

RAIRO

INFORMATIQUE THÉORIQUE

JEAN-PIERRE DUVAL

Mots de Lyndon et périodicité

RAIRO – Informatique théorique, tome 14, n° 2 (1980), p. 181-191.

http://www.numdam.org/item?id=ITA_1980__14_2_181_0

© AFCET, 1980, tous droits réservés.

L'accès aux archives de la revue « RAIRO – Informatique théorique » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/legal.php>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

MOTS DE LYNDON ET PÉRIODICITÉ (*)

par Jean-Pierre DUVAL (1)

Communiqué par J.-F. PERROT

Résumé. — *Nous montrons dans cet article comment, en ordonnant le monoïde libre X^* engendré par un alphabet X , nous pouvons établir une caractérisation de la périodicité des mots de ce monoïde.*

Abstract. — *We show in this note how, by ordering the free monoïd X^* , generated by an alphabet X , we can establish a characterisation of periodicity of the words on this alphabet.*

INTRODUCTION

Les problèmes de périodicité sont au centre de nombreuses questions sur les mots; que l'on s'intéresse, à la suite de [8], à l'étude des équations dans le monoïde libre, ou dans le groupe libre [9], ou bien aux problèmes algorithmes de « Word Matching » [7], on retrouve l'importance des résultats connus sur les périodes, tel le théorème de [6].

Plus récemment [11], posait la question d'une relation entre le nombre de couvertures disjointes d'un mot et sa période, et les résultats obtenus [3], à la suite de [1], permettent d'y apporter une réponse [5].

DÉFINITIONS : On désigne par X^* le monoïde libre engendré par un alphabet X , ses éléments sont appelés des mots, le mot vide est noté 1, et l'ensemble des mots non vides est noté X^+ . La longueur d'un mot f est notée $|f|$. Un mot g est un *facteur* d'un mot f ssi $f \in X^* g X^*$, c'est un *facteur gauche* si $f \in g X^*$, et un *facteur droit* si $f \in X^* g$.

On dit qu'un entier non nul λ est une *période* d'un mot $f \in X^+$ ssi chaque fois que deux lettres de f sont à distance λ elles sont identiques (c'est-à-dire ssi chaque fois que xgy est un facteur de f satisfaisant $x \in X, y \in X$, et $|xg| = \lambda$, alors $x = y$). Il n'est pas nécessaire, pour que $\lambda > 0$ soit une période de $f \in X^+$, que f soit une puissance exacte d'un mot de longueur λ , mais il suffit que f soit un facteur d'un mot de cette forme; ainsi, $u \in X^*, v \in X^+, k > 0$ le mot $f = (uv)^k u$ admet pour

(*) Reçu avril 1979, révisé novembre 1979.

(1) Laboratoire d'Informatique, Université de Rouen, Haute-Normandie, Mont-Saint-Aignan.

période $\lambda = |uv|$; plus avant, au sens de la définition tout entier supérieur ou égal à la longueur de f est une période de f . L'ensemble des périodes d'un mot f est donc non vide, et ce mot admet une *plus petite période* qui est inférieure ou égale à $|f|$.

Par exemple, pour $X = \{x, y\}$, la plus petite période du mot $xyxxy$ est 3, et celle du mot $xxxy$ est 4.

On appelle *point* d'un mot $f \in X^*$ tout couple (f', f'') de mots de X^* satisfaisant $f = f'f''$.

On dit qu'un mot non vide u est une *répétition* au point (f', f'') du mot $f = f'f'' \in X^+$ ssi :

$$uX^* \cap f''X^* \neq \emptyset \quad \text{et} \quad X^*u \cap X^*f' \neq \emptyset$$

[de façon plus intuitive si nous superposons les couples (f', f'') et (u, u) les lettres en coïncidence sont identiques].

Pour tout $g' \in X^*$, $g'' \in X^*$, et $u \in X^+$: le mot u est une répétition au point $(g'u, ug'')$; ce n'est cependant pas la seule forme possible et il n'est pas nécessaire que u^2 soit un facteur de f pour être une répétition en un point de f ; par exemple xyx est une répétition au point (y, xx) du mot yxx ; plus généralement en tout point (f', f'') , d'un mot non vide $f = f'f''$, le mot $u = f''f'$ est une répétition. Ainsi, en chaque point d'un mot, l'ensemble des répétitions est non vide et on appelle *répétition élémentaire* en un point (f', f'') l'unique répétition de plus courte longueur en ce point.

