JOURNAL OF COMPUTER AND SYSTEM SCIENCES 26, 14–33 (1983)

# Langages à un Compteur

MICHEL LATTEUX

*Université des Sciences et Techniques de Lille I*
*U.E.R. d'E.E.A., 59655 Villeneuve d'Ascq, France*

Received January 8, 1981

Nous étudions les langages à un compteur, c'est-à-dire, les langages appartenant à Rocl, le cône rationnel engendré par $D_1'^*$, le langage de semi-Dyck sur une lettre. Nous montrons que tout générateur de Rocl est bifidèle et qu'il existe un langage à un compteur qui domine tous les autres par transduction rationnelle préservant les longueurs. Nous établissons, ensuite, que le lemme d'itération des langages linéaires est *presque vrai* pour les langages à un compteur. Enfin, nous prouvons que Rocl ne contient aucune FAL non rationnelle et que, si Rocl est inclus dans la plus petite FAL contenant un cône rationnel $L$ clos par union, alors Rocl est inclus dans $L$. De même, si Rocl est inclus dans la plus petite FAL close par substitution contenant un cône rationnel algébrique $L$, alors Rocl est inclus dans $L$.

We study restricted one-counter languages, that is, languages belonging to Rocl, the full trio generated by $D_1'^*$, the restricted Dyck language over one pair of parentheses. We show that every generator of Rocl is a bifaithful one and that there exists a one-counter language which dominates the other ones by length-preserving rational transduction. Then, we establish that the pumping lemma for linear languages is *almost true* for restricted one-counter languages. Finally, we prove that Rocl contains no nonrational AFL and, if Rocl is included in the smallest full AFL containing a full semi-AFL $L$, then Rocl is included in $L$. If Rocl is included in the smallest full AFL closed under substitution containing a context-free full trio $L$, then Rocl is included in $L$.

## INTRODUCTION

Le rôle privilégié que joue, en théorie des langages, Alg, la famille des langages algébriques (*context-free languages*) vient des rapports très étroits qui lient cette famille aux diverses branches de l'informatique. L'étude des langages algébriques conduit naturellement à s'intéresser aux sous familles de Alg. Les plus connues sont obtenues en ajoutant des restrictions simples aux grammaires algébriques (c'est le cas pour Lin, la famille des langages linéaires) ou aux automates à pile. Ainsi, la famille des langages à un compteur, notée Rocl, est obtenue en considérant des automates dont l'alphabet de pile ne contient qu'un élément. Cette famille admet une autre définition très simple. Il a, en effet, été démontré que Rocl est le plus petit cône rationnel (famille fermée par transduction rationnelle) contenant $D_1'^*$, le langage de semi-Dyck sur une lettre, engendré par la grammaire $S \to SS$, $S \to aSb$, $S \to \varepsilon$.



0022-0000/83/010014-20$03.00/0
Copyright © 1983 by Academic Press, Inc.
All rights of reproduction in any form reserved.

L'intérêt de l'étude des langages à un compteur est renforcé par les rapports qui existent entre ces langages et les problèmes de synchronisation (cf. [20]). Un bon nombre de ces problèmes peut, en effet, s'exprimer par un langage qui est une intersection finie de langages à un compteur et il est bien connue que le fameux *reachability problem* pour les systèmes d'addition de vecteurs est décidable si et seulement si le plus petit cône rationnel clos par intersection contenant $D_1'^*$, est inclus dans la famille des langages récursifs.

Il n'est donc pas étonnant que de nombreux auteurs se soient intéressés aux langages à un compteur. Ainsi, il a été démontré que $D_1'^*$ était un langage d'index infini [30, 31, 33]; Greibach [15] a établi que tout générateur de Rocl possède un facteur itérant (cf. aussi [2]); Autebert [1] a prouvé la non-principalité de Rocl en tant que cylindre. On sait, aussi, que tout générateur de Rocl est effaçable [12] et que tout langage à un compteur est Parikh-borné [27], c'est-à-dire contient un langage borné de même image commutative. De nombreux autres résultats concernant la famille Rocl existent dans la littérature (cf., par exemple, [3, 6, 14–16, 18]).

Dans la section I, nous étudions les générateurs de Rocl et nous démontrons que, contrairement aux générateurs de Lin, ils sont tous bifidèles, c'est-à-dire qu'ils dominent par transduction rationnelle bifidèle tous les langages à un compteur. Ces transductions qui sont à la fois d'image finie (l'image de chaque mot est un langage fini) et fidèle (l'image inverse de chaque mot est finie) sont les seules qui puissent s'étendre aux mots infinis. Nous établissons, ensuite, que le langage $\check{L}' = D_1'^* \cup D_1'^*b$ domine tous les langages de Rocl par une transduction rationnelle préservant les longueurs.

Dans la deuxième section, nous montrons que le lemme d'itération des langages linéaires est *presque vrai* pour les langages à un compteur. Nous en déduisons que $\bar{D}_1^* = \{w \in \{a, b\}^* / |w|_a \neq |w|_b\}$ n'appartient pas à Rocl ce qui amène à conjecturer que tout langage à un compteur commutatif est rationnel.

La section III est consacrée aux produits de langages dans Rocl. Le résultat de base est le suivant: si $L_1$ et $L_2$ sont des langages non rationnels tels que $L_1 \# L_2 \in$ Rocl, alors $L_1 \in \mathscr{C}(I_1^*)$ et $L_2 \in \mathscr{C}(F_1^*)$, où $I_1^*$ (resp. $F_1^*$) est l'ensemble des facteurs gauches (resp. droits) des mots de $D_1'^*$. Ceci nous permet de montrer que Rocl est sans étoile et que Ocl $= \mathscr{F}$(Rocl), la plus petite FAL contenant Rocl est sans substitution.

Dans la dernière section, nous établissons pour $D_1'^*$, une propriété qui se révèle utile quand on doit comparer Rocl à d'autres familles de langages: si $\mathscr{L}$ est un cône rationnel clos par union, $D_1'^* \in \mathscr{F}(\mathscr{L})$ implique $D_1'^* \in \mathscr{L}$. Ce résultat peut être précisé si on se restreint aux langages algébriques. En établissant que tout langage algébrique $L$, vérifiant $L \subseteq D_1'^* \subseteq F(L)$, l'ensemble des facteurs de $L$, domine rationnellement $D_1'^*$, nous sommes en mesure d'en déduire: Si $D_1'^* \in \mathscr{F}_\sigma(\mathscr{L})$, la plus petite FAL close par substitution et contenant le cône rationel algébrique $\mathscr{L}$, alors $D_1'^* \in \mathscr{L}$. En prenant $\mathscr{L} =$ Lin, la famille des langages linéaires, nous retrouvons immédiatement que $D_1'^*$ n'appartient pas à $\mathscr{F}_\sigma$(Lin), la famille des langages quasirationnels. Enfin, nous montrons que $D_1'^*$ est égal au shuffle de $C_1^*$ par lui-même ($C_1 = \{a^n b^n / n \geqslant 0\}$), ce qui entraine que le résultat précédent n'est plus vrai si on remplace la substitution par l'intersection (même si on considère des cônes rationnels fidèles), puisque $D_1'^* \in \mathscr{F}^f_\cap(C_1) \backslash \mathscr{C}(C_1)$

*Définitions*

Nous noterons $\mathbb{N}_+$, l'ensemble des entiers positifs et $\mathbb{N} = \mathbb{N}_+ \cup \{0\}$. Pour tout alphabet $Y$ et tout mot $w \in Y^*$, $|w|$ désignera la longueur de $w$ et pour $y \in Y$, $|w|_y$ est le nombre d'occurrences de $y$ dans $w$. Le mot vide, de longueur zéro, sera noté $\varepsilon$. Pour $w, w' \in Y^*$, posons:

$FG(w) = \{w_1 / \exists w_2,\ w_1 w_2 = w\}$ l'ensemble des *facteurs gauches* de $w$,

$FD(w) = \{w_2 / \exists w_1,\ w_1 w_2 = w\}$ l'ensemble des *facteurs droits* de $w$,

$F(w) = \{w_2 / \exists w_1, w_3,\ w_1 w_2 w_3 = w\}$ l'ensemble des *facteurs* de $w$,

$$w \uparrow w' = \{w_1 w' \cdots w_k w' / w = w_1 \cdots w_k,\ \forall i \in \{1,\dots,k\},\ w_i \in Y\},$$

$$w \sqcup\!\sqcup\, w' = \{w_1 w_1' \cdots w_k w_k' / w = w_1 \cdots w_k,\ w' = w_1' \cdots w_k',\ w_i, w_i' \in Y^*\},$$

le *shuffle* de $w_1$ et $w_2$.

Ces opérations sont étendues aux langages de la manière suivante: Pour $L, L' \subseteq Y^*$,

$$FG(L)(\text{resp. } FD(L), F(L)) = \{w' \in FG(w)(\text{resp. } FD(w), F(w))/w \in L\},$$

$$L \sqcup\!\sqcup\, L' = \{w \sqcup\!\sqcup\, w' / w \in L,\ w' \in L'\},$$

$$L \uparrow L' = \{w_1 w_1' \cdots w_k w_k' / w = w_1 \cdots w_k \in L,\ \forall i \in \{1,\dots,k\},\ w_i \in Y,\ w_i' \in L'\}.$$

Nous dirons qu'un homomorphisme $h$ de $X^*$ dans $Y^*$ est

*alphabétique* si $h(X) \subseteq Y \cup \{\varepsilon\}$,

*strictement alphabétique* si $h(X) \subseteq Y$,

*non effaçant* si $h(X) \subseteq Y^+$,

*limité sur* $R \subseteq X^*$ s'il existe $k \in \mathbb{N}$ tel que $w \in F(R)$ et $h(w) = \varepsilon$ *entraine* $|w| \leqslant k$.

Pour les transductions rationnelles, nous prendrons comme définition la caractérisation qu'en a donné Nivat dans [30] (cf. [4]).

Une relation $\hat{\tau} \subseteq X^* \times Y^*$ est rationnelle s'il existe un alphabet $Z$, un langage rationnel $R \subseteq Z^*$ et deux homomorphismes (alphabétiques) $h$ et $g$ tels que

$$\hat{\tau} = \{(h(u), g(u))/u \in R\}.$$

La transduction $\tau$ de $X^*$ dans $Y^*$ canoniquement associée à $\hat{\tau}$ par $\forall w \in X^*$, $\tau(w) = \{w'/(w, w') \in \hat{\tau}\} = g(h^{-1}(w) \cap R)$ est une transduction rationnelle.

La transduction $\tau$ de $X^*$ dans $Y^*$ est dite

*fidèle* si $\forall v \in Y^*$, $\tau^{-1}(v) = \{w \in X^*/(w, v) \in \hat{\tau}\}$ est fini, *d'image finie* si $\forall w \in X^*$, $\tau(w)$ est fini,

*bifidèle* si elle est fidèle et d'image finie,

*croissante* si $\forall w \in X^*$, $\forall v \in \tau(w)$, $|v| \geqslant |w|$.

Une famille de langages $\mathscr{L}$ est un *cône rationnel* (resp. *fidèle*, *d'image finie*, *bifidèle*) si et seulement si $\mathscr{L}$ est fermé par transductions rationnelles (resp. fidèles, d'image finie, bifidèles).

Pour toute famille de langages $\mathscr{L}$, nous noterons $\mathscr{C}(\mathscr{L})$ (resp. $\mathscr{C}^{\mathrm{f}}(\mathscr{L})$, $\mathscr{C}^{\mathrm{if}}(\mathscr{L})$, $\mathscr{C}^{\mathrm{bf}}(\mathscr{L})$, $\mathscr{C}_{\cup}(\mathscr{L})$) le plus petit cône rationnel (resp. fidèle, d'image finie, bifidèle, clos par union) contenant $\mathscr{L}$ (nous conviendrons de remplacer $\mathscr{L}$ par $L$ si $\mathscr{L}$ ne contient que le langage $L$).

