UNIVERSITÉ DE OUAGADOUGOU

UFR -- Sciences Exactes et Appliquées

THÈSE

Discipline: Mathématiques

Spécialité: Mathématiques Discrètes

Titre

Combinatoire de mots et caractérisation de morphismes de certaines classes de mots

Présentée par:

Idrissa KABORÉ

Pour obtenir le diplôme de

DOCTORAT UNIQUE

Soutenue le 04 Mars 2004

JURY

Président:	Pr Akry KOULIBALY	Université Polytechnique de Bobo-Dsso
Membres:	Pr Julien CASSAIGNE	Université de la Méditerranée
	Pr Cyprien GNANVO	Université d'Abomey-Calavi
	Pr Moussa OUATTARA	Université de Ouagadougou
	Pr Théodore TAPSOBA	Université Polytechnique de Bobo-Dsso

REMERCIEMENTS

Je tiens à remercier le Professeur Akry KOULIBALY pour ses encouragements et qui me fait l'honneur de présider le Jury.

J'adresse un remerciement particulier au Professeur Julien CASSAIGNE pour le soin et l'attention avec lesquels il a mené la tâche de rapporteur et d'accepter de faire partie du jury. Je lui exprime ma reconnaissance pour m'avoir offert un séjour à l'I.M.L au sein de l'équipe Dynamique, Arithmétique et Combinatoire (DAC) et pendant lequel il m'a fait bénéficier de ses nombreuses remarques et suggestions.

Je remercie aussi le Professeur Cyprien GNANVO d'avoir accepté d'être rapporteur et membre du jury de ma thèse.

Je remercie également le Professeur Moussa OUATTARA pour son soutien et sa participation au jury.

Je tiens à manifester ma reconnaissance au Professeur Théodore TAPSOBA qui a assuré la direction de cette thèse. Son contact toujours agréable, ses encouragements, ses conseils et son soutien ont été stimulants et ont contribué à l'aboutissement de cette thèse.

Je remercie aussi tous les professeurs du département de mathématiques de l'UFR-SEA, tous les membres du laboratoire AGATA et mes camarades pour leurs encouragements.

Mes remerciements vont enfin à mes parents, à mes amis et aussi à tous ceux qui, d'une manière ou d'une autre, m'ont apporté un soutien dans mon parcours.

Table des matières

Introduction

Un mot est une suite de symboles pris dans un ensemble fini A appelé alphabet. On définit une structure de monoïde sur A^*, l'ensemble des mots finis de A, à l'aide de l'opération concaténation (ou juxtaposition): on parle de monoïde libre engendré par A. L'étude des mots infinis remonte au moins aux travaux de Thue ([36], [37]) où il abordait le problème de répétition dans les mots. Par la suite la notion de complexité a été introduite pour mesurer la diversité des motifs apparaissant dans les mots. Ainsi on appelle complexité d'un mot u, la fonction p_u qui compte le nombre de facteurs distincts de longueur donnée dans u.

Le plus souvent on étudie les mots en les regroupant par classe. Une classe peut être définie comme un ensemble de mots possédant en commun une certaine propriété combinatoire. La complexité est l'une des principales propriétés combinatoires dans la classification des mots. Par exemple, les mots ultimement périodiques sont les mots infinis dont la complexité est bornée.

Dans ce travail nous étudions des classes de mots ayant une certaine complexité. Il s'agit notamment des mots sturmiens qui sont de complexité $n+1$; des mots d'Arnoux-Rauzy qui sont de complexité $2n+1$ et certains mots quasi-sturmiens de complexité $n+2$. Nous nous intéressons aussi à l'étude des morphismes préservant respectivement les mots sturmiens et les mots d'Arnoux-Rauzy. On appelle morphisme toute application f de A^* dans lui-même telle que $f(uv) = f(u)f(v)$ pour tous u et v de A^*.

Cette thèse est structurée en quatre chapitres.

Dans le chapitre 1 nous rappelons des concepts de base en combinatoire de mots et certains résultats utiles pour les chapitres suivants.

Le chapitre 2 est consacré à l'étude des mots sturmiens et de leurs morphismes. Nous reprenons certaines propriétés combinatoires connues sur ces mots et nous montrons une nouvelle propriété caractéristique de ces derniers (théorème 2.2.7). Ensuite nous rappelons une caractérisation constructive des morphismes sturmiens dûe à F. Mignosi et P. Séébold ([28]) puis nous montrons que pour tout morphisme sturmien f et tout mot récurrent x, le mot $f(x)$ est sturmien si et seulement si x l'est.

Le chapitre 3 est consacré principalement aux morphismes d'Arnoux-Rauzy. Après avoir défini les mots d'Arnoux-Rauzy nous énonçons quelques propriétés combinatoires de ces mots qui nous serviront dans l'étude de leurs morphismes. Ensuite, nous obtenons une caractérisation de ces morphismes en montrant qu'ils sont engendrés par une famille de quatre morphismes.

Dans le dernier chapitre nous classifions tous les mots qui possèdent, pour tout entier n non nul, exactement $n+2$ facteurs de longueur n. Ensuite, nous étudions singulièrement une sous classe de ces mots que nous appelons mots quasi-sturmiens par insertion. Nous montrons qu'elle constitue une classe de mots 2-équilibrés puis nous étudions la palindromie et nous obtenons que pour ces mots $pal(n) = 0$ si n est pair et $pal(n) = 3$ sinon où pal est la fonction calculant le nombre de palindromes de longueur n.

Chapitre 1

Généralités

Dans ce chapitre nous rappelons des notions de base en combinatoire de mots, et certains résultats utiles pour le reste du travail. La plupart des notations peuvent être retrouvées dans le livre de M. Lothaire [27].

1.1 Mots

On appelle alphabet, et on note A, tout ensemble fini de symboles. Les éléments d'un alphabet sont appelés des lettres.
Un mot sur A est une suite finie ou infinie de lettres de A. La suite vide se définit comme le mot vide; on le note ε.
Considérons $A = \{a, b\}$ un alphabet à deux lettres; $ababba$ et $ababbabbba \cdots$ sont des mots sur A respectivement fini et infini. On désigne par A^* l'ensemble des mots finis sur A.
La concaténation ou le produit de deux mots finis u et v, noté $u \cdot v$, est le mot uv: c'est l'opération binaire sur A^* qui consiste à écrire le deuxième mot à la suite du premier. La concaténation est évidemment associative et admet le mot vide comme élément neutre. A^* est ainsi muni d'une structure de monoïde et on l'appelle le monoïde libre engendré par A. Il convient de remarquer que tout mot fini sur A, non vide, est une concaténation finie de lettres de A. On définit naturellemment les puissances de mots dans A^* comme dans tout monoïde. L'ensemble des mots finis non vides se note A^+ ($A^+ = A^* \setminus \{\varepsilon\}$)

et celui des mots infinis A^ω. Tout au long du travail on rencontrera diverses méthodes de construction de mots infinis. L'ensemble de tous les mots sur A est noté A^∞. On a donc $A^\infty = A^* \cup A^\omega$. Pour tout mot $u \in A^\infty$ on note $alph(u)$ l'ensemble des lettres de A présentes dans u.

Théorème 1.1.1 *Soit A un alphabet. Soient $(M, \cdot)$ un monoïde d'élément neutre e et h une application de A dans M. Alors il existe un et un seul homomorphisme $h^* : A^* \longrightarrow M$ tel que:*

$$\forall a \in A,\ h^*(a) = h(a).$$

Preuve: *Existence:* Posons $h^*(\varepsilon) = e$ et $h^*(a_1a_2\cdots a_n) = h(a_1)h(a_2)\cdots h(a_n)$. L'application h^* définit un homomorphisme de monoïdes.

Unicité: Soient g et g' deux homomorphismes de A^* dans M tels que $\forall a \in A,\ g(a) = g'(a)$. Alors, $g(\varepsilon) = g'(\varepsilon) = e$ et pour tout $u = a_1a_2\cdots a_n$ nous avons

$$g(u) = g(a_1)g(a_2)\cdots g(a_n) = g'(a_1)g'(a_2)\cdots g'(a_n) = g'(u).$$

$\square$

C'est par ce théorème que le monoïde A^* est appelé monoïde libre engendré par A. Soit u un mot dans A^* et a une lettre de A. La longueur de u, notée $|u|$, est le nombre de termes de la suite u. Par exemple la longueur du mot *abbaaab* est égale à 7. Par convention $|\varepsilon| = 0$. L'application $l : (A^*, \cdot, \varepsilon) \longrightarrow (\mathbb{N}, +, 0)$, $u \longmapsto |u|$ est un homomorphisme de monoïdes.

Le nombre d'occurrences de la lettre a dans le mot u , noté $|u|_a$ est le nombre d'apparitions de a dans u. En considérant $u = abbaaab$ nous avons $|u|_a = 4$.

1.2 Facteurs de mots

Définition 1.2.1 *Soit $u \in A^\infty$ et $v \in A^*$. On dit que v est un facteur de u s'il existe deux mots u_1 et u_2 tels que $u = u_1vu_2$. Si $u_1 = \varepsilon$ alors v est appelé préfixe de u. De même on dit que v est un suffixe de u si $u_2 = \varepsilon$. Si le facteur v n'est ni préfixe ni suffixe de u alors on l'appelle facteur propre.*

Un mot v est dit préfixe (resp. suffixe) *strict* d'un mot u lorsque v est un préfixe (resp. suffixe) de u tel que $v \neq u$.

Exemple: Sur l'alphabet binaire $A = \{a, b\}$, considérons le mot $u = ababbaaab$. Le mot aba est un préfixe de u tandis que $baaab$ est un suffixe. Le bba est un facteur propre de u.

Par convention le mot vide est facteur de tout mot.

On appelle langage d'un mot u l'ensemble de tous les facteurs de u. On le note $L(u)$. L'ensemble des facteurs de longueur n d'un mot infini u se note $L_n(u)$. Nous avons:

$$L(u) = \cup_{n\geq 0} L_n(u)$$

Soit v un facteur d'un mot infini $u = u_0u_1u_2\cdots$. Un entier i est la position d'une occurrence de v dans u si $u_iu_{i+1}\cdots u_{i+|v|-1} = v$

Définition 1.2.2 *Soit $u = u_0u_1u_2\cdots$ un mot infini sur un alphabet A.*

– On dit que u est périodique s'il existe un entier non nul T tel que $u_{i+T} = u_i$ pour tout $i \geq n_0$; la plus petite valeur pour T étant appelée longueur de la période.

– u est dit ultimement périodique s'il existe un entier non nul T et un entier n_0 tels que $u_{i+T} = u_i$ pour tout $i \geq 0$

En d'autres termes, u est dit périodique (resp. ultimement périodique) s'il existe deux mots finis v et w tels que $u = v^\omega$ (reps. $u = wv^\omega$).

On observe immédiatement que tout mot périodique est ultimement périodique.

Soit u un mot infini sur A et v un facteur de u. On dira qu'une lettre a est un prolongement à droite (resp. à gauche) de v dans u si va (resp. av) est dans $L(u)$.

On dit que v est *spécial* à droite (resp. à gauche) lorsqu'il admet plusieurs prolongements à droite (resp. à gauche). Un facteur spécial à la fois à droite et à gauche est appelé facteur *bispécial*.

Proposition 1.2.1 *Soit u un mot infini périodique. Alors, il existe un entier naturel n tel que u ne possède pas de facteur spécial de longueur n.*

Preuve: Soit u un mot infini périodique de période T. Soit v un facteur de longueur n avec $n \geq T$. Considérons l la première position de v dans u et soit m une autre occurrence de v dans u. Alors, pour tout $h \in \{0, 1, \ldots, n-1\}$: $u_{m+h} = u_{l+h}$.

Vérifions que les prolongements u_{l+n} et u_{m+n} de v respectivement à la position l et m sont identiques.

Nous avons:

$$\forall h \in 0, 1, \ldots, n-1 : \; u_{m+h} = u_{l+h+(m-l)} = u_{l+h}.$$

Il en résulte que T divise $m-l$ et donc il existe un entier k tel que $m = l+kT$. Par suite, $u_{m+n} = u_{l+kT+n} = u_{(l+n)+kT} = u_{l+n}$. Par conséquent v possède un unique prolongement à droite. De manière analogue on obtient que v possède un unique prolongement à gauche. □

Cette propriété s'étend naturellent aux mots ultimement périodiques.

Mot miroir. On définit récursivement une opération unaire sur A^*, notée $(\bar{\ })$ de la manière suivante: $\bar{\varepsilon} = \varepsilon$, $\overline{ua} = a\overline{u}$ pour tout $u \in A^*$ et $a \in A$. Cette opération s'appelle image miroir. L'image miroir $\overline{u}$ d'un mot u sera tout simplement appelée miroir de u. L'image miroir est une opération involutive (resp. contravariante) i.e $\overline{\overline{u}} = u$ (resp. $\overline{uv} = \overline{v}\,\overline{u}$) pour tous mots finis u et v.
En d'autres termes, le miroir de u est le mot obtenu en lisant u de la droite vers la gauche i.e $\overline{u} = u_n u_{n-1} \cdots u_1$ si $u = u_1 u_2 \cdots u_n$.

Définition 1.2.3 *Soit $u \in A^*$. On dit que u est un palindrome si $\overline{u} = u$.*

Par exemple, les mots "elle", "été", "ici" et "tôt" sont des palindromes en français.

Suites de mots et convergence. Soient (v_n) une suite de mots et u un mot infini. Pour tout entier naturel k on note $u^{[k]}$ le préfixe de u de longueur k. On dit que (v_n) converge vers le mot infini u et on pose $\lim v_n = u$ lorsque que pour tout entier k, il existe un entier n_0 tel que $u^{[k]}$ est préfixe de v_n pour tout $n \geq n_0$. Ceci est réalisé lorsque le mot v_n est préfixe propre de v_{n+1}.

3 Complexité

Intuitivement on reconnaît qu'un mot est d'autant plus "compliqué" qu'il possède "beaucoup" de facteurs distincts. Ainsi on comptera le nombre de facteurs (de même longueur) qui apparaissent dans un mot pour mesurer sa complexité.

Définition 1.3.1 *Soit u un mot infini. On appelle fonction de complexité de u que l'on note p_u, l'application de $\mathbb{N}$ dans $\mathbb{N}^*$ définie par $p_u(n) = \#L_n(u)$, où $\#L_n(u)$ désigne le cardinal de $L_n(u)$.*

Autrement dit la fonction de complexité de u est la fonction qui compte le nombre de facteurs de u de longueur donnée. Notons que la fonction de complexité est croissante et vérifie de plus la double inégalité :

$$1 \leq p_u(n) \leq (\#A)^n.$$

Dans tout ce qui suit la fonction de complexité sera simplement désignée par p lorsqu'aucune confusion n'est à craindre.

Exemple: Considérons le mot $u = abbbaabbbaabbbaabbb\cdots$. Nous avons $p(1) = 2, p(2) = 4, p(3) = 5, \ldots$.

Définition 1.3.2 *On appelle mesure de la prédictibilité ou fonction des facteurs spéciaux à droite d'un mot infini u la fonction s définie de $\mathbb{N}$ dans $\mathbb{N}$ par $s(n) = p(n+1) - p(n)$.*

Définition 1.3.3 *Soit u un mot infini et v un facteur de u. On appelle degré à droite (resp. à gauche) de w et on note $\partial^+ w$ (resp. $\partial^- w$) le nombre de prolongements à droite (resp. à gauche) de w dans u .*

Remarquons que pour tout facteur fini d'un mot infini u nous avons toujours:

$$\partial^- w \geq 0 \text{ et } \partial^+ w \geq 1.$$

Proposition 1.3.1 *Soit $n \in \mathbb{N}$ et $u \in A^\omega$. Les propriétés suivantes sont toujours vérifiées:*

(i) $p(n+1) = \sum_{w \in L_n(u)} \partial^+ w = \sum_{w \in L_n(u)} \partial^- w.$

(ii) $p(n+1) - p(n) = \sum_{w \in L_n(u)} (\partial^+ w - 1) = \sum_{w \in L_n(u)} (\partial^- w - 1).$

Preuve: A chaque facteur w de longueur n correspond $\partial^+ w$ (resp. $\partial^- w$) facteurs de longueur $n+1$ de u. D'où l'assertion *(i)*

Par ailleurs nous avons:

$$\sum_{w \in L_n(u)} (\partial^+ w - 1) = \sum_{w \in L_n(u)} \partial^+ w - \sum_{w \in L_n(u)} 1 = p(n+1) - p(n).$$

De même nous avons $p(n+1) - p(n) = \sum_{w \in L_n(u)} (\partial^- w - 1)$.

L'égalité (ii) fournit un puissant outil pour le calcul de la complexité ([12]).

Théorème 1.3.2 *([29]) Soit u un mot infini. Les assertions suivantes sont équivalentes:*

(i) u est ultimement périodique.

(ii) il existe un entier n tel que $p(n) = p(n+1)$.

(iii) il existe un entier n tel que $p(n) \leq n$.

(iv) $(p(n))_n$ est bornée.

Preuve: Soit u un mot infini.

Supposons u ultimement périodique. D'après la proposition 1.2.1 il existe un entier n tel qu'aucun facteur de longueur n de u ne soit spécial i.e

$$\exists n : \forall v \in L_n(u),\ \partial^{-} v = 1.$$

D'où, $\sum_{v \in L_n(u)} (\partial^{+} v - 1) = 0$ i.e $p(n+1) - p(n) = 0$ (cf. proposition 1.3.1 (ii)). Ainsi $(i) \Longrightarrow (ii)$.

Supposons qu'il existe un entier n tel que $p(n+1) = p(n)$. Donc

$$p(n+1) - p(n) = \sum_{v \in L_n(u)} (\partial^{-} v - 1) = 0.$$

Or comme u est infini on a

$$\forall v \in L_n(u) : \partial^{+} v \geq 1 \text{ ou encore } \partial^{+} v - 1 \geq 0.$$

D'où les implications

$$\sum_{v \in L_n(u)} (\partial^{+} v - 1) = 0 \Longrightarrow (\forall v \in L_n(u) : \partial^{+} v - 1 = 0) \Longrightarrow (\forall v \in L_n(u) : \partial^{+} v = 1)$$

Ainsi u n'admet pas de facteur spécial de longueur n. De plus tout facteur de u de longueur plus grande que n ne peut être spécial car sinon tout suffixe de longueur n d'un tel facteur serait aussi spécial. Donc

$$\forall m \geq n, \forall v \in L_m(u) : \partial^{+} v = 1.$$

D'où : $\forall m \geq n,\ p(m+1) - p(m) = 0$.

Donc $(p(n))_n$ est bornée.

D'où $(ii) \Longrightarrow (iv)$

L'implication $(iv) \Longrightarrow (iii)$ est évidente.

$(iii) \Longrightarrow (i)$. Par l'absurde supposons u non ultimement périodique. Alors pour tout entier n, u admet au moins un facteur spécial à droite de longueur n i.e

$$\forall n, \exists v \in L_n(u) : \partial^{+} v \geq 2 \text{ ou encore } \partial^{+} v - 1 \geq 1$$

Par conséquent

$$p(n+1) - p(n) = \sum_{v \in L_n(u)} (\partial^+ v - 1) \geq 1.$$

On en déduit que $p(n) \geq n + 1$. □

La fonction de complexité est une notion très utile pour l'étude des mots infinis et des systèmes symboliques ([2], [3], [12], [33], [35]). En effet, elle permet de caractériser en particulier les mots ultimement périodiques (voir théorème 1.3.2), ainsi que certains mots de basse complexité, comme ceux de complexité $n+1$. Ces derniers sont appelés des mots sturmiens et ont fait l'objet de nombreuses études tant du point de vue combinatoire, qu'arithmétique et géométrique (voir à ce sujet [16], [30], [32]).

1.4 Systèmes dynamiques symboliques

L'ensemble A est muni de la topologie discrète (*i.e* toute partie de A est à la fois ouverte et fermée), et l'ensemble A^ω de la topologie produit. L'application $d : A^\omega \times A^\omega \longrightarrow \mathbb{R}^+$ définie par

$$d(u,v) = \exp(-inf\,\{i \in \mathbb{N} : u_i \neq v_i\})$$

est une distance. A^ω munie de cette distance est un espace métrique compact.

Définition 1.4.1 *On appelle décalage sur A^ω l'application T de A^ω dans lui-même définie par $T((u_n)_{n\in\mathbb{N}}) = (u_{n+1})_{n\in\mathbb{N}}$*

Plus précisément l'application T consiste à effacer la première lettre du mot. Le décalage T est une application uniformément continue, surjective et non injective.

Définition 1.4.2 *Un système dynamique symbolique est un couple (Ω,S), où Ω est un fermé de A^ω invariant par le décalage, et où S est la restriction du décalage T à Ω.*

Définition 1.4.3 *Un système dynamique symbolique (Ω,S) est dit minimal s'il ne contient pas de fermé invariant non trivial, ou encore si pour tout point $v \in \Omega$ l'ensemble $\{S^n(v),\, n \in \mathbb{N}\}$ est dense dans Ω.*

Définition 1.4.4 *Soit u un mot infini sur un alphabet A. On appelle orbite de u l'ensemble $O(u) = \{T^n(u),\, n \in \mathbb{N}\}$ où T désigne le décalage.*

Posons $X_u = \overline{O(u)}$ où $\overline{O(u)}$ est l'adhérence de $O(u)$ dans A^ω. Observons que X_u est un fermé invariant par T. Ainsi à tout mot infini u on associe le système dynamique (X_u, T). Notons que X_u est fini si et seulement si u est ultimement périodique.

Proposition 1.4.1 *Soit u un mot infini sur un alphabet fini A. Alors*

$$X_u = \{v \in A^\omega : L(v) \subseteq L(u)\}.$$

Preuve: Soit v un mot infini sur A tel que $L(v) \subseteq L(u)$ et $v^{[n]}$ le préfixe de longueur n de v. Alors $v^{[n]}$ est un facteur de u et donc il existe un indice i_n tel que $v^{[n]}$ apparaît à cet indice dans u. Ainsi $v^{[n]}$ est un préfixe de $T^{i_n}(u)$. D'où $v = \lim_{n\to+\infty} T^{i_n}(u)$ et donc $v \in X_u$. Ce qui montre que $\{v \in A^\omega : L(v) \subseteq L(u)\} \subseteq X_u$.

