Ann. sc. math. Québec, 1986, Vol. 10, No 2, pp. 237-245

MOTS DE LYNDON ET UN THÉORÈME DE SHIRSHOV

Christophe Reutenauer

1. Introduction

Les mots de Lyndon ont été introduits dans [5] et [1] pour la construction des quotients de la série centrale descendante des groupes libres et le calcul différentiel libre. Par définition, un mot sur un alphabet totalement ordonné est de Lyndon s'il est minimum (pour l'ordre lexicographique) parmi tous ses conjugués (voir au paragraphe 2). Du point de vue de la combinatoire des mots, les mots de Lyndon sont d'une grande richesse, et constituent une factorisation du monoïde libre (voir chapitre 5 de [4]): tout mot s'écrit d'une et une seule manière comme un produit décroissant de mots de Lyndon. Par ailleur, l'étude de ces mots a mené à l'élaboration d'algorithmes très fins, dans le cadre de la recherche des facteurs d'un mot [3].

Un théorème de Shirshov [8], présenté dans le chapitre 7 de [4], affirme que tout mot assez long possède un facteur qui est soit une puissance p-ième, soit "n-divisé" (voir au paragraphe 3); ce théorème a servi à Shirshov à résoudre par l'affirmative le problème de Kurosh pour les algèbres à identité polynomiale (voir [8], [4] chapitre 7, [7] chapitre 4). Une autre application permet de donner le cadre d'une réponse positive au problème de Burnside pour les semi-groupes [6]. Ce théorème s'insère dans la problématique des "régularités inévitables" des mots comme étudiée dans [4]; par exemple, d'après un théorème de Van der Waerden, tout mot sur un alphabet fixé comporte des répétitions régulièrement espacées de la même lettre, autant de fois qu'on le veut, pourvu que ce mot soit assez long (voir [4] chapitre 3).

Nous démontrons ici une propriété de régularité inévitable liée aux mots de Lyndon; appelons *indice* de Lyndon d'un mot le nombre de mots de son unique factorisation décroissante en mots de Lyndon. Les mots de Lyndon sont donc ceux d'indice 1; remarquons qu'il existe de tels mots de longueur arbitraire. La propriété peut alors s'énoncer ainsi: *tout mot assez long contient en facteur un mot d'indice aussi grand qu'on le veut* (Théorème 1).

Ce résultat, outre son intérêt intrinsèque, nous permet une démonstration élégante du théorème de Shirshov cité plus haut: le facteur d'indice élevé contient soit une puissance p-ième d'un mot de Lyndon, soit assez de mots de Lyndon strictement décroissants, ce qui induit une n-division (voir le paragraphe 3).

2. Mots de Lyndon et régularités inévitables

A étant un alphabet totalement ordonné, on munit le monoïde libre A^* engendré par A de l'*ordre lexicographique* défini par: $u < v$ si et seulement si

(i) soit il existe un mot w non vide tel que $uw = v$;

(ii) soit il existe des mots x, y, z et des *lettres* (= éléments de A) a, b tels que $u = xay$, $v = xbz$ et $a < b$.

Autrement dit, $<$ est l'ordre du dictionnaire: par exemple, si $a < b$, on a: $aab < aaba < aba$.

Un mot u est *facteur* (resp. *facteur gauche*) d'un mot v s'il existe des mots x et y tels que $v = xuy$ (resp. $v = uy$). On notera que pour tous mots u, v, x, y, z:

(1) $\quad u < v \Rightarrow xu < xv$;

(2) $\quad u < v$, u pas facteur gauche de $v \Rightarrow uy < vz$.

Pour un mot non vide $w \in A^*$, les deux conditions suivantes sont équivalentes:

(3) pour toute factorisation $w = uv$, avec $u, v \neq 1$ (= le *mot vide*), on a $w < vu$.

(4) pour toute factorisation $w = uv$, avec $u,v \neq 1$, on a $w < v$ (voir [4] pp. 64-65).

Dans ce cas, w est appelé un *mot de Lyndon*. Par exemple, a, ab, aabb, aabab sont des mots de Lyndon.

Une propriété de base des mots de Lyndon est le

THÉORÈME (Chen, Fox, Lyndon [1], voir [4] th. 5.1.5). *Tout mot* w *s'écrit de manière unique*

$$w = \ell_1 \ell_2 \ldots \ell_n \quad (n \geq 0) \tag{5}$$

où les ℓ_i *sont des mots de Lyndon avec* $\ell_1 \geq \ell_2 \geq \ldots \geq \ell_n$.