Si $\mathscr{L} = \mathscr{C}(L)$ (resp. $\mathscr{C}^{\mathrm{f}}(L)$, $\mathscr{C}^{\mathrm{bf}}(L)$) nous dirons que $\mathscr{L}$ est un cône rationnel (resp. fidèle, bifidèle) *principal* et que $L$ est un *générateur* (resp. fidèle, bifidèle) de $\mathscr{L}$.

Le langage $D_1'^* = \{w \in \{a, b\}^* / |w|_a = |w|_b$ et $w' \in FG(w)$ entraine $|w'|_a \geqslant |w'|_b\}$, appelé langage de semi-Dyck sur une lettre est un générateur fidèle de la famille des langages à un compteur, $\mathrm{Rocl} = \mathscr{C}(D_1'^*) = \mathscr{C}^{\mathrm{f}}(D_1'^*)$. Nous poserons $I_1^* = FG(D_1'^*)$ et $F_1^* = FD(D_1'^*)$.

Nous désignerons, respectivement, par Rat, Lin, Alg les familles des langages rationnels, linéaires, et algébriques (ou context-free).

Soient $\mathscr{L}$, $\mathscr{L}'$ deux familles de langages. Une substitution $s$ définie sur un alphabet $Y$ est une *$\mathscr{L}'$-substitution* si pour tout $y \in Y$, le langage $s(y)$ appartient à $\mathscr{L}'$. Nous noterons

$\mathscr{L} \square \mathscr{L}' = \{s(L)/L \in \mathscr{L}$, $s$ est une $\mathscr{L}'$-substitution$\}$,

$\mathscr{F}(\mathscr{L}) = \mathrm{Rat} \square \mathscr{C}(\mathscr{L})$, la FAL engendrée par $\mathscr{L}$, c'est-à-dire le plus petit cône rationnel fermé par union, produit et étoile, contenant $\mathscr{L}$.

$\mathscr{F}_\sigma(\mathscr{L})$, la plus petite FAL fermée par ($\mathscr{L}$–) substitution, contenant $\mathscr{L}$.

$$\mathscr{F}_\sigma(\mathscr{L}) = \bigcup_{i \geqslant 1} \mathscr{L}_i \qquad \text{avec} \quad \mathscr{L}_1 = \mathscr{C}_{\cup}(\mathscr{L}) \quad \text{et} \quad \forall j \geqslant 1, \quad \mathscr{L}_{j+1} = \mathscr{L}_j \square \mathscr{L}_1.$$

$\mathscr{C}_{\cap}(\mathscr{L})$ (resp. $\mathscr{C}^{\mathrm{f}}_{\cap}(\mathscr{L})$, $\mathscr{F}^{f}_{\cap}(\mathscr{L})$), le plus petit cône rationnel (resp. fidèle, fidèle et clos par union et étoile) contenant $\mathscr{L}$ et fermé par intersection.

Pour deux langages $L$, $L'$ inclus dans $Y^*$, posons

$$LL'^{-1} = \{w \in Y^* / \exists w' \in L', \ ww' \in L\} = L/L',$$
$$L^{-1}L' = \{w \in Y^* / \exists w' \in L, \ w'w \in L'\}.$$

Nous dirons que $L$ est un *IRS-langage*, s'il ne possède pas de facteur itérant, c'est-à-dire si tout langage rationnel inclus dans $L$ est fini.

Un *transducteur séquentiel* est un 6-uple $M = (X, Y, Q, q_0, *, \cdot)$, où $X$ et $Y$ sont des alphabets, $Q$ un ensemble fini d'états, $q_0 \in Q$ est l'état initial, "$*$" et "$\cdot$" sont deux applications partielles de même domaine inclus dans $Q \times X$ à valeur, respectivement, dans $Y^*$ et $Q$. On étand ces deux applications sur $Q \times X^*$ en posant: $\forall q \in Q$, $q * \varepsilon = \varepsilon$ et $q \cdot \varepsilon = q$ et $\forall w \in X^*$, $x \in X$, $q \in Q$,

$$\begin{aligned} q * wx &= \text{non défini} && \text{si} \quad q * w \quad \text{non défini,} \\ &= (q * w)(q \cdot w) * x, && \text{sinon;} \\ q \cdot wx &= \text{non défini,} && \text{si} \quad q \cdot w \quad \text{non défini,} \\ &= (q \cdot w) \cdot x, && \text{sinon.} \end{aligned}$$

L'*application séquentielle s*, associée au transducteur, est définie par $\forall u \in X^*$, $s(u) = q_0 * u$.

Cette application est *fidèle* si $\forall v \in Y^*$, $s^{-1}(v)$ est fini. Elle est *alphabétique* si $\forall q \in Q$, $\forall x \in X$, soit $q * x$ est non défini, soit $q * x \in Y \cup \{\varepsilon\}$.

## I. Sur les generateurs de Rocl

Pour tout cône rationnel principal $\mathscr{L}$, on peut se demander si on a besoin de toute la puissance des transductions rationnelles pour engendrer ce cône à partir d'un seul langage. Plus précisément, pour un type particulier de transduction rationnelle, on peut se poser deux questions:

(i) Existe-t-il un générateur $L$ qui domine chaque langage de $\mathscr{L}$ par une transduction de ce type?

(ii) Tout générateur de $\mathscr{L}$ domine-t-il chaque langage de $\mathscr{L}$ par une transduction de ce type?

Plusieurs résultats des langages sont des réponses à ces questions. Ainsi, l'existence, pour Alg, d'un *hardest language* [14] et le fait que tout cône rationnel principal possède un générateur séquentiel [25] peuvent être reliés à la question (i). De même, en relation avec la question (ii), il a été montré que tout générateur algébrique est un générateur fidèle [7, 11] et que tout générateur de la famille, Lin, des langages linéaires est un générateur sous-séquentiel [25].

Nous allons, d'abord, établir que tout générateur de Rocl est un générateur bifidèle. Ce résultat a été démontré dans [25] pour les générateurs algébriques en exhibant un IRS-générateur bifidèle. Le fait que tout générateur de Rocl contienne un facteur itérant [15], interdit d'utiliser la même méthode. On peut, cependant, démontrer pour $D_1'^*$ un lemme que nous avions établi, dans [25], pour tout IRS-langage.

LEMME 1. *Si $D_1'^* \in \mathscr{C}(L)$ (resp. $\mathscr{C}^{\mathrm{f}}(L)$), alors $D_1'^* \in \mathscr{C}^{\mathrm{if}}(L)$ (resp. $\mathscr{C}^{\mathrm{bf}}(L)$).*

*Démonstration.* Si $D_1'^* \in \mathscr{C}(L)$, il existe un langage rationnel $R \subseteq Z^*$ et deux homomorphismes alphabétiques $h$ et $g$ tels que $D_1'^* = g(h^{-1}(L) \cap R)$ et $\forall z \in Z$ $h(z)\, g(z) \neq \varepsilon$. Soit $k$ le nombre d'états de l'automate déterministe minimal qui reconnait $R$. Posons $Z_0 = \{z \in Z / h(z) = \varepsilon\}$, $R' = R \backslash Z^* Z_0^{2k} Z^*$ et $K = \{a^k, b^k\}^*$.

Montrons, d'abord, que $L' = D_1'^* \cap K$ est inclus dans $L'' = g(h^{-1}(L) \cap R')$. Comme $R = R' \cup R''$ avec $R'' = R \cap Z^*Z_0^{2k}Z^*$, $D_1'^* = g(h^{-1}(L) \cap R') \cup g(h^{-1}(L) \cap R'')$ et il suffit de prouver que $K \cap g(h^{-1}(L) \cap R'') = \emptyset$. Pour cela, considérons $w = g(u_1 v u_2) \in g(h^{-1}(L) \cap R'')$ avec $v \in Z_0^{2k}$ et $u_1 v u_2 \in h^{-1}(L) \cap R''$. Alors, $v$ peut se factoriser en $v = v_1 v_2 v_3 v_4 v_5$ avec $v_2, v_4 \in Z_0^+$ et $u_1 v_1 v_2^+ v_3 v_4^+ v_5 u_2 \subseteq R''$. Comme $h(u_1 v_1 v_2^+ v_3 v_4^+ v_5 u_2) = h(u_1 v u_2) \in L$, nous en déduisons que $g(u_1 v_1)\, g(v_2^+)\, g(v_3)\, g(v_4^+)\, g(v_5 u_2) \subseteq g(h^{-1}(L) \cap R'') \subseteq D_1'^*$, avec $g(v_2)$ et $g(v_4) \neq \varepsilon$ vérifiant $|g(v_2)|_a = |g(v_2)|_b$ et $|g(v_4)|_a = |g(v_4)|_b$. Donc il existe $i < 2k$ tel que $ba^ib$ soit facteur de $g(v)$, ce qui entraine que $g(v)$ ne peut pas être facteur d'un mot de $K$ et $w = g(u_1 v u_2) \notin K$. Nous obtenons, donc, $L' = D_1'^* \cap K \subseteq L'' = g(h^{-1}(L) \cap R') \subseteq D_1'^*$, ce qui implique $L'' \cap K = D_1'^* \cap K$. Par construction de $R'$, $L'' \in \mathscr{C}^{\mathrm{if}}(L)$ et si, de plus $g$ est limité sur $R$, il est, à fortiori, limité sur $R'$ et $L'' \in \mathscr{C}^{\mathrm{bf}}(L)$. La démonstration se termine en remarquant que $D_1'^* \in \mathscr{C}^{\mathrm{bf}}(L'')$. En effet, prenons l'homomorphisme $\phi$ défini par $\phi(a) = a^{2k}$ et $\phi(b) = b^{2k}$. Alors, il est facile de vérifier que $D_1'^* = \phi^{-1}(D_1'^* \cap K) = \phi^{-1}(L'' \cap K)$. ∎

Montrons, maintenant, que, contrairement à $\mathrm{Sym} = \{w\bar{w}^R / w \in \{a, b\}^*\}$, $D_1'^*$ est un générateur bifidèle. Plus précisément, $z$ étant une nouvelle lettre, nous allons établir que $D_1'^*$ est un générateur séquentiel de Rocl; c'est-à-dire,

Lemme 2. *Pour tout langage* $L \in$ Rocl, *il existe une application séquentielle alphabétique et fidèle s telle que* $s(D_1'^*z) = L$.

*Démonstration.* Comme $L \in \mathscr{C}^{\mathrm{f}}(D_1'^*)$ [11], il suffit de montrer qu'on peut passer, par une application séquentielle alphabétique et fidèle, de $D_1'^*z$ à $L'z$ avec $L' = \{x, y\}^*\ (D_1'^* \uparrow \{x, y\}^*)$ [25]. Considérons l'homomorphisme $h$ défini sur $Z = \{a, b, x, y, z\}$ par $h(a) = a^3b^2$, $h(b) = b$, $h(x) = ab$, $h(y) = a^2b^2$ *et* $h(z) = z$. Alors, $V = h(Z)$ est un code préfixe qui vérifie $D_1'^*z \cap V^* = D_1'^*z \cap FG(V^*)$. D'après [25, lemme 3], il existe une application séquentielle alphabétique et fidèle $s$ telle que $s(D_1'^*z) = h^{-1}(D_1'^*z)$. Pour achever la démonstration, il suffit de vérifier que $h^{-1}(D_1'^*z)$ est égal à $L'z$. ∎

Remarquons que $z$, le marqueur de fin, est nécessaire dans l'énoncé du lemme précédent. On peut, en effet, montrer qu'il n'est par vérifié pour $D_1'^*$ ni pour les autres générateurs classique $D_1' = aD_1'^*b$ et $\check{L} = D_1'^*b$. D'autre part, c'est un problème ouvert de savoir si, comme pour les langages linéaires, pour tout générateur $L$ de Rocl et tout marqueur de fin $\#$, $L\#$ est un générateur séquentiel. De même, on ne sait pas si, comme pour les générateurs algébriques, tout générateur de Rocl est un générateur fonctionnel.