Réciproquement, soit $v \in X_u$. Alors, il existe une suite d'entiers $(i_n)_{n\geq 0}$ telle que $v = \lim_{n\to+\infty} T^{i_n}(u)$. Soit w un facteur de v; alors il existe un entier n tel que w est facteur du préfixe $v^{[n]}$ de longueur n du mot v. Comme $v \in X_u$ il existe un entier n_0 tel que $T^{i_{n_0}}(u)$ commence par $v^{[n]}$. Par conséquent w est facteur de $T^{i_{n_0}}(u)$ et donc de u. Ainsi nous avons $L(v) \subseteq L(u)$. D'où l'inclusion $X_u \subseteq \{v \in A^\omega : L(v) \subseteq L(u)\}$. □

Définition 1.4.5 *Soit u un mot infini sur un alphabet fini A.*

u est dit récurrent si tout facteur de u apparaît une infinité de fois dans u.

u est dit uniformément récurrent si tout facteur de u apparaît avec des lacunes bornées c'est-à-dire que pour tout $n \in \mathbb{N}$ il existe un entier N tel que tout facteur de longueur N contient tous les facteurs de longueur n de u.

On remarque immédiatement que tout mot périodique dans A^ω est récurrent et même uniformément récurrent.

Tout mot uniformément récurrent est récurrent mais la réciproque est fausse.

On trouvera dans [23] la preuve des résultats classiques suivants sur les mots récurrents et uniformément récurrents.

Proposition 1.4.2 *Soit u un mot infini sur un alphabet fini A. Les assertions suivantes sont équivalentes:*

(i) u est récurrent.

(ii) Il existe une suite strictement croissante $(n_k)_{n\geq 0}$ d'entiers naturels telle que: $u = \lim_{n\to+\infty} T^{n_k}(u)$.

(iii) L'application T est surjective sur X_u.

Proposition 1.4.3 *Soit u un mot infini sur un alphabet fini A. Les assertions suivantes sont équivalentes:*

(i) u est uniformément récurrent.

(ii) $\forall v \in X_u$; $X_v = X_u$.

(iii) $\forall v \in L(u)$; $L(v) = L(u)$.

Théorème 1.4.4 *Soit u un mot infini sur un alphabet fini A. Le système dynamique (X_u, T) est minimal si et seulement si u est uniformément récurrent.*

C'est en vertu de ce théorème qu'un mot infini uniformément récurrent est appelé mot minimal.

1.5 Graphes de mots

1.5.1 Graphes de De Bruijn

On considère un alphabet A de cardinal q non nul. Dans le but de construire les mots finis circulaires de longueur q^n tels que tout mot de longueur n apparaisse une et une seule fois De Bruijn a utilisé, dans [18], les graphes de mots. Ainsi nous appelons graphe de De Bruijn le graphe d'un mot de complexité maximale.

Définition 1.5.1 *Le graphe de De Bruijn d'ordre n sur l'alphabet A est le graphe orienté de q^n sommets de longueur n où deux sommets v et v' sont reliés par un arc de v vers v' si le suffixe de longueur $n-1$ de v est le préfixe de longueur $n-1$ de v'.*

Exemples de graphes de De Bruijn sur un alphabet à deux lettres $\{a, b\}$: voir la figure 1.1 à la page 14

1.5.2 Graphes de Rauzy

Introduits par G. Rauzy ils permettent une visualisation des enchaînements des facteurs d'un mot u donné.

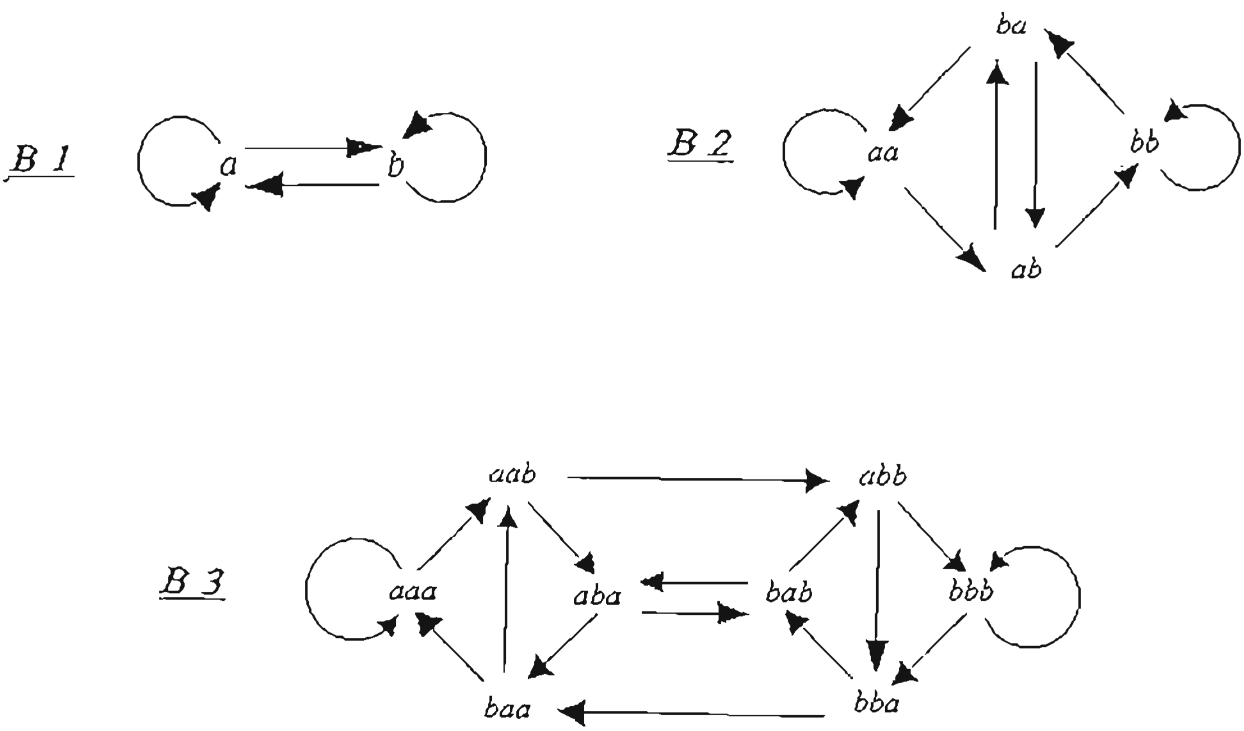

FIG. 1.1 *Les graphes de De Bruijn d'ordre 1, 2 et 3 sur $\{a, b\}$*

Définition 1.5.2 *Soit u un mot sur A. Pour tout $n \in \mathbb{N}$, on appelle graphe de Rauzy d'ordre n de u le graphe orienté, noté Γ_n, tel que:*

- Ses sommets sont les éléments de $L_n(u)$

- Pour tout $v, v' \in L_n(u)$ il existe un arc de v vers v' si et seulement s'il existe a et b, éléments de A, vérifiant $w = va = bv'$ et $w \in L_{n+1}(u)$. la lettre a est appelée étiquette de l'arc de v à v', on note $v \xrightarrow{a} v'$.

A tout arc de Γ_n correspond un et un seul facteur de longueur $n+1$ de u. Ainsi le graphe de Rauzy d'ordre n d'un mot u est une représentation pratique de l'ensemble des facteurs de longueur $n+1$ de u. Deux sommets reliés par un arc sont dits successifs car ils se suivent dans ledit mot.

Exemple: La figure 1.2 de la page 15 donne les graphes de Rauzy d'ordre 0, 1, 2 et 3 du mot de Thue-Morse défini dans la section 1.5 et dont les premières lettres sont:

$$M = abbabaabbaababbabaababbaabbabaabbaababba \cdots$$

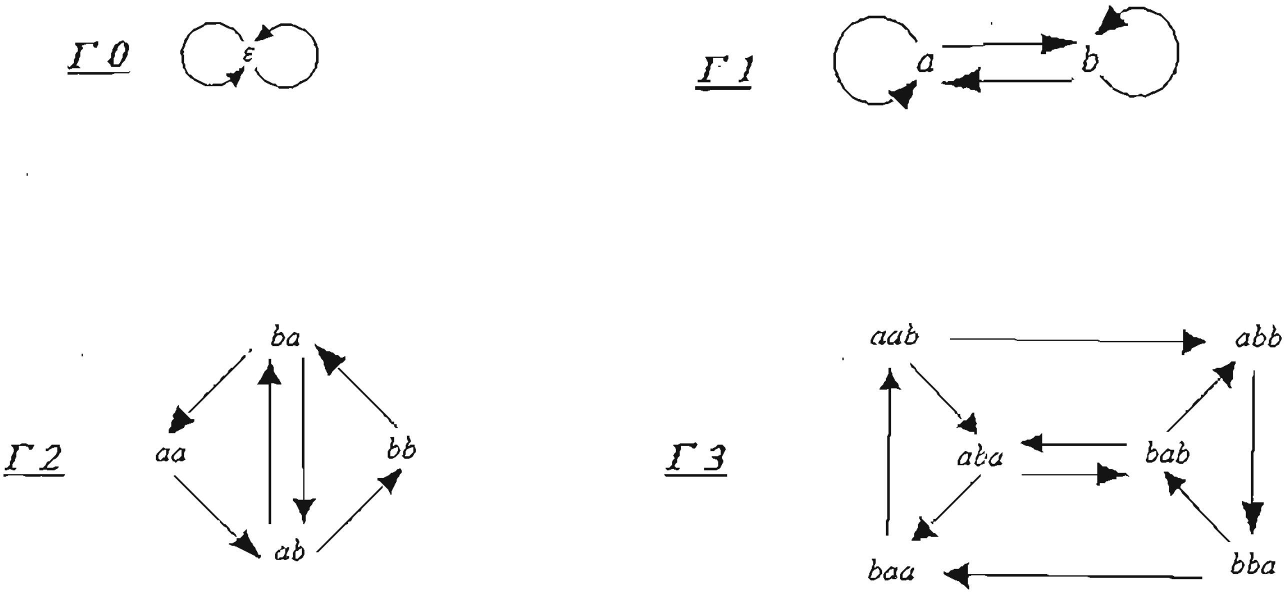

FIG. 1.2 *Les quatre premiers graphes de Rauzy du mot de Morse*

Théorème 1.5.1 *([30]) Soit u un mot infni sur un alphabet fini A. Les propriétés suivantes sont équivalentes:*

(i) u est récurrent.

(ii) Tout facteur de u admet au moins deux occurrences dans u.

(iii) Pour tout entier $n \in \mathbb{N}$, le graphe de Rauzy d'ordre Γ_n est fortement connexe.

(iv) Tout facteur de u est prolongeable à gauche.

1.5.3 Graphes dérivés des graphes de Rauzy

Soit u un mot infini sur un alphabet fini A.

Définition 1.5.3 *Soit Γ_n le graphe de Rauzy d'ordre n de u. Le graphe dérivé de Γ_n, noté $D(\Gamma_n)$, est le graphe orienté tel que:*

- Ses sommets sont les arcs de Γ_n, c'est-à-dire les facteurs de longueurs $n+1$ de u.

Il admet un arc de v vers v', lorsque dans Γ_n le sommet d'arrivée de l'arc v est le sommet de départ de l'arc v'; c'est-à -dire qu'il existe a et b de A tels que $va = bv'$.

Nous observons que les arcs de Γ_{n+1} correspondent à tous les éléments de $L_{n+2}(u)$ tandis que les arcs de $D(\Gamma_n)$ représentent tous les mots possibles de longueur $n+2$ construits à partir des éléments $L_{n+1}(u)$. Ainsi tout arc de Γ_{n+1} est un arc de $D(\Gamma_n)$ et donc Γ_{n+1} est un sous graphe de $D(\Gamma_n)$. Donc le graphe de Rauzy Γ_{n+1} est obtenu à partir du graphe $D(\Gamma_n)$ par retrait éventuel de certains arcs.

Proposition 1.5.2 *Soit u un mot récurrent sur un alphabet fini A et n un entier naturel. Si le mot u ne possède pas de facteur bispécial de longueur n, alors $\Gamma_{n+1} = D(\Gamma_n)$.*

Preuve: Soit u un mot récurrent sur un alphabet fini A et n un entier naturel tel que u n'admette pas de facteur bispécial de longueur n. Supposons $\Gamma_{n+1} \neq D(\Gamma_n)$. Alors il existe un arc atb ($a, b \in A$) de longueur $n+2$ qui n'appartient pas à Γ_{n-1} i.e atb n'est pas dans $L_{n+2}(u)$. Or, comme atb est un arc de $D(\Gamma_n)$ alors at et tb appartiennent à $L_{n+1}(u)$. Par suite, comme u est récurrent, le facteur tb (resp. at) est prolongeable à gauche (resp. à droite) par une lettre a' (resp. b'). Plus précisément $a'tb$ et atb' sont dans $L_{n+2}(u)$. De plus a' est distincte de a et b' distincte de b puisque atb n'est pas dans $L_{n+2}(u)$. Par conséquent t est bispécial car les mots at, tb, $a't$ et tb' sont tous facteurs de longueur $n+1$ de u. Absurde. □

1.6 Morphismes de mots

Les substitutions sont des applications définies sur A^* qui génèrent de façon naturelle des mots infinis de basse complexité.

Définition 1.6.1 *On appelle substitution sur un alphabet A toute application $f : A \longrightarrow A^*$. Toute substitution f se prolonge de manière naturelle en morphisme du monoïde A^* dans lui-même par concaténation (i.e $f(uv) = f(u)f(v)$, $\forall u, v \in A^*$ et $f(\varepsilon) = \varepsilon$), puis en une application de A^∞ dans lui-même.*

Point fixe d'un morphisme. On appelle point fixe d'un morphisme f un mot infini u vérifiant $f(u) = u$.

Définition 1.6.2 *Un morphisme f est dit:*

- *uniforme de module p si pour tout $a \in A$ on a: $|f(a)| = p$;*
- *croissant si pour tout $a \in A$ on a: $|f(a)| \geq 2$;*
- *non effaçant si pour tout $a \in A$ on a: $f(a) \neq \varepsilon$;*
- *prolongeable en a (une lettre de A) si $f(a) = au$ avec $u \in A^+$.*

Un morphisme de module 1 est appelé morphisme littéral.

Définition 1.6.3 *Soit f un morphisme sur A^*. L'image miroir de f est le morphisme noté $\overline{f}$ et défini sur A^* par $\overline{f}(a) = \overline{f(a)}$ pour tout $a \in A$.*

Proposition 1.6.1 *Soit f un morphisme prolongeable en $a \in A$. Alors pour tout $k \geq 1$ on a:*

$$f^k(a) = auf(u)f^2(u)\cdots f^{k-1}(u).$$

Preuve: Comme $f(a) = au$, on a $f^2(a) = f(f(a)) = f(au) = f(a)f(u) = auf(u)$. Supposons que l'on ait $f^k(a) = auf(u)f^2(u)\cdots f^{k-1}(u)$; on obtient alors:

$$\begin{aligned} f^{k+1}(a) &= f(f^k(a)) \\ &= f(auf(u)f^2(u)\cdots f^{k-1}(u)) \\ &= f(a)f(u)f(f(u))f(f^2(u))\cdots f(f^{k-1}(u)) \\ &= auf(u)f^2(u)\cdots f^k(u) \end{aligned}$$

□

La suite $a, f(a), f^2(a), \ldots, f^k(a), \ldots$ converge vers un mot infini u. Ce mot est dit engendré par f et est noté $f^\omega(a)$. On écrit alors $f^\omega(a) = \lim f^k(a)$.

Exemple: Considérons, sur l'alphabet binaire $\{a, b\}$, le morphisme de Thue-Morse μ défini par $\mu(a) = ab$ et $\mu(b) = ba$. μ est prolongeable en a et engendre donc un mot infini M appelé mot de Thue-Morse.

$$M = \mu^\omega(a) = abbabaabbaababbabaababbaabbabaabbaababba\cdots.$$

Le mot de Thue-Morse a fait l'objet de nombreuses études, tant arithmétiques que combinatoires (voir par exemple [11], [19]).

Proposition 1.6.2 *Soit f un morphisme prolongeable en a. Alors, pour tout $k \geq 1$, f^k est un morphisme prolongeable engendrant le mot infini $f^\omega(a)$.*

Preuve: f étant prolongeable en a on a $f(a) = au$ et $f^k(a) = auf(u)f^2(u)\cdots f^{k-1}(u)$ où u est un mot fini non vide. En posant $v = uf(u)f^2(u)\cdots f^{k-1}(u)$, on a $f^k(a) = av$ où v est un mot fini non vide car u l'est. Ainsi, f^k est un morphisme prolongeable en a. Le mot infini qu'il engendre est la limite de la suite a, $f^k(a)$, $f^{2k}(a)$, $\ldots$, $f^{nk}(a)$, $\ldots$ qui est une suite extraite de la suite $\left(f^k(a)\right)_k$ dont la limite est le mot engendré par f. Par conséquent f et f^k engendrent le même mot infini $f^\omega(a)$. □

Lorsque u est point fixe d'un morphisme prolongeable en a nous avons un critère de minimalité en fonction de f ([31]).

Théorème 1.6.3 *Soit u un point fixe d'un morphisme f sur un alphabet A. Si a est préfixe de u avec $|f(a)| \geq 2$ et si toutes les lettres de A sont dans u, alors les propriétés suivantes sont équivalentes:*

(i) u est minimal et $\lim |f^k(b)| = +\infty$ pour toutes les lettres b de A.

(ii) il existe $l \leq \#A$ tel que pour tout $b \in A$, $a \in L(f^l(b))$.

(iii) pour tout $b \in A$ il existe $k(b) \in \mathbb{N}$ tel que $a \in L(f^{k(b)}(b))$.

Lorsque $A = \{a,b\}$, T. Tapsoba a donné dans [35] un critère de minimalité très simple.

Théorème 1.6.4 *Soit u un mot infini point fixe d'un morphisme f sur un alphabet $A = \{a,b\}$ tel que a soit préfixe de u et $b \in alph(u)$.*

(i) Supposons f croissante, alors: u est minimale si et seulement si $a \in alph(f(b))$.

(ii) Supposons $|f(a)| \geq 2$ et $f(b) = b$, alors: u est minimale si et seulement si $f(a) \in aA^*a$.

Chapitre 2

Mots et morphismes sturmiens

2.1 Introduction

Un mot sturmien est un mot infini binaire qui possède exactement $n + 1$ facteurs distincts de longueur n. L'étude des mots sturmiens a été initiée par Hedlund et Morse à la fin des années 1930 dans un développement sur les systèmes dynamiques symboliques ([29], [30]). Depuis lors une importante littérature a été consacrée à ces mots, notamment au cours des deux dernières décennies ([5], [6], [7], [10], [21], [34]). On connaît aujourd'hui de multiples caractérisations de ces mots et diverses méthodes pour les générer.

En particulier Morse et Hedlund ont montré que ces mots s'interprètent comme le codage de l'orbite d'un point du cercle unité sous l'action d'une rotation d'angle α irrationnel lorsqu'on partitionne le cercle en deux intervalles de longueurs respectives α et $1 - \alpha$ ([30]). Une autre méthode de construction a été proposée par Arnoux et Rauzy qui ont montré que les mots sturmiens s'obtiennent également en itérant infiniment deux morphismes particuliers ([4], [32]).

Dans ce chapitre nous nous intéressons aux morphismes sturmiens *i.e* les morphismes tels que l'image de tout mot sturmien est un mot sturmien. Ces morphismes ont été largement étudiés ([7], [8], [9], [16], [17], [28], [34]).

Ce chapitre est organisé ainsi qu'il suit. Nous reprenons d'abord (section 2.2) certaines propriétés connues sur ces mots et nous montrons une nouvelle propriété caractéristique de ces derniers (théorème 2.2.7). Dans la section 2.3 nous rappelons une caractérisation constructive des morphismes sturmiens dûe a F. Mignosi et P. Séébold ([28]). Puis nous montrons que si l'image d'un mot par un morphisme sturmien est un mot sturmien alors ce mot l'est aussi.

2.2 Mots sturmiens

Soit u un mot infini sur un alphabet fini A. On a vu au chapitre 1 que la fonction de complexité p est croissante. De plus si u est ultimement périodique alors p est bornée et on peut trouver n tel que $p(n+1) = p(n)$. On en déduit que si u est non ultimement périodique alors sa fonction de complexité p vérifie $p(n) \geq n+1$ puisque p est strictement croissante et que $p(1) \geq 2$. Dans ce paragraphe nous nous intéressons aux mots infinis non ultimement périodiques de complexité minimale.

Définition 2.2.1 *Soit u un mot infini sur un alphabet A. On dit que u est sturmien si sa fonction de complexité p vérifie:*

$$\forall n \in \mathbb{N} : p(n) = n + 1.$$

Tout mot sturmien u est défini sur un alphabet de cardinal 2 car d'après la définition ci-dessus $p(1) = 2$. Pour cela dans ce chapitre, sauf mention contraire, nous travaillerons sur l'alphabet $\{a, b\}$.

Exemple: Le mot sturmien le plus connu est le célèbre mot de Fibonacci

$$F = abaababaabaababaababaabaababaaba\cdots$$

engendré par le morphisme Φ défini par $\Phi(a) = ab$ et $\Phi(b) = a$. Ce mot a été profondément étudié et nombreuses de ses propriétés peuvent être retrouvées dans [5] et [34].

La définition 2.2.1 conduit à:

Proposition 2.2.1 *Soit u un mot sturmien. On a:*

(i) u contient ab et ba.

(ii) u contient un et un seul des mots aa et bb.

Preuve: (i) Si ab n'est pas dans u, alors $u = b^\omega$ ou $u = b^k a^\omega \cdot (k \geq 0)$et u serait ultimement périodique; ce qui est contradictoire avec le caractère sturmien de u. De même on montre que ba est dans u.

(ii) Nous avons $p(2) = 3$. Ainsi u possède exactement 3 facteurs de longueur 2. D'après (i) ab et ba sont deux de ces trois facteurs de longueur 2 . Nécessairement le troisième facteur de longueur 2 de u est soit aa, soit bb. □

Définition 2.2.2 *Soit u un mot sur un alphabet fini A. On dit qu'une lettre a est isolée dans u lorsque $a \in alph(u)$ tandis que aa n'est pas dans u.*

Par exemple, dans le mot de Fibonacci la lettre b est isolée.
Comme conséquence de la proposition 2.2.1, tout mot sturmien possède une et une seule lettre isolée.

Définition 2.2.3 *Un mot sturmien u est dit a-sturmien (resp. b-sturmien) si a (resp. b) n'est pas isolée dans u.*

La remarque suivante vient de la proposition 2.2.1.

