Les mots de Sturm

Fathi BEN ARIBI

20 décembre 2008

1 Objectifs

Dans cette présentation, nous donnerons quelques résultats de combinatoire des mots. Avant tout, il est nécessaire d'introduire quelques notations :

Définition 1. Soit w un mot sur un alphabet A. On note $F(w)$ l'ensemble des facteurs finis de w, $F_n(w)$ l'ensemble des facteurs de w de longueur n, et $P(w,n) = |F_n(w)|$ le nombre de facteurs de w de longueur n. On dit qu'un ensemble de mots finis X est *factoriel* s'il existe un mot w tel que $X = F(w)$.

Définition 2. Un mot infini w est dit *de Sturm* si $\forall n \geq 1, P(w,n) = n+1$.

L'objectif de cette présentation est de montrer que cette propriété (1) est équivalente à chacune des deux propriétés suivantes :
(2) w est équilibré apériodique
(3) w est mécanique irrationnel
Ces termes seront définis par la suite.

Remarquons qu'un mot de Sturm a exactement deux facteurs de longueur 1, donc s'écrit avec deux lettres. Pour cette raison, par la suite, nous n'étudierons que des mots sur l'alphabet $A = \{0; 1\}$

2 Mots équilibrés - Mots périodiques

Nous allons tout d'abord montrer l'équivalence entre les propriétés (1) et (2).

Définition 3. Soit x un mot fini. Sa *hauteur* $h(x)$ est le nombre de 1 qu'il contient. L'*écart* entre deux mots finis x et y est $\delta(x,y) = |h(x) - h(y)|$. Un ensemble X de mots finis est dit *équilibré* si $\forall x, y \in X, |x| = |y| \Rightarrow \delta(x,y) \leqslant 1$. Un mot w est dit *équilibré* si $F(w)$ est équilibré.

Définition 4. Un mot infini w est dit *périodique* (ou *purement périodique*) s'il est de la forme $w = y^\omega$ pour un certain $y \in A^*$. Un mot infini w est dit *éventuellement périodique* s'il est de la forme $w = xy^\omega$ pour $x, y \in A^*$. Un mot infini w est dit *apériodique* s'il n'est pas éventuellement périodique.

Proposition 1. *Soit X un ensemble factoriel équilibré. Alors* $\forall n \geqslant 0$, $|X \cap A^n| \leqslant n+1$.

Démonstration. Le résultat est clair pour $n = 0$ et $n = 1$. Il est vrai pour $n = 2$, car 00 et 11 ne peuvent pas être tous deux dans X (car X est équilibré). Pour les autres cas, raisonnons par l'absurde et posons $n \geqslant 3$ le plus petit entier pour lequel $|X \cap A^n| > n+1$. Soit $Y = X \cap A^{n-1}$ et $Z = X \cap A^n$. On a donc $|Y| \leqslant n$ et $|Z| \geqslant n+2$. Chaque $z \in Z$ a son suffixe de longueur $n-1$ dans Y, donc, par le principe des tiroirs, il existe $y, y' \in Y, y \neq y'$, tels que $0y, 1y, 0y', 1y' \in Z$. $y \neq y'$, donc il existe x tel que $x0$ est préfixe de y et $x1$ est préfixe de y'. Mais alors $0x0$ et $1x1$ sont des éléments de X, ce qui contredit le fait que X est équilibré. □

Proposition 2. *Soit X factoriel. X est non équilibré si et seulement si il existe w palindrome tel que $0w0, 1w1 \in X$*