Notons que si $\lambda > 0$ est une période de $f \in X^+$, en tout point (f', f'') du mot f il existe au moins une répétition de longueur λ , et par suite, la plus petite période de f est supérieure ou égale à la longueur de la répétition élémentaire en chaque point (f', f'') de f . On dit que (f', f'') est un *point critique* de $f = f'f'' \in X^+$, ssi la longueur de la répétition élémentaire en ce point est exactement égale à la plus petite période de f .

Le résultat peut alors s'énoncer :

THÉORÈME DU POINT CRITIQUE : *Tout mot non vide admet au moins un point critique, c'est-à-dire, la plus petite période d'un mot non vide est égale au maximum des longueurs des plus courtes répétitions en chaque point du mot.*

Par exemple pour $X = \{x, y\}$ et $f = xyxxyxxyxxyxxy; f = (xyxxyxxyx)^1 xyxxy$, 8 est une période de f , c'est la plus petite période de f ; notons la longueur de la répétition élémentaire en chaque point de f :

x	y	x	x	y	x	y	x	x	y	x	x	y	
1	2	3	1	5	2	2	8	1	3	3	1	3	1

La répétition élémentaire au point $(xyxx, yxyxxxy)$ est $yxyxx$ de longueur 5, et la plus grande longueur de ces répétitions est 8, égale à la plus petite période de f et s'obtient au point $(xyxxxy, xxyxx)$; qui est le seul point critique du mot considéré.

Un premier résultat sur ce point a été établi par [1]; considérant un mot « suffisamment long » comme une application f de \mathbb{Z} dans X : on dit que $\lambda > 0$ est une période locale de f au point $i \in \mathbb{Z}$ ssi pour tout $j \in [i - \lambda, i]$, $f(j) = f(j + \lambda)$, c'est-à-dire que $f([i - \lambda, i + \lambda])$ est de la forme $xuxux$ avec $x \in X$ et $|ux| = \lambda$; leur résultat est alors le suivant : *soit g un facteur fini de f , si en chaque lettre de g il existe une période locale (de f), le maximum sur g des minimums en chaque lettre de g des périodes locales, est une période de g .*

Notons que l'existence d'une période locale λ , au sens de Cesari-Vincent, implique l'existence d'une répétition de longueur λ , mais la réciproque n'est pas vraie.

Exemple :

$$f = \dots yxx y_1 yxx_1 yxx y_1 yxx \dots,$$

$$g = yxx y_1 yxx_1 yxx y_1 yxx \quad \text{avec } y_1 = y \text{ et } x_1 = x.$$

La période locale de f au voisinage de y_1 n'est pas définie, sans connaissance du contexte de g dans f , elle serait au moins égale à 4 en supposant que la lettre de f précédent g est y . Il en est de même au point x_1 , et la période locale serait de 7 en supposant que la lettre de f précédent g est x , et serait au moins égale à 11 si nous supposons comme précédemment que cette lettre est y . Par contre les répétitions sont définies, yxx de longueur 4 est une répétition au point $(yxx, y_1 yxx y_1 yxx)$, y de longueur 1 est une répétition au point $(yxx y_1, yxx y_1 yxx)$, x de longueur 1 est une répétition au point $(yxx y_1 yx, x_1 yxx y_1 yxx)$, yxx de longueur 3 est une répétition au point $(yxx y_1 yxx_1, yxx y_1 yxx)$; aucun de ces points n'est un point critique de g dont la période élémentaire est 7; toutefois l'existence d'un point critique dans g qui est garantie par le théorème du point critique est confirmée par l'observation : $(yxx y_1 y, xxyxx y_1 yxx)$ est un point critique de g .