Remarque 2.2.1 *Tout mot sturmien u est soit a-sturmien soit b-sturmien.*

Théorème 2.2.2 *Soit u un mot récurrent sur l'alphabet $A = \{a, b\}$. Les assertions suivantes sont équivalentes:*
(i) u est sturmien.
(ii) $\forall n \in \mathbb{N}, \exists! (D_n, G_n) \in L_n^2(u) : \partial^+ D_n = \partial^- G_n = 2$.

Preuve: u sturmien $\Longleftrightarrow (\forall n \in \mathbb{N},\ p(n) = n + 1)$
$\Longleftrightarrow (\forall n \in \mathbb{N},\ p(n + 1) - p(n) = 1)$
$\Longleftrightarrow \left(\forall n \in \mathbb{N},\ p(n + 1) - p(n) = \sum_{w \in L_n(u)}(\partial^+ w - 1) = \sum_{w \in L_n(u)}(\partial^- w - 1) = 1\right)$
$\Longleftrightarrow (\forall n \in \mathbb{N},\ \exists! (D_n, G_n) \in L_n^2(u) : \partial^+ D_n = \partial^- G_n = 2)$ car u est récurrent.

Proposition 2.2.3 *([24]) Tout mot sturmien est uniformément récurrent.*

Définition 2.2.4 *Soit u un mot défini sur A. On dit que u est équilibré si pour tous facteurs v et w de même longueur et toute lettre a de A on a:*

$$||v|_a - |w|_a| \leq 1$$

Proposition 2.2.4 *Soit u un mot sur l'alphabet A. u est non équilibré s'il existe deux lettres distinctes a et b et un mot fini v tels que ava et bvb appartiennent à $L(u)$.*

Preuve: Supposons u non équilibré. Alors on peut considérer deux facteurs w et w' de longueur n minimale de u tels que $|(|w|_a - |w'|_a)| \geq 2$. Comme n est minimal tout couple $(t,\, t')$ de facteurs de u de longueur m inférieure à n vérifie $||t|_a - |t'|_b| \leq 1$. On peut encore écrire $w = xvy$ et $w' = x'v'y'$ avec $x, y\, x'\, y' \in A$. Comme $|xv| = |x'v'| < n$, par minimalité de n nous avons $||xv|_a - |x'v'|_a| \leq 1$. Par conséquent $y \neq y'$. De la même manière on montre que $x \neq x'$. Supposons $x \neq y$. Puisque $y \neq y'$, fixons pour continuer $x = a$ et $x' = b$. Alors $y = b$ et $y' = a$ et donc $w = avb$ et $w' = bv'a$. D'où:

$$||w|_a - |w'|_a| = ||avb|_a - |bv'a|_a| = ||v|_a - |v'|_a| \leq 1$$

Ce qui est absurde, car par hypothèse, $||w|_a - |w'|_a| \geq 2$. Donc $w = ava$ et $w' = bv'b$. Montrons pour terminer que $v = v'$. Si $v \neq v'$ alors nous aurons ($v = tat'$ et $v' = tbt''$) ou ($v = tbt'$ et $v' = tat''$) avec t, t' et t'' des mots non vides. Si $v = tat'$ et $v' = tbt''$ alors ata et btb sont dans u, ce qui est contradictoire avec la minimalité de n. Si $v = tbt'$ et $v' = tat''$ sont dans u alors nous avons:

$$||t'a|_a - |t''b|_a| = ||atbt'a|_a - |btat''b|_a| = ||w|_a - |w'|_a| \geq 2.$$

Ce qui est contradictoire avec la minimalité de n car $|t'a| = |t''b| < n$.

Conséquence: Tout mot u équilibré admet une lettre isolée.

Définition 2.2.5 *On appelle clôture de mots équilibrés, tout ensemble C de mots sur A vérifiant:*

- $\forall v \in C : L(v) \subseteq C$
- $\forall v, w \in C,\ |v| = |w| :\ ||v|_a - |w|_a| \leq 1$

Lemme 2.2.1 *Soit C une clôture de mots équilibrés. On a:*

$$p(n, C) \leq n + 1$$

où $p(n, C) = \#\,\{v \in C :\ |v| = n\}$.

Preuve: Soit C une clôture de mots équilibrés. Supposons qu'il existe n un entier naturel tel que $p(n, \mathrm{C}) \geq n + 2$. Prenons n le plus petit des entiers vérifiant cette condition i.e $p(n-1, \mathrm{C}) \leq n$. Nous avons:

$$s(n, \mathrm{C}) = p(n, \mathrm{C}) - p(n-1, \mathrm{C}) \geq (n+2) - p(n-1, \mathrm{C}) \geq (n+2) - n = 2$$

Ainsi C admet au moins deux éléments (de longueur $n-1$) possédant chacun deux prolongements dans C. Soit D_1 et D_2 deux de ces facteurs. Alors les mots

$$D_1a,\ D_1b,\ D_2a \text{ et } D_2b$$

sont dans C. Comme D_1 et D_2 sont distincts ils existe alors des mots D, D', D'' tels que $D_1 = D'aD$ et $D_2 = D''bD$. Ainsi nous observons que aDa et bDb sont dans C. D'où:

$$||aDa|_a - |bDb|_a| = 2 \geq 1.$$

Donc C n'est pas une clôture de mots équilibrés. Absurde. □

Corollaire 2.2.1 *Soit u un mot infini sur A. Si u est équilibré alors $p(n) \leq n+1$ pour tout entier n.*

Preuve: Il suffit de remarquer que pour tout mot u infini et équilibré $L(u)$ forme une clôture de mots équilibrés puis d'appliquer le lemme 2.2.1 ci-dessus. □

La caractérisation suivante est fort utile ([30]).

Théorème 2.2.5 *(Morse & Hedlund) u est un mot sturmien si et seulement si il est non ultimement périodique et équilibré.*

Théorème 2.2.6 *L'ensemble des facteurs d'un mot sturmien est stable par image miroir, i.e*

$$\forall v \in L(u) : \overline{v} \in L(u)$$

Preuve: Soient n un entier naturel et u un mot sturmien donné. Désignons par $\overline{L(u)}$ l'ensemble des images miroir des éléments de $L(u)$: $\overline{L(u)} = \{\overline{v} :\ v \in L(u)\}$. Posons

$$\mathbf{C} = L(u) \cup \overline{L(u)}$$

et montrons que C est une clôture de mots équilibrés. Soit v et w deux éléments de C de même longueur.

Supposons v et w dans $L(u)$. Alors

$$||v|_a - |w|_a| \leq 1$$

puisque u est sturmien (cf. théorème 2.2.5).

Supposons l'un des deux mots dans $L(u)$ et l'autre dans $\overline{L(u)}$. On peut, arbitrairement, considérer v dans $L(u)$ et w dans $\overline{L(u)}$. Alors $\overline{w}$ est dans $L(u)$ et comme $|w|_a = |\overline{w}|_a$ on a:

$$||v|_a - |w|_a| = ||v|_a - |\overline{w}|_a| \leq 1,$$

puisque u est équilibré.

Supposons v et w dans $\overline{L(u)}$. Alors $\overline{v}$ et $\overline{w}$ sont dans $L(u)$. D'où

$$||v|_a - |w|_a| = ||\overline{v}|_a - |\overline{w}|_a| \leq 1$$

puisque u est équilibré.

Ainsi v et w étant choisis quelconques dans $\mathbf{C}$ on en déduit que $\mathbf{C}$ est une clôture de mots équilibrés. Donc $p(n, \mathbf{C}) = \#\left(L_n(u) \cup \overline{L_n(u)}\right) \leq n + 1$ d'après le lemme 2.2.1. Par suite, $L_n(u) = L_n(u) \cup \overline{L_n(u)}$ puisque $\#L_n(u) = n + 1$ (car u est sturmien) et $L_n(u) \subseteq \left(L_n(u) \cup \overline{L_n(u)}\right)$. Ainsi $\overline{L_n(u)} \subseteq L_n(u)$ et par conséquent $\overline{L(u)} \subseteq L(u)$. □

Théorème 2.2.7 *Soit u un mot a-sturmien. Alors il existe un entier naturel non nul n tel que u s'écrit*

$$u = a^{n_0}ba^{n+\epsilon_1}ba^{n+\epsilon_2}ba^{n+\epsilon_3}b\cdots$$

avec $n_0 \leq n + 1$ et $(\epsilon_i)_{i\geq 1}$ une suite sturmienne sur l'alphabet $\{0, 1\}$.

Preuve: Soit u un mot a-sturmien. Il existe un entier non nul n tel que u est de la forme

$$u = a^{n_0}ba^{n+\epsilon_1}ba^{n+\epsilon_2}ba^{n+\epsilon_3}b\cdots$$

où $n_0 \leq n + 1$ et $(\epsilon_i)_{i\geq 1}$ une suite sur l'alphabet $\{0, 1\}$. En effet, comme u a-sturmien, alors b est isolé dans u. Donc u est de la forme

$$u = a^{n_0}ba^{n_1}ba^{n_2}ba^{n_3}b\cdots.$$

Supposons qu'il existe $i \geq 1$ et $j \geq 1$ tels que $n_j - n_i \geq 2$. Alors $ba^{n_i}b$ et a^{n_i+2} (facteur de a^{n_j}) sont dans $L(u)$, ce qui contredit l'équilibre. Donc pour tous $i \geq 1$ et $j \geq 1$, $|n_j - n_i| \leq 1$. Si on prend $n = \min\left\{(n_i)_{i\geq 1}\right\}$ et $\epsilon_i = n_i - n$, alors $(\epsilon_i) \in \{0, 1\}$ pour $i \geq 1$. Nous avons $n_0 \leq n + 1$ car sinon, a^{n+2} (facteur de a^{n_0}) et ba^nb sont dans $L(u)$, ce qui contredit encore l'équilibre. Il reste donc à prouver que la suite $(\epsilon_i)_{i\geq 1}$ est sturmienne.

Supposons $(\epsilon_i)_{i\geq 1}$ non sturmienne. Alors $(\epsilon_i)_{i\geq 1}$ est ultimement périodique ou non équilibrée.

Si $(\epsilon_i)_{i\geq 1}$ était ultimement périodique alors u le serait aussi, ce qui est impossible car u est sturmien. Par suite, $(\epsilon_i)_{i\geq 1}$ est non équilibrée, et possède de ce fait deux facteurs de la forme $0t0$ et $1t1$. Considérons les facteurs de $(\epsilon_i)_{i\geq 1}$; en les écrivant de telle sorte que chaque lettre précède et succède un b puis en remplaçant 0 par a^n et 1 par a^{n+1} on retrouve certains facteurs de u. Ainsi u admet deux facteurs de la forme ba^nTa^nb et $ba^{n+1}Ta^{n+1}b$. En posant $w = a^nTa^n$ on remarque que les mots awa et bwb sont dans u. Donc u est non équilibré. Absurde, car u est un mot sturmien. Par conséquent $(\epsilon_i)_{i\geq 1}$ est une suite sturmienne.

2.3 Morphismes sturmiens

Définition 2.3.1 *Un morphisme f de A^* dans lui-même est sturmien si $f(x)$ est sturmien pour tout mot sturmien x.*

Le morphisme identité Id et le morphisme E qui échange a et b sont évidemment des morphismes sturmiens.

Proposition 2.3.1 *Les morphismes α et $\overline{\alpha}$ ci-dessous définis sont des morphismes sturmiens.*

$$\begin{array}{llll} \alpha : & A^* \longrightarrow A^* & \qquad \overline{\alpha} : & A^* \longrightarrow A^* \\ & a \longmapsto a & & a \longmapsto a \\ & b \longmapsto ab & & b \longmapsto ba \end{array}$$

Preuve: Montrons que cette proposition est vraie pour α; la preuve se faisant de manière similaire pour $\overline{\alpha}$.

Soit u un mot sturmien.

Supposons $\alpha(u)$ non sturmien. Alors, $\alpha(u)$ est ultimement périodique ou non équilibré.

Si $\alpha(u)$ est ultimement périodique alors il existe deux mots finis v et w et deux facteurs r et s tels que $\alpha(u) = vw^\omega$, $v = \alpha(r)$ et $w = \alpha(s)$. Par suite le mot $\alpha(u)$ s'écrit $\alpha(rs^\omega)$; et comme α est injectif il vient que $u = rs^\omega$. Donc u est ultimement périodique. Absurde.

Si $\alpha(u)$ est non équilibré alors $\alpha(u)$ admet deux facteurs de la forme aza et bzb. Comme les mots $\alpha(a^2)$, $\alpha(ab)$, $\alpha(ba)$ et $\alpha(b^2)$ ne contiennent pas b^2 alors $\alpha(u)$ ne peut contenir b^2. Donc le mot z s'écrit $z = ata$ et $\alpha(u)$ admet comme facteurs les mots $v_1 = aataa$ et $v_2 = abatab$. Comme la première lettre et la troisième occurrence de a ne peut provenir que de la lettre a, on a $at = \alpha(y)$ où y est un facteur de u. Il s'en suit que $\alpha(aya)$ et $\alpha(byb)$ apparaissent dans $\alpha(u)$ et donc aya et byb sont des facteurs de u. Absurde, car u sturmien. □

Corollaire 2.3.1 *Les morphismes $E \circ \alpha$ et $E \circ \overline{\alpha}$ sont sturmiens.*

Preuve: Soit u un mot sturmien. $\alpha(u)$ est un mot sturmien puisque α est un morphisme sturmien (Proposition 2.3.1). Par suite $E(\alpha(u))$ est sturmien car le morphisme échange est aussi sturmien. On montre de même pour $E \circ \overline{\alpha}$. □

Définition 2.3.2 *Un morphisme sturmien f est dit a-sturmien (resp. b-sturmien) si $f(ab)$ ou $f(ba)$ contient a^2 (resp. b^2).*

En vertu de cette définition on fait la remarque suivante.

Remarque 2.3.1 *Tout morphisme sturmien, autre que Id et E, est soit a-sturmien, soit b-sturmien.*

Les morphismes α et $\overline{\alpha}$ sont a-sturmiens tandis que $E \circ \alpha$ et $E \circ \overline{\alpha}$ sont b-sturmiens.

Lemme 2.3.1 *Soient u un mot récurrent sur A et $\varphi \in \{\alpha, \overline{\alpha}\}$. Le mot $\varphi(u)$ est sturmien si et seulement si u est sturmien.*

Preuve: Soient u un mot récurrent sur A et $\varphi \in \{\alpha, \overline{\alpha}\}$. La démonstration étant similaire pour α et $\overline{\alpha}$ nous faisons la preuve pour $\varphi = \alpha$.

Supposons u sturmien alors $\alpha(u)$ est sturmien d'après la proposition 2.3.1 .

Réciproquement supposons u non sturmien et montrons que $\alpha(u)$ ne peut être sturmien. Comme u est non sturmien alors il possède deux facteurs de la forme ata et btb. Nous avons:

$$\alpha(ata) = \alpha(a)\alpha(t)\alpha(a) = a\alpha(t)a = aTa$$

$$\alpha(btb) = \alpha(b)\alpha(t)\alpha(b) = ab\alpha(t)ab = abTab$$

où $T = \alpha(t)$. u étant infini il existe une lettre x dans A telle que $atax$ soit un facteur de u. Par suite nous avons:

$$\alpha(atax) = \alpha(ata)\alpha(x) = aTaay$$

où $y \in \{\varepsilon, b\}$. Ainsi $\alpha(u)$ contient les deux mots awa et bwb avec $w = Ta$. Par conséquent $\alpha(u)$ est non équilibré et ne peut donc être sturmien. $\square$

Lemme 2.3.2 *Soit f un morphisme sturmien. Alors, il existe $g \in \{\alpha, \overline{\alpha}, E \circ \alpha, E \circ \overline{\alpha}\}$ et un morphisme h sturmien tels que $f = g \circ h$.*

Preuve: Soit f un morphisme sturmien.

– Supposons f a-sturmien. Alors $f(a)$ et $f(b)$ sont des facteurs d'un certain mot a-sturmien. Par conséquent $f(a)$ et $f(b)$ appartiennent à $\{a, ab\}^*$ ou à $\{a, ba\}^*$ puisque tout mot a-sturmien est dans $\{a, ab\}^* \cup \{a, ba\}^*$.

Cas 1: $f(a), f(b) \in \{a, ab\}^*$. Comme $\alpha(a) = a$ et $\alpha(b) = ab$ on en déduit que $f(a)$ et $f(b)$ sont dans $\{\alpha(a), \alpha(b)\}^*$. Il existe x et y dans A^* tels que $f(a) = \alpha(x)$ et $f(b) = \alpha(y)$. Soit h la substitution telle que $h(a) = x$ et $h(b) = y$. h se prolonge par concaténation en un morphisme sur A^*. Nous avons $f = \alpha \circ h$.

Le morphisme h est sturmien. En effet soit u un mot sturmien. Alors $f(u) = \alpha \circ h(u) = \alpha(h(u))$ est sturmien puisque f est sturmien. On en déduit par application du lemme 2.3.1, que $h(u)$ est sturmien. Par conséquent h est sturmien.

Cas 2: $f(a), f(b) \in \{a, ba\}^*$. Puisque $\overline{\alpha}(a) = a$ et $\overline{\alpha}(b) = ba$, on montre comme dans le cas précédent qu'il existe un morphisme h sturmien tel que $f = \overline{\alpha} \circ h$.

– Supposons f b-sturmien. Alors $E \circ f$ est a-sturmien puisque le morphisme E échange les rôles de a et b. Par suite il existe un élément g dans $\{\alpha, \overline{\alpha}\}$ et un morphisme h sturmien tels que $E \circ f = g \circ h$. Or $f = E \circ (E \circ f)$. D'où $f = E \circ (g \circ h) = g' \circ h$ avec $g' \in \{E \circ \alpha, E \circ \overline{\alpha}\}$. $\square$

Théorème 2.3.2 *Soit f un morphisme sur A^*. Alors, f est sturmien si et seulement si $f \in \{E, \alpha, \overline{\alpha}\}^*$*

Preuve: *Suffisance:* Si $f \in \{E, \alpha, \overline{\alpha}\}^*$ alors f est sturmien car la composition des morphismes préserve les morphismes sturmiens.

Nécessité: Soit f un morphisme sturmien. Montrons par récurrence sur $\|f\|$, la longueur de f ($\|f\| = \sum_{a\in A} |f(a)|$), que nous avons:

$$f \in \{E, \alpha, \overline{\alpha}\}^*$$

- Si $\|f\| = 2$ alors $f \in \{Id, E\}$, puisque f est non effaçant en tant que morphisme srurmien. Donc $f \in \{E, \alpha, \overline{\alpha}\}^*$

Supposons que $\|f\| \geq 3$ et que tout morphisme sturmien de longueur strictement inférieure à $\|f\|$ appartienne à $\{E, \alpha, \overline{\alpha}\}^*$. D'après le lemme 2.3.2 il existe $g \in \{\alpha, \overline{\alpha}, E \circ \alpha, E \circ \overline{\alpha}\}$ et h un morphisme sturmien tels que $f = g \circ h$. Vérifions que $\|h\| < \|f\|$. Pour ce faire, prenons arbitrairement $g = \alpha$. Comme $|\alpha(a)| = 1$ et $|\alpha(b)| = 2$, nous avons :

$$\begin{aligned}
\|f\| &= \Big(|h(a)|_a + 2\,|h(a)|_b\Big) + \Big(|h(b)|_a + 2\,|h(b)|_b\Big) \\
&= \Big(|h(a)| + |h(a)|_b\Big) + \Big(|h(b)| + |h(b)|_b\Big) \\
&= \Big(|h(a)| + |h(b)|\Big) + \Big(|h(a)|_b + |h(b)|_b\Big) \\
&= \|h\| + |h(ab)|_b
\end{aligned}$$

Comme h est sturmien, $h(ab)$ contient nécessairement un b. Donc $|h(ab)|_b \geq 1$ et par conséquent $\|h\| < \|f\|$.

Donc, par hypothèse de récurrence, $h \in \{E, \alpha, \overline{\alpha}\}^*$. Ainsi, $f = g \circ h \in \{E, \alpha, \overline{\alpha}\}^*$.

□

Théorème 2.3.3 *Soient x un mot récurrent et f un morphisme sturmien sur A^*. Alors, $f(x)$ est sturmien si et seulement si x est sturmien.*

Preuve: Ce théorème est une conséquence immédiate du lemme 2.3.1 et du théorème 2.3.2. □

Chapitre 3

Mots et morphismes d'Arnoux-Rauzy

3.1 Introduction

Un mot d'Arnoux-Rauzy est un mot (sur un alphabet à trois lettres) uniformément récurrent de complexité $2n+1$ caractérisé par une certaine condition combinatoire (voir définition plus loin). L'étude de ces mots a été introduite par G. Rauzy dans [32]. Plus tard P. Arnoux et G. Rauzy ont donné une caractérisation géométrique de ces mots ([4]), ce qui a valu d'ailleurs leurs noms à l'appellation actuelle de cette classe de mots. Depuis lors une importante littérature a été publiée sur leurs propriétés (voir par exemple [14], [38]).

Les graphes des mots d'Arnoux-Rauzy sont souvent vus comme une généralisation de ceux des mots sturmiens. Ceci a permis de nombreuses généralisations des propriétés combinatoires des mots sturmiens à ceux d'Arnoux-Rauzy.

F. Mignosi et P. Séébold ([28]) ont obtenu une caractérisation des morphismes sturmiens. Nous nous proposons dans ce chapitre de donner une caractérisation des morphismes d'Arnoux-Rauzy qui sont des morphismes tels que l'image de tout mot d'Arnoux-Rauzy est d'Arnoux-Rauzy.

Dans ce qui suit, après avoir défini et énoncé quelques propriétés des mots d'Arnoux-Rauzy nous étudions les morphismes d'Arnoux-Rauzy et nous montrons qu'ils sont engendrés par une famille de quatre morphismes.

Les principaux résultats de ce chapitre sont dans [25].

3.2 Mots d'Arnoux-Rauzy

Dans tout ce chapitre l'alphabet considéré est $A = \{a, b, c\}$.

Définition 3.2.1 *Soit u un mot sur un alphabet à trois lettres de complexité $p(n) = 2n+1$. On dit que u est un mot d'Arnoux-Rauzy si pour tout nombre entier naturel n, il existe un unique facteur de longueur n triprolongeable à droite D_n, et un unique facteur de longueur n triprolongeable à gauche G_n (i.e $\partial^- G_n = \partial^+ D_n = 3$). On dira simplement que u est un AR-mot.*

L'exemple le plus classique de mot d'Arnoux-Rauzy est le mot de Tribonacci ([15]), engendré par le morphisme σ défini sur l'alphabet A par $\sigma(a) = ab$, $\sigma(b) = ac$ et $\sigma(c) = a$.