Démonstration. Le sens retour est évident. Montrons le sens direct : supposons X non équilibré. Soient alors $u, v \in X$ tels que $|u| = |v| = n, \delta(u, v) \geqslant 2$, avec n minimal. Les premières et dernières lettres de u et v sont distinctes ; supposons que u commence par 0 et v par 1, alors $u = 0wau', v = 1wbv'$ avec a et b des lettres distinctes. On a $a = 0$ et $b = 1$ car sinon $\delta(u', v') = \delta(u, v)$ ce qui contredit la minimalité de n. De même, $u' = v' = \epsilon$, donc $u = 0w0$ et $v = 1w1$ sont dans X. Montrons que w est un palindrome, par l'absurde. Si w n'est pas un palindrome, alors il existe un mot z et une lettre a, tels que za est préfixe de w, $\tilde{z}$ est suffixe de w mais $a\tilde{z}$ n'est pas suffixe de w. Soit b l'autre lettre. Alors $b\tilde{z}$ est suffixe de w. Donc $0za$ est préfixe de u et $b\tilde{z}1$ est suffixe de v. Si $a = 0$ et $b = 1$ alors $\delta(0z0, 1\tilde{z}1) = 2$, et sinon $u = 0z1u"$ et $v = v"1\tilde{z}0$ avec $\delta(u", v") = \delta(u, v)$; les deux cas contredisent la minimalité de n. Donc w est un palindrome, ce qui complète la preuve.

□

Définition 5. Soit x un mot infini. Le *graphe des facteurs d'ordre* n de x $G_n(x)$ est le graphe de sommets dans $F_n(x)$ et d'arêtes étiquetées par 0 ou 1 ; $G_n(x) = \{(au, b, ub); u \in F_{n-1}(x)\}$. On dit qu'un facteur p est *conservatif* si une arête exactement part de ce sommet dans le graphe, et *spécial droit* sinon.

Proposition 3. *Soit x un mot infini. x est éventuellement périodique si et seulement si $\exists n \geqslant 1, P(x, n) < n + 1$.*

Démonstration. Si $x = uy^\omega$, alors $P(x,n) \leqslant |uy|$, car il y a au plus $|uy|$ possibilités pour la première lettre d'un facteur de x de longueur n. Donc $\exists n \geqslant 1, P(x, n) < n + 1$.

Si $P(x, n) < n + 1$, alors $\exists m \geqslant 1, P(x, m) = P(x, m + 1)$. En effet, sinon, $\forall m \geqslant 1, P(x, m) < P(x, m + 1)$ et donc $P(x, n) > P(x, 1) + n - 1 = n + 1$, ce qui contredit l'hypothèse. Soit donc m tel que $P(x, m) = P(x, m + 1)$. Alors, dans $G_m(x)$, il part exactement une arête de chaque sommet. Donc la lecture du mot x par blocs de n lettres correspond à un chemin dans le graphe qui est ici éventuellement cyclique car le graphe a un nombre fini de sommets. Donc x est éventuellement périodique. □

Proposition 4. *Soit x un mot infini, $n \geqslant 1$, et c le nombre de facteurs conservatifs de x de longueur n. Si x a un facteur de longueur $n + c$ dont tous les facteurs de longueur n sont conservatifs, alors x est éventuellement périodique.*

Démonstration. Soit $w = a_1a_2...a_{n+c}$ facteur de x, dont tous les facteurs de longueur n sont conservatifs ; posons $p_i = a_i...a_{i+n-1}$ pour $i = 0, ..., c$. Dans le graphe $G_n(x)$, le chemin $(p_0; ...; p_c)$ est un sous-chemin du chemin de x, de longueur $c+1$, donc, comme il n'y a que c sommets conservatifs, contient un cycle ; de plus, chaque sommet p_i n'a qu'une seule arête sortante, donc le chemin de x doit rester dans ce cycle indéfiniment. x est donc éventuellement périodique. □

On peut maintenant montrer l'équivalence entre les propriétés (1) et (2) :

Théorème 1. *Soit x un mot infini. x est de Sturm si et seulement si x est équilibré et apériodique.*

Démonstration. Si x est apériodique, alors $\forall n \geqslant 0,\ P(x,n) \geqslant n+1$ selon la Proposition 3. Et si x est équilibré, alors $\forall n \geqslant 0,\ P(x,n) \leqslant n+1$ selon la Proposition 1. Donc x est un mot de Sturm.

