HAL open science

Langages bornés arithmétiques

Philippe Butzbach

To cite this version:

Philippe Butzbach. Langages bornés arithmétiques. Modélisation et simulation. Université Joseph-Fourier - Grenoble I, 1968. Français. NNT : . tel-00280985

HAL Id: tel-00280985

https://theses.hal.science/tel-00280985v1

Submitted on 20 May 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

N° d'ordre

THESE

présentée à la Faculté des Sciences
de l'Université de Grenoble

pour obtenir
le grade de Docteur de Troisième Cycle
"MATHEMATIQUES APPLIQUEES"

par

Philippe BUTZBACH

LANGAGES BORNES ARITHMETIQUES

Thèse soutenue le 25 janvier 1968, devant la Commission d'examen :

Monsieur	J. KUNTZMANN	Président
Madame	F. BERTRANDIAS	Examinateur
Messieurs	B. VAUQUOIS	Examinateur
	M. NIVAT	Examinateur

N° d'ordre

THESE

présentée à la Faculté des Sciences
de l'Université de Grenoble

pour obtenir
le grade de Docteur de Troisième Cycle
"MATHEMATIQUES APPLIQUEES"

par

Philippe BUTZBACH

LANGAGES BORNES ARITHMETIQUES

Thèse soutenue le 25 janvier 1968, devant la Commission d'examen :

Monsieur	J. KUNTZMANN	Président
Madame	F. BERTRANDIAS	Examinateur
Messieurs	B. VAUQUOIS	Examinateur
	M. NIVAT	Examinateur

F A C U L T E D E S S C I E N C E S

LISTE DES PROFESSEURS

DOYENS HONORAIRES :

M. MORET

M. WEIL

DOYEN :

M. BONNIER E.

PROFESSEURS TITULAIRES :

MM.	NEEL Louis	Chaire de Physique Expérimentale
	HEILMANN René	Chaire de Chimie
	KRAVTCHENKO Julien	Chaire de Mécanique Rationnelle
	CHABAUTY Claude	Chaire de calcul différentiel et intégral
	BENOIT Jean	Chaire de Radioélectricité
	CHENE Marcel	Chaire de Chimie Papetière
	WEIL Louis	Chaire de Thermodynamique
	FELICI Noël	Chaire d'Electrostatique
	KUNTZMANN Jean	Chaire de Mathématiques Appliquées
	BARBIER Reynold	Chaire de Géologie Appliquée
	SANTON Lucien	Chaire de Mécanique des Fluides
	OZENDA Paul	Chaire de Botanique
	FALLOT Maurice	Chaire de Physique Industrielle
	KOSZUL Jean-Louis	Chaire de Mathématiques M.P.C.
	GALVANI O.	Mathématiques
	MOUSSA André	Chaire de Chimie Nucléaire
	TRAYNARD Philippe	Chaire de Chimie Générale

	SOUTIF Michel	Chaire de Physique Générale
	CRAYA Antoine	Chaire d'Hydrodynamique
	REULOS R.	Théorie des Champs
	BESSON Jean	Chaire de Chimie
	AYANT Yves	Physique Approfondie
	GALLISSOT	Mathématiques
Melle	LUTZ Elisabeth	Mathématiques
MM.	BLAMBERT Maurice	Chaire de Mathématiques
	BOUCHEZ Robert	Physique Nucléaire
	LLIBOUTRY Louis	Géophysique
	MICHEL Robert	Chaire de Minéralogie et Pétrographie
	BONNIER Etienne	Chaire d'Electrochimie et d'Electrométallurgie
	DESSAUX Georges	Chaire de Physiologie Animale
	PILLET E.	Chaire de Physique Industrielle et Electrotechnique
	YOCCOZ Jean	Chaire de Physique Nucléaire Théorique
	DEBELMAS Jacques	Chaire de Géologie Générale
	GERBER R.	Mathématiques
	PAUTHENET R.	Electrotechnique
	VAUQUOIS B.	Chaire de Calcul Electronique
	BARJON R.	Physique Nucléaire
	BARBIER Jean-Claude	Chaire de Physique
	SILBER R.	Mécanique des Fluides
	BUYLE-BODIN Maurice	Chaire d'Electronique
	DREYFUS B.	Thermodynamique
	KLEIN J.	Mathématiques
	VAILLANT F.	Zoologie et Hydrobiologie
	ARNOUD Paul	Chaire de Chimie M.P.C.
	SENGEL P.	Chaire de Zoologie
	BARNOUD F.	Chaire de Biosynthèse de la Cellulose
	BRISSONNEAU P.	Physique
	GAGNAIRE Didier	Chaire de Chimie Physique

Mme	KOFLER L.	Botanique
MM.	DEGRANGE Charles	Zoologie
	PEBAY-PEROULA J.C.	Physique
	RASSAT A.	Chaire de Chimie Systématique

PROFESSEURS SANS CHAIRE :