Nous appellerons *indice* de w l'entier n de la factorisation (5). Remarquons que les mots de Lyndon sont les mots d'indice 1, et qu'il en existe en toute longueur: si $a < b$, alors $a^i b^j$ est un mot de Lyndon $(i,j \neq 0)$.

Nous démontrerons le

THÉORÈME 1. *Il existe une fonction* $N(k,q)$ *telle que pour tout alphabet totalement ordonné* A *de cardinalité* k *et tout mot* w *dans* A* *de longueur au moins* $N(k,q)$, w *se factorise en*

$$w = u x_1 \ldots x_q v$$

où chaque x_i *est un mot de Lyndon avec*

$$x_1 \geq x_2 \geq \ldots \geq x_q.$$

Autrement dit, w contient en facteur un mot d'indice q.

Dans la suite de ce paragraphe, nous posons, pour tout alphabet fini totalement ordonné A:

$$a = \inf(A)$$

i.c. a est le plus petit élément de A. Nous notons B le sous-ensemble de A* défini par

$$B = a(A \backslash a)^*$$

i.e. B est l'ensemble des mots dont la première lettre est a et dont les lettres suivantes sont distinctes de a. On notera que tout mot dans B est un mot de Lyndon. Le sous-monoïde B* de A* engendré par B est un sous-monoïde libre, librement engendré par B (autrement dit, B est un *code*, [4] chapitre 1).

L'ensemble B est totalement ordonné par l'ordre lexicographique de A*, et on peut donc considérer B comme un alphabet totalement ordonné. Nous noterons $\prec$ l'ordre lexicographique sur B* défini par cet ordre total sur B. On peut donc parler de mots de Lyndon dans B* (pour l'ordre $\prec$).

Les deux lemmes techniques suivants montrent que l'ordre $\prec$ sur B* coïncide avec l'ordre induit par $<$ sur B* et que les mots de Lyndon sur B* (pour $\prec$) sont les mêmes si on les considère comme mots de A* avec l'ordre $<$.

LEMME 1. *Soient* u, v *dans* B* *tels que* $u \prec v$. *Alors* $u < v$.

PREUVE. Si u est facteur gauche de v dans B*, alors il l'est aussi dans A*, et l'on a $u < v$. Dans le cas contraire, on a par définition de l'ordre lexicographique $\prec$ sur B*: $u = u_1\alpha u_2$, $v = u_1\beta v_2$ avec $\alpha,\beta \in B$, $u_1,u_2,v_2 \in B^*$, et $\alpha \prec \beta$. Donc $\alpha < \beta$. Si α n'est pas facteur gauche de β dans A*, alors $\alpha = xby$, $\beta = xcz$ avec $b,c \in A$ et $b < c$; par suite, $u = u_1xbyu_2$, $v = u_1xczv_2$, d'où l'on déduit que $u < v$. Dans le cas contraire (qui est le cas subtil), on a $\beta = \alpha\beta'$, avec $\beta' \neq 1$. Comme $\beta \in B = a(A\setminus a)^*$, la première lettre de β' est $b \neq a$, donc $a < b$. Autrement dit, $\beta' = b\beta''$ et par suite: $u = u_1\alpha u_2$, $v = u_1\alpha b\beta''v_2$. De plus, $u_2 \in B^*$ et u_2 n'est pas vide (sinon u est facteur gauche de v dans B*), donc u_2 commence par un a: $u_2 = au_2' \Rightarrow u = u_1\alpha au_2'$, d'où l'on conclut que $u < v$. □

LEMME 2. *Soit* x *dans* B* *un mot de Lyndon sur* B* *pour l'ordre* $\prec$. *Alors* x *est un mot de Lyndon sur* A*.

PREUVE. On a $x = x_1 \ldots x_n$ pour des mots x_i dans $B = a(A\setminus a)^*$. Soient u, v des mots non vide dans A* tels que $x = uv$. Nous montrons que $x < vu$ (donc x est un mot de Lyndon).

Il existe un entier i dans $\{1,\dots,n\}$ tel qu'on ait un des deux cas suivants:

(i) soit $u = x_1\dots x_{i-1}$, $v = x_i\dots x_n$ et $i \neq 1$. Dans ce cas, u et v sont dans B^*, et comme x est un mot de Lyndon sur B^*, on obtient $x \prec vu$. Par le Lemme 1, on a alors $x < vu$.

