 est un préfixe strict d'image de lettre.

Comme u est spécial à droite, il se prolonge au moins de deux manières ua et ub. Les mots u_3a et u_3b ne peuvent être préfixes de la même image de lettre, donc u_2 se prolonge par au moins deux lettres différentes x et y, avec u_3a préfixe de $f(x)$ et u_3b préfixe de $f(y)$, c'est-à-dire que u_2 est spécial à droite dans L; son ordre est supérieur ou égal à celui de u.

Réciproquement, si v est un facteur spécial à droite de L qui se prolonge en vx et vy, soit p le plus grand préfixe commun à $f(x)$ et $f(y)$ (qui est un préfixe strict de ces deux mots). Alors $f(v)p$ est spécial à droite dans L'. Plus précisément, si v est d'ordre m, il se prolonge de $m+1$ manières différentes vx_0, vx_1, ..., vx_m. L'ensemble des préfixes des $f(x_i)$ forme un arbre à droite fini, dans lequel chaque embranchement correspond à un plus grand préfixe commun à au moins deux des $f(x_i)$. Si on note p_1, ..., p_r ces préfixes, et pour tout j compris entre 1 et r on note H_j l'ensemble des lettres $a \in \Sigma'$ telles que $p_j a$ est préfixe de l'un des $f(x_i)$, on trouve que $f(v)p_j$ est spécial à droite d'ordre au moins $\#H_j - 1$ dans L'. La somme des ordres des r facteurs spéciaux obtenus est donc supérieure ou égale à l'ordre de v. Il peut ne pas y avoir égalité si $f(v)$ est facteur de l'image par f d'un mot qui ne contient pas v.

On voit donc que, connaissant les facteurs spéciaux à droite dans L, on peut calculer les facteurs spéciaux à droite dans L' en appliquant f et en ajoutant un ou plusieurs suffixes p qui ne dépendent que des lettres par lesquelles le facteur considéré se prolonge. On obtient ainsi tous les facteurs spéciaux à droite dans L', sauf peut-être ceux de longueur inférieure à l, qu'il faut déterminer à la main, et que nous appellerons *facteurs spéciaux exceptionnels*.

On obtient de la même manière les facteurs spéciaux à gauche si f est un morphisme suffixe, et tous les types de facteurs bispéciaux si f est bipréfixe.

7.5 Application : suites de complexité affine

7.5.1 Suites de complexité ultimement affine

Revenons au problème initial : on cherche à construire des suites de complexité ultimement affine $p(n) = \alpha n + \beta$ sur un alphabet binaire $\Sigma = \{\alpha, \beta\}$. Pour que ce soit possible, il faut que $(\alpha, \beta) \in \{0,1\} \times (\mathbb{N} \setminus \{0\}) \cup (\mathbb{N} \setminus \{0,1\}) \times \mathbb{Z}$. On va voir que cette condition est suffisante.

Théorème 7.1. *Soient $j \geqslant 0$, $k \geqslant 1$ et $l \geqslant 1$ trois entiers. Soit $\Sigma_k = \{c_1, \ldots, c_k\}$ un alphabet à k lettres. On considère les deux morphismes f et g définis ainsi :*

$$f: \begin{array}{ccl} \Sigma_k^* & \longrightarrow & \Sigma_k^* \\ c_i & \longmapsto & c_1 c_{i+1} \text{ si } i \neq k \\ c_k & \longmapsto & c_1 \end{array} \qquad \text{et} \qquad g: \begin{array}{ccl} \Sigma_k^* & \longrightarrow & \Sigma^* \\ c_i & \longmapsto & a^l b^{i+j} \end{array}$$

pour tout i compris entre 1 et k.

La suite $\mathbf{u} = g(f^\omega(c_1))$ a pour complexité

$$p(n) = (k-1)n - \frac{k(k-1)}{2} - j(k-2) + l + 1$$

pour $n \geqslant \max(l, j+k-1, j+1)$.

Preuve : On construit tout d'abord la base des arbres des facteurs spéciaux à gauche de $FL(f)$, qui ne contient qu'un seul arbre, de type Σ_k, puisque chaque facteur spécial peut être prolongé par chacune des k lettres de l'alphabet : comme f est un morphisme circulaire suffixe, cet arbre est un point fixe du procédé décrit à la section 7.4, qui ne peut être que filiforme puisqu'il n'y a pas de facteurs spéciaux exceptionnels. Il n'y a donc aucun facteur bispécial d'ordre non nul. En quelque sorte, $f^\omega(c_1)$ est une suite de Fibonacci généralisée, et sa complexité est $(k-1)n+1$.