Nous avons donné en [3] une preuve complète du théorème du point critique qui utilise une récurrence sur la longueur des mots différente de celle de Cesari-Vincent. En [5] nous avons développé l'application au problème de M. P. Schützenberger. Le théorème du point critique peut également se démontrer en faisant appel à la théorie des mots de Lyndon, qui a pour origine l'étude des factorisations des mots du monoïde libre : c'est cette preuve, conceptuellement plus intéressante, qui fait l'objet du présent article.

Rappelons que, étant donné un alphabet totalement ordonné $(X, <)$, et le monoïde libre ordonné lexicographiquement $(X^*, <)$, on dit qu'un mot non

vide $x_1 \dots x_n$ de X^+ est un *mot de Lyndon* si et seulement si il est strictement inférieur à chacun de ses conjugués propres, c'est-à-dire, pour $i=2, \dots, n$, on a

$$x_1 \dots x_n < x_i \dots x_n x_1 \dots x_{i-1} \quad [10].$$

On sait, [2], que : *tout mot de X^+ se factorise de manière unique en une suite non croissante de mots de Lyndon, c'est-à-dire, que l'ensemble des mots de Lyndon forme une factorisation complète du monoïde libre* [12]. Nous avons montré, en outre, que cette factorisation s'obtient en temps linéaire en la longueur du mot [4].

La preuve s'obtient très facilement à partir des propriétés classiques des mots de Lyndon dans le cas particulier d'un mot f « suffisamment long » par rapport à sa période c'est-à-dire $|f| \geq 3\lambda$, où λ est la plus petite période de f . f est de la forme $f = (v)^k w$ avec $|v| = \lambda$, $k \geq 3$, $v \in w X^*$. λ étant la plus petite période de f , v est un mot primitif, il admet donc un conjugué $v'' v'$, avec $v = v' v''$, qui est un mot de Lyndon d'où $f = v' (v'' v')^{k-1} v'' w$, or le mot $v'' v'$ étant un mot de Lyndon il n'admet pas de facteur gauche propre qui soit facteur droit et par conséquent la répétition élémentaire au point $(v' v'' v', (v'' v')^{k-2} v'' w)$ est exactement $v'' v'$, avec $|v'' v'| = |v| = \lambda$, le point en question est ainsi un point critique de f .

Dans le cas général nous obtiendrons le résultat annoncé comme conséquence d'un énoncé plus précis, à savoir :

DÉFINITION : On appelle *facteur résiduel* d'un mot $f \in X^+$ le plus long facteur droit de f qui est de la forme $(uv)^k u$, satisfaisant $u \in X^*$, $v \in X^+$, $k > 0$ et uv est un mot de Lyndon.

THÉORÈME : Soient $f \in X^+$, r le facteur résiduel de f , et, $g \in X^*$ le facteur gauche « complémentaire » tel que $f = gr$:

- (a) si g n'est pas facteur de r , (g, r) est un point critique de f ;
- (b) si g est facteur de r , (f', f'') est un point critique de f , où f'' est le facteur droit maximal de r et f' satisfait $f = f' f''$.

Après quelques propriétés techniques des mots de Lyndon et des factorisations en mots de Lyndon, nous étudierons successivement le cas où f est un mot de Lyndon (proposition 5) et le cas général où f est un produit de mots de Lyndon.

1. RAPPELS

Rappelons que l'on considère un alphabet X et X^* le monoïde libre engendré par X . On dit qu'un facteur g d'un mot $f \in X^+$ est un *facteur propre* ssi $g \neq f$ et

$g \neq 1$; on dit que c'est un *facteur strict* ssi $g \neq f$. On dit que deux mots f et g sont *conjugués* s'ils sont de la forme $f = h' h''$ et $g = h'' h'$, si de plus $h' \neq 1$ et $h'' \neq 1$ on dit qu'ils sont *conjugués propres*. Un mot qui n'est puissance stricte d'aucun autre mot de X^* est dit *primitif*, il est alors différent de chacun de ses conjugués propres [8], de plus chaque conjugué est lui-même primitif.

Nous supposons l'alphabet X totalement ordonné d'une manière quelconque par $<$; on définit alors l'ordre lexicographique sur X^* par $f \leq g$ si et seulement si l'une des deux propositions suivantes est satisfaite :

(1.1) $g \in f X^*$;

(1.2) f est de la forme hxh' , g de la forme hyh'' avec $x, y \in X$ et $x < y$.