Prenons, maintenant, un générateur $L$ de Rocl et un langage à un compteur $L'$. Comme $D_1'^* \in \mathscr{C}^{\mathrm{f}}(L)$ [12], le lemme 1 entraine $D_1'^* \in \mathscr{C}^{\mathrm{bf}}(L)$ et donc $D_1'^*z \in \mathscr{C}^{\mathrm{bf}}(L)$, mais d'après le lemme précédent, $L' \in \mathscr{C}^{\mathrm{bf}}(D_1'^*z) \subseteq \mathscr{C}^{\mathrm{bf}}(L)$. Nous pouvons donc énoncer

Proposition 3. *Tout générateur $L$ de* Rocl *est un générateur bifidèle, c'est-à-dire que pour tout $L' \in$* Rocl, *il existe une transduction rationnelle bifidèle $\tau$ telle que $L' = \tau(L)$.*

En utilisant les résultats de Leguy sur les transductions rationnelles décroissantes [28, 29], on peut préciser ce résultat.

Proposition 4. *Soient $L$ un générateur de* Rocl *et $L' \in$* Rocl. *Alors, il existe un langage rationnel $R$, un homomorphisme strictement alphabétique $h$ et un homomorphisme alphabétique $g$, limité sur $R$ tels que $L' = g(h^{-1}(L) \cap R)$.*

Montrons, maintenant, l'existence d'un générateur particulier de Rocl qui domine tous les langages à un compteur par transductions rationnelles conservant les longueurs. Clairement, $D_1'^*$, le langage de semi-Dyck sur une lettre engendré par la grammaire: $S \to SS$, $S \to aSb$, $S \to \varepsilon$, ne contient que des mots de longueur paire et donc, ne peut pas convenir. Il en est de même du langage du Lukasiewicz, $\check{L} = D_1'^*b$, engendré par $S \to aSS$, $S \to b$. Par contre, en posant $\check{L}' = \check{L} \cup D_1'^*$, nous obtenons

Proposition 5. *Pour tout langage à un compteur $L$, il existe une transduction rationnelle $\tau$ conservant les longueurs telle que $L = \tau(\check{L}')$.*

*Démonstration.* Montrons, d'abord, que $D_1'^*$ est un générateur croissant, c'est-à-dire qu'il domine tout langage à un compteur par une transduction rationnelle croissante. Prenons $L' \in \text{Rocl} = \mathscr{C}^{\mathrm{f}}(D_1'^*)$. Alors, il existe un langage rationnel $R \subseteq Z^*$ et deux homomorphismes non effaçant $h$ et $g$ tels que $L' = g(h^{-1}(D_1'^*) \cap R)$. Clairement, il suffit de montrer qu'on peut passer de $D_1'^*$ à $L = h^{-1}(D_1'^*)$ par une transduction rationnelle croissante. Posons $X = \{a, b\}$, $Z = \{z_1, ..., z_n\}$, $\forall w \in X^*$, $d(w) = |w|_a - |w|_b$, $\bar{d}(w) = \inf\{d(w')/w = w'w''\}$. Pour tout $i \in \{1, ..., n\}$, notons $d_i = d(h(z_i))$, $\bar{d}_i = -\bar{d}(h(z_i))$ et $r_i = |h(z_i)|$. Soient $r = \sup\{r_i/i \in \{1, ..., n\}\}$, $Q = \{q_0, ..., q_{2r}\}$, $\Delta = Q \times (X \cup \{\varepsilon\}) \times Z \times Q$ et pour $i \in \{1, ..., 4\}$ définissons l'homomorphisme $\Pi_i$ sur $\Delta$ par $\Pi_i(y) = y_i$ si $y = (y_1, y_2, y_3, y_4) \in \Delta$.

Posons

$$Y_1 = \{(q_i, \varepsilon, z_t, q_j) \in \Delta/i \geqslant \bar{d}_t \text{ et } j = i + d_t < r\},$$

$$Y_2 = \{(q_i, a, z_t, q_j) \in \Delta/j = i + d_t - r \geqslant 0\},$$

$$Y_3 = \{(q_i, b, z_t, q_j) \in \Delta/i < \bar{d}_t \text{ et } j = i + d_t + r\},$$

$$Y = Y_1 \cup Y_2 \cup Y_3.$$

Considérons, maintenant, le langage rationnel sur $\Delta^*$.

$$K = \{y_1 \cdots y_k/\forall s \in \{1, ..., k\},\ y_s \in Y,\ \Pi_4(y_s) = \Pi_1(y_{s+1}) \text{ si } s < k \text{ et } \Pi_1(y_1) = \Pi_4(y_k) = q_0\}.$$

La transduction rationnelle $\tau$ de $X^*$ dans $Z^*$ définie par: $\forall u \in X^*$, $\tau(u) = \Pi_3(\Pi_2^{-1}(u) \cap K)$ est une transduction croissante puisque $\Pi_3$ est strictement alphabétique et $\Pi_2$ est alphabétique.

Vérifions l'égalité $h^{-1}(D_1'^*) = \tau(D_1'^*)$. Pour cela, montrons d'abord la propriété ($I_1^*$ désigne $FG(D_1'^*)$).

($*$) Si $y \in FG(K)$ avec $\Pi_2(y) \in I_1^*$, alors $h \circ \Pi_3(y) \in I_1^*$ et $d \circ h \circ \Pi_3(y) = r \times d \circ \Pi_2(y) + l$ avec $l = o$ si $y = \varepsilon$, $\Pi_4(y'')$ si $y = y'y''$ avec $y'' \in Y$.

Raisonnons par induction sur la longueur de $y$. Si $y = \varepsilon$, il est clair que la propriété ($*$) est verifiée. Prenons, donc, $y = y'(q_i, x, z_t, q_j) \in FG(K)$ avec $\Pi_2(y) \in I_1^*$. Alors $y' \in FG(K)$ et $\Pi_2(y') \in I_1^*$ et l'hypothèse de récurrence implique que $w' = h \circ \Pi_3(y') \in I_1^*$ et que $d(w') = r \times d \circ \Pi_2(y') + i \geqslant i$. Posons $w = h \circ \Pi_3(y) = w'h(z_t)$ et distinguons trois cas pour $x$.

(1) $x = \varepsilon$. Par définition de $Y$ et $K$, $i \geqslant \bar{d}_t$ et $j = i + d_t$. Donc $\bar{d}(w) = \inf\{\bar{d}(w'), d(w') + \bar{d}_t\} = 0$ puisque $\bar{d}(w') = 0$ et $\bar{d}_t - d(w') \leqslant d_t - i \leqslant 0$, ce qui entraine $w \in I_1^*$. En outre, $d(w) = d(w') + d_t = r \times d \circ \Pi_2(y') + i + d_t = r \times d \circ \Pi_2(y) + j$.

(2) $x = a$. Alors, $j = i + d_t - r \geqslant 0$, ce qui implique $i \geqslant \bar{d}_t$ et donc $\bar{d}(w) = o$ et $w \in I_1^*$. De plus, $d(w) = d(w') + d_t = r \times d \circ \Pi_2(y') + i + d_t = r \times d \circ \Pi_2(y) + i + d_t - r$ puisque $\Pi_2(y) = \Pi_2(y')a$ et $d(w) = r \times d \circ \Pi_2(y) + j$.

(3) $x = b$. Alors $i < \bar{d}_t$, $j = i + d_t + r$, et comme $\Pi_2(y) = \Pi_2(y')b \in I_1^*$, $d \circ \Pi_2(y') \geqslant 1$, donc $d(w') \geqslant r + i$. Comme $\bar{d}_t \leqslant r$, on en déduit que $w = w'h(z_t) \in I_1^*$. En outre, $d(w) = d(w') + d_t = r \times d \circ \Pi_2(y') + i + d_t = r \times d \circ \Pi_2(y) + i + d_t + r = r \times d \circ \Pi_2(y) + j$.

Prenons, maintenant, $v \in \tau(D_1'^*) = \Pi_3(\Pi_2^{-1}(D_1'^*) \cap K)$. Alors, $v = \Pi_3(y)$ avec $y \in K$ et $\Pi_2(y) \in D_1'^*$. D'après la propriété ($*$), $h(v) \in I_1^*$ et $d \circ h(v) = r \times d \circ \Pi_2(y) + l$ avec $d \circ \Pi_2(y) = 0$ puisque $\Pi_2(y) \in D_1'^*$ et $l = 0$ puisque $y \in K$, donc $d \circ h(v) = o$ et $h(v) \in D_1'^*$, ce qui entraine $v \in h^{-1}(D_1'^*)$ et $\tau(D_1'^*) \subseteq h^{-1}(D_1'^*)$. Pour établir l'inclusion inverse, nous utiliserons la propriété.

($**$) Soit $z \in Z^*$ tel que $h(z) \in I_1^*$, alors il existe $y \in FG(K)$ tel que $\Pi_2(y) \in I_1^*$, $\Pi_3(y) = z$ et $d \circ h(z) = r \times d \circ \Pi_2(y) + l$ avec $l = o$ si $y = \varepsilon$, $\Pi_4(y'')$ si $y = y'y''$ avec $y'' \in Y$.

Raisonnons par induction sur la longueur de $z$. Si $z = \varepsilon$, la propriété ($**$) est vérifiée en choisissant $y = \varepsilon$. Prenons, donc, $z = z'z_t$ avec $h(z) \in I_1^*$, ce qui entraine $h(z') \in I_1^*$ et l'hypothèse de récurrence implique l'existence de $y' \in FG(K)$ tel que $\Pi_2(y') \in I_1^*$, $\Pi_3(y') = z'$ et $d \circ h(z') = r \times d \circ \Pi_2(y') + i$. Distinguons trois cas:

(1) $i \geqslant \bar{d}_t$ et $j = i + d_t < r$. Alors $y = y'(q_i, \varepsilon, z_t, q_j) \in FG(K)$ et vérifie $\Pi_2(y) = \Pi_2(y') \in I_1^*$, $\Pi_3(y) = \Pi_3(y')$ $z_t = z$ et $d \circ h(z) = r \times d \circ \Pi_2(y') + i + d_t = r \times d \circ \Pi_2(y) + j$.

(2) $j = i + d_t - r \geqslant o$. Alors $y = y'(q_i, a, z_t, q_j) \in FG(K)$, $\Pi_2(y) = \Pi_2(y')\, a \in I_1^*$, $\Pi_3(y) = \Pi_3(y')\, z_t = z'z_t = z$ et $d \circ h(z) = d \circ h(z') + d_t = r \times d \circ \Pi_2(y) - r + i + d_t = r \times d \circ \Pi_2(y) + j$.

(3) $i < \bar{d}_t$. Prenons $j = i + d_t + r$ et $y = y'(q_i, b, z_t, q_j) \in FG(K)$. Alors, $\Pi_3(y) = \Pi_3(y')\, z_t = z$ et $d \circ h(z) = r \times d \circ \Pi_2(y) + r + i + d_t = r \times d \circ \Pi_2(y) + j$. D'autre part, comme $h(z) = h(z')\, h(z_t) \in I_1^*$, on a $d \circ h(z') \geqslant \bar{d}_t$, ce qui entraine $d \circ \Pi_2(y') \geqslant 1$ puisque $\bar{d}_t > i$. Nous en déduisons que $\Pi_2(y) = \Pi_2(y')\, b \in I_1^*$.

Prenons, maintenant, $z \in h^{-1}(D_1'^*)$. Alors $h(z) \in I_1^*$ et la propriété $(**)$ entraine qu'il existe $y \in FG(K)$ avec $\Pi_3(y) = z$, $\Pi_2(y) \in I_1^*$ et $d \circ h(z) = o = r \times d \circ \Pi_2(y) + l$. Nous en déduisons que $d \circ \Pi_2(y) = o$, donc $\Pi_2(y) \in D_1'^*$ et $l = 0$, donc $y \in K$ et $z = \Pi_3(y) \in \Pi_3(\Pi_2^{-1}(D_1'^*) \cap K) = \tau(D_1'^*)$. Nous avons, donc, démontré l'égalité entre $\tau(D_1'^*)$ et $h^{-1}(D_1'^*)$, ce qui entraine que $D_1'^*$ est un générateur croissant.