Réciproquement, remarquons que x ne peut être de Sturm et éventuellement périodique, par la Proposition 3 ; supposons donc x de Sturm et non équilibré, et montrons qu'il est éventuellement périodique ; ceci fournira une contradiction, et on en déduira que si x est de Sturm, alors il est apériodique, mais aussi équilibré. x est non équilibré, donc il existe un palindrome w tel que $0w0$ et $1w1$ sont des facteurs de x. En particulier, w est spécial droit. Posons $n = |w| + 1$. x est de Sturm, donc il existe un seul facteur spécial droit de longueur n, qui est soit $0w$ soit $1w$. Supposons que $0w$ est spécial droit, donc $1w$ ne l'est pas, et donc $0w1$ est facteur de x mais $1w0$ ne l'est pas, car $1w1$ l'est déjà. Toute occurence de $1w$ dans le mot x est suivie d'un 1. Soit v un mot de longueur $n-1$ tel que $u = 1w1v \in F(x)$. u est de longueur $2n$, donc, selon la Proposition 4, il suffit de montrer que tous les facteurs de u de longueur n sont conservatifs pour montrer que x est éventuellement périodique. Pour cela, il suffit de montrer que le seul facteur spécial droit de longueur n de x, à savoir $0w$, n'est pas facteur de u. Raisonnons par l'absurde, et supoposons que $0w$ est facteur de $u = 1w1v$.

<table>
<tr><td colspan="7">u</td></tr>
<tr><td>1</td><td colspan="3">w</td><td>1</td><td colspan="2">v</td></tr>
<tr><td colspan="2"></td><td>0</td><td colspan="3">w</td><td></td></tr>
<tr><td>1</td><td>s</td><td>0</td><td>t</td><td>1</td><td>y</td><td>z</td></tr>
</table>

On a donc les factorisations $w = s0t$, $v = yz$ et $w = t1y$. w est un palindrome, donc $w = \tilde{t}0\tilde{s}$, et donc le préfixe t de w est suivi à la fois d'un 0 et d'un 1, d'où la contradiction. □

3 Pente d'un mot

Définition 6. Si x est fini, on appelle *pente* de x le rationnel $\pi(x) = \dfrac{h(x)}{|x|}$.

Remarquons que $\pi(xy) = \dfrac{|x|}{|xy|}\pi(x) + \dfrac{|y|}{|xy|}\pi(y)$.

Proposition 5. *Soit X un ensemble factoriel de mots finis. X est équilibré si et seulement si* $\forall x, y \in X, x, y \neq \epsilon, |\pi(x) - \pi(y)| < \dfrac{1}{|x|} + \dfrac{1}{|y|}$.

Démonstration. Si l'inégalité est vérifiée, alors en particulier pour $x, y \in X$ de même longueur, on a $|h(x) - h(y)| < 2$, donc X est équilibré.

Réciproquement, supposons X équilibré, soit $x, y \in X, x, y \neq \epsilon$. Si $|x| = |y|$, alors l'inégalité est vérifiée. Supposons $|x| > |y|$ et montrons le résultat par récurrence sur $|x| + |y|$. Le résultat est clair pour $|x| + |y| = 3$. Dans le cas général, posons $x = zt$ avec $|z| = |y|$. Par hypothèse de récurrence, $|\pi(t) - \pi(y)| < \dfrac{1}{|t|} + \dfrac{1}{|y|}$; de plus, X est équilibré, donc $|h(z) - h(y)| \leqslant 1$, donc $|\pi(z) - \pi(y)| < \dfrac{1}{|y|}$. D'où : $|\pi(x) - \pi(y)| = |\dfrac{|z|}{|x|}\pi(z) + \dfrac{|t|}{|x|}\pi(t) - \pi(y)| \leqslant \dfrac{|z|}{|x|}|\pi(z) - \pi(x)| + \dfrac{|t|}{|x|}|\pi(t) - \pi(y)| < \dfrac{1}{|x|} + \dfrac{|t|}{|x|}(\dfrac{1}{|y|} + \dfrac{1}{|t|}) = \dfrac{1}{|x|} + \dfrac{1}{|y|}$. □

Corollaire 1. *Soit x un mot infini équilibré. Notons x_n le préfixe de longueur n de x. Alors la suite de rationnels $(\pi(x_n))_{n>0}$ converge vers une limite α.*

On appelle cette limite la *pente* du mot infini x.