Les mots d'Arnoux-Rauzy sont tous infinis et non ultimement périodiques car leur fonction de complexité est non bornée.

Remarque 3.2.1 *Tout facteur v d'un AR-mot u vérifie*

$$\partial^- v = 1 \text{ ou } 3 \text{ et } \partial^+ v = 1 \text{ ou } 3.$$

En d'autres termes aucun facteur d'un AR-mot ne peut être biprolongeable. Dans [24] P. Hubert a démontré le théorème suivant.

Théorème 3.2.1 *Tout mot d'Arnoux-Rauzy est uniformément récurrent.*

Proposition 3.2.2 *Tout mot d'Arnoux-Rauzy admet exclusivement l'un des trois mots a^2, b^2 et c^2 en facteur.*

Preuve: Soit u un mot d'Arnoux-Rauzy. Alors les trois lettres sont présentes dans u car $p(1) = 3$. Il existe donc une et une seule lettre x dans $\{a, b, c\}$ telle que $\partial^+ x = 3$. On peut prendre sans restreindre la généralité $x = a$. Ainsi a^2, ab, et ac sont facteurs de u.

Supposons b^2 et c^2 dans u. Alors les cinq facteurs de longueur 2 sont a^2, b^2, c^2, ab et ac. Ainsi, le seul prolongement á droite de b est b; donc u s'écrit de la forme wb^ω et ultimement périodique. Absurde, car u est un AR-mot .

Supposons b^2 dans u et pas c^2. Alors les cinq facteurs de longueur 2 sont soit a^2, b^2, ab, ac et cb, soit a^2, b^2, ab, ac et ca.

Dans le premier cas le seul prolongement á droite de b est b, donc on conclut comme dans la première supposition. Dans le deuxième cas on a $\partial^- b = 2$. Ce qui est impossible en vertu de la remarque 3.2.1. Les mêmes arguments restent valables si c^2 dans u et pas b^2.

Finalement u contient un seul carré de lettre. □

Cette proposition nous permet de scinder les mots d'Arnoux-Rauzy en trois classes à travers la définition suivante.

Définition 3.2.2 *Un mot d'Arnoux-Rauzy u est un AR_a-mot (resp. AR_b-mot, AR_c-mot) si a^2 (resp. b^2, c^2) est facteur de u.*

Le mot de tribonacci est un AR_a-mot.

Proposition 3.2.3 *Soit u un mot d'Arnoux-Rauzy. Les cinq facteurs de longueur 2 sont du type*

$$x^2,\ xy,\ xz,\ yx \text{ et } zx$$

où $\{x, y, z\} = \{a, b, c\}$.

preuve: On peut prendre sans restreindre la généralité $x = a$, $y = b$ et $z = c$. u contient a^2 et donc ne contient ni b^2 ni c^2, en vertu de la proposition 3.2.2. De plus $\partial^+ a = 3$, $\partial^+ b = 1$ et $\partial^+ c = 1$. Donc ba ou bien bc (resp. ca ou bien cb) est dans u.
Si bc (resp. cb) est dans u alors $\partial^- c = 2$ (resp. $\partial^- b = 2$). Ce qui est contradictoire car u est un AR-mot. Par conséquent ba et ca sont facteurs de u. Donc a^2, ab, ac, ba et ca sont les cinq facteurs de longueur 2 de u. □

La proposition suivante nous donne une forme générale des mots d'Arnoux-Rauzy.

Proposition 3.2.4 *Pour tout AR_a-mot u, il existe un entier naturel non nul n tel que u s'écrit sous la forme*

$$u = a^{n_0} x_1 a^{n+\epsilon_1} x_2 a^{n+\epsilon_2} x_3 a^{n+\epsilon_3} \cdots$$

où $0 \leq n_0 \leq n+1$, $x_i \in \{b, c\}$ et $\epsilon_i \in \{0, 1\}$ pour $i = 1, 2, 3, \ldots$.

Preuve: Soit u un AR_a-mot. u contient a^2 et comme il est uniformément récurrent et non ultimement périodique il admet une puissance maximale de a (laquelle puissance étant supérieure ou égale à 2). Ainsi il existe un entier naturel non nul n tel que cette puissance soit égale à $n + 1$. Alors $xa^{n+1}y$ est dans $L_{n+3}(u)$ avec $x \neq a$ et $y \neq a$. Il s'en suit que

$$a^{k+1},\ xa^k \text{ et } a^k y$$

sont dans $L_{n+1}(u)$ pour tout $k \in \{1, 2, \dots, n\}$. D'où

$$\partial^- a^k \geq 2 \text{ et } \partial^+ a^k \geq 2$$

et par la remaque 3.2.1 on conclut que

$$\partial^- a^k = \partial^+ a^k = 3,\ k = 1, 2, \dots, n.\ (\star)$$

u admet un facteur de la forme $xa^m y$ avec $m < n+1$ et $x, y \in \{b, c\}$. En effet dans le cas contraire u s'écrit

$$u = a^{n_0} x_1 a^{n+1} x_2 a^{n+1} x_3 a^{n+1} \cdots$$

où $0 \leq n_0 \leq n+1$ et $x_i \in \{b, c\}$ pour $i = 1, 2, 3, \dots$. Or u étant un AR-mot, les lettres b et c sont dans u et donc $a^{n+1}b$ et $a^{n+1}c$ sont dans $L(u)$. Par conséquent $\partial^+ a^{n+1} = 2$ car $a^{n+2} \notin L(u)$ puisque $n+1$ est la puissance maximale de a dans u. Ce qui est contradictoire car u étant un AR-mot, nécessairement $\partial^+ a^{n+1} = 1$ ou $\partial^+ a^{n+1} = 3$ (cf. Remarque 3.2.1).

Supposons $m < n$. Alors $m + 1 \leq n$ et donc

$$\partial^- a^{m+1} = \partial^+ a^{m+1} = 3,$$

d'après $(\star)$. Par conséquent xa^{m+1} est dans $L_{m+2}(u)$ et comme $xa^m y$ est dans $L_{m+2}(u)$ il s'en suit que $\partial^+ xa^m \geq 2$ d'où $\partial^+ xa^m = 3$ (Remarque 3.2.1). Ce qui est impossible car u étant un AR-mot, a^{m+1} est l'unique facteur spécial à droite de u de longueur $m+1$. Par conséquent $m = n$.

En définitive u s'écrit sous la forme

$$u = a^{n_0} x_1 a^{n+\epsilon_1} x_2 a^{n+\epsilon_2} x_3 a^{n+\epsilon_3} \cdots$$

où $0 \leq n_0 \leq n+1$, $x_i \in \{b, c\}$ et $\epsilon_i \in \{0, 1\}$ pour $i = 1, 2, 3, \dots$. □

Proposition 3.2.5 *Soit u un AR_a-mot. Alors a^{n+1} où $n+1$ est la puissance maximale de a, admet la même lettre comme prolongement à gauche et à droite dans u.*

Preuve: Comme $n+1$ est la puissance maximale de a dans u on a:

$$\partial^- a^{n+1} = \partial^+ a^{n+1} = 1.$$

Par conséquent il existe une et une seule lettre dans $\{b, c\}$ précédant (resp. suivant) a^{n+1} dans u. Il en résulte que

$$xa^{n+1}y \in L_{n+3}(u)$$

où $x, y \in \{b, c\}$.

Supposons $x \neq y$. Alors ya^{n+1} et $a^{n+1}x$ ne sont pas dans $L_{n+2}(u)$. Cependant ya^n et $a^n x$ sont dans $L_{n+1}(u)$ puisque $\partial^- a^n = \partial^+ a^n = 3$. Par suite,

$$\partial^+ xa^n = \partial^- a^n y = 3$$

car l'unique facteur spécial à droite (resp. à gauche) de longueur $n + 1$ de u finit (resp. commence) par a^n alors que $\partial^+ a^{n+1} = \partial^- a^{n+1} = 1$. Il en résulte que les mots

$$xa^n x, xa^n y, xa^{n+1}, ya^n y \text{ et } a^{n+1} y$$

appartiennent à $L_{n+2}(u)$.

D'autre part, $\partial^+ xa^n = 3$ implique $\partial^+ ya^n = 1$. Par suite y est l'unique prolongement de ya^n dans u et donc $ya^n x \notin L_{n+2}(u)$. Par conséquent ya^n est nécessairement suivi de ya^n et alors u est ultimement périodique. Absurde, car u est un AR-mot. □

3.3 Morphismes d'Arnoux-Rauzy

Définition 3.3.1 *Un morphisme f sur A^* est dit d'Arnoux-Rauzy, si pour tout mot d'Arnoux-Rauzy u, $f(u)$ est un mot d'Arnoux-Rauzy. Plus simplement on dira que f est un AR-morphisme.*

Les six morphismes ci-dessous définis sont évidemment des AR-morphismes.

$$\begin{array}{llll} Id: & A^* \longrightarrow A^* & E: & A^* \longrightarrow A^* & E': & A^* \longrightarrow A^* \\ & a \longmapsto a & & a \longmapsto c & & a \longmapsto b \\ & b \longmapsto b & & b \longmapsto a & & b \longmapsto c \\ & c \longmapsto c & & c \longmapsto b & & c \longmapsto a \\ E_a: & A^* \longrightarrow A^* & E_b: & A^* \longrightarrow A^* & E_c: & A^* \longrightarrow A^* \\ & a \longmapsto a & & a \longmapsto c & & a \longmapsto b \\ & b \longmapsto c & & b \longmapsto b & & b \longmapsto a \\ & c \longmapsto b & & c \longmapsto a & & c \longmapsto c \end{array}$$

Pour la suite nous désignerons par AR-morphisme trivial sur A^* tout élément de l'ensemble $\{Id, E, E', E_a, E_b, E_c\}$.

Lemme 3.3.1 *Les relations suivantes sont vérifiées:*

$$Id = E \circ E \circ E,$$

$$E' = E \circ E,$$

$$E_b = E \circ E_a,$$

$$E_c = E_a \circ E.$$

Preuve:

$$E \circ E \circ E(a) = E \circ E(c) = E(b) = a, \tag{3.1}$$

$$E \circ E \circ E(b) = E \circ E(a) = E(c) = b, \tag{3.2}$$

$$E \circ E \circ E(c) = E \circ E(b) = E(a) = c. \tag{3.3}$$

(3.1), (3.2) et (3.3) $\Longrightarrow Id = E \circ E \circ E$.

$$E \circ E(a) = b = E'(a), \tag{3.4}$$

$$E \circ E(b) = c = E'(b), \tag{3.5}$$

$$E \circ E(c) = a = E'(c). \tag{3.6}$$

(3.4), (3.5) et (3.6) $\Longrightarrow E' = E \circ E$.

$$E \circ E_a(a) = E(a) = c = E_b(a), \tag{3.7}$$

$$E \circ E_a(b) = E(c) = b = E_b(b), \tag{3.8}$$

$$E \circ E_a(c) = E(b) = a = E_b(c). \tag{3.9}$$

(3.7), (3.8) et (3.9) $\Longrightarrow E_b = E \circ E_a$.

$$E_a \circ E(a) = E_a(c) = b = E_c(a), \tag{3.10}$$

$$E_a \circ E(b) = E_a(a) = a = E_c(b), \tag{3.11}$$

$$E_a \circ E(c) = E_a(b) = c = E_c(c). \tag{3.12}$$

(3.10), (3.11) et (3.12) $\Longrightarrow E_c = E_a \circ E$.
Il en résulte que l'ensemble des AR-morphismes triviaux est engendré par $\{E, E_a\}$. $\square$

Remarque 3.3.1 *L'ensemble des AR-morphismes triviaux étant isomorphe à S_3, nous retrouvons ici une propriété bien connue sur ce groupe de permutations; il est engendré par une permutation circulaire et une transposition.*

Proposition 3.3.1 *Pour tout AR-morphisme non trivial f, un et un seul des mots a^2, b^2 et c^2 est facteur de $f(abcacba)$.*

Preuve:
- $f(abcacba)$ contient a^2 ou b^2 ou c^2.
Supposons qu'aucun des mots a^2, b^2 et c^2 ne soit facteur de $f(abcacba)$.
Soit u un AR_a-mot. Les cinq facteurs de longueur 2 de u sont

$$a^2,\ ab,\ ac,\ ba \text{ et } ca.$$

Comme le mot

$$f(abcacba) = f(a)f(bc)f(ac)f(ba) = f(ab)f(ca)f(cb)f(a)$$

ne contient aucun carré de lettre alors il en est de même pour les mots

$$f(ab),\ f(ac),\ f(ba) \text{ et } f(ca).$$

Or f étant un AR-morphisme, alors $f(u)$ est un AR-mot et contient par conséquent un carré de lettre. Un tel carré est donc dans $f(a^2)$ car tout facteur de longueur 2 de $f(u)$ est nécessairement dans l'image d'un facteur de longueur 2 de u. Par suite $f(a)$ commence et finit par la même lettre puisque $f(a)$ ne contient pas de carré de lettre.
En considérant un AR_b-mot (resp. AR_c-mot) on montre de même que $f(b)$ (resp. $f(c)$) commence et finit par la même lettre.
$f(a) = x$ ou $f(a) = xrx$, $f(b) = y$ ou $f(b) = ysy$ et $f(c) = z$ ou $f(c) = ztz$
où r,s et t sont des mots finis et x, y et z sont des lettres différentes.
Comme f est non trivial nous avons, à une permutation près d'images, l'un des trois cas suivants:

Cas 1: $f(a) = xrx$, $f(b) = ysy$ et $f(c) = ztz$

Cas 2: $f(a) = x$, $f(b) = ysy$ et $f(c) = ztz$

Cas 3: $f(a) = x$, $f(b) = y$ et $f(c) = ztz$

Chacun des trois cas conduit à une contradiction

En effet:

Cas 1: $f(a) = xrx$, $f(b) = ysy$ et $f(c) = ztz$

Considérons un AR-mot u dont les facteurs de longueur 2 sont a^2, ab, ba, ac et ca. Il s'en suit que

$$xrx^2rx,\ xrxysy,\ ysyxrx,\ xrxztz \text{ et } ztzxrx$$

sont facteurs de $f(u)$ qui est un AR-mot. $f(u)$ contient x^2 et donc ne contient ni yz ni zy (cf. Proposition 3.2.3). Par conséquent

$$t = x \text{ ou } t = xt'x \quad (\star)$$

où t' est dans A^*

Considérons maintenant un AR-mot u' dont les facteurs de longueur 2 sont b^2, ba, ab, bc et cb. Alors les mots

$$ysy^2sy,\ ysyxrx,\ xrxysy,\ ysyztz \text{ et } ztzysy$$

sont présents dans $f(u')$ qui est du reste un AR-mot. Comme y^2 est dans $f(u')$ alors les mots xz et zx ne sont pas dans $f(u')$. Il en résulte que

$$t = y \text{ ou } t = yt'y \quad (\star\star)$$

où t' est dans A^*

$(\star)$ et $(\star\star)$ impliquent que t commence (ou finit) par deux lettres différentes en l'occurrence x et y. Absurde.

Cas 2: $f(a) = x$, $f(b) = ysy$ et $f(c) = ztz$ Considérons un AR-mot u dont les facteurs de longueur 2 sont a^2, ab, ba, ac et ca. Alors les mots

$$x^2,\ xysy,\ ysyx,\ xztz \text{ et } ztzx$$

sont facteurs du AR-mot $f(u)$. Les mots yz et zy ne sont pas facteurs de $f(u)$ puisque x^2 est dans $f(u)$. Donc

$$t = x \text{ ou } t = xt'x \quad (\star\star\star)$$

où t' est dans A^*

De même en prenant un AR-mot u' contenant b^2 on obtient

$t = y$ ou $t = yt'y$. $(\star\star\star\star)$

$(\star\star\star)$ et $(\star\star\star\star)$ impliquent que t commence (ou finit) par par deux lettres différentes en l'occurrence x et y. Absurde.

Cas 3: $f(a) = x$, $f(b) = y$ et $f(c) = ztz$

De manière analogue aux cas précédents on montre que t commence ou finit par deux lettres différentes qui sont x et y. Absurde.

Dans tous les cas $f(abcacba)$ contient a^2 ou b^2 ou c^2 en facteur.

– $f(abcacba)$ contient un et un seul des trois mots a^2, b^2 et c^2.

Supposons a^2 et b^2 facteurs de $f(abcacba)$. Comme

$$f(abcacba) = f(a)f(bc)f(ac)f(ba) = f(ab)f(ca)f(cb)f(a),$$

alors il existe deux paires de lettres $\{x, y\}$ et $\{x', y'\}$ telles que $f(xy)$ contienne a^2 et $f(x'y')$ contienne b^2. Ces deux paires étant issues de $\{a, b, c\}$ elles comportent au moins une lettre commune. Donc les mots xy et $x'y'$ vont appartenir à un même mot d'Arnoux-Rauzy (cf. Proposition 3.2.3). Considérons un AR-mot u contenant xy et $x'y'$. Il en résulte que a^2 et b^2 sont facteurs de $f(u)$. Par suite $f(u)$ n'est pas un AR-mot en vertu de la proposition 3.2.3. Absurde, car f est un AR-morphisme.

Donc $f(abcacba)$ contient un et un seul des mots a^2, b^2 et c^2. □

La proposition ci-dessus nous fournit donc une condition nécessaire pour qu'un morphisme non trivial soit un AR-morphisme.

Définition 3.3.2 *Un morphisme d'Arnoux-Rauzy f est un AR_a-morphisme (resp. AR_b-morphisme, AR_c-morphisme) si $f(u)$ est un AR_a-mot (resp. AR_b-mot, AR_c-mot) pour tout AR-mot u.*

Proposition 3.3.2 *(i) Tout AR-morphisme non trivial f est un AR_a-morphisme (resp. AR_b-morphisme, AR_c-morphisme) si $f(abcacba)$ contient a^2 (resp. b^2, c^2).*

(ii) Tout AR-morphisme non trivial est soit un AR_a-morphisme, soit un AR_b-morphisme, soit un AR_c-morphisme.

Preuve: (i) Soit f un AR-morphisme non trivial. Sans perdre la généralité on peut supposer b^2 facteur de $f(abcacba)$. Alors $f(a)$, $f(b)$ et $f(c)$ ne contiennent ni a^2 ni c^2.

Soit u un AR_a-mot tel que $f(u)$ ne soit pas un AR_b-mot. On a:

$$L_2(u) = \{a^2, ab, ba, ac, ca\}$$

et

$$L_2(f(u)) = \{x^2, xy, yx, xb, bx\}$$

où $x, y \in \{a, c\}$.

D'une part les mots $f(a^2)$, $f(ab)$, $f(ba)$, $f(ac)$ et $f(ca)$ ne contiennent pas b^2 car sinon b^2 serait dans $L_2(f(u))$. Par conséquent b^2 est dans $f(bc)$ ou $f(cb)$ puisque b^2 est dans $f(abcacba)$.

D'autre part $f(a)$, $f(ab)$, $f(ba)$, $f(ac)$ et $f(ca)$ ne contiennent pas x^2 puisque x^2 n'est pas dans $f(abcacba)$ qui contient déjà b^2 (cf. Proposition 3.3.1). Par suite $f(a)$ commence et finit par x i.e $f(a) = x$ ou $f(a) = xrx$ car x^2 est nécessairement dans $f(a^2)$.

Considérons u' un AR_z-mot où z est quelconque dans $\{b,c\}$ On a:

$$L_2(u') = \{z^2, az, za, zy, yz\} \text{ avec } y \in \{b, c\},$$

et

$$L_2(f(u')) = \{b^2, ba, ab, bc, cb\}.$$

Il en résulte que les mots $f(a)$, $f(az)$ et $f(za)$ (où z est quelconque dans $\{b, c\}$) ne contiennent pas ac. Par ailleurs

$$f(a^2) = x^2 \text{ ou } f(a^2) = xrx^2rx$$

car

$$f(a) = x \text{ ou } f(a) = xrx$$

où $x \in \{a, c\}$. Par conséquent ac n'est pas dans $f(a^2)$ puisque $x^2 \neq ac$.

En somme $f(a^2)$, $f(ab)$, $f(ba)$, $f(ac)$ et $f(ca)$ ne contiennent pas ac; ce qui est contradictoire car ac est dans $f(u)$ par supposition. Donc $f(u)$ est un AR_b-mot pour tout AR_a-mot u.

De même pour tout AR_b-mot ou AR_c-mot u', $f(u')$ est un AR_b-mot. On obtient ainsi l'assertion (i).

(ii) Soit f un AR-morphisme non trivial. Alors $f(abcacba)$ contient un et un seul des carrés a^2, b^2 et c^2 (Proposition 3.3.1). Par suite d'après (i) f est soit un AR_a-morphisme, soit un AR_b-morphisme, soit un AR_c-morphisme. $\square$

Considérons les six morphismes ci-dessous qui ont été introduits par G. Rauzy dans [29].

$$\begin{array}{llllll}
\alpha : & A \longrightarrow A^* & \overline{\alpha} : & A \longrightarrow A^* & \beta : & A \longrightarrow A^* \\
& a \longmapsto a & & a \longmapsto a & & a \longmapsto ba \\
& b \longmapsto ab & & b \longmapsto ba & & b \longmapsto b \\
& c \longmapsto ac & & c \longmapsto ca & & c \longmapsto bc \\
\overline{\beta} : & A \longrightarrow A^* & \gamma : & A \longrightarrow A^* & \overline{\gamma} : & A \longrightarrow A^* \\
& a \longmapsto ab & & a \longmapsto ca & & a \longmapsto ac \\
& b \longmapsto b & & b \longmapsto cb & & b \longmapsto bc \\
& c \longmapsto cb & & c \longmapsto c & & c \longmapsto c
\end{array}$$

Comme nous allons le montrer les six morphismes ci-dessus sont des AR-morphismes.

Lemme 3.3.2 *Soient $\varphi \in \{\alpha, \overline{\alpha}, \beta, \overline{\beta}, \gamma, \overline{\gamma}\}$ et u un AR-mot. Alors, pour tout entier naturel n, $\varphi(u)$ admet un facteur de longueur n triprolongeable à droite et un facteur de longueur n triprolongeable à gauche.*

Preuve: La démonstration étant similaire pour chacun des six morphismes nous faisons la preuve pour $\varphi = \alpha$.

Comme u est un AR-mot il existe un unique facteur D_n dans $L_n(u)$ tel que $\partial^+ D_n = 3$. Ainsi $D_n a$, $D_n b$ et $D_n c$ sont dans $L_{n+1}(u)$. Il en résulte que les mots

$$\alpha(D_n)a,\ \alpha(D_n)ab \text{ et } \alpha(D_n)ac$$

sont dans $\alpha(u)$.