En utilisant la caractérisation des transductions rationnelles croissantes de Leguy [28], il est facile de voir que, pour toute transduction rationnelle croissante $\theta$ et tout langage $L'$, il existe une transduction rationnelle $\theta'$ préservant les longueurs telle que $\theta'(L) = \theta(L')$ avec $L = c^*(L' \uparrow c^*)$, où $c$ est une nouvelle lettre. Pour achever notre démonstration, il nous reste donc, à démontrer qu'on peut passer de $\check{L}'$ à $L = c^*(D_1'^* \uparrow c^*)$ par transductions rationnelles préservant les longueurs.

Considérons les homomorphismes $h_1$, $h_2$ définis sur $\{a, b, z\}$ par $h_1(a) = h_2(a) = a$, $h_1(b) = h_2(b) = b$, $h_1(z) = ab$, $h_2(z) = cc$. Alors, $L' = h_2(h_1^{-1}(D_1'^*)) = (c^2)^* (D_1'^* \uparrow (c^2)^*)$. Prenons, maintenant, $X' = \{a', b'\}$, $\Delta = X \cup X' \cup \{c, c'\}$ et le langage rationnel $B = (X \cup \{c\})^* \sqcup\!\sqcup (X'c')^*$. Définissons les homomorphismes $g_1$ et $g_2$ sur $\Delta$ par:

$\forall x \in X \cup \{c\}$, $g_1(x) = g_2(x) = x$, $g_1(x') = x$, et $g_2(a') = ca$, $g_2(b') = cb$, $g_2(c') = \varepsilon$. Pour tout mot $u \in B$, $g_1(u)$ et $g_2(u)$ ont la même longueur et il est facile de vérifier que $L_1 = g_2(g_1^{-1}(L') \cap B)$ est égal à $\{w \in L / |w|_c \text{ est pair}\}$.

On pourrait montrer de la même façon qu'on peut passer de $\check{L}$ à $L_2 = \{w \in L / |w|_c$ est impair$\}$ par transductions rationnelles préservant les longueurs. Comme $D_1'^* = \check{L}' \cap (X^2)^*$, $\check{L} = \check{L}' \cap (X^2)^* b$, et que $L = L_1 \cup L_2$, la démonstration est terminée. ∎

En utilisant la caractérisation de ces transductions rationnelles due à Eilenberg [10] (cf. aussi [28]), on obtient immédiatement

COROLLAIRE 6. *Pour tout langage à un compteur $L$, il existe un langage rationnel $R$ et deux homomorphismes strictement alphabétiques $h$, $g$ tels que $L = g(h^{-1}(\check{L}') \cap R)$.*

Le même raisonnement permet de montrer un résultat analogue pour les langages de $\mathscr{C}(D_1^*)$, le cône rationnel engendré par $D_1^*$, le langage de Dyck sur une lettre. On prend, alors, comme générateur le langage $D_1^* \cup D_1^* b$. Par contre, pour Alg et Lin, on ne sait pas, pour l'instant, si on peut obtenir la même propriété.

## II. Un lemme d'itération pour les langages à un compteur

Le premier lemme d'itération pour les langages de Rocl est dû à Boasson [6] et a permis, entre autre, de montrer que $D_1^*$, le langage de Dyck sur une lettre n'appartient pas à la famille Rocl. Nous allons établir un autre lemme d'itération qui est loin de caractériser les langages de Rocl (cf. par exemple [4]). Cependant, ce résultat, très simple à énoncer, se révèle très utile.

PROPOSITION 7. *Soit $L$ un langage de* Rocl. *Alors, il existe un entier positif $N$ tel que tout mot $w \in L$, de longueur $|w| \geqslant N$ se factorise en $w = xuyvz$ avec $uv \neq \varepsilon$, $|xuvz| \leqslant N$, et $\forall n \geqslant 1$, $xu^n yv^n z \in L$.*

*Démonstration.* Si $L \in$ Rocl, d'après le corollaire 6, il existe un langage rationnel $R \subseteq Z^*$ et deux homomorphismes strictement alphabétiques $h$, $g$ tels que $L = g(h^{-1}(\check{L}') \cap R)$. Soient $M = (Q, Z, *, q_0, F)$ un automate fini déterministe qui reconnait $R$ et $k$ le nombre de ses états. Posons $N = 2k^3$ et prenons $w \in L$ avec $|w| \geqslant N$. Alors, $w = g(\alpha)$ avec $\alpha \in R$, $|\alpha| = |w|$, et $h(\alpha) \in \check{L}'$. Posons $\alpha = \beta\alpha'\gamma$ avec $|\beta| = |\gamma| = k^3$ et distinguons trois cas:

(1) Pour tout facteur gauche $\beta'$ de $\beta$, $d \circ h(\beta') = |h(\beta')|_a - |h(\beta')|_b < k^2$. Posons $\beta = \beta_1 \cdots \beta_t$ avec $t = k^3$ et $\forall i \in \{1,..., t\}$, $\beta_i \in Z$, $q_i = q_0 * \beta_1 \cdots \beta_i$ et $d_i = d \circ h(\beta_1,..., \beta_i)$. Comme $Q$ possède $k$ éléments et que chaque $d_i$ est inférieur à $k^2$, il existe $o < r < s < t$ tels que $(q_r, d_r) = (q_s, d_s)$. Posons $\beta' = \beta_1 \cdots \beta_r$, $\delta = \beta_{r+1} \cdots \beta_s$, et $\beta'' = \beta_{s+1} \cdots \beta_t \alpha'\gamma$. Alors, $\beta'\delta^*\beta'' \subseteq R$ et, comme $d \circ h(\delta) = 0$, $h(\beta'\delta^*\beta'') \subseteq L'$. On obtient le résultat annoncé en prenant $x = g(\beta')$, $u = g(\delta)$, $y = g(\beta'')$, et $v = z = \varepsilon$.

(2) Pour tout facteur droit $\gamma'$ de $\gamma$, $d \circ h(\gamma') > -k^2$. Ce cas se traite de la même façon que le cas précédent.

(3) Sinon $\beta$ se factorise en $\beta(1) \cdots \beta(k^2)\beta'$, $\gamma$ se factorise en $\gamma'\gamma(1) \cdots \gamma(k^2)$ avec $\forall i \in \{1,..., k^2\}$, $d \circ h(\beta(i)) = 1$ et $d \circ h(\gamma(i)) = -1$. Posons $q'_0 = q_0 * \beta\alpha'\gamma'$ et $\forall i \in \{1,..., k^2\}$, $q_i = q_0 * \beta(1) \cdots \beta(i)$, $q'_i = q'_0 * \gamma(1) \cdots \gamma(i)$. Il existe $o < r < s \leqslant k^2$ tels que $(q_r, q'_r) = (q_s, q'_s)$. Posons $\gamma_1 = \beta(1) \cdots \beta(r)$, $\gamma_2 = \beta(r+1) \cdots \beta(s)$, $\gamma_3 = \beta(s+1) \cdots \beta(k^2)\beta'\alpha'\gamma' \quad \gamma(1) \cdots \gamma(r)$, $\gamma_4 = \gamma(r+1) \cdots \gamma(s)$, et $\gamma_5 = \gamma(s+1) \cdots \gamma(k^2)$. Alors $\gamma_1\gamma_2^*\gamma_3\gamma_4^*\gamma_5 \subseteq R$ et, comme $d \circ h(\gamma_2) = -d \circ h(\gamma_4) \geqslant 0$, $\forall n \geqslant 1$, $h(\gamma_1\gamma_2^n\gamma_3\gamma_4^n\gamma_5) \in \check{L}'$. On obtient, donc, le résultat annoncé en prenant $x = g(\gamma_1)$, $u = g(\gamma_2)$, $y = g(\gamma_3)$, $v = g(\gamma_4)$, et $z = g(\gamma_5)$. ∎

Si on impose, en plus, que $xyz \in L$, on retrouve le lemme d'itération pour les langages linéaires qui, clairement, n'est pas vérifié par les langages de Rocl (il est utilisé pour montrer que $D_1'^*$ n'est pas un langage linéaire). On voit, cependant, qu'il est facile de déduire de la proposition précédente que $D_1^* \notin$ Rocl. En fait, on peut obtenir un résultat plus fort. Considérons, en effet, $\bar{D}_1^*$, le complémentaire de $D_1^*$, le langage de Dyck sur une lettre. On obtient

COROLLAIRE 8. *$\bar{D}_1^* = \{w \in \{a, b\}^* / |w|_a \neq |w|_b\}$ n'est pas un langage à un compteur.*

*Démonstration.* Nous allons établir, en fait, que $L = \{a^n b^p a^q / n + q \neq p\} = \bar{D}_1^* \cap a^*b^*a^* \notin$ Rocl. Supposons que $L \in$ Rocl et soit $N$ l'entier de la proposition 7.

Alors $w = a^N b^{2N+N!} a^N \in L$ et $w$ se factorise en $xuyvz$ avec $|xuvz| < N$, ce qui entraine $xuvz \in a^*$ et $uv = a^t$ avec $1 \leqslant t < N$. Prenons $n = 1 + N!/t \in \mathbb{N}$. Il est clair que $xu^nyv^nz = a^ib^{2N+N!}a^j$ avec $i + j = 2N + N!$, d'où la contradiction. ∎

Comme $\bar{D}_1^*$ est, parmi les langages commutatifs, un langage minimal pour la transduction rationnelle [24], le corollaire 8 nous incite à énoncer

CONJECTURE 1. *Tout langage commutatif de* Rocl *est rationnel.*

Supposons, maintenant, que $L^2$ appartienne à Rocl avec $L \subseteq a^*b^*$. Comme tout langage algébrique non rationnel inclus dans $a^*b^*$ domine rationnellement l'un des langages $L_> = \{a^nb^p/n > p \geqslant o\}$, $L_{\neq} = \{a^nb^p/n \neq p\}$, $L_< = \{a^nb^p/o \leqslant n < p\}$ [5], en vérifiant au moyen de la proposition 7 qu'aucun des langages $L_>^2$, $L_{\neq}^2$, $L_<^2$ n'appartient à Rocl, on obtient

COROLLAIRE 9. *Soit $L$ un langage inclus dans $a^*b^*$. Alors $L^2 \in$* Rocl *implique $L \in$* Rat.

## III. ROCL EST SANS ÉTOILE

Une famille de langage $\mathscr{L}$ est *sans produit* si $L_1 \# L_2 \in \mathscr{L}$ implique $L_1$ ou $L_2$ rationnel; de même $\mathscr{L}$ est *sans étoile* si $(L_1\#)^* \in \mathscr{L}$ implique $L_1 \in$ Rat, la famille des langages rationnels ($\#$ est un marqueur qui n'apparait pas dans $L_1$ et $L_2$).

Greibach a montré que Lin est sans produit. De même, $\mathscr{C}(D_1^*)$, le cône rationnel engendré par $D_1^*$, est sans produit [22]. Par contre, ce n'est pas le cas de Rocl. Considérons, en effet, les langages $I_1^*$ et $F_1^*$ ensembles des facteurs respectivement gauches et droits des mots de $D_1'^*$. Il est, alors, facile de vérifier

PROPOSITION 10. *Le langage $I_1^* \# F_1^*$ est égal à $D_1'^* \sqcup\!\sqcup\, a^* \# b^*$ et donc appartient à* Rocl.