(ii) soit $x_i = st$ $(s,t \neq 1)$ et $u = x_1\dots x_{i-1}s$, $v = tx_{i+1}\dots x_n$. Alors, d'après (i), on a $x \leq x_i\dots x_n x_1\dots x_{i-1}$. De plus, x_i est dans $a(A\backslash a)^*$, donc x_i commence par a, et t commence par une lettre $\neq a$, donc $> a$. Par suite $x_i < t$. Ceci implique, par (2) (puisque t est plus court que x_i), que $x_i\dots x_n x_1\dots x_{i-1} < tx_{i+1}\dots x_n x_1\dots x_{i-1}s = vu$. D'où, par transitivité, $x < vu$. □

PREUVE du Théorème 1. On peut prendre $N(1,q) = q$ et $N(k,1) = 1$. Donc $N(1,q)$ existe, et $N(k,1)$ existe pour tous entiers k et q. L'hypothèse de récurrence sera la suivante: $N(k-1,q)$ *existe et* $N(k',q-1)$ *existe pour tout entier* k'.

Nous montrons qu'alors on peut prendre $N(k,q) = N$ avec

$$N = N(k-1,q)(1 + N(k^{N(k-1,q)} + k^{N(k-1,q)-1} + \dots + k, q-1)).$$

Ce nombre existe bien sûr, d'après l'hypothèse de récurrence.

Soit A un alphabet totalement ordonné de cardinalité k, et w un mot sur A^* de longueur $\geq N$. On a

$$w = ux_1\dots x_n$$

avec $u \in (A\backslash a)^*$ et $x_i \in a(A\backslash a)^* = B$.

Si la longueur $|u|$ de u est $\geq N(k-1,q)$ ou si un des x_i vérifie $|x_i| > N(k-1,q)$, on peut conclure par l'hypothèse de récurrence, puisqu'on trouve alors dans w un facteur de longueur $\geq N(k-1,q)$ écrit sur l'alphabet $A\backslash a$ de cardinalité $k-1$. On peut donc supposer que $|u| < N(k-1,q)$ et pour tout i, $|x_i| \leq N(k-1,q)$. Par suite

$$|x_1 \ldots x_n| > N(k-1,q)N(k^{N(k-1,q)} + \ldots + k, q-1)$$

et on en déduit que

$$n > N(k^{N(k-1,q)} + \ldots + k, q-1). \tag{3}$$

Soit B_1 l'ensemble des mots de B de longueur $\leq N(k-1,q)$. Alors la cardinalité de B_1 est $\leq k^{N(k-1,q)} + \ldots + k$ et chaque x_i est dans B_1. Alors $x_1 \ldots x_{n-1}$ peut être considéré comme un mot sur l'alphabet B_1, et par (3) et l'hypothèse de récurrence, $x_1 \ldots x_{n-1}$ se factorise dans B^* comme $uy_1 \ldots y_{q-1}v$ où les y_i sont des mots de Lyndon décroissants sur B^*, donc sur A^* (Lemme 2). Mais y_{q-1} est suivi dans w par le mot non vide $vx_n \in B^*$. Ce mot commence donc par un a, donc w contient en facteur le mot $y_1 \ldots y_{q-1}a$, ce qui conclut la preuve, puisque a est un mot de Lyndon et $y_{q-1} \geq a$. □

3. Le théorème de Shirshov

Soit A un alphabet totalement ordonné et A^* le monoïde libre sur A ordonné lexicographiquement. Un mot w dans A^* est dit *n-divisé* s'il admet une factorisation dans A^*

$$w = x_1 \ldots x_n$$

telle que

$$\forall \sigma \in S_n \backslash \mathrm{id}, \quad w > x_{\sigma(1)} x_{\sigma(2)} \ldots x_{\sigma(n)}.$$

EXEMPLE. Le mot $w = baba$ admet la 3-division $w = b \cdot ab \cdot a$ comme on le vérifie aisément.

Une *puissance* p-ième est un mot de la forme $x^p = xx \ldots x$ (p fois) pour un mot x non vide.

THÉORÈME 2 (Shirshov [8], [7] p. 205, [4] p. 145). *Il existe une fonction* $M(k,p,n)$ *telle que pour tout alphabet totalement ordonné* A *de cardinal* k, *tout mot sur* A^* *de longueur au moins* $M(k,p,n)$ *possède en facteur soit un mot* n-*divisé, soit une puissance* p-*ième.*

Nous allons déduire ce théorème du Théorème 1. Auparavant, nous énonçons un lemme, suggéré par D. Perrin pour simplifier la démonstration.