Il en résulte immédiatement :

PROPOSITION 1 : Soient f, g, h trois éléments de X^* :

(1) si $f \leq g \leq h$: le plus long facteur gauche commun à f et h est un facteur gauche de g ;

(2) si f n'est pas facteur gauche de g et $f \leq g$: pour tout $v \in X^*$, on a $fv < g$.

2. PROPRIÉTÉS DES MOTS DE LYNDON

On dit qu'un mot f de X^+ est un *mot de Lyndon* [10] ou *mot lexicographique standard* [12] si, et seulement si, il satisfait l'une des trois propositions suivantes qui sont équivalentes :

(2.1) f est un mot primitif et il est le plus petit de sa classe de conjugués;

(2.2) f est strictement inférieur à chacun de ses conjugués propres;

(2.3) f est strictement inférieur à chacun de ses facteurs droits propres.

Rappelons brièvement la preuve de l'équivalence.

Un mot primitif est différent de chacun de ses conjugués propres; il est immédiat que (2.1) et (2.2) sont équivalentes.

Soient f' et f'' non vides tels que $f = f' f''$; $f'' < f'' f'$; il est clair que (2.3) implique (2.2).

Soient f' et f'' non vides tels que $f = f' f''$, et supposons : $f'' \leq f$ et $f < f'' f'$. Il en résulte, d'après la proposition 1, que f'' est un facteur gauche de f ; posons $f = f'' h$; $f = f'' h < f'' f'$, d'où $h < f'$ et $hf'' < f' f'' = f$, ce qui contredit (2.2); et par suite (2.2) implique (2.3).

L'équivalence des propositions (2.1), (2.2), et (2.3) est établie.

REMARQUE 1 : Il s'en déduit que tout mot primitif admet un conjugué qui est un mot de Lyndon, puisque tout conjugué d'un mot primitif est un mot primitif.

PROPOSITION 2 : Soit f un mot de Lyndon :

- (1) aucun facteur gauche propre de f n'est facteur droit de f ;
- (2) si h est un facteur de f : ou bien h est un facteur gauche de f , ou bien $f < h$.

La preuve en est immédiate. \square

LEMME 1 : Le produit fg , de deux mots de Lyndon f et g , est un mot de Lyndon si et seulement si $f < g$.

Preuve : Si f et g sont deux mots de Lyndon ainsi que $fg : f < fg < g$. Réciproquement, si $f < g$, g n'est pas facteur gauche de f , et ne saurait être facteur gauche fg sans que g admette un facteur propre gauche et droit, ce qui contredirait la proposition 2 sur g , et par suite $fg < g$; soit h un facteur droit propre de fg : ou bien h est un facteur droit de g et $fg < g \leq h$, ou bien h est de la forme $h'g$ et h' est un facteur droit propre de f d'où $f < h'$ et d'après la proposition 1 $fg < h'g$; il en résulte que (2.3) est satisfaite : fg est un mot de Lyndon. \square

REMARQUE 2 : Plus généralement si f et g sont deux mots de Lyndon satisfaisant $f < g$; pour tout entier k :

- (a) $f^k g$ est un mot de Lyndon; de plus si u est un facteur gauche non vide de f^k ;
- (b) $u < g$; et, si u est facteur de g : u est un facteur gauche de g .

LEMME 2 : Soit f un mot de Lyndon et s un facteur gauche de f , distinct de f ; soit $x \in X$:

- (a) si $f < sx$: pour tout entier k , $f^k sx$ est un mot de Lyndon;
- (b) si sx n'est pas facteur gauche de f et $sx < f$: pour tout entier non nul k , et pour tout $h \in X^*$, le plus long facteur gauche de $f^k sxh$ qui soit un mot de Lyndon est f .

Preuve : Pour établir (a), il suffit de monter, d'après le lemme 1, que sx est un mot de Lyndon. Posons $f = sv$ et considérons un facteur droit propre hx de sx ; hv est un facteur droit propre de sv , d'où, $sv < hv < hx$, et par suite $s < hx$, or $|s| \geq |hx|$ d'où $sx < hx$. Ce qui permet de conclure que sx satisfait (2.3), c'est donc un mot de Lyndon.