*Démonstration.* Comme tous cône rationnel est clos par shuffle avec un langage rationnel, $D_1'^* \sqcup\!\sqcup\, a^* \# b^* \in$ Rocl et il suffit d'établir l'égalité $I_1^* \# F_1^* = D_1'^* \sqcup\!\sqcup\, a^* \# b^*$. Il est clair que $I_1^* = D_1'^* \sqcup\!\sqcup\, a^*$ et que $F_1^* = D_1'^* \sqcup\!\sqcup\, b^*$. Donc $I_1^* \sqcup\!\sqcup\, F_1^* = (D_1'^* \sqcup\!\sqcup\, a^*) \# (D_1'^* \sqcup\!\sqcup\, b^*) \subseteq (D_1'^*)^2 \sqcup\!\sqcup\, a^* \# b^* = D_1'^* \sqcup\!\sqcup\, a^* \# b^*$. Réciproquement, si $w \in D_1'^* \sqcup\!\sqcup\, a^* \# b^*$, il existe $i, j \in \mathbb{N}$ et $u \in D_1'^*$ tels que $w \in u \sqcup\!\sqcup\, a^i \# b^j$. Donc, il existe $u_1, u_2$ tels que $u = u_1u_2$ et $w \in (u_1 \sqcup\!\sqcup\, a^i) \# (u_2 \sqcup\!\sqcup\, b^j)$ avec $u_1 \in I_1^*$ et $u_1 \sqcup\!\sqcup\, a^i \in I_1^* \sqcup\!\sqcup\, a^* = I_1^*$, $u_2 \in F_1^*$, et $u_2 \sqcup\!\sqcup\, b^j \in F_1^*$, ce qui entraine $w \in I_1^* \# F_1^*$. ∎

Montrons maintenant, que le langage $I_1^* \# F_1^*$ domine tout produit de deux langages non rationnels qui appartient à Rocl. Plus précisément, on obtient

PROPOSITION 11. *Soient $L_1 \subseteq X_1^*$ et $L_2 \subseteq X_2^*$ deux langages non rationnels avec $X_1 \cap X_2 = \varnothing$. Alors, le produit $L_1L_2$ est un langage à un compteur si et seulement si $L_1 \in \mathscr{C}(I_1^*)$ et $L_2 \in \mathscr{C}(F_1^*)$.*

*Démonstration.* Si $L_1L_2 \in$ Rocl, il existe un langage rationnel $R \subseteq Z^*$ et deux homomorphismes alphabétiques $h$ et $g$ tels que $L_1L_2 = g(h^{-1}(D_1'^*) \cap R)$. Posons $X = \{a, b\}$, $Z_i = \{z \in Z/g(z) \in X_i\}$, pour $i = 1, 2$. Alors $L_1L_2 = g(h^{-1}(D_1'^*) \cap R')$ avec $R' = R \cap Z_1^*Z_2^*$. On en déduit que $L_1L_2$ est une union finie de langages de la forme $L = g(h^{-1}(D_1'^*) \cap R_1R_2)$, avec $g(R_i) \subseteq X_i^*$ pour $i = 1, 2$. Notons, pour $i = 1, 2$, $\Pi_i$, la projection de $X_1 \cup X_2$ sur $X_i$ définie par $\Pi_i(x) = x$ si $x \in X_i$, $\varepsilon$ sinon. Nous allons montrer que $L = L' \cup L''$, avec $\Pi_2(L') \in$ Rat et $\Pi_1(L'') \in \mathscr{C}(I_1^*)$.

Pour $w \in X^*$, posons $d(w) = |w|_a - |w|_b$ et soit $k$ le nombre d'états d'un automate fini déterministe qui reconnait $h(R_2)$. Définissons les langages rationnels

$$\Delta = \{w \in X^*/\text{si } w' \in FD(w), \text{ alors } -k \leqslant d(w') \leqslant o\},$$

$$\Delta' = \{w \in \Delta/d(w) = -k\},$$

$$\Delta'' = X^*b\Delta'.$$

Alors, $L = L' \cup L''$ avec $L' = g(h^{-1}(D_1'^*) \cap R_1R_2')$, $L'' = g(h^{-1}(D_1'^*) \cap R_1R_2'')$, où $R_2' = R_2 \cap h^{-1}(\Delta)$ et $R_2'' = R_2 \cap h^{-1}(\Delta'')$. En effet, comme $R_2'$ et $R_2''$ sont inclus dans $R_2$, $L' \cup L'' \subseteq L$. Réciproquement, si $w \in L$, $w = g(u_1u_2)$ avec $u_1 \in R_1$, $u_2 \in R_2$, et $h(u_1u_2) \in D_1'^*$, ce qui entraine $h(u_2) \in F_1^*$. Comme $F_1^* \subseteq \Delta \cup \Delta''$, $u_2 \in R_2' \cup R_2''$ et $w \in L' \cup L''$.

Montrons que $\Pi_2(L') \in$ Rat. Pour $i \in \{0,..., k\}$, posons $\Delta_i = \{w \in \Delta/d(w) = -i\}$, $R_{2,i} = R_2 \cap h^{-1}(\Delta_i)$, et $L_i' = g(h^{-1}(D_1'^*) \cap R_1R_{2,i})$. Prenons $j \in \{0,..., k\}$ et distinguons deux cas.

(i) Il existe $u_1 \in R_1$ tel que $h(u_1) \in I_1^*$ et $d \circ h(u_1) = j$. Alors $h(u_1R_{2,j}) = h(u_1)$ $h(R_{2,j}) \subseteq h(u_1)\Delta_j \subseteq D_1'^*$. Donc $g(h^{-1}(D_1'^*) \cap u_1R_{2,j}) = g(u_1R_{2,j})$ et $\Pi_2(L_j') = g(R_{2,j}) \in$ Rat.

(ii) Sinon, montrons que $L_i' = \varnothing$. Supposons le contraire. Alors, il existe $u_1 \in R_1$, $u_2 \in R_{2,j}$ tels que $h(u_1u_2) \in D_1'^*$. On en déduit que $h(u_1) \in I_1^*$ et $d \circ h(u_1) = -d \circ h(u_2) = j$, d'où la contradiction, ce qui implique $\Pi_2(L_i') = \varnothing \in$ Rat.

Comme $L' = \bigcup_{j=0}^{k} L_j'$, $\Pi_2(L') = \bigcup_{j=0}^{k} \Pi_2(L_j')$, union finie de langages rationnels, est un langage rationnel.

Montrons, maintenant, que $\Pi_1(L'') \in \mathscr{C}(I_1^*)$. Comme $\Pi_1(L'') = g(h^{-1}(D_1'^*/h(R_2'')) \cap R_1)$, il suffit de démontrer que $D_1'^*/h(R_2'') \in \mathscr{C}(I_1^*)$. Pour cela, établissons l'égalité $D_1'^*/h(R_2'') = (D_1'^*/(b^{k!})^*)/h(R_2'')$. L'inclusion dans un sens se déduit immédiatement du fait que $D_1'^* \subseteq D_1'^*/(b^{k!})$. Pour démontrer l'inclusion inverse, prenons $w \in (D_1'^*/(b^{k!})^*)/h(R_2'')$. Alors, il existe $w_1 \in h(R_2'')$ et $j \in \mathbb{N}$ tels que $ww_1(b^{k!})^j \in D_1'^*$. Mais, $w_1 \in h(R_2) \cap \Delta'' = h(R_2'')$ et $w_1 = w_1'bw_1''$ avec $w_1'' \in \Delta'$. D'après le choix de $k$ et la définition de $\Delta'$, $w_1''$ se factorise en $xuy$ avec $w_1'bxu^*y \subseteq h(R_2)$ et $-k \leqslant d(u) = s < 0$. Comme $xuy \in \Delta' \subseteq F_1^*$ et que $d(u) < o$, on a $xu^+y \subseteq F_1^*$. Pour tout $z \in xu^+y$, $d(z) \leqslant d(xuy)$, donc $bxu^+y \subseteq \Delta''$ et $w_1'bxu^+y \subseteq h(R_2) \cap \Delta'' = h(R_2'')$. Posons $t = jk!/s \in \mathbb{N}$ et considérons le mot $z' = ww_1'bxuu^ty$. Etant donné que $d(u^t) = jk! = d(b^{k!j})$ et que $ww_1(b^{k!})^j = ww_1'bxuy(b^{k!})^j \in D_1'^*$, il est clair que

$d(z')=0$ et comme $xu^+y\subseteq F_1^*$, $z'\in D_1'^*$. Nous en déduisons que $w\in D_1'^*/h(R_2'')$ et donc $D_1'^*/h(R_2'')=(D_1'^*/(b^{kl})^*)/h(R_2'')$.

Il suffit maintenant, pour établir que $\Pi_1(L'')\in\mathscr{C}(I_1^*)$, de démontrer que $D_1'^*/(b^{kl})^*\in\mathscr{C}(I_1^*)$. Or, pour tout entier positif $i$, $D_1'^*/(b^i)^*=(D_1'^*/b^*)\cap B_i$, où $B_i=\{w\in X^*/d(w)=\lambda i$ avec $\lambda\in\mathbb{Z}\}$ est un langage rationnel. De l'égalité $D_1'^*/b^*=I_1^*$, on en déduit que $\forall i\in\mathbb{N}_+$, $D_1'^*/(b^i)^*=I_1^*\cap B_i\in\mathscr{C}(I_1^*)$, donc $\Pi_1(L'')\in\mathscr{C}(I_1^*)$.

Nous avons, donc, montré que $L_1L_2$ est égal à une finie de langage $M$ vérifiant soit $\Pi_2(M)\in$ Rat, soit $\Pi_1(M)\in\mathscr{C}(I_1^*)$. Comme Rat et $\mathscr{C}(I_1^*)$ sont clos par union, on en déduit que $L_1L_2=C\cup D$ avec $\Pi_2(C)\in$ Rat et $\Pi_1(D)\in\mathscr{C}(I_1^*)$. Pour $i\in\{1,2\}$, notons $C_i=\Pi_i(C)$ et $D_i=\Pi_i(D)$. Comme $C\subseteq C_1C_2$, $D\subseteq D_1D_2$, et que $\forall i\in\{1,2\}$, $C_i\subseteq L_i$ et $D_i\subseteq L_i$, on obtient $L_1L_2=C\cup D\subseteq C_1C_2\cup D_1D_2\subseteq L_1L_2$ et donc $L_1L_2=C_1C_2\cup D_1D_2$. Alors, nécessairement, soit $L_2=C_2\in$ Rat, soit $L_1=D_1\in\mathscr{C}(I_1^*)$.

Nous avons, donc, montré que $L_1L_2\in$ Rocl implique $L_1\in\mathscr{C}(I_1^*)$ ou $L_2\in$ Rat. De façon symétrique, on peut démontrer que $L_1L_2\in$ Rocl entraine $L_1\in$ Rat ou $L_2\in\mathscr{C}(F_1^*)$. Donc, si $L_1$ et $L_2$ sont des langages non rationnels, $L_1L_2\in$ Rocl implique bien $L_1\in\mathscr{C}(I_1^*)$ et $L_2\in\mathscr{C}(F_1^*)$. La réciproque découle immédiatement de la proposition 10. ▮

Considérons, de nouveau, les langages $L_>=\{a^nb^p/n>p\}$, $L_<=\{a^nb^p/p\geq n\}$, et $L_{\neq}=\{a^nb^p/n\neq p\}$.

D'après le corollaire 9, aucun des langages $L_>^2$, $L_<^2$, $L_{\neq}^2$ n'appartient à Rocl. La proposition précédente entraine que $L_>\notin\mathscr{C}(F_1^*)$, $L_<\notin\mathscr{C}(I_1^*)$, et $L_{\neq}\notin\mathscr{C}(I_1^*)\cap\mathscr{C}(F_1^*)$, ce qui implique, par symétrie, $L_{\neq}\notin\mathscr{C}(I_1^*)$ et $L_{\neq}\notin\mathscr{C}(F_1^*)$. Nous pouvons, donc, énoncer

COROLLAIRE 12. *Soit $L\subseteq Y^*$ avec $Y\cap\{a,b\}=\varnothing$. Alors pour $L'=L_<$ (resp. $L_{\neq}$) et $L''=L_>$ (resp. $L_{\neq}$) on a*

(i) $L'L\in$ Rocl *si et seulement si* $L\in$ Rat.

(ii) $LL''\in$ Rocl *si et seulement si* $L\in$ Rat.

En utilisant le fait que $\mathscr{C}(I_1^*)$ et $\mathscr{C}(F_1^*)$ sont sans produit [23], on peut montrer que Rocl ne peut pas contenir le produit de trois langages non rationnels.

COROLLAIRE 13. *Soient $L_1$, $L_2$, $L_3$ des langages définis sur des alphabets disjoints deux à deux. Alors, si $L_1L_2L_3\in$ Rocl, l'un des trois langages $L_1,L_2,L_3$ est rationnel.*

*Démonstration.* Si $L_1L_2L_3\in$ Rocl, d'après la proposition 11, $L_1L_2\in\mathscr{C}(I_1^*)$ ou $L_3\in$ Rat. Mais, comme $\mathscr{C}(I_1^*)$ est sans produit [23], $L_1L_2\in\mathscr{C}(I_1^*)$ implique, soit $L_1\in$ Rat, soit $L_2\in$ Rat, d'où le résultat. ▮

En particulier, Rocl ne peut contenir un langage non rationnel à la puissance trois.