Par ailleurs il existe une lettre x dans A telle que:

$$D_n a x \in L_{n+2}(u).$$

Donc

$$\alpha(D_n a x) = \alpha(D_n) a a y \in L(\alpha(u))$$

avec $y \in \{\varepsilon, b, c\}$. D'où

$$\alpha(D_n) a a \in L(\alpha(u)).$$

Ce qui prouve que
$$\partial^+\alpha(D_n)a = 3.$$
Remarquant que $|\alpha(D_n)a| > n$ on déduit que le suffixe D'_n de longueur n du mot $\alpha(D_n)a$ vérifie $\partial^+ D'_n = 3$ dans $\alpha(u)$.
De la même manière on montre que $\alpha(u)$ admet un facteur G'_n de longueur n tel que $\partial^- G'_n = 3$. $\square$

Lemme 3.3.3 *Soient $\varphi \in \{\alpha, \overline{\alpha}, \beta, \overline{\beta}, \gamma, \overline{\gamma}\}$ et u un AR-mot. Alors, pour tout entier naturel n, $\varphi(u)$ admet un unique facteur de longueur n triprolongeable à droite et un unique facteur de longueur n triprolongeable à gauche.*

Preuve: Sans perdre de vue la généralité nous faisons la preuve avec $\varphi = \alpha$ comme au lemme 3.3.2.
Soit $F_n \in L_n(\alpha(u))$, (n assez grand) avec $\partial^+ F_n = 3$ tel que F_n non suffixe de $\alpha(D_n)a$ où $D_n \in L_n(u)$ et $\partial^+ D_n = 3$. Alors on peut écrire
$$F_n = FxF_0a \text{ et } \alpha(D_n)a = DyF_0a$$
où $D, F, F_0 \in L(\alpha(u))$ et $x, y \in \{a, b, c\}$ avec $x \neq y$.
- Supposons $x, y \in \{b, c\}$. On peut prendre $x = b$ et $y = c$. Ainsi:
$$F_n = F_1abF_0a = F_1(ab)F_0a = F_1\alpha(b)\alpha(D_0)a = F_1\alpha(bD_0)a$$
et
$$\alpha(D_n)a = D_1acF_0a = D_1(ac)F_0a = D_1\alpha(c)\alpha(D_0)a = D_1\alpha(cD_0)a$$
avec D_0 suffixe de D_n, D_1 préfixe de $\alpha(D_n)$ et F_1 préfixe de F_n.
D'une part F_na, F_nb et F_nc sont dans $L_{n+1}(\alpha(u))$ car $\partial^+\alpha(F_n) = 3$. D'où:
$F_na = F_1\alpha(bD_0)aa = F_1\alpha(bD_0)\alpha(a)a = F_1\alpha(bD_0a)a$,
$F_nb = F_1\alpha(bD_0)ab = F_1\alpha(bD_0)\alpha(b) = F_1\alpha(bD_0b)$,
$F_nc = F_1\alpha(bD_0)ac = F_1\alpha(bD_0)\alpha(c) = F_1\alpha(bD_0c)$.
Ainsi bD_0a, bD_0b et bD_0c sont dans $L(u)$ et donc $\partial^+bD_0 = 3$. D'autre part $\alpha(D_n)aa$, $\alpha(D_n)ab$ et $\alpha(D_n)ac$ sont dans $L_{n+1}(\alpha(u))$ car $\partial^+\alpha(D_n)a = 3$. D'où:
$\alpha(D_n)aa = D\alpha(cD_0)aa = D_1\alpha(cD_0)\alpha(a)a = D_1\alpha(cD_0a)a$,
$\alpha(D_n)ab = D\alpha(cD_0)ab = D_1\alpha(cD_0)\alpha(b) = D_1\alpha(cD_0b)$,
$\alpha(D_n)ac = D\alpha(cD_0)ac = D_1\alpha(cD_0)\alpha(c) = D_1\alpha(cD_0c)$.

Ainsi cD_0a, cD_0b et cD_0c sont dans $L(u)$ et donc $\partial^+ cD_0 = 3$.
Il ressort que le AR-mot u contient deux facteurs spéciaux à droite de même longueur, en l'occurrence bD_0 et cD_0. Absurde, donc F_n est suffixe de $\alpha(D_n)a$.

- Supposons $x = a$ et $y = b$ ou c. On peut prendre $x = a$ et $y = b$ sans perdre la généralité et on obtiendra les égalités:

$$\alpha(D_n)a = DabF_0a = D(ab)(F_0)a = Dab\alpha(D_0)a = D\alpha(bD_0)a$$

et

$$F_1aF_0a = F_1a(F_0)a = F_1a\alpha(D_0)a = F_1\alpha(aD_0)a.$$

De la même manière que dans la supposition précédente on montre que les facteurs aD_0 et bD_0 vérifient

$$\partial^+ aD_0 = 3 \text{ et } \partial^+ bD_0 = 3$$

dans u. En d'autres termes le AR-mot u admet deux facteurs spéciaux à droite de même longueur. Absurde.

Finalement $\alpha(u)$ admet un et un seul facteur de longueur n triprolongeable à droite pour tout entier naturel n.
De même on vérifie que $\alpha(u)$ contient un et un seul facteur de longueur n triprolongeable à gauche. □

Lemme 3.3.4 *Soient $\varphi \in \{\alpha, \overline{\alpha}, \beta, \overline{\beta}, \gamma, \overline{\gamma}\}$ et u un AR-mot. Alors, $\varphi(u)$ n'admet pas de facteur biprolongeable.*

Preuve: Pour les mêmes raisons qu'au lemme 3.3.2 nous faisons la preuve pour $\varphi = \alpha$. Soit $F_n \in L_n(\alpha(u))$ avec $\partial^+\alpha(F_n) = 2$. Alors il existe $x, y \in \{a, b, c\}$ avec $x \neq y$ tels que :

$$F_nx, F_ny \in L_{n+1}(\alpha(u)).$$

- Supposons $x = a$ et $y = b$ i.e $F_na, F_nb \in L_{n+1}(\alpha(u))$. D'une part F_n finit par a car b ne succède qu' à la lettre a dans $\alpha(u)$. D'autre part on peut prendre F_n de sorte qu'il commence par a. Par suite

$$F_na = aFaa = \alpha(Da)a \text{ et } F_nb = aFab = \alpha(Db)$$

où $aF = \alpha(D)$ avec $Da, Db \in L(u)$.
D'où $\partial^+\alpha(D) \geq 2$ et comme u est un AR-mot alors $\partial^+\alpha(D) = 3$ et donc $Dc \in L(u)$. Ainsi:

$$\alpha(Dc) = \alpha(D)ac = F_nc \in L_{n+1}(\alpha(u)).$$

Ce qui prouve que

$$\partial^+\alpha(F_n) = 3.$$

Ce qui est contradictoire car par supposition $\partial^+\alpha(F_n) = 2$.

Supposons $F_nb, F_nc \in L_{n+1}(\alpha(u))$. Comme précédemment on peut prendre F_n tel que sa première lettre soit a. De plus F_n finit par a pour des raisons évoquées ci-dessus. D'où

$$F_nb = aFab = \alpha(Db) \text{ et } F_nc = aFac = \alpha(Dc)$$

où $aF = \alpha(D)$ avec $Db, Dc \in L(u)$. D'où $\partial^+D \geq 2$ et comme u est un AR-mot alors $\partial^+D = 3$ et donc $Da \in L(u)$. Par suite il existe une lettre x telle que $Dax \in L(u)$. Ainsi:

$$\alpha(Dax) = \alpha(Da)ay \in L(\alpha(u))$$

avec $y \in \{\varepsilon, b, c\}$. Or

$$\alpha(Da)ay = F_nay \in L_{n+1}(\alpha(u)).$$

Donc $F_na \in L(\alpha(u))$ et par conséquent

$$\partial^+\alpha(F_n) = 3;$$

contradiction avec $\partial^+\alpha(F_n) = 2$.
Finalement $\alpha(u)$ n'admet pas de facteur biprolongeable à droite. De même $\alpha(u)$ n'admet pas de facteur biprolongeable à gauche. □

Théorème 3.3.3 *Les morphismes α, $\overline{\alpha}$, β, $\overline{\beta}$, γ et $\overline{\gamma}$ sont des AR-morphismes.*

Preuve: Soient $\varphi \in \{\alpha, \overline{\alpha}, \beta, \overline{\beta}, \gamma, \overline{\gamma}\}$ et u un AR-mot. Les lemmes 3.3.2, 3.3.3 et 3.3.4 nous assurent que $\varphi(u)$ est un AR-mot. □.
Pour le AR-morphisme α nous avons $\alpha(abcacba) = aabacaacaba$. Ainsi, d'après la proposition 3.3.2, α est un AR_a-morphisme. En faisant de même pour $\varphi \in \{\overline{\alpha}, \beta, \overline{\beta}, \gamma, \overline{\gamma}\}$ nous avons la remarque suivante:

Remarque 3.3.2 *Les morphismes α et $\overline{\alpha}$ (resp. β et $\overline{\beta}$, γ et $\overline{\gamma}$) sont des AR_a-morphismes (resp. AR_b-morphismes, AR_c-morphismes).*

Proposition 3.3.4 *(i) Tout facteur d'un AR_a-mot débutant (resp. finissant) par b ou c et finissant (resp. débutant) par a est image par $\overline{\alpha}$ (resp. α) d'un mot fini sur A.*
(ii) Tout facteur d'un AR_a-mot commençant et finissant par a est une image par $\overline{\alpha}$ (resp. α) d'un mot fini sur A.

Preuve: (i) Soit u un AR_a-mot, x_1t_1 et x_2t_k deux facteurs finis de u avec x_1 et x_k dans $\{b,c\}$. En vertu de la proposition 3.2.4 on a:

$$x_1t_1 = x_1a^{n+\epsilon_1}x_2a^{n+\epsilon_2}x_3a^{n+\epsilon_3}\cdots x_ka^{n_0},$$

où $x_i \in \{b,c\}$ et $\epsilon_i \in \{0,1\}$ pour $i = 1,2,\ldots,k-1$ et $1 \leq n_0 \leq n+1$.
Ainsi:

$$\begin{aligned} x_1t_1 &= (x_1a)a^{m+\epsilon_1}(x_2a)a^{m+\epsilon_2}\cdots(x_ka)a^{m_0}, \text{ avec } m = n-1 \text{ et } m_0 = n_0 - 1 \\ &= \overline{\alpha}(x_1)a^{m+\epsilon_1}\overline{\alpha}(x_2)a^{m+\epsilon_2}\cdots\overline{\alpha}(x_k)a^{m_0} \\ &= \overline{\alpha}(x_1)\overline{\alpha}(a^{m+\epsilon_1})\overline{\alpha}(x_2)\overline{\alpha}(a^{m+\epsilon_2})\cdots\overline{\alpha}(x_k)\overline{\alpha}(a^{m_0}) \\ &= \overline{\alpha}(x_1a^{m+\epsilon_1}x_2a^{m+\epsilon_2}\cdots x_ka^{m_0}) \\ &= \overline{\alpha}(x_1t'). \end{aligned}$$

De la même façon on montre que t_2x_k s'écrit sous la forme $\alpha(t''x_k)$.
(ii) Soit $a^{n_1}x_1tx_2a^{n_2}$ un facteur de u, avec $1 \leq n_1 \leq n+1$, $1 \leq n_2 \leq n+1$ et $x_1,x_2 \in \{b,c\}$. Les égalités obtenues dans la preuve de l'assertion (i) ci-dessus nous permettent d'écrire:

$$x_1tx_2a^{n_2} = \overline{\alpha}(x_1t') \text{ et } a^{n_1}x_1tx_2 = \alpha(t''x_2).$$

D'où :

$$a^{n_1}x_1tx_2a^{n_2} = \overline{\alpha}(a^{n_1}x_1t') = \alpha(t''x_2a^{n_2}).$$

□

Remarque 3.3.3 *Tout facteur d'un AR_a-mot débutant (resp. finissant) par a est image par α (resp. $\overline{\alpha}$) d'un mot fini sur A.*

Proposition 3.3.5 *Soit f un AR_a-morphisme et $x,y \in A$. Les assertions suivantes sont vraies:*

(i) $f(x) \in aA^* \cup A^*a$, pour tout $x \in A$.

(ii) Si $f(x) \in bA^* \cup cA^*$ alors $f(y) \in A^*a$, pour tout $y \in A$.

(iii) Si $f(x) \in A^*b \cup A^*c$ alors $f(y) \in aA^*$, pour tout $y \in A$.

Preuve: (i) Si $f(x) \in yA^*z$ ou si $f(x) = y$ avec $y, z \in \{b, c\}$ alors l'un au moins des mots b^2, c^2, bc et cb est dans $f(x^2)$. Or aucun de ces cinq mots n'est dans un AR_a-mot. Contradiction car f transforme les AR-mots en des AR_a-mots.

(ii) Soient $x, y, z \in A$ tels que $f(x) \in bA^* \cup cA^*$ et $f(y) \in A^*z$. Il s'en suit que $f(yx)$ contient zb ou zc. Si $z = b$ ou c on obtient une contradiction comme en (i). Donc $z = a$.

(iii) Même raisonnement qu'en (ii) en considérant $f(xy)$. □

Proposition 3.3.6 *Soit f un AR_a-morphisme. Alors f se décompose sous la forme $f = \varphi \circ g$ où $\varphi \in \{\alpha, \overline{\alpha}\}$ et g est un morphisme sur A^*.*

Preuve: Les mots $f(a)$, $f(b)$ et $f(c)$ sont présents dans un AR_a-mot puisque l'image par f de tout AR-mot est un AR_a-mot. Par conséquent nous avons l'un des trois cas suivants:

(1) $f(a)$, $f(b)$ et $f(c)$ commencent et finissent par a.

(2) L'un au moins des mots $f(a)$, $f(b)$ et $f(c)$ commence par b ou c,

(3) L'un au moins des mots $f(a)$, $f(b)$ et $f(c)$ finit par b ou c.

Cas 1: $f(a)$, $f(b)$ et $f(c)$ commencent et finissent par a. En vertu de l'assertion (ii) de la proposition 3.3.4 il existe y_1, y_2 et y_3 dans A^* et φ dans $\{\alpha, \overline{\alpha}\}$ tels que

$$f(a) = \varphi(y_1),\ f(b) = \varphi(y_2) \text{ et } f(c) = \varphi(y_3).$$

Cas 2: L'un au moins des mots $f(a)$, $f(b)$ et $f(c)$ commence par b ou c. Alors $f(a)$, $f(b)$ et $f(c)$ finissent par a d'après la proposition 3.3.5 (ii). Par conséquent, selon la remarque 3.3.2, il existe des mots y_1, y_2 et y_3 dans A^* tels que

$$f(a) = \varphi(y_1),\ f(b) = \varphi(y_2) \text{ et } f(c) = \varphi(y_3) \text{ avec } \varphi = \overline{\alpha}.$$

Cas 3: L'un au moins des mots $f(a)$, $f(b)$ et $f(c)$ finit par b ou c. Alors $f(a)$, $f(b)$ et $f(c)$ commencent par a d'après la proposition 3.3.5 (iii). Par conséquent, selon la remarque 3.3.2, il existe des mots y_1, y_2 et y_3 dans A^* tels que

$$f(a) = \varphi(y_1),\ f(b) = \varphi(y_2) \text{ et } f(c) = \varphi(y_3) \text{ avec } \varphi = \alpha.$$

Dans tous les cas, il existe $\varphi \in \{\alpha, \overline{\alpha}\}$ et des mots y_1, y_2 et y_3 dans A^* tels que $f(a) = \varphi(y_1)$, $f(b) = \varphi(y_2)$ et $f(c) = \varphi(y_3)$. Posons $g(a) = y_1$, $g(b) = y_2$ et $g(c) = y_3$. g est une substitution sur A qui, par prolongement naturel, définit un morphisme sur A^* tel que $f = \varphi \circ g$. □

Remarque 3.3.4 *Cette proposition s'adapte aux AR_b-morphismes et aux AR_c-morphismes en considérant les paires $\{\beta, \overline{\beta}\}$ et $\{\gamma, \overline{\gamma}\}$.*

Théorème 3.3.7 *Soient u un mot récurrent sur A et $\varphi \in \{\alpha, \overline{\alpha}, \beta, \overline{\beta}, \gamma, \overline{\gamma}\}$. Le mot $\varphi(u)$ est un AR-mot si et seulement si u l'est.*

Preuve: Sans restreindre la généralité on peut prendre $\varphi = \alpha$.
Nécessité: Soit u un mot récurrent tel que $\alpha(u)$ est un AR_a-mot avec u non d'Arnoux-Rauzy. Alors u admet, soit deux facteurs de même longueur triprolongeables à droite ou à gauche, soit un facteur biprolongeable à droite ou à gauche, soit aucun facteur spécial.

Supposons u sans facteur spécial pour une longueur n donnée. u est donc nécessairement ultimement périodique et par conséquent $\alpha(u)$ ne peut être un AR_a-mot. Absurde.

– Supposons que u admette pour un entier n donné un facteur F_n,de longueur n, biprolongeable à droite i.e $\partial^+ F_n = 2$.

Cas 1: $F_n b, F_n c \in L_{n+1}(u)$. Alors les mots

$$\alpha(F_n b) = \alpha(F_n)ab \text{ et } \alpha(F_n c) = \alpha(F_n)ac$$

sont deux facteurs de même longueur de $\alpha(u)$. On en déduit que $\partial^+\alpha(F_n)a = 3$ puisque $\alpha(u)$ est un AR_a-mot. D'où $\alpha(F_n)aa = \alpha(F_n a)a \in L(\alpha(u))$ et donc $F_n a \in L_{n+1}(u)$. Ce qui prouve que F_n est triprolongeable; absurde puisque F_n est biprolongeable par supposition.

Cas 2: $F_n a, F_n b \in L_{n+1}(u)$. On a $\partial^+ F_n a \geq 1$. Il existe donc $x \in A$ tel que $F_n ax \in L_{n+2}(u)$. Alors $\alpha(F_n ax) = \alpha(F_n)aay$ (où $y \in \{\varepsilon, b, c\}$) et $\alpha(F_n b) = \alpha(F_n)ab$ sont des facteurs de $\alpha(u)$. On en déduit que $\partial^+\alpha(F_n)a = 3$ puisque $\alpha(u)$ est un AR-mot. D'où $\alpha(F_n)ac = \alpha(F_n c) \in L(\alpha(u))$ et donc $F_n c \in L_{n+1}(u)$. Absurde.

Cas 3: $F_n a, F_n c \in L_{n+1}(u)$. Il suffit d'interchanger les rôles de b et c dans le cas précédent pour aboutir à la même conclusion.

Ainsi u n'admet pas de facteur biprolongeable à droite. De même on montre que u

n'admet pas de facteur biprolongeable à gauche.

Supposons que u admette deux facteurs de même longueur triprolongeables à droite. Soit D_1 et D_2 de tels facteurs i.e

$$|D_1| = |D_2| \text{ et } \partial^+ D_1 = \partial^+ D_2 = 3.$$

On sait que

$$\partial^+ \alpha(D_1)a = \partial^+ \alpha(D_2)a = 3$$

dans $\alpha(u)$.

Comme $D_1 \neq D_2$ on peut trouver $x, y \in A$ avec $x \neq y$ et des mots D_0, D' et D'' dans A^* tels que $D_1 = D'xD_0$ et $D_2 = D''yD_0$. On a alors les implications suivantes:

$$\begin{aligned} x \neq y &\implies \alpha(x) \neq \alpha(y) \\ &\implies \alpha(xD_0) \neq \alpha(yD_0) \\ &\implies \alpha(D_1) \neq \alpha(D_2) \\ &\implies \alpha(D_1)a \neq \alpha(D_2)a. \end{aligned}$$

Mais dans u xD_0 yD_0 sont triprolongeables à droite, donc $\alpha(D_1)a$ et $\alpha(D_2)a$ le sont dans $\alpha(u)$. Absurde, car $\alpha(u)$ est un AR-mot. De même on montre que u n'admet pas deux facteurs de même longueur triprolongeables à gauche.

Suffisance: cf. théorème 3.3.3. □

Théorème 3.3.8 *Pour tout AR-morphisme non trivial f il existe $\varphi \in \{\alpha, \overline{\alpha}, \beta, \overline{\beta}, \gamma, \overline{\gamma}\}$ et un AR-morphisme g tels que $f = \varphi \circ g$.*

Preuve: Soit f un AR-morphisme non trivial. D'après la proposition 3.3.6, il existe $\varphi \in \{\alpha, \overline{\alpha}, \beta, \overline{\beta}, \gamma, \overline{\gamma}\}$ et g un morphisme sur A^* tels que $f = \varphi \circ g$. Le morphisme g est un AR-morphisme. En effet, soit u un AR-mot; alors $f(u) = (\varphi \circ g)(u) = \varphi[g(u)]$ est un AR-mot. Par application du théorème 3.3.7, $g(u)$ est un AR-mot et par conséquent g est un AR-morphisme. □

Notation: Si H est un ensemble de morphismes sur A^*, on désignera par H^* l'ensemble de tous les morphismes de A^* engendrés par les éléments de H.

Lemme 3.3.5 *Les égalités suivantes sont vérifiées.*

$$(1)\ \beta = E_a E \alpha E_a E \text{ et } \overline{\beta} = E_a E \overline{\alpha} E_a E;$$

(2) $\gamma = EE_a\alpha EE_a$ et $\overline{\gamma} = EE_a\overline{\alpha}EE_a$.

Preuve: (1)

$$\begin{aligned} E_aE\alpha E_aE(a) &= E_aE\alpha E_a(c) = E_aE\alpha(b) = E_aE(ab) = E_a(ca) = ba = \beta(a), \\ E_aE\alpha E_aE(b) &= E_aE\alpha E_a(a) = E_aE\alpha(a) = E_aE(a) = E_a(c) = b = \beta(b), \\ E_aE\alpha E_aE(c) &= E_aE\alpha E_a(b) = E_aE\alpha(c) = E_aE(ac) = E_a(cb) = bc = \beta(c). \end{aligned}$$

D'où, $\beta = E_aE\alpha E_aE$. De même, on vérifie que $\overline{\beta} = E_aE\overline{\alpha}E_aE$

(2)

$$\begin{aligned} EE_a\alpha EE_a(a) &= EE_a\alpha E(a) = EE_a\alpha(c) = EE_a(ac) = E(bc) = ca = \gamma(a), \\ EE_a\alpha EE_a(b) &= EE_a\alpha E(c) = EE_a\alpha(b) = EE_a(ab) = E(ac) = cb = \gamma(b), \\ EE_a\alpha EE_a(c) &= EE_a\alpha E(b) = EE_a\alpha(a) = EE_a(a) = E(a) = c = \gamma(c). \end{aligned}$$

D'où, $\gamma = EE_a\alpha EE_a$. De même on vérifie que $\overline{\gamma} = EE_a\overline{\alpha}EE_a$. □

Théorème 3.3.9 *Un morphisme f sur A^* est un AR-morphisme si et seulement si $f \in \{E, E_a, \alpha, \overline{\alpha}\}^*$.*

Preuve: *Suffisance:* Si $f \in \{E, E_a, \alpha, \overline{\alpha}\}^*$ alors f est un AR-morphisme puisque la composition de morphismes préserve les AR-morphismes.