Établissons (b) : Sachant que sx n'est pas facteur gauche de f et $sx < f$, pour tout $h : sxh < f^k sxh$; le plus long facteur gauche de $f^k sxh$ qui soit un mot de Lyndon est un facteur gauche de $f^k s$, et d'après (2.3) ce ne peut-être que f . \square

Du lemme 2, il se déduit immédiatement la proposition suivante :

PROPOSITION 3 : Soit $u \in X^+$; si u est facteur gauche d'un mot de Lyndon, u s'écrit de manière unique sous la forme $u = t^k s$ où t est un mot de Lyndon, k un entier non nul, et s un facteur gauche distinct de t .

Preuve : Il suffit de poser t , le plus long facteur gauche de u qui soit un mot de Lyndon, et de choisir k et s tels que $t^k s$ soit le plus long facteur gauche de u de cette forme, il résulte du lemme 2 que, sachant que u est facteur gauche d'un mot de Lyndon et que t est le plus long facteur gauche de u qui est un mot de Lyndon, nécessairement $u = t^k s$. \square

REMARQUE 3 : Soit f et g deux mots de Lyndon tels que $g < f$; soit k un entier non nul et s un facteur gauche strict de f : $gf^k s$ est de la forme $f' s'$ où f' est un mot de Lyndon et s' un facteur gauche strict de f' .

Preuve : Il résulte du lemme 1 que gf^{k+1} est un mot de Lyndon, et d'après la proposition 3 nous pouvons conclure. \square

3. PROPRIÉTÉS DES FACTORISATIONS EN MOTS DE LYNDON

Rappelons le :

THÉORÈME DE CHEN FOX ET LYNDON : *Tout mot $f \in X^+$ se factorise de manière unique en une suite non croissante de mots de Lyndon, c'est-à-dire, qu'il existe une suite unique, $f_1 \geq f_2 \geq \dots \geq f_n$, de mots de Lyndon telle que $f = f_1 f_2 \dots f_n$.*

Nous en donnons une démonstration pour la commodité du lecteur.

Preuve : La proposition est vérifiée pour tout mot de longueur 1, supposons la vraie pour tout mot de longueur au plus p et montrons qu'elle est vraie pour tout mot de longueur $p+1$.

Soit $f \in X^+$ tel que $|f| = p+1$; soit h le plus long facteur droit de f qui soit un mot de Lyndon et soit g le facteur gauche complémentaire de h , tel que $f = gh$. D'après l'hypothèse de récurrence g se factorise de manière unique en une suite non croissante de mots de Lyndon : $g = g_1 \dots g_n$; $g_n h$ n'est pas un mot de Lyndon et d'après le lemme 1 il en résulte $g_n \geq h$. Ce qui établit l'existence d'une factorisation de f : $f = g_1 \dots g_n h$.

Supposons l'existence d'une autre factorisation de f , $f = f_1 \dots f_m$, en mots de Lyndon. Par choix de h : $|h| \geq |f_m|$; si nous supposons $|h| > |f_m|$ alors h s'écrit $h = f'_j f_{j+1} \dots f_m$ où f'_j est un facteur droit non vide de f_j , il en résulte, puisque h est un mot de Lyndon, que $f'_j < f_m$, et, puisque f_j est un mot de Lyndon, que $f_j \leq f'_j$, d'où $f_j < f_m$, ce qui est contradictoire; et par suite $|h| = |f_m|$ et $h = f_m$. Ce qui établit l'unicité de la factorisation.

La proposition est donc vraie au rang $p+1$. Ce qui achève la preuve. \square

NOTATION : La factorisation d'un mot f en mots de Lyndon est un élément de $(X^+)^*$ que nous noterons : $F(f) = (f_1) \dots (f_n)$.

REMARQUE 4 : Si $F(f) = (f_1) \dots (f_n)$ alors f_1 est le plus long facteur gauche de f qui est un mot de Lyndon, f_n est le plus long facteur droit de f qui est un mot de Lyndon, et c'est le facteur droit de f minimal dans l'ordre lexicographique.