COROLLAIRE 14. *Soient* $L \subseteq X^*$ *et* $\# \notin X$ *Alors* $(L\#)^3 \in \mathrm{Rocl}$ *si et seulement si* $L$ *est rationnel.*

En fait, le corollaire 9 nous incite à énoncer

*Conjecture* 2. $L\#L \in \mathrm{Rocl}$ entraine $L \in \mathrm{Rat}$.

Cette conjecture peut s'écrire autrement.

*Conjecture* 2 *bis*. $\mathscr{C}(I_1^*) \cap \mathscr{C}(F_1^*) = \mathrm{Rat}$.

Remarquons que si cette conjecture était vérifiée, on aurait le premier exemple de deux cônes rationnels contenant strictement Rat et dont l'intersection se réduit à Rat.

Montrons, maintenant, l'équivalence de ces deux énoncés.

Prenons $L \# L \in \mathrm{Rocl}$. D'après la proposition 11, $L \in \mathscr{C}(I_1^*) \cap \mathscr{C}(F_1^*)$ et la conjecture 2 bis entraine $L \in \mathrm{Rat}$.

Réciproquement, prenons $L \in \mathscr{C}(I_1^*) \cap \mathscr{C}(F_1^*)$. Alors, $L \# L \in \mathscr{C}(I_1^* \# F_1^*)$ et d'après la proposition 10, $L \# L \in \mathrm{Rocl}$ et la conjecture 2 entraine $L \in \mathrm{Rat}$.

Du corollaire 14 découle immédiatement

COROLLAIRE 15. Rocl *est sans étoile, c'est-à-dire que* $(L\#)^* \in \mathrm{Rocl}$ *implique* $L \in \mathrm{Rat}$.

De ce corollaire, nous allons pouvoir déduire un résultat pour la famille Ocl, la plus petite FAL contenant Rocl, famille étudiée par Greibach dans [13]. Pour cela, rappellons un résultat qui est une extension du lemme syntaxique de Greibach [13].

PROPOSITION 16 [26]. *Soient* $\mathscr{L}_1$ *un cône rationnel bifidèle,* $\mathscr{L}_2$ *un cône rationnel et les languages* $L_1 \subseteq X_1^*$, $L_2 \subseteq X_2^*$ *avec* $X_1 \cap X_2 = \varnothing$. *Alors* $L_1 \uparrow L_2 \in \mathscr{L}_1 \square \mathscr{L}_2$ *implique* $L_1 \in \mathscr{L}_1$ *ou* $(L_2\#)^* \in \mathscr{L}_2$.

Nour dirons qu'une famille de langage $\mathscr{L}$ est *sans substitution* si $L_1 \uparrow L_2 \in \mathscr{L}$, avec $L_1, L_2$ définis sur des alphabets disjoints, implique $L_1$ ou $L_2 \in \mathrm{Rat}$. La proposition précédente entraine alors immédiatement

COROLLAIRE 17. *Si* $\mathscr{L}$ *est un cône rationnel sans étoile,* $\mathrm{Rat} \square \mathscr{L}$, *la* FAL *engendrée par* $\mathscr{L}$ *est sans substitution.*

COROLLAIRE 18. *La famille* $\mathrm{Ocl} = \mathrm{Rat} \square \mathrm{Rocl}$ *est sans substitution.*

## IV. INDÉPENDANCE DE $D_1'^*$

Dans cette section, nous montrons d'abord, pour $D_1'^*$, un résultat qui est vérifié par les CIL-langages (cf. [21]) et qui peut être utilisé pour établir la non appartenance de $D_1'^*$ à certaines familles de langages. Nous allons établir que $D_1'^*$ est *indépendant de l'étoile*; c'est-à-dire que, pour tout langage $L$, $D_1'^* \in \mathscr{F}(L) =$

$\mathrm{Rat} \square \mathscr{C}(L)$ implique $D_1'^* \in \mathscr{C}(L)$. Remarquons que cette propriété n'est pas vérifiée par $D_1^*$, le langage de Dyck sur une lettre puisque $D_1^* \in \mathrm{Rat} \square \mathrm{Rocl}$ et $D_1^* \notin \mathrm{Rocl}$ [6]. Posons $X = \{a, b\}$ et $\forall w \in X^*$, $d(w) = |w|_a - |w|_b$, $\bar{d}(w) = \inf\{d(w')/w' \in FG(w)\}$. Nous utiliserons deux lemmes intermédiaires qui font intervenir les langages $A_k = \{x \in X^*/d(w) = 0, \bar{d}(w) \geqslant -k\} = \bigcup_{i=0}^{k} (D_1'^* \sqcup\!\sqcup\, b^i a^i)$.

Lemme 19 [21]. *Soit $\mathscr{L}$ un cône rationnel fermé par union. Si $D_1'^* \in \mathscr{F}(\mathscr{L}) = \mathrm{Rat} \square \mathscr{L}$, il existe $k \in \mathbb{N}$ et $L \in \mathscr{L}$ tels que $D_1'^* \subseteq L \subseteq A_k$.*

On ne sait pas, pour l'instant, si ce résultat reste vérifié pour un cône rationnel (non nécessairement clos par union). D'après le lemme suivant, cette question est équivalente à savoir si $D_1'^*$ est premier, c'est-à-dire, si $D_1'^* = L_1 \cup L_2$ implique $D_1'^* \in \mathscr{C}(\{L_1, L_2\})$. Montrons, maintenant, que tout langage compris entre $D_1'^*$ et $A_k$ domine rationnellement $D_1'^*$. Plus généralement, en posant $D_1' = aD_1'^*b$, et $A_k' = \{w \in X^*/-k \leqslant \bar{d}(w) \leqslant d(w) \leqslant k\}$, on peut énoncer

Lemme 20. *Soit $L$ un langage vérifiant $D_1' \subseteq L \subseteq A_k'$ avec $k \in \mathbb{N}$. Alors $D_1'^* \in \mathscr{C}^{\mathrm{bf}}(L)$.*

*Démonstration.* Considérons le langage rationnel $R = a\{a^{k+2}, b^{k+2}\}^* b$. Il est facile de vérifier que $A_k' \cap R = D_1' \cap R$.

En effet, par hypothèse, $D_1' \cap R \subseteq A_k' \cap R$. Réciproquement, prenons $w = aw'b \in A_k' \cap R$. Alors, $d(w) = \lambda(k+2)$ avec $\lambda \in \mathbb{Z}$ et comme $-k \leqslant d(w) \leqslant k$, $\lambda = 0$ et $d(w) = 0$. Supposons, maintenant, que $w$ n'appartienne pas à $D_1'$. Il existe, alors, une factorisation de $w'$ en $w_1' w_2'$ telle que $d(aw_1') = \bar{d}(w) \leqslant 0$. On en déduit que $w_1' = w_1'' b$ et $w_2' = aw_2''$. Donc $w_1' \in \{a^{k+2}, b^{k+2}\}^*$ et $\bar{d}(w) = d(aw_1') = 1 + \lambda(k+2) \leqslant 0$ avec $\lambda \in \mathbb{Z}$, ce qui entraine $\lambda \leqslant -1$ et $\bar{d}(w) < -k$, d'où la contradiction. Donc, $D_1' \cap R = L \cap R = A_k' \cap R$ appartient à $\mathscr{C}^{\mathrm{bf}}(L)$.

Prenons, maintenant, $Z = X \cup \{a', b'\}$ et définissons les homomorphismes $h$ et $g$ sur $Z$ par $h(a') = a$, $h(b') = b$, $h(a) = a^{k+2}$, $h(b) = b^{k+2}$, $g(a') = g(b') = \varepsilon$ $g(a) = a$ et $g(b) = b$. Alors, clairement, $D_1'^* = g(h^{-1}(D_1' \cap R) \cap a'X^*b') \in \mathscr{C}^{\mathrm{bf}}(L)$. ∎

Prenons un langage $L$ tel que $D_1'^* \in \mathscr{F}(L)$. Alors, comme $\mathscr{C}(L)$ est clos par union, les deux lemmes précédents entrainent immédiatement

Proposition 21. *$D_1'^*$ est indépendant de l'étoile; c'est-à-dire que, pour tout langage $L$, $D_1'^* \in \mathscr{F}(L)$ implique $D_1'^* \in \mathscr{C}(L)$.*

En se restreignant aux langages algébriques, on peut obtenir un résultat plus fort en montrant que $D_1'^*$ est indépendant de la substitution, relativement à Alg. Pour cela, introduisons les langages rationnels

$$B_j = \{w \in X^*/w' \in FD(w) \Rightarrow -j \leqslant d(w') \leqslant 0\}, \qquad j \in \mathbb{N}.$$

$$B_j' = \{w \in X^*/w' \in FG(w) \Rightarrow 0 \leqslant d(w') \leqslant j\}, \qquad j \in \mathbb{N}.$$

Etablissons d'abord

LEMME 22. *Soit $L$ un langage algébrique inclus dans $D_1'^*$. Alors,* $\forall i \in \mathbb{N}$, $\exists j \in \mathbb{N}$ *tel que:*

(i) *si* $w \in FG(L)$ *et* $d(w) \leqslant i$, *alors* $w \in LB_j'^{-1}$,

(ii) *si* $w \in FD(L)$ *et* $d(w) \geqslant -i$, *alors* $w \in B_j'^{-1}L$.

*Démonstration.* Montrons, d'abord, que $\forall i$, $\exists j$ tel que (i) soit vérifié. Soient $G = (X, N, P)$ une grammaire algébrique réduite, sous forme normale de Chomsky et $A \in N$ tels que $L(G, A) = L$. Prenons $i \in \mathbb{N}$, $w \in FG(L)$ avec $d(w) \leqslant i$. Il existe $z$ tel que $wz \in L$ et donc une dérivation gauche $u_0 = A \Rightarrow^g u_1 \cdots \Rightarrow^g u_n = wz$. Soit $t$ le plus petit entier tel que $w \in FG(u_t)$. Alors $u_t = wC_1 \cdots C_s$ avec $s \in \mathbb{N}$ et $\forall m \in \{1,..., s\}$, $C_m \in N$. Comme $G$ est réduite, $\forall C \in N$, il existe $x, y \in X^*$ tels que $A \overset{*}{\Rightarrow} xCy$. Donc $xL(G, C)y \subseteq L \subseteq D_1'^*$ et il existe $k \in \mathbb{N}$ tel que $\forall C \in N$, $L(G, C) \subseteq A_k'$. Il existe, aussi, $l \in \mathbb{N}$ tel que $\forall C \in N$, $\exists x \in L(G, C)$ avec $|x| \leqslant l$. Pour tout $m \in \{1,..., s\}$, prenons $x_m \in L(G, C_m)$ avec $|x_m| \leqslant l$ et posons $x = x_1 \cdots x_s$. Alors, $wx \in L$. Montrons que $x \in B_j$ avec $j = i + 2k + l$. Prenons $x' \in FD(x)$. Comme $wx \in D_1'^*$, $x' \in F_1^*$ et $d(x') \leqslant 0$. Si $w = \varepsilon$, $t = 0$ et $s = 1$. Donc $x = x_1$ et $d(x') \geqslant -|x'| \geqslant -|x_1| \geqslant -l \geqslant -j$. Si $w \neq \varepsilon$, $t > o$. Il existe $m \in \{1,..., s\}$ tel que $x' = x''x_{m+1} \cdots x_s$ avec $|x''| \leqslant l$. Si $m = s$, $x' = x''$ et $d(x') \geqslant -j$. Sinon, prenons $\alpha$ le plus petit entier positif tel que $\forall \beta \in \{\alpha,..., t\}$, $C_{m+1} \cdots C_s \in FD(u_\beta)$. Comme $G$ est sous forme normale de Chomsky et que $w \notin FG(u_{\alpha-1})$, il existe $w_1 \in FG(w)$, $B$, $C \in N$ tels que $u_{\alpha-1} = w_1BC_{m+2} \cdots C_s \overset{g}{\Rightarrow} w_1CC_{m+1} \cdots C_s$, et $C \overset{*}{\Rightarrow} w_2C_1 \cdots C_m \overset{*}{\Rightarrow} w_2x_1 \cdots x_m = w'$ avec $w = w_1w_2$. Alors, $d(w') \leqslant k$ et $d(w_2) \geqslant d(w') \geqslant -k$. Comme $d(w) = d(w_1) + d(w_2) \leqslant i$, $d(w_1) \leqslant i + k$ et $d(w_1w') \leqslant i + 2k$. De plus, $w_1w'x_{m+1} \cdots x_s \in L \subseteq D_1'^*$, donc $d(w_1w'x_{m+1} \cdots x_s) = 0$, ce qui implique $d(x_{m+1} \cdots x_s) = 0$, ce qui implique $d(x_{m+1} \cdots x_s) \geqslant -i - 2k$ et $d(x') \geqslant d(x'') - i - 2k \geqslant -l - i - 2k = -j$. Nous en déduisons immédiatement que $x \in B_j$ et $w \in LB_j^{-1}$.