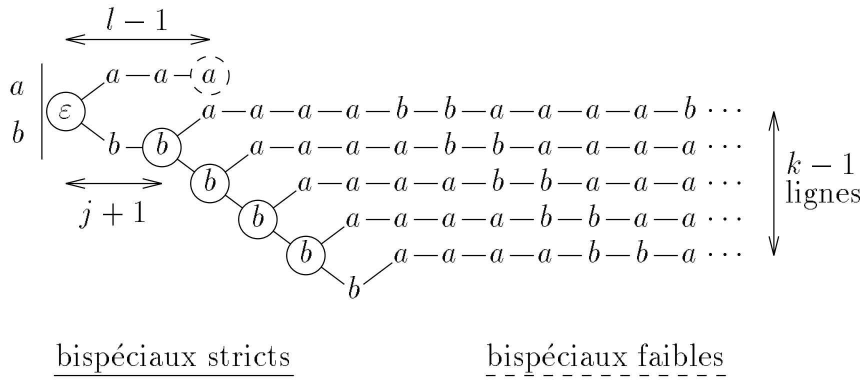

FIG. 7.6 - *Arbre des facteurs spéciaux à gauche de $g(f^\omega(c_1))$ pour $j=1$, $k=6$, $l=4$*

On applique ensuite le morphisme g, qui est circulaire et suffixe, pour obtenir l'arbre des facteurs spéciaux à gauche de $F(g(FL(f)))$ (figure 7.6). Cet arbre comporte $k-1$ branches infinies (chaque facteur spécial pour $FL(f)$, d'ordre $k-1$, donne un élément dans chacune des branches) et une branche finie de longueur $l-1$, contenant les facteurs spéciaux exceptionnels. Il ne reste plus qu'à compter les facteurs bispéciaux pour obtenir la complexité cherchée : il y a un bispécial faible, a^{l-1}, et $k-1$ bispéciaux stricts, ε, b^{j+1}, b^{j+2}, ..., b^{j+k-2}. ■

Corollaire 7.1. *Pour tout couple d'entiers $(\alpha,\beta) \in \{0,1\} \times (\mathbb{N} \setminus \{0\}) \cup (\mathbb{N} \setminus \{0,1\}) \times \mathbb{Z}$, il existe une suite binaire de complexité ultimement $\alpha n+\beta$.*

Preuve : La suite $\mathbf{u}$ définie dans le théorème 7.1 fournit une solution dans tous les cas, avec $k=\alpha+1 \geqslant 1$ et j et l choisis en fonction de α et β.

Si $k=1$, la suite $\mathbf{u}$ a pour complexité ultimement $j+l+1$. Si on prend $j=\left\lfloor\frac{\beta-1}{2}\right\rfloor$ et $l=\left\lfloor\frac{\beta}{2}\right\rfloor$, on a $p(n)=\beta$ pour $n \geqslant \left\lfloor\frac{\beta+1}{2}\right\rfloor$.

Si $k=2$, la suite $\mathbf{u}$ a pour complexité ultimement $n+l$. On prend donc $l=\beta$ et (par exemple) $j=0$, et on a $p(n)=n+\beta$ pour $n \geqslant \beta$.

Si $k > 2$, soit par exemple

$$j = \max\left(0, \left\lfloor \frac{-k^2 + 3k - 2 - 2\beta}{2k - 4} \right\rfloor\right)$$

et

$$l = \beta + \frac{k(k-1)}{2} + j(k-2) - 1 \geqslant 1 \ .$$

La suite $\mathbf{u}$ a alors pour complexité $p(n) = \alpha n + \beta$ pour

$$n \geqslant \max\left(\alpha, \frac{\alpha}{2} - \frac{\beta}{\alpha - 1}, \beta + \frac{\alpha(\alpha+1)}{2} - 1\right) \ .$$

■

7.5.2 Complexité $\alpha n + \beta$ pour $n \geqslant 1$

La question suivante, variante du problème ci-dessus, a été posée par Pascal Alessandri : pour quelles valeurs de (α, β) existe-t-il une suite de complexité $p(n) = \alpha n + \beta$ pour tout $n \geqslant 1$? Elle peut être résolue par une construction similaire.

Théorème 7.2. *Soit $(\alpha, \beta) \in \mathbb{N} \times \mathbb{Z}$. Il existe une suite de complexité $p(n) = \alpha n + \beta$ pour tout $n \geqslant 1$ (et $p(0) = 1$) si et seulement si $\alpha + \beta \geqslant 1$ et $2\alpha + \beta \leqslant (\alpha + \beta)^2$.*

Preuve : Ces conditions sont évidemment nécessaires : $p(1) = \alpha + \beta$ est le nombre de lettres distinctes apparaissant dans la suite, donc est un nombre strictement positif; et si la suite contient $\alpha + \beta$ lettres distinctes, elle ne peut contenir plus de $(\alpha + \beta)^2$ facteurs de longueur 2, donc $2\alpha + \beta = p(2) \leqslant p(1)^2 = (\alpha + \beta)^2$.