Nécessité: Montrons que si f est un AR-morphisme alors $f \in \{E, E_a, \alpha, \overline{\alpha}, \beta, \overline{\beta}, \gamma, \overline{\gamma}\}^*$.
Pour ce faire nous menons une récurrence sur $\|f\|$ la longueur de f,

$$\|f\| = \sum_{x \in A} |f(x)| = |f(abc)|.$$

Si $\|f\| = 3$ alors $f \in \{E, E_a\}^+ \subset \{E, E_a, \alpha, \overline{\alpha}, \beta, \overline{\beta}, \gamma, \overline{\gamma}\}^*$.
Supposons que tout AR-morphisme de longueur strictement inférieure à n ($n \geq 4$) soit dans $\{E, E_a, \alpha, \overline{\alpha}, \beta, \overline{\beta}, \gamma, \overline{\gamma}\}^*$.
Considérons un AR-morphisme f tel que $\|f\| = n$. D'après le théorème 3.3.8 il existe $\varphi \in \{\alpha, \overline{\alpha}, \beta, \overline{\beta}, \gamma, \overline{\gamma}\}$ et un AR-morphisme g tels que $f = \varphi \circ g$.
Nous avons: $\|g\| < n$. En effet, prenons arbitrairement $\varphi = \alpha$. Comme $|\alpha(a)| = 1$,

$|\alpha(b)| = 2$ et $|\varphi(c)| = 2$, nous avons:

$$\begin{aligned} \|f\| &= \left|\alpha\left[g(abc)\right]\right| \\ &= |g(abc)|_a + 2\,|g(abc)|_b + 2\,|g(abc)|_c \\ &= \|g\| + \Big(|g(abc)|_b + |g(abc)|_c \Big) \end{aligned}$$

Comme g est AR-morphisme, les lettres b et c sont nécessairement dans $g(abc)$ car sinon g transformerait les mots ternaires en mots unaires ou binaires et ne préservera pas dans ce cas les AR-mots. Ainsi, $|g(abc)|_b + |g(abc)|_c > 1$ et par conséquent $\|g\| < n$. Ainsi, par hypothèse de récurrence, $g \in \left\{E, E_a, \alpha, \overline{\alpha}, \beta, \overline{\beta}, \gamma, \overline{\gamma}\right\}^*$. D'où

$$\varphi \circ g = f \in \left\{E, E_a, \alpha, \overline{\alpha}, \beta, \overline{\beta}, \gamma, \overline{\gamma}\right\}^*.$$

Par ailleurs, en vertu du lemme 3.3.5, nous savons que les morphismes β, $\overline{\beta}$, γ et $\overline{\gamma}$ sont dans $\{E, E_a, \alpha, \overline{\alpha}\}^*$. D'où l'inclusion:

$$\left\{E, E_a, \alpha, \overline{\alpha}, \beta, \overline{\beta}, \gamma, \overline{\gamma}\right\}^* \subseteq \{E, E_a, \alpha, \overline{\alpha}\}^*.$$

L'inclusion

$$\{E, E_a, \alpha, \overline{\alpha}\}^* \subseteq \left\{E, E_a, \alpha, \overline{\alpha}, \beta, \overline{\beta}, \gamma, \overline{\gamma}\right\}^*$$

est évidente.

D'où l'égalité:

$$\left\{E, E_a, \alpha, \overline{\alpha}, \beta, \overline{\beta}, \gamma, \overline{\gamma}\right\}^* = \{E, E_a, \alpha, \overline{\alpha}\}^*.$$

□

Ce théorème nous permet de construire à souhait des morphismes d'Arnoux-Rauzy. Pour la suite nous noterons $\Phi(AR)$ le monoïde des AR-morphismes:

$$\Phi(AR) = \{E, E_a, \alpha, \overline{\alpha}\}^*.$$

Corollaire 3.3.1 *Soient u un mot récurrent, f et g deux morphismes sur A^*.*

(1) Si $f \in \Phi(AR)$ alors $f(u)$ est un AR-mot si et seulement si u est un AR-mot.

(2) Si $f \in \Phi(AR)$ et $f \circ g \in \Phi(AR)$ alors $g \in \Phi(AR)$. On dit alors que le monoïde des AR-morphismes est unitaire à gauche.

Preuve: Soit f un AR-morphisme, selon le théorème 3.3.9 f s'écrit

$$f = \varphi_1\varphi_2\cdots\varphi_r$$

où $\varphi_i \in \{E, E_a, \alpha, \overline{\alpha}\}$ pour $i = 1, 2, \ldots, r$. Le mot

$$f(u) = \varphi_1\varphi_2\cdots\varphi_r(u) = \varphi_1(\varphi_2\cdots\varphi_r(u))$$

est un AR-mot si et seulement si le mot

$$\varphi_2\cdots\varphi_r(u)$$

l'est (Théorème 3.3.7). On réitère ce procédé $r-1$ fois à partir de $\varphi_2\varphi_3\cdots\varphi_r(u)$ et on obtient l'assertion (1).

Soient u un AR-mot, f un AR-morphisme et g un morphisme sur A^* tel que $f \circ g \in \Phi(AR)$. Alors $f(g(u)) = f \circ g(u)$ est un AR-mot. Par conséquent, d'après l'assertion (1) $g(u)$ est un AR-mot puisque f est un AR-morphisme et u est récurrent. Ainsi g est un AR-morphisme. □

Remarque 3.3.5 *Si nous avons montré que $\Phi(AR)$ est unitaire à gauche il faut noter que les arguments utilisés ne nous permettent pas de prouver que $\Phi(AR)$ est unitaire à droite. Il nous paraît donc intéressant d'étudier cette question.*

Avant d'énoncer un dernier corollaire, montrons d'abord deux propriétés générales sur l'opération "miroir de morphismes". Rappelons que le morphisme miroir d'un morphisme f est noté $\overline{f}$ et est défini par $\overline{f}(a) = \overline{f(a)}$ pour tout $a \in A$.

Proposition 3.3.10 *Soient v un mot fini, f et g deux morphismes sur A^*. Alors:*

(1) $\overline{f}(v) = \overline{f(\overline{v})}$;

(2) $\overline{f \circ g} = \overline{f} \circ \overline{g}$.

Preuve: Soit v un mot fini, f et g deux morphismes sur A^*.

(1) Si $v \in A \cup \{\varepsilon\}$ il est évident que $\overline{f}(v) = \overline{f(\overline{v})}$. Si v est de longueur n, supérieure ou égale à 2 alors v s'écrit $v = v_1v_2\cdots v_n$ avec $v_i \in A$, $i = 1, 2, \ldots, n$.

Ainsi:

$$\begin{aligned}
\overline{f}(v) &= \overline{f}(v_1)\overline{f}(v_2)\cdots\overline{f}(v_n) \\
&= \overline{f(v_1)}\,\overline{f(v_2)}\cdots\overline{f(v_n)} \\
&= \overline{f(v_n)f(v_{n-1})\cdots f(v_1)} \\
&= \overline{f(v_nv_{n-1}\cdots v_1)} \\
&= \overline{f(\overline{v})}.
\end{aligned}$$

(2) Soit a une lettre quelconque de A. Alors:

$$\begin{aligned}
\overline{f\circ g}(a) &= \overline{f\circ g(a)} \\
&= \overline{f\Big(g(a)\Big)} \\
&= \overline{f}\Big(\overline{g(a)}\Big) \\
&= \overline{f}\Big(\overline{g}(a)\Big) \\
&= \overline{f}\circ\overline{g}(a).
\end{aligned}$$

□

D'où le corollaire suivant.

Corollaire 3.3.2 *Soit f un morphisme sur A^*. $\overline{f}$ est un AR-morphisme si et seulement si f l'est.*

Preuve: Soit f un AR-morphisme. Alors d'après le théorème 3.3.8 f s'écrit sous la forme $f = \varphi_1\varphi_2\cdots\varphi_r$ où $\varphi_i \in \{E, E_a, \alpha, \overline{\alpha}\}$ pour $i = 1, 2 \ldots, r$. Par passage au miroir on obtient, en utilisant la proposition 3.3.10, les égalités suivantes:

$$\begin{aligned}
\overline{f} &= \overline{\varphi_1\varphi_2\cdots\varphi_r} \\
&= \overline{\varphi_1}\,\overline{\varphi_2}\cdots\overline{\varphi_r}.
\end{aligned}$$

Or pour tout élément φ dans $\{E, E_a, \alpha, \overline{\alpha}\}$, $\overline{\varphi}$ reste dans $\{E, E_a\, \alpha, \overline{\alpha}\}$. Donc

$$\overline{f} \in \{E, E_a, \alpha, \overline{\alpha}\}^*.$$

Ainsi $\overline{f}$ est un AR-morphisme d'après le théorème 3.3.9.

La réciproque est immédiate puisque l'opération "miroir" est involutive sur les morphismes. □

Chapitre 4

Combinatoire des mots de complexité $n+2$

4.1 Introduction

La fonction de complexité p, qui calcule le nombre de facteurs de longueur donnée dans un mot, est souvent utilisée pour caractériser certaines familles de mots. Ainsi on sait qu'un mot dont la fonction de complexité vérifie $p(n) \leq n$ pour un certain entier n, est ultimement périodique [29] et que les mots sturmiens sont les mots de complexité minimale parmi les mots infinis non ultimement périodiques. De nombreuses études menées sur les mots sturmiens ont conduit à de nombreuses généralisations parmi lesquelles les mots quasi-sturmiens.

La notion de mot quasi-sturmien a été introduite au milieu des années 1990 pour caractériser les mots infinis dont les propriétés se rapprochent des mots sturmiens.

Un mot infini est dit quasi-sturmien s'il existe des entiers n_0 et k tels que $p(n) = n + k$ pour tout $n \geq n_0$. Les mots quasi-sturmiens ont une combinatoire assez proche des mots sturmiens et leurs caractérisations font intervenir ces derniers([13], [20], [22]). Dans ce chapitre nous classifions tous les mots qui possèdent, pour tout entier n non nul, exactement $n+2$ facteurs de longueur n. Ensuite, nous étudions précisément une sous classe de ces mots que nous appelons mots quasi-sturmiens par insertion. Nous montrons d'abord un résultat permettant de construire ces mots à partir des mots sturmiens puis nous

étudions certaines de leurs propriétés combinatoires.
Les principaux résultats de ce chapitre sont dans [26].

4.2 Une sous classe de mots de complexité $n+2$

Soit $A = \{a, b, c\}$ un alphabet fixé.

Théorème 4.2.1 *Soit $u = a^{n_0}ba^{n+\epsilon_1}ba^{n+\epsilon_2}ba^{n+\epsilon_3}b\cdots$ un mot sturmien sur $\{a, b\}$ et soit v le mot sur $\{a, b, c\}$ défini par $v = bx_1bx_2bx_3\cdots$ où pour tout entier non nul i, $x_i = a$ si $\epsilon_i = 0$ et $x_i = c$ si $\epsilon_i = 1$. Alors v est de complexité $n+2$ pour tout $n \geq 1$.*

Preuve: La suite $(\epsilon_i)_{i\geq 1}$ est sturmienne d'après le théorème 2.2.7. Par conséquent le mot $x = x_1x_2x_3\cdots$ est sturmien sur $\{a, c\}$ car a et c s'identifient respectivement à 0 et 1 dans la suite $(\epsilon_i)_{i\geq 1}$.
Comme v ne contient pas de carré de lettre alors tout facteur non vide de v est non triprolongeable à droite. Il nous suffit donc de montrer que v admet un et un seul facteur biprolongeable à droite de longueur n, pour tout $n \geq 1$.
Pour $n = 1$, b est le seul facteur biprolongeable à droite. Supposons la propriété vérifiée jusqu'au rang n ($n \geq 1$). Comme x est sturmien alors v est récurrent et non ultimement périodique. Par conséquent

$$\#\left\{D \in L_{n+1}(v) : \partial^+ D = 2\right\} \geq 1.$$

Supposons que v admette deux facteurs biprolongeables à droite de longueur $n+1$. Soient D_1 et D_2 les dits facteurs. D'après l'hypothèse de récurrence, v n'admet qu'un unique facteur biprolongeable à droite D de longueur n. Par suite, on peut écrire $D_1 = aD$ et $D_2 = cD$. Il en résulte que les mots

$$aDa,\ aDc,\ cDa \text{ et } cDc$$

sont dans $L_{n+2}(v)$. En effaçant les lettres b dans ces quatre facteurs de v, on obtient quatre facteurs de x de la forme

$$ata,\ atc,\ cta \text{ et } ctc.$$

Par suite x n'est pas équilibré puisqu'il contient deux facteurs de la forme ata et ctc. Absurde, car x est sturmien. Ainsi, v admet un unique facteur biprolongeable à droite

de longueur $n+1$.
Finalement, pour tout $n \geq 1$ v admet un unique facteur biprolongeable à droite de longueur n. D'où $p(n+1) = p(n) + 1$ et comme $p(1) = 3$ il en résulte que $p(n) = n+2$ pour tout $n \geq 1$. □

Corollaire 4.2.1 *Soit v un mot infini sur A de la forme*

$$v = x_0 z x_1 z x_2 z x_3 \cdots$$

où $x_0 \in \{\varepsilon, x, y\}$ et $x_i \in \{x, y\}$ pour $i = 1, 2, 3, \ldots$ avec $\{x, y, z\} = \{a, b, c\}$.
v est un mot de complexité $n+2$ $(n \geq 1)$ si et seulement si le mot extrait

$$v' = x_0 x_1 x_2 x_3 \cdots$$

est un mot sturmien sur $\{x, y\}$.

Preuve: D'après la preuve du théorème 4.2.1 ci-dessus la condition est suffisante. Montrons qu'elle est nécessaire. Pour ce faire supposons $v' = x_0 x_1 x_2 x_3 \cdots$ non sturmien. Alors v' est ultimement périodique ou non équilibré. Si v' est non ultimement périodique, alors v, qui est image morphique de v', l'est aussi. Ainsi, v est non ultimement périodique. Donc v' est non équilibré et possède de ce fait deux facteurs spéciaux à droite de longueur n de la forme

$$xt = xt_1 t_2 \cdots t_{n-1} \text{ et } yt = yt_1 t_2 \cdots t_{n-1}.$$

Par conséquent les mots

$$xtx,\ xty,\ ytx \text{ et } yty$$

sont des facteurs de v'. Par suite les mots

$$xrx,\ xry,\ yrx \text{ et } yry \text{ où } r = zt_1 zt_2 z \cdots zt_{n-1} z$$

sont des facteurs de même longueur de v. Ainsi xr et yr sont deux facteurs spéciaux à droite de v de même longueur. Absurde, car v étant de complexité $n+2$ $(n \geq 1)$, il admet un et un seul facteur biprolongeable à droite pour toute longueur n donnée. Donc v' est sturmien. □

Définition 4.2.1 *Soit u un mot récurrent sur A. On dira que u est un mot quasi-sturmien par insertion s'il est de complexité $n+2$ et de la forme*

$$v = x_0 z x_1 z x_2 z x_3 \cdots$$

où $x_0 \in \{\varepsilon, x, y\}$ *et* $x_i \in \{x, y\}$ *pour* $i = 1, 2, 3, \ldots$ *avec* $\{x, y, z\} = \{a, b, c\}$.

Exemple: Considérons le mot de Fibonacci

$$F = abaababaabaababaababaabaababaaba\cdots$$

Le mot quasi-sturmien par insertion issu de F et commençant par a est

$$Q = acbcacacbcacbcacacbcacacbcacbcacacbcacbcacacbcacacbcacbcacacbca\cdots$$

Remarque 4.2.1 *Soit u un mot quasi-sturmien par insertion. u ne comporte aucun carré de lettre i.e toute lettre de alph(u) est isolée dans u.*

Théorème 4.2.2 *Soit u un mot quasi-sturmien par insertion.*

$$\forall n \in \mathbb{N}^*, \exists! \,(D_n, G_n) \in L_n^2(u) : \partial^+ D_n = \partial^- G_n = 2).$$

Preuve: u quasi-sturmien par insertion $\Longrightarrow (\forall n \in \mathbb{N}^*, p(n) = n+2)$
$\Longrightarrow (\forall n \in \mathbb{N}^*, p(n+1) - p(n) = 1)$
$\Longrightarrow \left(\forall n \in \mathbb{N}^*, p(n+1) - p(n) = \sum_{w \in L_n(u)}(\partial^+ w - 1) = \sum_{w \in L_n(u)}(\partial^- w - 1) = 1\right)$
$\Longrightarrow (\forall n \in \mathbb{N}^* \exists! \,(D_n, G_n) \in L_n^2(u) : \partial^+ D_n = \partial^- G_n = 2).$ □

Dans toute la suite on dira qu'un facteur est bispécial lorsqu'il est biprolongeable à la fois à gauche et à droite dans le mot donné.

Définition 4.2.2 *Un mot u, quasi-sturmien par insertion est dit a-spécial (resp. b-spécial, c-spécial) si la lettre a (resp. b, c) est biprolongeable dans u.*

Par unicité de la lettre biprolongeable tout mot u quasi-sturmien par insertion est soit a-spécial, soit b-spécial, soit c-spécial.

Définition 4.2.3 *Soit u un mot quasi-sturmien par insertion. On appelle mot extrait de u et on note u' le mot sturmien obtenu de u en effaçant la lettre biprolongeable.*

Proposition 4.2.3 *Soit u un mot quasi-sturmien par insertion a-spécial et w un facteur non vide de u de longueur paire. Alors w se termine par la lettre a si et seulement si il ne commence pas par a. De plus:*

$$|w|_a = \frac{|w|}{2} \text{ et } |w|_b + |w|_c = \frac{|w|}{2}.$$

Preuve: Soit $w = w_1 w_2 \cdots w_{2p-1} w_{2p}$ un facteur non vide de u de longueur paire. Si w commence par a i.e $w_1 = a$ comme u est a-spécial alors tous les termes de w d'indices impairs correspondent à la lettre a. Ainsi w_{2p} est différente de a car u ne contient pas aa. En d'autres termes w ne finit pas par a. De même on montre que w finit par a lorsqu'il ne débute pas par a. Plus précisément on a

$$w_1 = w_3 = \cdots = w_{2p-1} = a$$

ou bien

$$w_2 = w_4 = \cdots = w_{2p} = a.$$

D'où

$$|w|_a = \frac{|w|}{2}$$

et comme $|w|_a + |w|_b + |w|_c = |w|$ il en résulte que

$$|w|_b + |w|_c = \frac{|w|}{2}.$$

□

Proposition 4.2.4 *Soit u un mot quasi-sturmien par insertion a-spécial et v un facteur de u de longueur impaire. Alors on a l'une des assertions suivantes:*

(i) v débute et finit par a,

(ii) v ne débute pas par a et ne finit pas par a.

Preuve: (i) Supposons que v débute par a: v peut s'écrire $v = wx$ avec w un facteur de longueur paire de u et x une lettre de A. Ainsi w finit par b ou c puisqu'il commence par a (Proposition 4.2.3). Par suite la lettre x suivant w est nécessairement a car dans u les lettres b et c sont toujours suivies de a. Donc v finit par a.

(ii) Supposons que v débute par b ou c. v peut s'écrire $v = wx$ avec w un facteur de longueur paire de u commençant par b ou c et x une lettre de A. Ainsi w finit par a puisqu'il commence par b ou c (Proposition 4.2.3). Par conséquent la lettre x suivant w est nécessairement distincte de a puisque u ne contient pas aa. Ainsi, v finit par b ou c.
□

4.3 Classification des mots de complexité $n+2$

La classification des mots de complexité $n+2$ a été établie par P. Alessandri dans un rapport de recherche ([1]). Nous en donnons une formulation légèrement différente.

Lemme 4.3.1 *Soit u un mot de l'une des formes suivantes, à une permutation de lettres près:*

(1) cv *avec v sturmien sur* $\{a,b\}$.

(2) $T^{k_0}(a^{k+\epsilon_1}bca^{k+\epsilon_2}bca^{k+\epsilon_3}bc\cdots)$

(3) $T^{k_0}((ab)^{k+\epsilon_1}c(ab)^{k+\epsilon_2}c(ab)^{k+\epsilon_3}c\cdots)$

où T est le décalage, $k_0\geq 0$, $k>0$ et $(\epsilon_i)_{i\geq 1}$, une suite sturmienne sur $\{0,1\}$. Alors, u est de complexité $n+2$.

Preuve: (1) Pour tout $n\geq 1$, tout préfixe de longueur n n'apparaît qu'une seule fois dans u. Donc $p_u(n)=1+p_v(n)=1+(n+1)$ puisque v est sturmien. D'où $p_u(u)=n+2$.

(2) u n'est pas ultimement périodique car il est image du mot sturmien

$$v=T^{k_0}(a^{k+\epsilon_1}ba^{k+\epsilon_2}ba^{k+\epsilon_3}b\cdots)$$

par le morphisme non périodique: $(a\longmapsto a,\ b\longmapsto bc)$. Donc u possède au moins un facteur spécial pour toute longueur donnée n . Nous observons que a est la seule lettre spéciale dans u:la lettre a se prolonge par a et c à gauche et par a et b à droite. Donc tout facteur spécial de u est biprolongeable. Supposons que u possède deux facteurs T_1 et T_2 spéciaux à droite pour une certaine longueur n. On peut prendre n minimal et nous aurons ($T_1=aT$ et $T_2=bT$) ou ($T_1=bT$ et $T_2=cT$) ou ($T_1=aT$ et $T_2=cT$) avec $T\in A^+$. Si bT est dans u alors T commence nécessairement par c puisque b ne précède que c dans u. Par suite, comme a ne précède pas c dans u, le premier cas ne peut se produire. Il en est de même pour le second cas puisque cc n'est pas dans u. Ainsi, le seul cas possible est $T_1=aT$ et $T_2=cT$. Il en résulte que $aTa, cTb\in L(u)$ et donc $aTa, bcTb\in L(u)$. En effaçant c dans aTa et $bcTb$ On retrouve deux facteurs ata et btb du mot v. Donc v n'est pas équilibré. Absurde, car v est sturmien. Donc $p_u(n+1)-p_u(n)=1$, pour tout $n\geq 1$ et comme $p_u(1)=3$ on déduit que $p_u(n)=n+2$. La forme (3) se montre de la même manière qu'en (2). □

Théorème 4.3.1 *Soit u un mot infini sur l'alphabet $\{a,b,c\}$. Le mot u est de complexité*

$n+2$ *si et seulement si il a l'une des formes suivantes à une permutation de lettres près:*

(1) cv *avec* v *sturmien sur* $\{a,b\}$.