En raisonnant de la même façon, on montrerait que $\forall i \in \mathbb{N}$, $\exists j = i + 2k + l$ tel que $w \in FD(L)$ et $d(w) \geqslant -i$ implique $w \in B_j'^{-1}L$, d'où le résultat. ∎

Nous pouvons, alors, énoncer

PROPOSITION 23. *Soit $L$ un langage algébrique vérifiant $L \subseteq D_1'^* \subseteq F(L)$. Alors,* $D_1'^* \in \mathscr{C}(L)$.

*Démonstration.* Soient $G = (X, N, P)$ une grammaire algébrique réduite, sous forme normale de Chomsky et $A \in N$ tels que $L(G, A) = L$. Il existe $k \in \mathbb{N}_+$ tel que $\forall B \in N$, $\exists x, y \in X^*$ vérifiant $A \overset{*}{\Rightarrow} xBy$ et $|x|, |y| < k$. Il est, alors, clair que pour tout $B \in N$, $L(G, B) \subseteq A_{k-1}'$. Prenons $w \in D_1'^*$ et posons $w' = a^kwb^k$. Comme $w' \in D_1'^* \subseteq F(L)$, il existe $z_1, z_2 \in X^*$ tels que $A \overset{*}{\Rightarrow} z_1w'z_2$. Alors, il existe $B_1, B_2 \in N$ tels que $B_1 \overset{*}{\Rightarrow} z_1'w_1'$, $B_2 \overset{*}{\Rightarrow} w_2'z_2'$ avec $w' = w_1'w_2'$, $z_1' \in FD(z_1)$ et $z_2' \in FG(z_2)$. Supposons que $w_1' = a^kw_1$ et $w_2' = w_2b^k$ avec $w = w_1w_2$. Alors, $w_2 \in F_1^*$, donc $d(w_2) \leqslant 0$ et $\bar{d}(w_2'z_2') \leqslant d(w_2') \leqslant -k$, ce qui entraine $w_2'z_2' \notin A_{k-1}'$, d'où la contradiction, puisque $w_2'z_2' \in L(G, B_2) \subseteq A_{k-1}'$. Nous en déduisons que, soit $w_1' = a^kwb^s$ avec $s \leqslant k$, soit $w_2' = a^twb^k$ avec $t \leqslant k$. Soient $i = 2k - 1$, $j$, l'entier correspondant du lemme 22, $R = \{xa^s \in X^*/|x| < k,\ 0 \leqslant s \leqslant k\}$, et $R' = \{b^sy \in X^*/$

$|y|<k,\ 0\leqslant s\leqslant k\}$. Alors, le langage $L'=R^{-1}LB_j^{-1}\cup B_j'L^{-1}R'^{-1}\in\mathscr{C}(L)$ puisque $R$, $R'$, $B_j$, $B_j'$ sont des langages rationnels et que tout cône rationnel est clos par quotient rationnel à gauche et à droite. Montrons que $w\in L'$. Si $w_1'=a^kwb^s$, il existe $x_1, y_1\in X^*$ tels que $|x_1|, |y_1|<k$ et $A\overset{*}{\Rightarrow}x_1B_1y_1\overset{*}{\Rightarrow}x_1z_1'a^kwb^sy_1$. Alors, $wb^sy_1\in FD(L)$ avec $d(wb^sy_1)=d(b^sy_1)\geqslant -i$ et le lemme précédent entraine que $wb^sy_1\in B_j'^{-1}L$. Comme $b^sy_1\in R'$, $w\in B_j'^{-1}LR'^{-1}\subseteq L'$. On montrerait de la même façon que si $w_2'=a^twb^k$ avec $t\leqslant k$, $w\in R^{-1}LB_j^{-1}\subseteq L$. Nous en déduisons, donc, que $w\in L'$ et que $D_1'^*\subseteq L'$.

Montrons, maintenant, que $L'$ est inclus dans $A_n'$ avec $n=\sup(i,j)$. Prenons, en effet, $w\in R^{-1}LB_j^{-1}$. Alors, il existe $x\in R$, $y\in B_j$ tels que $xwy\in L\subseteq D_1'^*$ avec $0\leqslant d(x)\leqslant i$ et $-j\leqslant d(y)\leqslant 0$, ce qui entraine bien $d(w)\geqslant -i>-n$ et $d(w)=-d(xy)\in\{-i,..., j\}\subseteq\{-n,..., +n\}$. De même $B_j'^{-1}LR'\subseteq A_n'$, et $L'\subseteq A_n'$.

Nous avons, donc, $L'\in\mathscr{C}(L)$ avec $D_1'^*\subseteq L'\subseteq A_n'$. Le lemme 20 entraine, alors, $D_1'^*\in\mathscr{C}^{\mathrm{bf}}(L')\subseteq\mathscr{C}(L)$. ∎

Nous pouvons en déduire

COROLLAIRE 24. *Soit $\mathscr{L}$ une famille de langages algébriques. Si $D_1'^*\in\mathscr{C}_\cup(\mathscr{L})$, alors $D_1'^*\in\mathscr{C}(\mathscr{L})$.*

*Démonstration.* Clairement, il suffit d'établir que $D_1'^*=L_1\cup L_2$ avec $L_1, L_2\in$ Alg implique $D_1'^*\in\mathscr{C}(\{L_1, L_2\})$. Si $D_1'^*\subseteq F(L_1)$, la proposition précédente entraine $D_1'^*\subseteq\mathscr{C}(L_1)$. Dans le cas contraire, il existe $w\in D_1'^*\backslash F(L_1)$ et $\forall w'\in D_1'^*$, $w'w\in D_1'^*\backslash L_1\subseteq L_2$, ce qui entraine $w'\in L_2w^{-1}$ et $D_1'^*\subseteq L_2w^{-1}$. Il est clair que $L_2w^{-1}\subseteq D_1'^*w^{-1}=D_1'^*$. Donc $D_1'^*=L_2w^{-1}\in\mathscr{C}(L_2)$. ∎

Pour obtenir le résultat annoncé, c'est-à-dire que $D_1'^*$ est indépendant de la substitution, relativement à Alg, nous utiliserons un dernier lemme.

LEMME 25. *Soit $\mathscr{L}_1$, $\mathscr{L}_2$ des cônes rationnels. Si $D_1'^*\in\mathscr{L}_1\,\square\,\mathscr{L}_2\backslash\mathscr{L}_1$, alors il existe $L\in\mathscr{L}_2$ tel que $L\subseteq D_1'^*\subseteq F(L)$.*

*Démonstration.* Il existe $L\in\mathscr{L}_1$, $L\subseteq Y^*$ avec $Y=\{y_1,..., y_n\}$ et $s$ une $\mathscr{L}_2$-substitution tels que $D_1'^*=s(L)$. On peut supposer que $\forall i\in\{1,..., n\}$, $L_i=s(y_i)\neq\varnothing$ et qu'il existe $\alpha_i, \beta_i\in Y^*$ tels que $\alpha_iy_i\beta_i\in L$. Alors, $\forall i\in\{1,..., n\}$, il existe $\alpha_i'\in s(\alpha_i)$, $\beta_i'\in s(\beta_i)$ tels que $\alpha_i'L_i\beta_i'\subseteq s(L)=D_1'^*$. Supposons que, $\forall i\in\{1,..., n\}$, il existe $w_i\in D_1'^*\backslash F(L_i)$. Alors, $w=w_1\cdots w_n$ est tel que $\forall i\in\{1,..., n\}$, $w\in D_1'^*\backslash F(L_i)$. Posons $k=2|w|$ et considérons la substitution finie $s'$ définie sur $Y$ par $s'(y_i)=s(y_i)\cap(X^*\backslash X^kX^*)\subseteq L_i$. Alors, $L'=s'(L)\in\mathscr{L}_1$ et vérifie $s'(L)\subseteq s(L)=D_1'^*$. Prenons, maintenant, $w'\in D_1'^*$. Alors, $w'\uparrow w\in D_1'^*=s(L)$ et il existe $u_1,..., u_t\in Y$ tels que $u=u_1\cdots u_t\in L$ et $w'\uparrow w\in s(u)$. Donc $w'\uparrow w=v_1\cdots v_t$ avec $\forall i\in\{1,..., t\}$ $v_i\in s(u_i)$. Pour tout $i\in\{1,..., t\}$, $w\notin F(s(u_i))$, donc $|v_i|>k$ et $v_i\in s'(u_i)$. Nous en déduisons que $L'$ vérifie $L'\subseteq D_1'^*$ et $D_1'^*\uparrow w\subseteq L'$. Posons $Z=X\cup\bar{X}$ avec $\bar{X}=\{\bar{a}, \bar{b}\}$ et définissons sur $Z$ les homomorphismes $h$ et $g$ par $h(x)=h(\bar{x})=x$, $g(x)=\varepsilon$, $g(\bar{x})=x$ pour $x\in X$. Alors, il est clair que $g(h^{-1}(L')\cap(\bar{X}w)^*)$ est égal à $D_1'^*$. Donc, $D_1'^*\in\mathscr{C}(L')\subseteq\mathscr{L}_1$, d'où la contradiction avec l'hypothèse. Il existe donc,

$L_i \in \mathscr{L}_2$ tel que $D_1'^* \subseteq F(L_i)$. Alors, le langage $L = \alpha_i' L_i \beta_i' \in \mathscr{C}(L_i) \subseteq \mathscr{L}_2$ et vérifie $L \subseteq D_1'^* \subseteq F(L_i) \subseteq F(L)$. ▮

En particulier, si $\mathscr{L}_2 \subseteq \mathrm{Alg}$, le lemme précédent et la proposition 23 entrainent immédiatement

LEMME 26. *Soient $\mathscr{L}_1$, $\mathscr{L}_2$ des cônes rationnels. Si $D_1'^* \in \mathscr{L}_1 \square \mathscr{L}_2$ avec $\mathscr{L}_2 \subseteq \mathrm{Alg}$, alors $D_1'^* \in \mathscr{L}_1 \cup \mathscr{L}_2$.*

Prenons maintenant un cône rationnel $\mathscr{L} \subseteq \mathrm{Alg}$ et supposons que $D_1'^* \in \mathscr{F}_\sigma(\mathscr{L})$, la plus petite FAL close par substitution contenant $\mathscr{L}$. Comme $\mathscr{F}_\sigma(\mathscr{L}) = \bigcup_{i \geqslant 1} \mathscr{L}_i$ avec $\mathscr{L}_1 = \mathscr{L}$ et $\forall t \geqslant 1$, $\mathscr{L}_{t+1} = \mathscr{L}_t \square \mathscr{L}$, le lemme précédent permet d'obtenir par induction

PROPOSITION 27. *Soit $\mathscr{L}$ un cône rationnel algébrique. Alors $D_1'^* \in \mathscr{F}_\sigma(\mathscr{L})$ implique $D_1'^* \in \mathscr{L}$.*

Une conséquence immédiate de cette proposition est que, puisque $D_1'^*$ n'appartient pas à Lin, la famille des langages linéaires, $D_1'^*$ n'appartient pas non plus à $\mathscr{F}_\sigma(\mathrm{Lin})$, la famille des langages quasirationnels (cf. [30, 31, 33]). D'autre part, comme $D_1^* = \{w \in \{a, b\}^* / |w|_a = |w|_b\}$ appartient à Com, la famille des langages commutatifs et que $D_1'^* \notin \mathscr{C}(\mathrm{Com})$ [21], nous obtenons

COROLLAIRE 28. *$D_1'^*$ n'appartient pas à $\mathscr{F}_\sigma(D_1^*)$.*

Montrons, maintenant, que dans la proposition 27, on ne peut pas remplacer la substitution par l'intersection, même si on se restreint aux homomorphismes non effaçant. Considérons le langage $C_1 = \{a^n b^n / n \geqslant 0\}$.