Supposons maintenant que (α, β) satisfait les deux conditions, et que $\alpha > 0$, le cas $\alpha = 0$ étant facile à résoudre avec des suites périodiques. Soit $k = \alpha + 1$ et $l = \alpha + \beta$: on va construire un HDOL-système $G = (\Sigma_k, f, c_1, \Sigma_l, h)$ où $\Sigma_k = \{c_1, \ldots, c_k\}$ et $\Sigma_l = \{d_1, \ldots, d_l\}$, de telle sorte que la suite associée $\mathbf{u} = h(f^\omega(c_1))$ aura la complexité voulue.

Le DOL-système sous-jacent $G^0 = (\Sigma_k, f, c_1)$ est le même que pour le théorème 7.1. La suite $f^\omega(c_1)$ est donc une suite de Fibonacci généralisée, de complexité $\alpha n + 1$ pour tout n.

La construction du morphisme h dépend du signe de β. Supposons tout d'abord que $l \geqslant k$, c'est-à-dire que $\beta > 0$. Si $l = k$, la suite $f^\omega(c_1)$ a déjà la complexité cherchée et il suffit de prendre $h = Id$. Sinon, on définit h par $h(c_1) = d_1 d_{k+1} d_{k+2} \ldots d_l$ et $h(c_i) = d_i$ pour $2 \leqslant i \leqslant k$: le morphisme h

est circulaire et bipréfixe. On voit facilement que la base des arbres de facteurs spéciaux à gauche de $F(\mathbf{u})$ a deux éléments, un arbre filiforme de type $\{d_1, \ldots, d_k\}$ (image par h de l'arbre de $FL(f)$) et un arbre de type Σ_l contenant le seul nœud ε, seul facteur spécial exceptionnel, et seul facteur bispécial d'ordre non nul, $k-l$. On en déduit immédiatement que la complexité est bien $p(n) = \alpha n + \beta$ pour $n \geqslant 1$.

Supposons maintenant que $l < k$, c'est-à-dire $\beta \leqslant 0$. On définit h de la manière suivante : $h(c_1) = d_1$, $h(c_i) = d_1 d_i$ pour $2 \leqslant i \leqslant l$, et $h(c_{l+1})$, ..., $h(c_k)$ sont $k-l$ éléments pris parmi l'ensemble $E = d_1\{d_2, \ldots, d_l\}^2$, qui en compte $(l-1)^2$; comme $k-l = 1-\beta \leqslant (\alpha+\beta)^2 - 2\alpha - 2\beta + 1 = (l-1)^2$, c'est toujours possible. L'ordre des images n'a pas d'importance.

Le morphisme h est suffixe et circulaire, de délai de synchronisation 2, grâce au rôle de marqueur joué par d_1. Nous pouvons donc déterminer les facteurs spéciaux à gauche de $F(\mathbf{u})$ en fonction de ceux de $FL(f)$.

Soit u un facteur spécial à gauche pour $FL(f)$: il se prolonge par chacune des lettres de Σ_k. Les suffixes communs des $f(c_i)$ sont ε et toutes les lettres d_i telles que $E \cap d_1 \Sigma_k d_i \neq \emptyset$; les facteurs spéciaux à gauche pour $F(\mathbf{u})$ obtenus sont donc $f(u)$, d'ordre $l-1$ (il se prolonge par chacune des lettres de $H_1 = \Sigma_l$), et $d_i f(u)$, d'ordre $\#(E \cap d_1 \Sigma_k d_i)$ quand cet ensemble est non vide (il se prolonge par $H_i = \{d_1\} \cup \{\, d_j \mid d_1 d_j d_i \in E \,\}$). Les seuls mots non synchronisés sont ε et certains d_i, mais ils ne donnent pas de facteurs spéciaux exceptionnels. Finalement, la base contient un arbre pour chaque ensemble H_i, avec une ou plusieurs branches filiformes partant de ε (une branche pour chaque j tel que $H_i \subset H_j$). Comme précédemment, ε est le seul facteur bispécial d'ordre non nul, et son ordre vaut $k-l$ (l'ordre de ε est toujours égal à $p(2) - 2p(1) + 1$). ■

7.6 Quelques exemples de calculs de complexité de suites définies par morphisme

7.6.1 Suite de Thue-Morse

La suite de Thue-Morse est le point fixe du morphisme

$$\begin{array}{rccc} \theta: & \Sigma^* & \longrightarrow & \Sigma^* \\ & a & \longmapsto & ab \\ & b & \longmapsto & ba \end{array} .$$

La figure 7.4 montre l'arbre des facteurs spéciaux à droite de cette suite. Les facteurs bispéciaux sont

<u>bispéciaux stricts :</u> ε, $\theta^m(ab)$, $\theta^m(ba)$ pour $m \geqslant 0$;

<u>bispéciaux ordinaires :</u> a, b;

bispéciaux faibles : $\theta^m(aba)$, $\theta^m(bab)$) pour $m \geqslant 0$.