(2) $T^{k_0}(a^{k+\epsilon_1}bca^{k+\epsilon_2}bca^{k+\epsilon_3}bc\cdots)$

(3) $T^{k_0}((ab)^{k+\epsilon_1}ac(ab)^{k+\epsilon_2}ac(ab)^{k+\epsilon_3}ac\cdots)$

(4) $T^{k_0}((ab)^{k+\epsilon_1}c(ab)^{k+\epsilon_2}c(ab)^{k+\epsilon_3}c\cdots)$

où T *est le décalage,* $k_0 \geq 0$, $k > 0$ *et* $(\epsilon_i)_{i\geq 1}$, *une suite sturmienne sur* $\{0,1\}$.

Preuve: *Nécessité:* Soit u un mot de complexité $n+2$ sur $\{a,b,c\}$. Alors, u est soit récurrent, soit non récurrent.

– Supposons u non récurrent. Il existe un entier n_0 non nul tel que le préfixe de longueur n_0 n'apparaisse qu'une seule fois dans u. En prenant $v = T(u)$, nous avons $p_v(n) = p_u(n) - 1 = n+1$ pour tout $n \geq n_0$. Ainsi, la complexité de v est non bornée. Par conséquent v est non ultimement périodique et donc sa fonction de complexité p est strictement croissante et vérifie $p_v(n) > n$, pour tout n. Il en résulte que $p_v(n) = n+1$ pour tout n.Ce qui prouve que v est sturmien. Donc, v est binaire et s'écrit par exemple avec a et b. Par suite, comme u est ternaire, nous avons $u = cv$.

– Supposons u récurrent. Alors pour tout n, le mot u possède un unique facteur spécial à droite et un unique facteur spécial à gauche de longueur n. Par suite, u possède soit une lettre bispéciale dont son carré, soit une lettre bispéciale sans son carré, soit enfin une lettre spéciale à gauche et une autre lettre spéciale à droite:

Cas 1: Prenons a la lettre bispéciale avec aa dans u. Par suite, comme u est récurrent nous avons:

$$L_2(u) = \{aa, ab, ca, bc\} \text{ ou } L_2(u) = \{aa, ac, ba, cb\}$$

On peut supposer pour continuer que $L_2(u) = \{aa, ab, ca, bc\}$. Définissons les morphismes suivants:

$$\begin{array}{llll} f_1: & \{a,b\} \longrightarrow \{a,b,c\} & \text{et} \quad g_1: & \{a,b,c\} \longrightarrow \{a,b\} \\ & a \longmapsto a & & a \longmapsto a \\ & b \longmapsto bc & & b \longmapsto b \\ & & & c \longmapsto \varepsilon \end{array}$$

Remarquons que $g \circ f = Id_{\{a,b\}}$. Posons $v = g(u)$. Nous avons $u = f(v)$ ou $u = cf(v)$. Montrons que v est un mot a-sturmien. bb n'est pas dans v puisque bcb n'est pas dans u. Le mot v est non ultimement périodique puisque $f(v)$ est non ultimement périodique.

Donc v possède au moins un facteur spécial pour toute longueur donnée. Considérons w un facteur spécial à droite de v. Nous avons $wa, wb \in L(v)$. Donc $f(wa), f(wb) \in L(u)$ i.e $f(w)a, f(w)bc \in L(u)$. Ainsi $f(w)$ est spécial à droite dans u.

Soit w et w' deux facteurs spéciaux à droite de même longueur dans v. Alors $f(w)$ et $f(w')$ sont spéciaux à droite dans u. Soit N un entier tel que $N \geq \max(|f(w)|, |f(w')|)$. u possède un unique facteur spécial z de longueur N. Le suffixe de longueur $|f(w)|$ de z est spécial à droite donc égal à $f(w)$. De même le suffixe de longueur $|f(w')|$ de z est égal $f(w')$. Par suite $f(w)$ et $f(w')$ sont suffixes de z. Donc les mots $g \circ f(w) = w$ et $g \circ f(w') = w'$ sont suffixes de $g(z)$. Par conséquent $w = w'$ puisqu'ils sont de même longueur. Finalement v possède un et un seul facteur spécial pour toute longueur donnée. Ainsi, v est sturmien. Comme bb n'est pas dans v alors v est a-sturmien. Donc, il existe un entier naturel non nul k tel que u s'écrit

$$v = a^{k_0}ba^{k+\epsilon_1}ba^{k+\epsilon_2}ba^{k+\epsilon_3}b\cdots$$

avec $k_0 \leq k+1$ et $(\epsilon_i)_{i\geq 1}$ une suite sturmienne sur l'alphabet $\{0,1\}$. En définitive u vérifie (2).

Cas 2: Prenons a la lettre bispéciale. Comme aa n'est pas dans u nous avons:

$$L_2(u) = \{ab, ba, ac, ca\}.$$

Dans ce cas en posant

$$\begin{array}{llll} f_2: & \{a,b\} \longrightarrow \{a,b,c\} & \text{et} \quad g_2: & \{a,b,c\} \longrightarrow \{a,b\} \\ & a \longmapsto ac & & a \longmapsto \varepsilon \\ & b \longmapsto ab & & b \longmapsto b \\ & & & c \longmapsto c \end{array}$$

puis en raisonnant comme au cas 1 on obtient que u a la forme (3) du théorème.

Cas 3: Prenons a la lettre spéciale à gauche et b la lettre spéciale à droite . Nous avons

$$L_2(u) = \{ab, ac, ba, bc\}.$$

Dans ce cas aussi en posant

$$\begin{array}{llll} f_3: & \{a,b\} \longrightarrow \{a,b,c\} & \text{et} \quad g_3: & \{a,b,c\} \longrightarrow \{a,b\} \\ & a \longmapsto ab & & a \longmapsto b \\ & b \longmapsto c & & b \longmapsto \varepsilon \\ & & & c \longmapsto b \end{array}$$

puis en suivant une démarche analogue au cas 1 on vérifie que u est de la forme (4).
Suffisance: On applique le lemme 4.3.1 pour les formes (1), (2) et (4) puis le corollaire 4.2.1 pour la forme (3)

4.4 Combinatoire des mots quasi-sturmiens par insertion

Dans ce paragraphe nous généralisons la notion d'équilibre vue au chapitre 2 puis nous étudions la palindromie.

Définition 4.4.1 *Un mot u est dit k-équilibré si k est le plus petit entier tel que pour tout couple (v, w) de facteurs de u de même longueur et toute lettre a de A on a:*

$$\Big| |v|_a - |w|_a \Big| \leq k.$$

Théorème 4.4.1 *Tout mot u quasi-sturmien par insertion est 2-équilibré.*

Preuve: Soit u un mot quasi-sturmien par insertion. On peut supposer sans perdre la généralité que u est a-spécial.

1$^{\text{ère}}$ **étape:** Soient v et w deux facteurs de longueur égale et paire. De v et w on peut extraire deux mots v' et w' de longueur commune $\frac{|v|}{2}$ ne contenant pas a, la lettre spéciale de u. Les mots v' et w' sont en réalité deux facteurs du mot sturmien u' extrait de u. Par conséquent

$$\Big| |v'|_x - |w'|_x \Big| \leq 1$$

pour tout $x \in \{b, c\}$. De plus, remarquons que:

$$|v'|_x = |v|_x \text{ et } |w'|_x = |w|_x .$$

Par ailleurs nous avons aussi:

$$\Big| |v'|_a - |w'|_a \Big| = \left| \frac{|v|}{2} - \frac{|w|}{2} \right| = 0 \leq 1.$$

Il en résulte que:

$$\Big| |v|_x - |w|_x \Big| \leq 1$$

pour tout $x \in \{a, b, c\}$.

$2^{\text{ème}}$ étape: Soient r et s deux facteurs de longueur égale et impaire. On peut écrire $r = vx$ et $s = wy$ où (x, y) est un couple de lettres et (v, w) un couple de facteurs de longueur égale et paire de u. On a:

$$\begin{aligned}
\Big||r|_z - |s|_z\Big| &= \Big|\Big(|v|_z + |x|_z\Big) - \Big(|w|_z + |y|_z\Big)\Big| \\
&= \Big|\Big(|v|_z - |w|_z\Big) + \Big(|x|_z - |y|_z\Big)\Big| \\
&\leq \Big|\Big(|v|_z - |w|_z\Big)\Big| + \Big|\Big(|x|_z - |y|_z\Big)\Big| \\
&\leq 1 + 1 = 2
\end{aligned}$$

(la dernière majoration est obtenue à partir du résultat de la $1^{\text{ère}}$ étape).

$3^{\text{ème}}$ étape: Le majorant 2 est atteint puisque u contient l'un des mots bab et cac. En effet l'unique facteur biprolongeable de longueur 2 de u est l'un des mots ba et ca. Alors bab est dans u si ba est biprolongeable; dans le cas contraire cac serait dans u puisque ca serait le facteur biprolongeable.

Supposons pour continuer que u contienne bab. On sait que aca est dans u. D'où

$$\Big||bab|_b - |aca|_b\Big| = 2.$$

En définitive u est 2-équilibré. □

Soit $v \in A^+$, On appelle vecteur des occurrences des lettres dans v et on note $\mathbf{c}(v)$ le triplet $(|v|_a, |v|_b, |v|_c)$.

Soit u un mot infini sur A et n un entier naturel non nul. On désigne par $V_n(u)$ l'ensemble $\{\mathbf{c}(v) : v \in L_n(u)\}$.

Pour tout mot sturmien u, E. Coven et G. A. Hedlund ont montré dans [16] que $\#V_n(u) = 2$ pour tout $n \geq 1$.

Théorème 4.4.2 *Un mot récurrent u est quasi-sturmien par insertion si et seulement si les conditions suivantes sont satisfaites:*

(i) $\#V_1(u) = 3$,

(ii) $\#V_n(u) = 2$ pour tout n pair non nul,

(iii) $\#V_n(u) = 4$ pour tout n impair et différent de 1.

Preuve: Soit u un mot récurrent.

Supposons u quasi-sturmien par insertion sur A. On peut supposer u a-spécial sans perdre la généralité.

(i) Nous avons $V_1(u) = \{(1,0,0);\ (0,1,0);\ (0,0,1)\}$ et donc $\#V_1(u) = 3$.

(ii) Soit n un entier pair non nul et soit u' le mot sturmien extrait de u. Soit θ_n, l'application de $L_n(u)$ dans $L_{\frac{n}{2}}(u')$ telle que pour tout $v \in L_n(u)$, $\theta_n(v) = v'$ est le mot obtenu en effaçant la lettre spéciale a dans v. Cette application est trivialement surjective mais non injective. En effet, soit $v' = x_1x_2\cdots x_{\frac{n}{2}}$ un élément de $L_{\frac{n}{2}}(u')$. Alors les mots

$$v_1 = ax_1ax_2\cdots ax_{\frac{n}{2}} \text{ et } v_2 = x_1ax_2a\cdots x_{\frac{n}{2}}a$$

sont deux éléments distincts de $L_n(u)$ tels que $\theta_n(v_1) = \theta_n(v_2) = v'$.

Pour tout élément v de $L_n(u)$ on a:

$$\mathsf{c}(v) = \left(\frac{n}{2}, |v'|_b, |v'|_c\right) \text{ où } v' = \theta_n(v).$$

Comme θ est surjectif on en déduit:

$$\{\mathsf{c}(v) : v \in L_n(u)\} = \left\{\left(\frac{n}{2}, |v'|_b, |v'|_c\right) : v' \in L_{\frac{n}{2}}(u')\right\}.$$

D'où

$$\#V_n(u) = \#\left\{\left(\frac{n}{2}, |v'|_b, |v'|_c\right) : v' \in L_{\frac{n}{2}}(u')\right\} = \#V_{\frac{n}{2}}(u') = 2$$

car u' est sturmien.

(iii) Soit n un entier impair différent de 1 et $v \in L_n(u)$. On peut écrire $v = wx$ où x est une lettre et w un facteur de longueur $n-1$ de u. Soit F_1 (resp. F_2) l'ensemble des éléments de $L_n(u)$ vérifiant l'assertion (i) (resp.(ii)) de la proposition 4.2.4. D'après la même proposition nous avons

$$L_n(u) = F_1 \cup F_2 \text{ avec } F_1 \cap F_2 = \emptyset.$$

Soit φ l'application de F_1 dans $L_{\frac{n-1}{2}}(u')$ consistant à effacer la lettre spéciale a. On a l'égalité vectorielle

$$\mathsf{c}(v) = \left(\frac{n+1}{2}, |v'|_b, |v'|_c\right) \text{ où } v' = \theta_n(v).$$

À tout élément $v' = x_1x_2\cdots x_{\frac{n-1}{2}}$ de $L_{\frac{n-1}{2}}(u')$ correspond l'unique élément $v = ax_1ax_2\cdots ax_{\frac{n-1}{2}}a$ de F_1. Il en résulte que φ est bijective. D'où

$$\{\mathsf{c}(v) : v \in F_1\} = \left\{\left(\frac{n+1}{2}, |v'|_b, |v'|_c\right) : v' \in L_{\frac{n-1}{2}}(u')\right\}.$$

Ainsi

$$\#\{\mathsf{c}(v) : v \in F_1\} = \#\left\{\left(\frac{n+1}{2}, |v'|_b, |v'|_c\right) : v' \in L_{\frac{n-1}{2}}(u')\right\} = \#V_{\frac{n-1}{2}}(u') = 2$$

car u' est sturmien. De même soit ψ l'application de F_2 dans $L_{\frac{n+1}{2}}(u')$ qui consiste à effacer la lettre spéciale a. On a

$$\mathsf{c}(v) = \left(\frac{n-1}{2}, |v'|_b, |v'|_c\right)$$

où $v' = \psi_n(v)$. À tout élément $v' = x_1x_2\cdots x_{\frac{n+1}{2}}$ de $L_{\frac{n+1}{2}}(u')$ correspond l'unique élément $v = ax_1ax_2\cdots ax_{\frac{n+1}{2}}a$ de F_2. Il en résulte que ψ est bijective. D'où

$$\{\mathsf{c}(v) : v \in F_2\} = \left\{\left(\frac{n-1}{2}, |v'|_b, |v'|_c\right) : v' \in L_{\frac{n+1}{2}}(u')\right\}.$$

Ainsi

$$\#\{\mathsf{c}(v) : v \in F_2\} = \#\left\{\left(\frac{n-1}{2}, |v'|_b, |v'|_c\right) : v' \in L_{\frac{n+1}{2}}(u')\right\} = \#V_{\frac{n+1}{2}}(u') = 2$$

car u' est sturmien. D'autre part, comme F_1 et F_2 forment une partition de $L_n(u)$, on a:

$$V_n(u) = \{\mathsf{c}(v) : v \in F_1\} \cup \{\mathsf{c}(v) : v \in F_2\}.$$

De plus la réunion est disjointe car

$$\left\{\left(\tfrac{n+1}{2}, |v'|_b, |v'|_c\right) : v' \in L_{\frac{n-1}{2}}(u')\right\} \cap \left\{\left(\tfrac{n-1}{2}, |v'|_b, |v'|_c\right) : v' \in L_{\frac{n+1}{2}}(u')\right\} = \emptyset.$$

Par suite

$$\#V_n(u) = \#\{\mathsf{c}(v) : v \in F_1\} + \#\{\mathsf{c}(v) : v \in F_2\}.$$

Ainsi,

$$\#V_n(u) = 2 + 2 = 4.$$

Réciproquement supposons les conditions (i), (ii) et (iii) satisfaites par u. On déduit de (i) que les trois lettres sont présentes dans u.

De la condition (ii) on a $\#V_2(u) = 2$. Si $(2,0,0)$ est dans $V_2(u)$ alors le deuxième vecteur est soit $(1,1,0)$, soit $(1\,0,1)$. En prenant $V_2(u) = \Big\{(2,0,0);\ (1,1,0)\Big\}$, on déduit qu'aucun facteur de longueur 2 de u ne contient la lettre c. En d'autres termes, c n'est pas dans u; contradiction. En prenant $V_2(u) = \Big\{(2,0,0);\ (1,0,1)\Big\}$ on aboutit à la même conclusion. Donc $(2,0,0)$ n'est pas dans $V_2(u)$. De même on montre que $(0,2,0)$ et $(0,2,0)$ ne sont pas dans $V_2(u)$. Ainsi, les mots aa, bb et cc ne sont pas facteurs de u. Par conséquent $V_2(u)$ est l'une des trois paires:

$$\Big\{(1,1,0);\ (1,0,1)\Big\},\ \Big\{(1,1,0);\ (0,1,1)\Big\} \text{ et } \Big\{(1,0,1);\ (0,1,1)\Big\}$$

On peut supposer pour continuer la preuve que

$$V_2(u) = \Big\{(1,1,0);\ (1,0,1)\Big\}$$

Par suite, les facteurs de longueur 2 de u, formés chacun de lettres distinctes contiennent toujours la lettre a. Donc, les facteurs de longueur 2 de u sont ab, ac, ba et ca. Plus précisément les lettres b et c n'admettent qu'un seul prolongement à droite et à gauche. Par conséquent u est de la forme

$$u = x_0 a x_1 a x_2 a x_3 \cdots$$

où $x_0 \in \{\varepsilon, b, c\}$ et $x_i \in \{b, c\}$ pour $i = 1, 2, 3, \ldots$.

Le mot u est non ultimement périodique. En effet, en supposant le contraire, u est alors périodique puisqu'il est récurrent. Par suite, $\#V_m(u) = 1$ où m est la période de u. Ce qui est impossible car selon les conditions (i), (ii) et (iii) $\#V_n(u) \neq 1$ pour tout entier naturel n non nul.

Montrons à présent que u est de complexité $n+2$, $(n \geq 1)$. Pour cela nous allons établir que pour tout naturel non nul n, u admet un unique facteur biprolongeable de longueur n. Par l'absurde, supposons que u admette deux facteurs biprolongeables à droite. Alors u contient deux facteurs biprolongeables à droite de la forme xr et yr où $x, y \in \{b, c\}$

et $r \in L(u)$. Par suite, les mots :

$$xrx,\ xry,\ yrx \text{ et } yry$$

sont des facteurs de u.

Posons:

$$F = \{xrx, xry, yrx, yry\}.$$

On remarque aisément que:

$$\#\{\mathsf{c}(w) : w \in F\} = 3.$$

Par ailleurs

$$F' = \{axr, rxa, ayr, rya\}$$

est un ensemble de facteurs de u de même longueur que les éléments de F.

De plus

$$\#\{\mathsf{c}(w) : w \in F'\} = 2.$$

Les éléments de F' contiennent un a de plus que ceux de F. Donc

$$\{\mathsf{c}(w) : w \in F\} \cap \{\mathsf{c}(w) : w \in F'\} = \emptyset$$

Considérons m la longueur commune des éléments de F et F'. Nous avons:

$$(F \cup F') \subset L_m(u).$$

Par conséquent

$$\begin{aligned}
\#V_m(u) &\geq \#\{\mathsf{c}(w) : w \in F \cup F'\} \\
&= \#\{\mathsf{c}(w) : w \in F\} + \#\{\mathsf{c}(w) : w \in F'\} \\
&= 3 + 2 = 5
\end{aligned}$$

Absurde, car selon (i), (ii) et (iii) on a:

$$\#V_m(u) \leq 4.$$

Exemple: Dressons les ensembles $L_k(u)$ pour $k = 1, 2, \ldots, 6, \ldots$

$$\begin{aligned}
L_1(Q) &= \{a, b, c\} \\
L_2(Q) &= \{ac, bc, ca, cb\} \\
L_3(Q) &= \{aca, acb, bca, cac, cbc\} \\
L_4(Q) &= \{acac, acbc, bcac, caca, cacb, cbca\} \\
L_5(Q) &= \{acacb, acbca, bcaca, bcacb, cacac, cacbc, cbcac\} \\
L_6(Q) &= \{acacbc, acbcac, bcacac, bcacbc, cacacb, cacbca, cbcaca, cbcacb\} \\
&\ldots\ldots
\end{aligned}$$

On en déduit les ensembles $V_k(Q)$, $k = 1, 2, \ldots, 6 \ldots$:

$$V_1(Q) = \{(1,0,0);\ (0,1,0);\ (0,0,1)\}$$

$$V_2(Q) = \{(1,0,1);\ (0,1,1)\}$$

$$V_3(Q) = \{(2,0,1);\ (1,0,2);\ (1,1,1);\ (0,1,2)\}$$

$$V_4(Q) = \{(2,0,2);\ (1,1,2)\}$$

$$V_5(Q) = \{(2,1,2);\ (1,2,2);\ (2,0,3);\ (1,1,3)\}$$

$$V_6(Q) = \{(2,1,3);\ (1,2,3)\}$$

$$\ldots\ldots\ldots$$

Par dénombrement direct on obtient:

$\#V_1(Q) = 3$, $\#V_2(Q) = 2$, $\#V_3(Q) = 4$, $\#V_4(Q) = 2$, $\#V_5(Q) = 4$ et $\#V_6(Q) = 2, \cdots$

Lemme 4.4.1 *L'ensemble des facteurs d'un mot quasi-sturmien par insertion est stable par image miroir.*

Preuve: Soit u un mot quasi-sturmien par insertion et u' son mot sturmien extrait. Soit m un entier non nul et $v \in L_{2m}(u)$. On peut supposer sans perdre la généralité que u est a-spécial. Par suite, d'après la proposition 4.2.3 on a:

$$v = ax_1ax_2 \cdots ax_m \text{ ou } v = x_1ax_2 \cdots ax_ma.$$

Ainsi le mot $v' = x_1 x_2 \cdots x_m$ est un facteur de u' et puisque u' est un mot sturmien alors $\overline{v'} = x_m x_{m-1} \cdots x_1$ est dans u' (Théorème 2.2.6) . Pour tout $y_1 y_2 \ldots y_n$ dans u', les deux mots $ay_1 ay_2 a \ldots ay_n$ et $y_1 ay_2 a \ldots ay_n a$ sont dans u car le mot $ay_1 ay_2 a \ldots ay_n a$ est dans u. Par conséquent les mots

$$x_m a x_{m-1} a \cdots x_1 a \text{ et } a x_m a x_{m-1} \cdots a x_1$$

sont dans u. D'où $\overline{v} \in L_{2m}(u)$ et le miroir préserve les facteurs de longueur paire de u. Par ailleurs comme tout facteur de u de longueur impaire de u est préfixe d'un facteur de longueur paire plus grande on déduit que les facteurs de u de longueur impaire sont aussi préservés par le miroir. $\square$

Comme l'ensemble des facteurs de longueur 2 d'un mot de complexité $n+2$ de la forme (1), (2) ou (4) n'est pas préservé par le miroir on fait la remarque suivante:

Remarque 4.4.1 *L'ensemble des facteurs d'un mot de complexité $n+2$ de la forme (1), (2) ou (4) du théorème 4.3.1 n'est pas stable par le miroir.*

Lemme 4.4.2 *Soit u un mot quasi-sturmien par insertion. Le mot u ne contient pas de palindrome de longueur paire.*

Preuve: D'après la proposition 4.2.3 tout facteur de u de longueur paire commence par une lettre par laquelle il ne finit pas et ne peut donc être un palindrome. $\square$

Proposition 4.4.3 *Soit u un mot quasi-sturmien par insertion. u ne contient pas de facteur bispécial de longueur paire.*

Preuve: Soit u un mot quasi-sturmien par insertion. On peut supposer que u est a-spécial sans perte de généralité. Alors tout facteur de u débutant (resp. finissant) par b ou c ne peut être biprolongeable à gauche (resp. à droite). Ainsi, d'après la proposition 4.2.3, aucun facteur de longueur paire ne peut être biprolongeable à la fois à gauche et à droite. $\square$

En vertu de cette proposition il convient de faire la remarque suivante.