PROPOSITION 29. *Le langage $D_1'^*$ est égal à $C_1^* \sqcup\!\sqcup\, C_1^*$, le shuffle de $C_1^*$ par lui-même.*

*Démonstration.* Comme $C_1^*$ est inclus dans $D_1'^*$, $C_1^* \sqcup\!\sqcup\, C_1^* \subseteq D_1'^* \sqcup\!\sqcup\, D_1'^* = D_1'^*$. Pour démontrer l'inclusion inverse, établissons la propriété

$(*)$ Si $a^i w \in D_1'^*$ avec $i \in \mathbb{N}$, il existe $w_1, w_2$ tels que $w \in w_1 \sqcup\!\sqcup\, w_2$, $a^i w_1 \in C_1^*$, et $w_2 \in C_1^*$.

Raisonnons par récurrence sur la longueur de $w$.
Si $|w| = i$, alors $w = b^i$ et on prend $w_1 = b^i$, $w_2 = \varepsilon$.
Sinon, $|w| > i$ et $w \neq \varepsilon$. Distinguons deux cas.

(i) Si $i = 0$, $w \in D_1'^* \backslash \{\varepsilon\}$, donc $w = aw'$ avec $|w'| < |w|$. D'après l'hypothèse de récurrence, il existe $w_1'$, $w_2'$ tels que $w' \in w_1' \sqcup\!\sqcup\, w_2'$ avec $aw_1', w_2' \in C_1^*$. Alors, $w = aw' \in aw_1' \sqcup\!\sqcup\, w_2'$, et il suffit de prendre $w_1 = aw_1'$, $w_2 = w_2'$.

(ii) Si $i > 0$, $w = w'w''$ avec $|w'|_b = i$ et $|w''| < |w|$. Posons $p = |w'|_a$. Alors, $w' \subseteq b^i \sqcup\!\sqcup\, a^p$ et $a^p w'' \in D_1'^*$. D'après l'hypothèse de récurrence, il existe $w_1''$, $w_2''$ tels que $w'' \in w_1'' \sqcup\!\sqcup\, w_2''$ avec $a^p w_1'', w_2'' \in C_1^*$. Alors, $w = w'w'' \in (b^i \sqcup\!\sqcup\, a^p)(w_1'' \sqcup\!\sqcup\, w_2'') \subseteq$

$b^i w_2'' \sqcup\!\sqcup\, a^p w_1''$, d'où le résultat en posant $w_1 = b^i w_2''$ et $w_2 = a^p w_1''$, puisque $a^i w_1 = a^i b^i w_2'' \in C_1^*$.

Prenons, maintenant, $w \in D_1'^*$. D'après la propriété (∗), il existe $w_1, w_2 \in C_1^*$ tels que $w \in w_1 \sqcup\!\sqcup\, w_2$, ce qui entraine $D_1'^* \subseteq C_1^* \sqcup\!\sqcup\, C_1^*$. ∎

Comme tout cône rationnel (fidèle) clos par intersection est clos par shuffle, $D_1'^* \in \mathscr{F}^{\mathrm{f}}_{\cap}(C_1)$, alors que $D_1'^* \notin \mathscr{C}^{\mathrm{f}}(C_1) = \mathscr{C}(C_1)$. Enfin, la proposition précédente permet de préciser les rapports entre certaines familles de langages fermées par intersection, construites à partir des langages $C_1, D_1^*, D_1'^*$, $\mathrm{Copy} = \{ww/w \in X^*\}$ et $\mathrm{Sym} = \{ww^R/w \in X^*\}$.

PROPOSITION 30. $\mathscr{C}^{\mathrm{f}}_{\cap}(C_1) = \mathscr{C}_{\cap}(C_1) = \mathscr{C}^{\mathrm{f}}_{\cap}(D_1^*) \subsetneq \mathscr{C}^{\mathrm{f}}_{\cap}(D_1'^*) \subseteq \mathscr{F}^{\mathrm{f}}_{\cap}(C_1^*) = \mathscr{F}^{\mathrm{f}}_{\cap}(D_1^*) = \mathscr{F}^{\mathrm{f}}_{\cap}(D_1'^*) \subsetneq \mathscr{C}^{\mathrm{f}}_{\cap}(\mathrm{Copy}) = \mathscr{F}^{\mathrm{f}}_{\cap}(\mathrm{Copy}) \subseteq \mathscr{C}^{\mathrm{f}}_{\cap}(\mathrm{Sym}) \subseteq \mathscr{F}^{\mathrm{f}}_{\cap}(\mathrm{Sym})$.

*Démonstration.* Les égalités

$$\mathscr{C}^{\mathrm{f}}_{\cap}(C_1) = \mathscr{C}_{\cap}(C_1) = \mathscr{C}^{\mathrm{f}}_{\cap}(D_1^*) = \mathscr{C}_{\cap}(D_1^*) \tag{1}$$

ont été démontrées dans [17, 21, 24]. Comme $C_1 = D_1'^* \cap a^*b^*$, $\mathscr{C}^{\mathrm{f}}_{\cap}(C_1) \subseteq \mathscr{C}^{\mathrm{f}}_{\cap}(D_1'^*)$. L'inclusion stricte découle du fait que $D_1'^* \notin \mathscr{C}_{\cap}(\mathrm{Bor})$ [21], où Bor désigne la famille des langages bornés. L'inclusion

$$\mathscr{C}^{\mathrm{f}}_{\cap}(D_1'^*) \subseteq \mathscr{F}^{\mathrm{f}}_{\cap}(C_1) \tag{2}$$

est une conséquence de la proposition précédente. Les relations (1) et (2) entrainent immédiatement les égalités $\mathscr{F}^{\mathrm{f}}_{\cap}(C_1) = \mathscr{F}^{\mathrm{f}}_{\cap}(D_1^*) = \mathscr{F}^{\mathrm{f}}_{\cap}(D_1'^*)$. L'égalité entre $\mathscr{C}^{\mathrm{f}}_{\cap}(\mathrm{Copy})$ et $\mathscr{F}^{\mathrm{f}}_{\cap}(\mathrm{Copy})$ est établie dans [8], ainsi que l'inclusion de $\mathscr{F}^{\mathrm{f}}_{\cap}(\mathrm{Copy})$ dans $\mathscr{C}^{\mathrm{f}}_{\cap}(\mathrm{Sym})$. Enfin, le fait que Copy n'appartient pas à $\mathscr{F}^{\mathrm{f}}_{\cap}(C_1)$ énoncé dans [8], est une conséquence d'un résultat plus général de Turakainen qui démontre dans [32] que $\mathrm{Copy} \notin \mathscr{F}^{\mathrm{f}}_{\cap}(\mathrm{Bor})$. ∎

Notons que, pour chacune des inclusions qui apparaissent dans l'énoncé de la proposition précédente, la strictitude reste un problème ouvert (cf. [8, 9, 19, 20]).

## RÉFÉRENCES


1. J. M. AUTEBERT, Non-principalité du cylindre des langages à compteur, *Math. Systems Theory* **11** (1977), 157–167.
2. J. BEAUQUIER, "Contribution à l'étude de la complexité *structurelle des langages algébriques*," Thèse Sc. Math., Paris, 1977.
3. J. BEAUQUIER, Independence of linear and one-counter generators, *in* "Fundamentals of Computation Theory," pp. 45–51, Akademie-Verlag, Berlin, 1979.
4. J. BERSTEL, "Transductions and context-free languages," Teubner Verlag, Stuttgart, 1979.
5. J. BERSTEL ET L. BOASSON, Une suite décroissante de cônes rationnels, *in* "Automata Languages and Programming" (Loeckx, Ed.), Lecture Notes in Computer Science No. 14, pp. 383–397, 1974.
6. L. BOASSON, Two iteration theorems for some families of langages, *J. Comput. System Sci.* **7** (1973), 583–596.

7. L. Boasson et M. Nivat, Sur diverses familles de langages fermées par transductions rationnelles, *Acta Inform.* **2** (1973), 180–188.
8. R. Book, S. Greibach, et C. Wrathall, Reset machines, *J. Comput. System Sci.* **19** (1979), 256–276.
9. R. Book et M. Nivat, Linear languages and the intersection closures of classes of languages, *SIAM J. Comput.* **7** (1978), 167–177.
10. S. Eilenberg, "Automata Languages and Machines," Vol. A, Academic Press, New York, 1974.
11. S. Ginsburg, J. Goldstine, et S. Greibach, Uniformly erasable AFL, *J. Comput. System Sci.* **10** (1975), 165–182.
12. S. Ginsburg, J. Goldstine, et S. Greibach, Some uniformly erasable families of languages, *Theoret. Comput. Sci.* **2** (1976), 29–44.
13. S. Greibach, An infinite hierarchy of context-free languages, *J. Assoc. Comput. Mach.* **16** (1969), 91–106.
14. S. Greibach, The hardest context-free language, *SIAM J. Comput.* **2** (1973), 304–310.
15. S. Greibach, One-counter languages and the IRS condition, *J. Comput. System Sci.* **10** (1975), 237–247.
16. S. Greibach, A note on the recognition of one-counter languages, *RAIRO Inform. Théor.* **R2** (1975), 5–12.
17. S. Greibach, Remarks on blind and partially blind one-way multicounter machines, *Theoret. Comput. Sci.* **7** (1978), 311–324.
18. S. Greibach, One-counter languages and the chevron operation, *RAIRO Inform. Théor.* **13** (1979), 189–194.
19. M. Jantzen, On the hierarchy of Petri Net languages, *RAIRO Inform. Théor.* **13** (1979), 19–30.
20. M. Jantzen, "The Power of Synchronizing Operations on Strings," Univ. California, Santa Barbara, 1980.
21. M. Latteux, Cônes rationnels commutativement clos, *RAIRO Inform. Théor.* **11** (1977), 29–51.
22. M. Latteux, Produit dans le cône rationnel engendré par $D_1^*$, *Theoret. Comput. Sci.* **5** (1977), 129–134.
23. M. Latteux, "Langages Commutatifs," Thèse Sc. Math., Lille, France, 1978.
24. M. Latteux, Cônes rationnels commutatifs, *J. Comput. System Sci.* **18** (1979), 307–333.
25. M. Latteux, Sur les générateurs algébriques et linéaires, *Acta Inform.* **13** (1980), 347–363.
26. M. Latteux, A propos du lemme de substitution, *Theoret. Comput. Sci.* (1980), à paraître.
27. M. Latteux et J. Leguy, Une propriété de la famille GRE, *in* "Fundamentals of Computation Theory," pp. 255–261, Akademie-Verlag, Berlin, 1979.
28. J. Leguy, Transductions rationnelles décroissantes, *RAIRO Inform. Théor.* (1979), à paraître.
29. J. Leguy, "Transduction rationnelle décroissante et substitution," Thèse, Lille, France, 1980.
30. M. Nivat, "Transductions des langages de Chomsky," Thèse Sc. Math., Grenoble, France, 1967.
31. A. Salomaa, On the index of a context-free grammar and language, *Inform. and Control* **14** (1969), 474–477.
32. P. Turakainen, On some bounded semi AFLs and AFLs, *Inform. Sci.* (1980), à paraître.
33. M. K. Yntema, Inclusion relations among families of context-free languages, *Inform. and Control* **10** (1967), 572–597.