Étant donné un mot f , soit $t^k s$ le plus long facteur droit de f qui satisfasse : $k \geq 1$, t est un mot de Lyndon et s un facteur gauche strict de t ; on dit que $r = t^k s$ est le *facteur résiduel* de f ; posons

$$f = gt^k s \quad \text{et} \quad F(g) = (g_1) \dots (g_p), \quad F(s) = (s_1) \dots (s_q).$$

Les assertions suivantes sont satisfaites :

$$(3.1) \quad s_1 < t;$$

$$(3.2) \quad g_p > t;$$

$$(3.3) \quad F(f) = F(g) (t)^k F(s);$$

$$(3.4) \quad \text{pour tout } h \in X^* : F(fh) = F(g) F(t^k sh).$$

Preuve : L'assertion (3.1) résulte immédiatement de la définition de s ; $t^k s$ est le plus long facteur droit de f qui est de cette forme et il résulte de la remarque 3 que $g_p > t$; (3.1) et (3.2) permettent de conclure à (3.3). De plus g_p est un mot de Lyndon et $t^k s$ n'est pas un facteur gauche de g , on a donc pour tout $h \in X^*$, $g_p > t^k sh$, d'où il résulte (3.4). \square

Il s'en déduit immédiatement :

PROPOSITION 4 : Soit f un élément de X^+ , r le facteur résiduel de f , et $g \in X^*$ tel que $f = gr$; pour tout $h \in X^*$: $F(fh) = F(g) F(rh)$.

4. POINT CRITIQUE D'UN MOT DE LYNDON

Considérons un mot de Lyndon f de longueur ≥ 2 ; soit f'' le facteur droit maximal de f , et f' tel que $f = f' f''$;

$$(4.1) \quad \text{Aucun facteur droit non vide de } f' \text{ n'est facteur gauche de } f''.$$

Preuve : Soit $u \in X^*$ le plus long facteur droit de f' qui soit facteur gauche de f'' . Posons $f'' = uv$, $uf'' = uuv$ est un facteur droit de f , de même que v ; f'' est le facteur droit maximal de f d'où $uuv \leq uv$ et $v \leq uv$; il en résulte $uv \leq v \leq uv$ d'où $u = 1$; ce qui permet de conclure. \square

$$(4.2) \quad \text{Aucun facteur gauche non vide de } f' \text{ n'est facteur droit de } f''.$$

Preuve : En effet f est un mot de Lyndon et n'admet pas de facteur propre qui soit facteur gauche et facteur droit. \square

$$(4.3) \quad f' \text{ n'est pas facteur de } f''.$$

Preuve : Supposons que f'' est de la forme $f'' = uf'v$, alors $f = f'uf'v$ est un mot de Lyndon et $f'uf'v < f'v$, d'où $uf'v < v$ ce qui est contradictoire puisque $f'' = uf'v$ est le facteur droit maximal de f . \square

(4.4) f'' n'est pas facteur de f' .

Preuve : Supposons f' de la forme $f' = uf''v$, alors $f = uf''vf''$ et $f'' < f''vf''$; ce qui est contradictoire puisque f'' est le facteur droit maximal de f . \square

Des assertions 4.1, 4.2, 4.3, 4.4, nous pouvons déduire la proposition suivante :

PROPOSITION 5 : Soient f un mot de Lyndon, f'' le facteur droit maximal de f , et f' le facteur gauche de f tel que $f = f'f''$:

- (a) (f', f'') est un point critique de f ;
- (b) (f'', f') est un point critique de $f''f'$.

Preuve : D'après les propriétés, la répétition élémentaire en (f', f'') est $f''f'$ et $|f''f'| = |f|$ est la plus petite période de f . De même $f'f''$ est la répétition élémentaire de (f'', f') et $|f|$ est la plus petite période de $f''f'$. \square

5. POINT CRITIQUE D'UN PRODUIT DE MOTS DE LYNDON

Considérons une suite non croissante de mots de Lyndon, f_1, \dots, f_p , et, t un mot de Lyndon strictement inférieur à f_p , k un entier non nul, et s un facteur gauche strict de t , satisfaisant : $t^k s$ n'est pas facteur gauche de f_p .

Notons t'' le facteur droit maximal de t et posons t' tel que $t : t = t' t''$.