On a donc :

$$BS(X) - BF(X) = 1 + 2\sum_{n=0}^{\infty}\left(X^{2.2^n} - X^{3.2^n}\right)$$

et on en déduit immédiatement les séries $S(X)$ et $P(X)$. On retrouve ainsi les formules classiques [11] :

$$s(n) = \begin{cases} 1 & \text{si } n = 0 \\ 2 & \text{si } 0 < n \leqslant 2 \\ 4 & \text{si } 2.2^m < n \leqslant 3.2^m \\ 2 & \text{si } 3.2^m < n \leqslant 4.2^m \end{cases}$$

$$p(n) = \begin{cases} 1 & \text{si } n = 0 \\ 2 & \text{si } n = 1 \\ 4 & \text{si } n = 2 \\ 4n - 2.2^m - 4 & \text{si } 2.2^m < n \leqslant 3.2^m \\ 2n + 4.2^m - 2 & \text{si } 3.2^m < n \leqslant 4.2^m \end{cases}$$

où m est un entier positif ou nul.

7.6.2 Un exemple non primitif

Considérons le morphisme suivant :

$$\begin{array}{rrcl} f: & \Sigma^* & \longrightarrow & \Sigma^* \\ & a & \longmapsto & aba \\ & b & \longmapsto & bb \end{array} .$$

Ce morphisme n'est pas primitif, car les images successives de b ne contiennent aucun a (un morphisme f est dit primitif s'il existe un entier n tel que l'image de chaque lettre par f^n contient toutes les lettres). La suite qu'il engendre,

$$f^\omega(a) = ababbababbbbababbababbbbbbbbababbababbbbababbababbbbbbbbbbbbbbbbab\ldots$$

est récurrente (tout facteur apparaît une infinité de fois), mais pas uniformément récurrente puisqu'il existe des facteurs arbitrairement longs qui ne contiennent pas de a.

Les facteurs bispéciaux sont

<u>bispéciaux stricts :</u> $f^m(b)$, de longueur 2^m, pour $m \geqslant 0$;

<u>bispéciaux ordinaires :</u> ε;

bispéciaux faibles : $f^m(bab)$, de longueur $2^m(3+m/2)$, pour $m \geqslant 0$.

On peut donc exprimer BS et BF comme des sommes. Il est également possible d'écrire des équations fonctionnelles :

$$BS(X) = BS(X^2) + X \ ;$$

$$BF(X) = \varphi(X,X) \text{ avec } \varphi(X,Y) = \varphi(XYX, YY) + YXY \ .$$

On peut enfin écrire que $s(n) = 1 + q + r$, où $q = \lceil \log_2 n \rceil$ est le nombre de facteurs bispéciaux stricts de longueur inférieure à n, et $r = \left\lceil \frac{W(128n \log 2)}{\log 2} - 6 \right\rceil$ est le nombre de facteurs bispéciaux faibles de longueur inférieure à n ($W(x)$ est l'unique fonction analytique réelle sur $]-1/e, +\infty[$ telle que $W(x)e^{W(x)} = x$). On calcule ensuite :

$$\begin{aligned} p(n) &= 1 + \sum_{m=0}^{n-1} s(m) \\ &= 1 + n + \sum_{m=0}^{q-1} (n - 1 - 2^m) - \sum_{m=0}^{r-1} (n - 1 - 2^m(3 + m/2)) \\ &= 1 + n + (n-1)q - (2^q - 1) - (n-1)r + (2^{r+1} + 2^{r-1}r - 2) \ . \end{aligned}$$

Dans cette expression, les termes dominants sont nq et nr, soit

$$p(n) = n \frac{\log n - W(128n \log 2)}{\log 2} + O(n) \ .$$

On peut alors, en utilisant le fait que $W(x) = \log x - \log \log x + O\left(\frac{\log \log x}{\log x}\right)$ obtenir un équivalent de la complexité de la suite étudiée : $p(n) \sim n \log_2 \log_2 n$.