Démonstration. $(\pi(x_n))_{n>0}$ est de Cauchy, selon la proposition précédente, donc converge. □

Proposition 6. *Soit x un mot infini équilibré de pente α. Tout facteur u non vide de x vérifie :* $|\pi(u) - \alpha| \leqslant \dfrac{1}{|u|}$.

Et, plus précisément, on a
$\forall u \in F(x), \alpha|u| - 1 < h(u) \leqslant \alpha|u| + 1$
ou
$\forall u \in F(x), \alpha|u| - 1 \leqslant h(u) < \alpha|u| + 1$.
Bien sûr, ces inégalités sont toutes strictes si α est irrationnel.

Démonstration. Soit x_n le préfixe de longueur n de x. Soit $\eta > 0$, et n_0 tel que $\forall n \geqslant n_0,\ |\pi(x_n) - \alpha| \leqslant \eta$. Par la Proposition 5, on a donc

$$|\pi(u) - \alpha| \leqslant |\pi(u) - \pi(x_n)| + |\pi(x_n) - \alpha| < \frac{1}{|u|} + \frac{1}{n} + \eta$$

Pour $n \longrightarrow \infty$ et $\eta \longrightarrow 0$, on trouve la première inégalité. Cette inégalité implique que $\alpha|u| - 1 \leqslant h(u) \leqslant \alpha|u| + 1$. Si le deuxième résultat était faux, il existerait alors $u, v \in F(x)$ tels que $\alpha|u| - 1 = h(u)$ et $\alpha|v| + 1 = h(v)$. Mais alors $|\pi(u) - \pi(v)| = \dfrac{1}{u} + \dfrac{1}{v}$, ce qui contredit la Proposition 5.

□

Nous avons enfin le résultat suivant : Pour les mots équilibrés, la périodicité et la rationnalité de la pente sont liées. Ce résultat sera utile pour démontrer l'équivalence entre les propriétés (2) et (3).

Théorème 2. *Soit x un mot infini équilibré, de pente α. Alors α est rationnel si et seulement si x est éventuellement périodique.*

Démonstration. Si $x = uy^\omega$, alors

$$\pi(uy^n) = \frac{h(u) + nh(y)}{|u| + n|y|} \longrightarrow \pi(y)$$

donc la pente de x est rationnelle.

Réciproquement, par la Proposition 6, on suppose que $\forall u \in F(x)$, $\alpha|u| - 1 < h(u) \leqslant \alpha|u| + 1$ (l'autre cas est symétrique). Posons $\alpha = p/q$ avec p et q premiers entre eux. Par l'inégalité supposée, tout facteur u de x de longueur q a une hauteur p ou $p+1$. Il n'y a qu'un nombre fini d'occurences de tels facteurs de hauteur $p+1$, car sinon il existerait un facteur $w = uzv$ de x avec $|u| = |v| = q$ et $h(u) = h(v) = p + 1$, et donc, par l'inégalité supposée,

$$2 + 2p + h(z) = h(uzv) \leqslant 1 + \alpha q + \alpha|z| + \alpha q = 1 + 2p + \alpha|z|$$

donc $h(z) \leqslant \alpha|z| - 1$, ce qui contredit l'inégalité supposée.

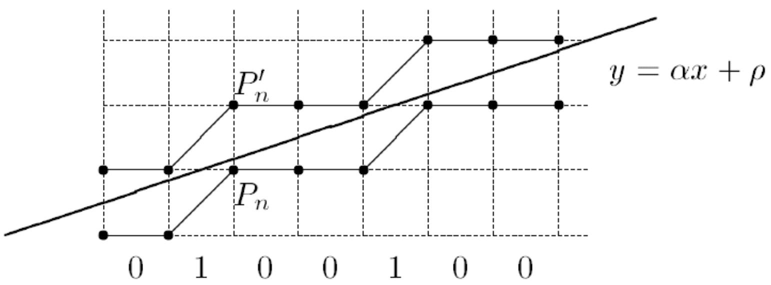
Ainsi, il exise une factorisation $x = ty$ telle que chaque mot de $F_q(y)$ a la même hauteur p. Considérons maintenant une occurence azb d'un facteur de y de longueur $q+1$, avec a et b des lettres. Comme $h(az) = h(zb)$, on a $a = b$. Donc y est périodique de période q. Par conséquent, x est éventuellement périodique. □

4 Mots mécaniques

Définition 7. Soient α et ρ deux réels, avec $0 \leqslant \alpha \leqslant 1$.