Les assertions suivantes sont satisfaites :

- (5.1) $t^k s < f_p$;
- (5.2) pour $i = 1, \dots, p$: $t^k s$ n'est pas facteur gauche de f_i ;
- (5.3) pour $i = 1, \dots, p$: $t^k s$ n'est pas facteur de f_i ;
- (5.4) $t^k s$ n'est pas facteur de $f_1 \dots f_p$;
- (5.5) aucun facteur gauche non vide de $t^k s$ n'est facteur droit de $f_1 \dots f_p$;
- (5.6) si $f_1 \dots f_p$ est facteur de $t^k s$: $f_1 \dots f_p$ est facteur de t ;
- (5.7) si $f_1 \dots f_p$ n'est pas facteur de t : $(f_1 \dots f_p, t^k s)$ est un point critique de $f_1 \dots f_p t^k s$;
- (5.8) si $f_1 \dots f_p$ est facteur droit de t : $(f_1 \dots f_p t', t'' t^{k-1} s)$ est un point critique de $f_1 \dots f_p t^k s$;
- (5.9) si $f_1 \dots f_p$ est facteur de t , sans être facteur droit de t : t'' n'est pas facteur de $f_1 \dots f_p$;
- (5.10) si $f_1 \dots f_p$ est facteur de t , sans être facteur droit de t : $(f_1 \dots f_p t', t'' t^{k-1} s)$ est un point critique de $f_1 \dots f_p t^k s$.

Preuve : t et f_p étant deux mots de Lyndon satisfaisant $t < f_p$, il résulte de la remarque 2 (b) que l'assertion (5.1) est satisfaite.

Par hypothèse pour $i=1, \dots, p, f_p \leq f_i$, et d'après (5.1), $t^k s < f_p \leq f_i$, $t^k s$ n'étant pas facteur gauche de f_p , la proposition 1 nous permet de conclure à (5.2).

Il s'en déduit d'après la proposition 2 (2) que l'assertion (5.3) est vérifiée.

D'après (5.3), si nous supposons que $t^k s$ est facteur de $f_1 \dots f_p$, il en résulte qu'un facteur gauche de $t^k s$ est facteur droit de l'un des f_i ce qui est impossible, sachant que $t^k s < f_i$ et que f_i est un mot de Lyndon, il en résulte que (5.4) est vérifiée.

De la même façon nous pouvons conclure à (5.5).

L'assertion (5.6) se déduit alors immédiatement de (5.5).

Considérons la répétition élémentaire u en $(f_1 \dots f_p, t^k s)$, il résulte de l'assertion (5.5) que u ne saurait être facteur droit de $f_1 \dots f_p$ et facteur gauche de $t^k s$; il résulte de l'assertion (5.4) que u ne saurait être facteur droit de $f_1 \dots f_p$ et $t^k s$ facteur gauche de u . Il résulte alors de (5.6) que si $f_1 \dots f_p$ n'est pas facteur de t , nécessairement $t^k s$ est un facteur gauche de u , par suite $|u|$ est une période de $f_1 \dots f_p t^k s$, c'est la plus petite période, ce qui établit (5.7).

Si $f_1 \dots f_p$ est un facteur droit de t , il est clair que $|t|$ est une période de $f_1 \dots f_p t^k s$, et t étant un mot de Lyndon il résulte de la proposition 5 que la répétition élémentaire en $(f_1 \dots f_p t', t'' t^{k-1} s)$ est $t'' t'$, ce qui permet de conclure à l'assertion (5.8).

Supposons que t est de la forme $at''b$, sachant que t'' est le facteur droit maximal de t , nous avons nécessairement $b=1$, d'où il résulte que l'assertion (5.9) est vérifiée.