LEMME 3. *Soient* x, y *des mots de Lyndon tels que* $x > y$ *et* i, j *des entiers* ≥ 1. *Alors* $y^j x^i$ *est un mot de Lyndon.*

PREUVE. 1) Nous utilisons le résultat suivant [4] (prop. 5.1.3): si y, x sont des mots de Lyndon tels que $y < x$ alors yx est un mot de Lyndon et

$$y < yx < x.$$

2) Nous montrons que yx^i est un mot de Lyndon et que $yx^i < x$. C'est clair pour $i = 1$ d'après 1). Le passage de i à i+1 se déduisant immédiatement de 1), l'assertion est prouvée.

3) Nous montrons que $y^j x^i$ est un mot de Lyndon, par récurrence sur j (i fixé). Le cas $j = 1$ se déduit de 2), et la récurrence est immédiate d'après 1), puisque $y < y^j x^i$ dans l'ordre lexicographique. □

PREUVE du Théorème 2. Soit $M(k,p,n) = N(k,pn)$, où N est la fonction du Théorème 1. Si w est un mot de longueur au moins M(k,p,n) sur A ($|A| = k$), alors w possède un facteur de la forme

$$x_1 x_2 \ldots x_{pn}$$

où les x_i sont des mots de Lyndon avec $x_1 \geq x_2 \geq \ldots \geq x_{pn}$. Rassemblant les x_i qui se répètent, on obtient donc que w possède un facteur de la forme

$$y_1^{j_1} y_2^{j_2} \ldots y_r^{j_r} \qquad (j_i \geq 1, \ \sum j_i = pn)$$

où les y_i sont des mots de Lyndon avec $y_1 > y_2 > \ldots > y_r$. S'il y a un $j_i \geq p$, on a trouvé la puissance p-ième y_i^p comme facteur de p. Sinon, $j_i < p$ pour tout p, et $\sum j_i = pn$ implique $r \geq n$. Il suffit donc de montrer que

$$(y_1^{j_1})(y_2^{j_2}) \ldots (y_n^{j_n})$$

est une n-division. Mais si $k < \ell$, on a $y_k > y_\ell$, donc par le Lemme 3, $y_\ell^{j_\ell} y_k^{j_k}$ est un mot de Lyndon, donc

$$y_\ell^{j_\ell} y_k^{j_k} < y_k^{j_k} y_\ell^{j_\ell}.$$

Il suffit de raisonner par récurrence sur le nombre d'inversions de la permutation σ pour en déduire que

$$y_1^{j_1} \cdots y_n^{j_n} > y_{\sigma(1)}^{j_{\sigma(1)}} \cdots y_{\sigma(n)}^{j_{\sigma(n)}}.$$

Remerciements

Le présent travail est issu d'une discussion entre David Haussler et l'auteur. Il suggérait d'étendre le théorème de Shirshov à des préordres plus généraux que l'ordre lexicographique, dans le cadre des fonctions de moyenne sur le monoïde libre, étudiées dans [2].

Références

[1] CHEN, K.T., FOX, R.H., LYNDON, R.C., *Free differential calculus* IV. *The quotient groups of the lower central series*, Ann. Math. 58(1958), 81-95.

[2] EHRENFEUCHT, A., HAEMER, J., HAUSSLER, D., *Quasi-monotonic sequences: Theory, Algorithms and Applications*, Tech. Report, Univ. Colorado (Boulder), Computer Sci. Dept. (1984).

[3] DUVAL, J.-P., *Factorising words over an ordered alphabet*, J. Algorithms 4(1983), 363-381.

[4] LOTHAIRE, M., *Combinatorics on Words*, Addison Wesley (1983).

[5] LYNDON, R.C., *On the Burnside problem* 1, Trans. Amer. Math. Soc. 77(1954), 202-215.

[6] RESTIVO, A., REUTENAUER, C., *On the Burnside problem for semigroups*, J. Algebra 89(1984), 102-104.

[7] ROWEN, L.H., *Polynomial Identities in Ring Theory*, Acad. Press (1980).

[8] SHIRSHOV, A.I., *On rings with identity relations*, Math. Sb. 43(1957), 277-283 (en russe).

Département de mathématiques
et d'informatique
Université du Québec à Montréal
C.P. 8888, Succ. "A"
Montréal, Québec
Canada H3C 3P8

et CNRS (Paris)