Le *mot mécanique inférieur de pente α et d'ordonnée à l'origine ρ* est le mot infini $s_{\alpha,\rho}$ défini par $s_{\alpha,\rho}(n) = \lfloor \alpha(n+1) + \rho \rfloor - \lfloor \alpha n + \rho \rfloor$.

Le *mot mécanique supérieur de pente α et d'ordonnée à l'origine ρ* est le mot infini $s'_{\alpha,\rho}$ défini par $s'_{\alpha,\rho}(n) = \lceil \alpha(n+1) + \rho \rceil - \lceil \alpha n + \rho \rceil$.



Intuitivement, les mots mécaniques correspondent à la discrétisation d'une droite de pente α et d'ordonnée à l'origine ρ.

Remarquons que pour tous réels x et y,

$$y - x - 1 < \lfloor y \rfloor - \lfloor x \rfloor < y - x + 1$$

Cette inégalité sera utile par la suite.

Lemme 1. *Soit s mot mécanique de pente α. Alors s est équilibré de pente α.*

De plus, si α est rationnel, alors s est purement périodique, et si α est irrationnel, alors s est apériodique.

Démonstration. Soit $s = s_{\alpha,\rho}$ un mot mécanique inférieur (la preuve est similaire si on prend x mécanique supérieur). Si $u = s(n)..s(n+p-1)$ est un facteur, alors $h(u) = \lfloor \alpha(n+p)+\rho \rfloor - \lfloor \alpha n + \rho \rfloor$, donc

$$\alpha|u| - 1 < h(u) < \alpha|u| + 1$$

Ainsi, $h(u)$ ne prend que les valeurs $\lfloor \alpha|u| \rfloor$ et $\lfloor \alpha|u| \rfloor + 1$ quand u parcourt les facteurs de s d'une longueur fixée. s est donc équilibré. De plus, par la Proposition 6,

$$|\pi(u) - \alpha| \leqslant \frac{1}{|u|}$$

donc $\pi(u) \longrightarrow \alpha$ quand $|u| \longrightarrow \infty$ et α est la pente de s en tant que mot équilibré, ce qui montre la première assertion.

Si α est irrationnel, s est apériodique selon le Théorème 2. Si $\alpha = p/q$ est rationnel, alors $\forall n \geqslant 0, \lfloor \alpha(n+q)+\rho \rfloor = p + \lfloor \alpha n + \rho \rfloor$.
Donc $\forall n \geqslant 0,\ s(n+q) = s(n)$, donc s est purement périodique.

□

Lemme 2. *Soit s un mot infini équilibré. Si s est apériodique, alors s est mécanique irrationnel. Si s est purement périodique, alors s est mécanique rationnel.*

Démonstration. Par le Corollaire 1, s a une pente, notons-la α. Notons h_n la hauteur du préfixe de longueur n de s.

Pour tout réel τ, on a $\forall n \geqslant 0,\ h_n \leqslant \lfloor \alpha n + \tau \rfloor$ ou $\forall n \geqslant 0,\ h_n \geqslant \lfloor \alpha n + \tau \rfloor$, car sinon il existerait τ, n, k tels que $h_n < \lfloor \alpha n + \tau \rfloor$ et $h_{n+k} > \lfloor \alpha(n+k)+\tau \rfloor$ (ou la relation symétrique). Mais alors $h_{n+k} - h_n \geqslant 2 + \lfloor \alpha(n+k)+\tau \rfloor - \lfloor \alpha n + \tau \rfloor > 1 + \alpha k$, ce qui contredit la Proposition 6. Posons

$$\rho = inf\{\tau | \forall n \geqslant 0,\ h_n \leqslant \lfloor \alpha n + \tau \rfloor\}$$