Supposons que $f_1 \dots f_p$ est facteur de t sans être facteur droit et soit u la répétition élémentaire en $(f_1 \dots f_p t', t'' t^{k-1} s)$: t est un mot de Lyndon et t'' le facteur droit maximal de t , il résulte de la proposition 5 que u est de la forme $u = t'' vt'$, et de l'assertion (5.9) que v est non vide et que $f_1 \dots f_p t'$ est un facteur droit de $t'' vt'$. De plus les hypothèses de (5.10) impliquent que $|f| \geq 2$ et par suite $t' \neq 1$, or t est un mot de Lyndon et (t', t'') un point critique de t , il en résulte que vt' ne saurait être facteur gauche de $t^{k-1} s$ sans que v soit une puissance de t , et il s'en déduirait que $f_1 \dots f_p$ est un facteur droit de v et par suite de t ce qui contredirait l'hypothèse : par conséquent $t'' vt'$ n'est pas un facteur gauche de $t'' t^{k-1} s$ et il en résulte que $t'' t^{k-1} s$ est un facteur gauche de $t'' vt'. f_1 \dots f_p t'$ est un facteur droit de $u = t'' vt'$, et $t'' t^{k-1} s$ est un facteur gauche. $|u|$ est la plus petite période de $f_1 \dots f_p t^k s$ et $(f_1 \dots f_p t', t'' t^{k-1} s)$ un point critique ce qui établit (5.10), et achève la preuve des assertions (5.1) à (5.10). \square

Fin de la démonstration du théorème : Soient $f \in X^+$, $r = t^k s$ le facteur résiduel de f , et $g \in X^+$ tel que $f = gr$. Posons $F(g) = (f_1) \dots (f_p)$. Il résulte de l'assertion (3.2) que $f_1 \geq \dots \geq f_p > t$.

Si g n'est pas facteur de r l'assertion (5.7) nous permet de conclure que (g, r) est un point critique de f , ce qui établit (a).

Notons f'' le facteur droit maximal de $t^k s$ et t'' le facteur droit maximal de t , et posons $f = f' f''$, $t = t' t''$, nécessairement $f'' = t'' t^{k-1} s$ et par conséquent si g est facteur de r , il résulte de l'assertion (5.6) que g est facteur de t , et des assertions (5.8) et (5.10) que (f', f'') est un point critique de f ce qui établit (b), et achève la preuve du théorème.

BIBLIOGRAPHIE

1. Y. CESARI et M. VINCENT, *Une caractérisation des mots périodiques*, C. R. Acad. Sc., t. 286, série A, 1978, p. 1175.
2. K. T. CHEN, R. H. FOX et R. C. LYNDON, *Free Differential Calculus IV*, Ann. Math., vol. 68, 1958, p. 81-95.
3. J. P. DUVAL, *Sur la périodicité des mots*, Thèse de 3^e cycle, Université de Rouen, 1978.
4. J. P. DUVAL, *Algorithme de factorisation d'un mot en mots de Lyndon*, Actes du premier colloque A.F.C.E.T.-S.M.F. de Math. appliquées, t. II, 1978, p. 15-26.
5. J. P. DUVAL, *Périodes et répétitions des mots du monoïde libre*, Theoretical Computer Sciences, vol. 9, 1979, p. 17-26.
6. M. J. FINE et H. S. WILF, *Uniqueness Theorems for Periodic Functions*, Proc. Amer. Math. Soc., vol. 16, 1965, p. 109-114.
7. D. E. KNUTH, J. H. MORRIS et V. R. PRATT, *Fast Pattern Matching in Strings*, S.I.A. M.J. Comput., vol. 6, (2), 1977, p. 321-349.
8. A. LENTIN, *Équations dans le monoïde libre*, Gauthier-Villars et Mouton, Paris-La Haye, 1972.
9. R. C. LYNDON et P. E. SCHUPP, *Combinatorial Group Theory*, Springer-Verlag, 1977.
10. M. P. SCHUTZENBERGER, *Sur une propriété combinatoire des algèbres de Lie libres pouvant être utilisés dans un problème de Mathématiques appliquées*, Séminaire Dubreil-Pisot, 1958/1959, Institut Henri-Poincaré, Paris, 1958.
11. M. P. SCHUTZENBERGER, *A Property of Finitely Generated Submonoids*, in Algebraic Theory of Semi-Groups, G. POLLAK, éd., North Holland, 1979, p. 545-576.
12. G. VIENNOT, *Bases des algèbres de Lie libres et factorisation des monoïdes libres*, Lecture Notes in Mathematics, n° 691, Springer-Verlag, 1979.