Remarque 4.4.2 *Les facteurs bispéciaux des mots quasi-sturmiens par insertion sont de longueurs impaires.*

Proposition 4.4.4 *Le miroir d'un facteur biprolongeable à droite est biprolongeable à gauche et inversement.*

preuve: Soit u un mot quasi-sturmien par insertion que l'on peut supposer a-spécial sans perte de généralité. Soient n un entier non nul et $D_n, G_n \in L_n(u)$ tels que

$$\partial^- G_n = \partial^+ D_n = 2.$$

Alors nous avons:

$$D_n b,\ D_n c,\ bG_n,\ cG_n \in L_{n+1}(u).$$

Par suite, en vertu du lemme 4.2.1, nous avons:

$$b\overline{D_n},\ c\overline{D_n} \in L_{n+1}(u)$$

et donc $\partial^- \overline{D_n} = 2$. Ainsi $\overline{D_n}$ est biprolongeable à gauche.
De même on montre que $\overline{G_n}$ est biprolongeable à droite. □

Remarque 4.4.3 *Soient u un mot quasi-sturmien par insertion et P un palindrome de u. Si P est biprolongeable à droite alors P est biprolongeable à gauche et inversement.*

Proposition 4.4.5 *Soient u un mot quasi-sturmien par insertion. Tout facteur bispécial de u commence et finit par la lettre spéciale de u.*

Preuve: Soit s un facteur bispécial de u. Par la remarque 4.4.3 on sait que s est de longueur impaire. Par suite, le résultat découle de la proposition 4.2.4 puisque tout facteur de u ne finissant ou ne commençant pas par la lettre spéciale de u ne peut être bispécial. □

Lemme 4.4.3 *Soit u un mot quasi-sturmien par insertion. Alors, u possède, pour tout entier n impair, exactement trois palindromes de longueur n.*

Preuve: Soit u un mot quasi-sturmien par insertion.
Pour $n = 1$ la propriété est vraie car les trois lettres a, b et c présentes dans u sont des palindromes. Supposons la propriété vraie pour $n = 2m - 1$ supérieur à 1: Il existe exactement trois palindromes P_1, P_2 et P_3 de longueur $2m - 1$ dans u . Verifions que la propriété reste vraie pour la longueur $2m + 1$. Deux éventualités se présentent: soit l'un

de ces trois palindromes est spécial, soit aucun d'eux ne l'est.

Cas 1: Les palindromes P_1, P_2 et P_3 ne sont pas spéciaux. Montrons qu'il existe des lettres x_1, x_2 et x_3 dans $\{a,b,c\}$ telles que $x_iP_ix_i \in L_{2m+1}(u)$ pour $i = 1,2,3$.

Comme P_i $(i = 1,2,3)$ est non spécial alors P_i admet un unique prolongement à droite (resp. à gauche) dans u. Soit x_i l'unique prolongement à droite de P_i. On a $P_ix_i \in L_{2m}(u)$ et d'après le lemme 4.4.1 $x_iP_i = \overline{P_ix_i} \in L_{2m}(u)$. Ainsi x_i est l'unique prolongement à gauche de P_i dans u. Par conséquent $x_iP_ix_i \in L_{2m+1}(u)$ pour $i = 1,2$ ou 3. Ainsi il existe des lettres x_1, x_2 et x_3 dans $\{a,b,c\}$ telles que $x_iP_ix_i \in L_{2m+1}(u)$ pour $i = 1,2,3$. Ces trois mots

$$x_1P_1x_1,\ x_2P_2x_2 \text{ et } x_3P_3x_3$$

(qui sont distincts puisque P_1, P_2 et P_3 le sont) sont des palindromes sur u.

Cas 2: L'un des palindromes P_1, P_2 et P_3 est spécial. On peut supposer qu'il s'agit de P_1 pour continuer la preuve. Donc les palindromes P_2 et P_3 ne sont pas spéciaux. Ainsi on montre comme dans le cas 1 qu'il existe des lettres x_2 et x_3 dans $\{a,b,c\}$ telles que $x_iP_ix_i \in L_{2m+1}(u)$ avec $i = 2,3$.

Considérons maintenant le palindrome biprolongeable P_1. On a $\partial^- P_1 = \partial^+ P_1 = 2$ et il existe donc deux lettres x et y telles que $P_1x, P_1y, xP_1, yP_1 \in L_{2m}(u)$. L'unique facteur biprolongeable à gauche de longueur $2m$ est soit P_1x soit P_1y. On peut supposer sans perte de généralité que P_1x est ce facteur biprolongeable et donc P_1y et son miroir yP_1 ne sont pas biprolongeables dans u. Par suite, $xP_1x, yP_1x \in L_{2m-1}(u)$ et $yP_1y \notin L_{2m+1}(u)$. Ce qui prouve qu'il existe des lettres x_1, x_2 et x_3 dans $\{a,b,c\}$ telles que $x_iP_ix_i \in L_{2m+1}(u)$ pour $i = 1,2,3$.

Finalement u possède au moins trois palindromes de longueur $2m+1$.

Soit R un palindrome de longueur $2m+1$ sur u. Alors R est de la forme xPx avec $P \in L_{2m-1}(u)$ et $x \in A$. Comme R est un palindrome on a $R = \overline{R}$ i.e $xPx = \overline{xPx} = x\overline{P}x$. D'où $P = \overline{P}$ et comme P est de longueur $2m-1$ (impaire) alors par hypothèse de récurrence $P \in \{P_1, P_2, P_3\}$. Par conséquent $R \in \{x_1P_1x_1, x_2P_2x_2, x_3P_3x_3\}$. Ainsi u admet exactement trois palindromes. □

Définition 4.4.2 *La fonction de complexité palindromique pal_u d'un mot u est l'application de $\mathbb{N}$ dans $\mathbb{N}$ qui compte le nombre de palindromes de longueur n contenus dans u i.e $pal_u(n) = \#\{v \in L_n(u) : \overline{v} = v\}$.*

Proposition 4.4.6 *Pour tout mot infini périodique u, il existe un entier n_0 tel que $pal_u(n) \leq 2$ pour tout $n \geq n_0$.*

Preuve: Soient u un mot périodique de longueur T et I_n, l'ensemble des palindromes de longueur n de u avec $n \geq T$. Par la périodicité de u nous avons:

$$I_n = \{i < T : u_i u_{i+1} \cdots u_{i+n-1} = u_{i+n-1} \cdots u_{i+1} u_i\}.$$

Supposons I_n non vide et posons $i_0 = \min(I_n)$.
Soit $j \in I_n$. Nous avons alors:

$$u_{i_0+h} = u_{i_0+n-1-h} \text{ et } u_{j+h} = u_{j+n-1-h}$$

pour tout $h \in \{0, 1, \ldots, n-1\}$. D'où

$$\begin{aligned}
u_{i_0+n-1-h} &= u_{j+n-1-(h+j-i_0)} \\
&= u_{j+(h+j-i_0)} \text{ si } h \leq n-1-(j-i_0) \\
&= u_{i_0+h+2(j-i_0)} \\
&= u_{i_0+h}
\end{aligned}$$

Pour n suffisamment grand on a $n-1-(j-i_0) \geq T$ car $j - i_0 \leq T$. Soit n_0 le plus petit entier vérifiant cette condition. Alors pour $n \geq n_0$ on a $u_{i_0+h+2(j-i_0)} = u_{i_0+h}$ pour tout $h \in \{0, 1, \ldots, n-1-(j-i_0)\}$ et donc pour tout $h \in \{0, 1, \ldots, T-1\}$. Par conséquent la période T divise $2(j-i_0)$. Ainsi $2(j-i_0) = 0$ ou T puisque $j - i_0 < T$. Par suite, $j = i_0$ ou $j = i_0 + \frac{T}{2}$ et donc $\# I_n \leq 2$. □

Théorème 4.4.7 *Un mot récurrent u est quasi-sturmien par insertion si et seulement si les assertions suivantes sont vérifiées:*

(i) Le miroir préserve les facteurs de u i.e

$$\forall v \in L(u) : \ \overline{v} \in L(u).$$

(ii) $pal_u(2) = 0$.

(iii) $pal_u(n) = 3$ pour tout entier n impair.

(iv) Une et une seule des lettres présentes dans u est biprolongeable.

Preuve: Soit u un mot récurrent.

Supposons u quasi-sturmien par insertion. Les lemmes 4.4.1, 4.4.2 et 4.4.3 nous assurent respectivement les assertions (i), (ii), (iii). Le mot u vérifie (iv) comme tout mot quasi-sturmien par insertion.

Réciproquement supposons les assertions (i), (ii), (iii) et (iv) vraies pour u.

1[ère] **étape:** Montrons que u est de la forme

$$u = x_0 z x_1 z x_2 z x_3 \cdots$$

où $x_0 \in \{\varepsilon, x, y\}$ et $x_i \in \{x, y\}$ pour $i = 1, 2, 3, \ldots$ avec $\{x, y, z\} = \{a, b, c\}$. Les trois palindromes de longueur 1 (conformément à l'assertion (iii)) de u sont les lettres a, b et c. En vertu de l'assertion (ii), les mots aa, bb et cc ne sont pas dans u. En d'autres termes, u ne contient pas de carré de lettre. Par conséquent u ne possède pas de lettre triprolongeable et ainsi toute lettre spéciale dans u est biprolongeable à gauche ou à droite. De l'assertion (i) on déduit que toute lettre spéciale dans u est bispéciale. Donc, par (iv) on conclut que u ne possède qu'une seule lettre spéciale: soit z cette lettre. Alors, les deux autres lettres désignées par x et y sont non spéciales. Par suite, les facteurs de longueur 2 de u sont nécessairement zx, zy, xz et yz. Par conséquent u est de la forme

$$u = x_0 z x_1 z x_2 z x_3 \cdots$$

où $x_0 \in \{\varepsilon, x, y\}$ et $x_i \in \{x, y\}$ pour $i = 1, 2, 3, \ldots$ avec $\{x, y, z\} = \{a, b, c\}$.

2[ème] **étape:** Montrons que u est de complexité $n + 2$ ($n \geq 1$). Le mot u est non ultimement périodique car sinon il serait périodique puisqu'il est récurrent et ainsi il contiendrait au plus 2 palindromes pour toute longueur donnée à partir d'un certain rang (cf. proposition 4.4.6) et ceci contredirait l'assertion (iii). Par suite, u possède au moins un facteur biprolongeable à droite de longueur n pour tout naturel non nul n. Ainsi il nous suffira de montrer que pour tout entier n non nul, u admet un et un seul facteur de longueur n biprolongeable à droite. En effet, par l'absurde, supposons que u admette deux facteurs biprolongeables à droite de même longueur. Considérons R et R' de tels facteurs de longueur n minimale. Par minimalité de n nous aurons $R = xD$ et $R' = yD$ où x et y sont deux lettres différentes et D l'unique facteur biprolongeable à droite de longueur $n-1$. Par suite xDx, xDy, yDx et yDy sont des facteurs de u. Par l'assertion (i), il vient que les mots $x\overline{D}x$, $y\overline{D}x$, $x\overline{D}y$ et $y\overline{D}y$ sont aussi des facteurs de u. Ainsi, les facteurs D et $\overline{D}$ sont bispéciaux et à fortiori biprolongeables à droite de longueur $n-1$ dans u. Ce qui prouve que $D = \overline{D}$ puisque le facteur biprolongeable de longueur $n-1$ est unique. En d'autres termes D est un palindrome. Par suite, $n-1$ est impair et selon (iii) on peut donc considérer P_1 et P_2 les deux autres palindromes de longueur $n-1$. Les palindromes P_1 et P_2 sont non biprolongeables car par minimalité de n, D est l'unique facteur biprolongeable de longueur $n-1$ de u. Par (i) on montre que P_k $(k=1,2)$ se prolonge à gauche et à droite par la même lettre: il existe des lettres y_1 et y_2 telles que $y_1P_1y_1$ et $y_2P_2y_2$ soient deux facteurs de u. En somme u contient au total quatre palindromes de longueur $n+1$ en l'occurence xDx, yDy, $y_1P_1y_1$ et $y_2P_2y_2$ puisque D, P_1 et P_2 sont distincts. Ce qui est contradictoire avec (iii). $\square$

Bibliographie

[1] P. ALESSANDRI, *Classification et représentation des mots de complexité $n + 2$*, Preprint, LMD, Marseille, 1995.

[2] P. ALESSANDRI, *Codage de rotations et basses complexités*, Université Aix-Marseille II, Thèse 1996.

[3] J.-P. ALLOUCHE, *Sur la complexité des suites infinies*, Bull. Belg. Math. Soc. 1 (1994) 133-143.

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

[5] J. BERSTEL, *Recents results in Sturmian words*, in: J. Dassow (ed.), Proc. 2nd internat. conf. Developments in Language Theory (DLT). World scientific, Singapore, 1995.

[6] J. BERSTEL and A. de LUCA, *Sturmian words, Lyndon words and trees*, Theoret. Comput. Sci. 178 (1-2) (1997) 171-203.

[7] J. BERSTEL et P. SÉÉBOLD, *A characterization of Sturmian morphisms*, in: A. Borziskowski. S. Sokolowski (Eds.). MFCS'93. Lecture Notes in Computer Science 711 (1993) 281-290.

[8] J. BERSTEL et P. SÉÉBOLD, *Morphismes de Sturm*, Bull. Belg. Math. Soc. Simon Stevin 1 (1994) 175-189.

[9] J. BERSTEL et P. SÉÉBOLD, *A remark on morphic Sturmian words*, Informatique théorique et applications 28 (1994) 255-263..

[10] V. BERTHÉ, *Fréquences des facteurs des suites sturmiennes*, Theoret. Comput. Sci. 165 (1996) 295-309.

[11] S. BRLEK, *Enumeration of factors in Thue-Morse word*, Discrete Appl. Math. 24 (1989) 83-96.

[12] J. CASSAIGNE, *Complexité et facteurs spéciaux*, Bull. Belg. Math. Soc. 4 (1997), 67-88.

[13] J. CASSAIGNE, *Sequences with grouped factors*, in Developments in Language Theory III (DLT'97) pp-211-222, Aristotle University of Thessaloniki, 1998.

[14] J. CASSAIGNE, S. FERENCZI and L. Q. ZAMBONI, *Imbalances in Arnoux-Rauzy sequences*; Ann. Inst. Fourier, Grenoble 50, 4 (2000), 1265-1276.

[15] N. CHEKHOVA, P. HUBERT et A. MESSAOUDI, *Propriétés combinatoires, ergodiques et arithmétiques de la substitution de Tribonacci*, Preprint 98-24, Institut de Mathématiques de Luminy, Marseille (1998).

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

[17] D. CRISP, W. MORAN, A. POLLINGTON and P. SHIUE, *Substitution invariant cutting sequences*, J. Théorie des nombres de Bordeaux 5 (1993) 123-138.

[18] N. G. de BRUIJN, *A combinatorial problem*, Nederl. Akad. Wetensch. Proc. 49 (1946) 758-764 Indagationes Math. 8, 461-467 (1946).

[19] A. de LUCA and S. VARRICCHIO, *Some combinatorial properties of the Thue-Morse Sequence and a problem in semigroups*. Theoret. Comput. Sci. 63 (1989) 333-348.

[20] G. DIDIER, *Caractérisation des N-écritures et application à l'étude des suites de complexité ultimement $n + C^{ste}$*, Theoret. Comput. Sci. 215 (1999), 31-49.

[21] S. DULUCQ, D. GOUYOU-BEAUCHAMPS, *Sur les facteurs des suites de Sturm*, Theoret. comput. Sci. 71 (1990) 381-400.

[22] S. FERENCZI et C. MAUDUIT, *Transcendency of numbers with a low complexity expansion*, J. Number Theory 67 (1997) 146-161.

[23] N. P. FOGG, *Substitutions in Dynamics, Arithmetics and Combinatorics*, Lecture Notes in Mathematics (LNM) 1794 (2002).

[24] P. HUBERT, *Dynamique symbolique des billards polygonaux rationnels*; Thèse de doctorat Université d'Aix-Marseille II (1995).

[25] I. KABORÉ et T. TAPSOBA, *Caractérisation des morphismes d'Arnoux-Rauzy*, Preprint, 2003.

[26] I. KABORÉ et T. TAPSOBA, *Combinatoire de mots récurrents de complexité* $n+2$, Preprint, 2003.

[27] M. LOTHAIRE, *Algebraic combinatorics on words*, Cambridge University Press, 2002.

[28] F. MIGNOSI et P. SÉÉBOLD, *Morphismes sturmiens et règles de Rauzy*, J. Théorie des nombres de Bordeaux 5 (1993), 221-233.

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

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

[31] M. QUEFFELEC: *Contribution à l'étude spectrale de suites arithmétiques*. Paris-Nord. Thèse d'Etat, 1984.

[32] G. RAUZY, *Suites à termes dans un alphabet fini*. Sémin. Théorie des nombres (1982-1983) 25-01. 25-16. Bordeaux.

[33] G. ROTE, *Sequences with subword complexity* $2n$, J. Number Theory 46, 196-216 (1994).

[34] P. SÉÉBOLD, *Fibonacci morphisms and Sturmians words*, Theoret. Comput. Sci. 88 (1991) 365-384.

[35] T. TAPSOBA, *Automate calculant la complexité de suites automatiques*, J. Théorie des nombres de Bordeaux 6 (1994), 127-134.

[36] A. THUE, *Über unendliche Zeichenreihen*, Norske Vid. Skr. I. Kl. Christiana 7 (1906), 1-22.

[37] A. THUE, *Über die gegeneitige Lage gleicher Teile gewisser Zeichenreihen*, Norske Vid. Selsk. Skr. I. Math. Nat. Kl. Christiana 1 (1912), 1-67.

[38] N. WOZNY, L. Q. ZAMBONI, *Frequencies of factors in Arnoux-Rauzy sequences*, Acta Arithmetica 96 (2001) 261-278.

RÉSUMÉ: On appelle *complexité* d'un mot infini u, la fonction, habituellement notée p_u ou simplement p, qui compte le nombre de facteurs distincts de longueur donnée dans u. Nous étudions dans ce mémoire certaines classes de mots de basse complexité. Il s'agit notamment des mots de complexité $n+1$ appelés *mots sturmiens*, les *mots d'Arnoux-Rauzy* qui sont de complexité $2n+1$ et les *mots de complexité* $n+2$. Dans le chapitre 1 nous rappelons les concepts de base en combinatoire de mots. Le chapitre 2 est consacré aux mots sturmiens et les morphismes qui les préservent. Dans le chapitre 3 nous étudions les *morphismes d'Arnoux-Rauzy* *i.e* les morphismes qui préservent la classe des mots d'Arnoux-Rauzy. Nous y énonçons quelques propriétés combinatoires des mots d'Arnoux-Rauzy puis nous montrons que l'ensemble des morphismes d'Arnoux-Rauzy est un *monoïde* engendré par quatre morphismes particuliers. Nous montrons ensuite que ce monoïde est stable par l'opération "miroir" et unitaire à gauche. Dans le dernier chapitre nous classifions les mots ternaires de complexité $n+2$. Ensuite, nous étudions singulièrement une sous classe de ces mots que nous appelons mots *quasi-sturmiens par insertion*. Nous montrons que les mots de cette classe sont 2-*équilibrés* puis nous déterminons la fonction de *complexité palindromique*: ces mots possèdent 3 palindromes pour toute longueur impaire donnée mais ne contiennent pas de palindrome non trivial de longueur paire.

ABSTRACT: We call *complexity function* of an infinite word u, the function, usually denoted by p_u or simply p, which counts the number of distinct factors for given length in u. In this thesis, we study some classes of infinite words of low complexty: words with complexity $n+1$ called *sturmian words*, *Arnoux-Rauzy sequences* which are words with complexity $2n+1$ and *words with complexity* $n+2$. In the first chapter we recall some basic concepts in combinatorics on words. The second chapter is devoted to sturmian words and morphisms which preserve the class of these words. In the third chapter we study *Arnoux-Rauzy morphisms* *i.e* morphisms which preserve the class of Arnoux-Rauzy sequences. We give some combinatorial properties of Arnoux-Rauzy sequences and we show that the set of Arnoux-Rauzy morphisms is a *monoid* generated by a familly of four particular morphisms. We next show that this monoid is preserved by morphisms miror operator and is left unitary. In the last chapter we classify the words with complexity $n+2$, for n non-zero integer. Next, we particulary study one sub class of these words that we call *quasi-sturmian words by inserting*. We show that these words are 2-*balanced* and we determine their *palindromic complexity function*: These words have 3 palindromes for any given odd length but no empty palindrome for even length.

Mots clés: mots infinis, facteur de mots, complexité, facteurs spéciaux, équilibre, palindrome, morphismes.