Selon la Proposition 6, on a $\rho \leqslant 1$, et $\rho < 1$ si α est irrationnel. On a donc

$$\forall n \geqslant 0,\ h_n \leqslant \alpha n + \rho \leqslant h_n + 1$$

car sinon il existerait n tel que $h_n + 1 < \alpha n + \rho$, et donc $\sigma = h_n + 1 - \alpha n < \rho$ et $\alpha n + \sigma = h_n + 1 > h_n$, ce qui contredit la définition de ρ.

Si s est apériodique, alors α est irrationnel par le Théorème 2, et donc $\alpha n + \rho$ est entier pour au plus un n. Ainsi, on a soit $h_n = \lfloor \alpha n + \rho \rfloor$ pour tout n, et alors $s = s_{\alpha,\rho}$, soit $h_n = \lfloor \alpha n + \rho \rfloor$ pour tout n sauf un n_0, et alors $h_{n_0} + 1 = \alpha n_0 + \rho$, donc $h_n = \lceil \alpha n + \rho - 1 \rceil$ pour tout n et $s = s'_{\alpha,\rho-1}$.

Si $s = u^\omega$ est purement périodique, alors $\alpha = p/q$ avec $p = h(u) = h_q$. De nouveau, $h_n = \lfloor \alpha n + \rho \rfloor$ pour tout n si $\alpha n + \rho$ n'est jamais entier (ce qui dépend de ρ), et alors $s = s_{\alpha,\rho}$.

Mais si $h_n = \alpha n + \rho$ pour un certain n, alors $h_n = \lfloor \alpha n + \rho \rfloor$ pour tout n, et $s = s_{\alpha,\rho}$. Montrons cela par l'absurde : supposons $\alpha m + \rho = h_m + 1$ pour un m, qu'on peut supposer tel que $n < m < n + q$. Considérons alors

$y = s(n+1)...s(m)$ et $z = s(m+1)...s(n+q)$. Alors $\pi(y) = \dfrac{h_m - h_n}{m-n} = \alpha - \dfrac{1}{|y|}$ et $\pi(z) = \dfrac{h_{n+p} - h_m}{n+p-m} = \alpha + \dfrac{1}{|z|}$, d'où $|\pi(y) - \pi(z)| = \dfrac{1}{y} + \dfrac{1}{z}$, ce qui contredit la Proposition 5. De même, si $1+h_n = \alpha n+\rho$ pour un certain n, alors $h_n = \lceil \alpha n+\rho \rceil$ pour tout n, et $s = s'_{\alpha,\rho}$.

□

Remarquons que si s est équilibré éventuellement périodique, alors il n'est pas forcément mécanique : par exemple, 01^ω est de pente 1, donc n'est pas mécanique, car pour tout ρ, $s_{1,\rho} = s'_{1,\rho} = 1^\omega$. Les mots mécaniques sont soit totalement périodiques, soit pas du tout.

Nous pouvons maintenant démontrer le théorème cherché : l'équivalence entre les propriétés (1), (2) et (3).

Théorème 3. *Soit s un mot infini. Les conditions suivantes sont équivalentes :*
(1) *s est de Sturm.*
(2) *s est équilibré apériodique.*
(3) *s est mécanique irrationnel.*

Démonstration. On sait déjà par le Théorème 1 que (1) et (2) sont équivalentes. D'autre part, (2) $\Longrightarrow$ (3) par le Lemme 1 et (3) $\Longrightarrow$ (2) par le Lemme 2. □

Références

[1] J. Bestel and D. Perrin. Finite and infinite words. In *Algebraic Combinatorics on Words*, chapter 1, pages 0–39. Cambridge University Press, 2002.

[2] J. Bestel and P. Séébold. Sturmian words. In *Algebraic Combinatorics on Words*, chapter 2, pages 40–97. Cambridge University Press, 2002.