MM.	GIDON P.	Géologie et Minéralogie
	GIRAUD P.	Géologie
	PERRET R.	Servomécanismes
Mme	BARBIER M.J.	Electrochimie
Mme	SOUTIF J.	Physique
MM.	COHEN J.	Electrotechnique
	DEPASSEL R.	Mécanique des Fluides
	GASTINEL N.	Mathématiques Appliquées
	ANGLES-d'AURIAC P.	Mécanique des Fluides
	DUCROS P.	Minéralogie et Cristallographie
	GLENAT R.	Chimie
	LACAZE A.	Thermodynamique
	BARRA J.	Mathématiques Appliquées
	COUMES A.	Electronique
	PERRIAUX J.	Géologie et Minéralogie
	ROBERT A.	Chimie Papetière
	BIAREZ J.P.	Mécanique Physique
	BONNET G.	Electronique
	CAUQUIS G.	Chimie Générale
	BONNETAIN L.	Chimie Minérale
	DEPOMMIER P.	Etude Nucléaire et Génie Atomique
	HACQUES Gérard	Calcul Numérique
	POLOUJADOFF M.	Electrotechnique

MAITRES DE CONFERENCES :

MM.	DODU J.	Mécanique des Fluides
	LANCIA Roland	Physique Automatique
Mme	KAHANE J.	Physique
MM.	DEPORTES C.	Chimie
Mme	BOUCHE L.	Mathématiques
MM.	SARROT-RAYNAUD J.	Géologie Propédeutique
Mme	BONNIER M.J.	Chimie
MM.	KAHANE A.	Physique Générale
	DOLIQUE J.M.	Electronique
	BRIERE G.	Physique M.P.C.
	DESPRE P.	Chimie S.P.C.N.
	LAJZEROWICZ J.	Physique M.P.C.
	VALENTIN P.	Physique M.P.C.
	BERTRANDIAS J.P.	Mathématiques Appliquées T.M.P.
	LAURENT P.	Mathématiques Appliquées T.M.P.
	CAUBET J.P.	Mathématiques Pures
	PAYAN J.J.	Mathématiques
Mme	BERTRANDIAS F.	Mathématiques Pures M.P.C.
MM.	LONGEQUEUE J.P.	Physique
	NIVAT M.	Mathématiques Appliquées
	SOHM J.C.	Electrochimie
	ZADWORNY F.	Electronique
	DURAND F.	Chimie Physique
	CARLIER G.	Biologie Végétale
	AUBERT G.	Physique M.P.C.
	DELPUECH J.J.	Chimie Organique
	PFISTER J.C.	Physique C.P.E.M.
	CHIBON P.	Biologie Animale
	IDELMAN S.	Physiologie Animale
	BLOCH D.	Electrotechnique
	BRUGEL L.	I.U.T.
	SIBILLE R.	I.U.T.

Le présent travail a été entièrement effectué à l'Institut de Mathématiques Appliquées de Grenoble sous la direction bienveillante et attentionnée de Monsieur NIVAT

Je tiens à remercier Monsieur le Professeur KUNTZMANN qui me fait l'honneur de présider le Jury de cette thèse,

Monsieur le Professeur VAUQUOIS et Madame BERTRIANDIAS Maître de Conférences qui ont bien voulu être membre du Jury,

Monsieur NIVAT, Maître de Conférences, membre du Jury et sans qui ce travail n'aurait jamais vu le jour.

Je remercie tous ceux qui ont participé à la réalisation matérielle de cette thèse.

LANGAGES BORNES ARITHMETIQUES

INTRODUCTION

X^* étant le monoïde libre construit sur X alphabet fini, les langages bornés sur X sont les parties de X^* contenues dans un produit fini de sous-monoïdes monogènes de X^* ayant en outre la propriété d'être des langages de CHOMSKY ou encore langages context free.
Quelques généralités sur les langages context free et la notion d'ambiguité qui s'y attache :

X et V étant des ensembles finis et disjoints, on appelle G grammaire context free en les variables $v_i \in V$. La donnée :

1) d'un nombre fini de règles de la forme $v_i \to p_i$ où p_i est une partie finie de $(X \cup V)^*$

2) d'un élément distingué v_o de V

3) du procédé récursif d'énumération suivant :
   v_o est une production de G, si $x\, v_i\, y \in (X \cup V)^*$ est une production de G alors $x\, z\, y$ est une production de G pour tout $z \in p_i$.

L'ensemble des productions de G qui sont dans X^* est le langage context free L(G) engendré par G. L partie de X^* est un langage context free si et seulement si il existe G grammaire context free et L = L(G).
On ne restreint pas la classe des langages context free en se limitant aux grammaires propres, c'est-à-dire vérifiant en outre :

pour tout i $\quad p_i \cap (V \cup \{e\}) = \emptyset \quad$ où e est le mot vide.

Nous aurons recours à Chomsky et Schutzenberger (1963) pour donner une définition de l'ambiguité d'une grammaire context free propre.

Introduisons d'abord la notion de série formelle de X^* :

à toute application r de X^* dans Z on peut faire correspondre la série formelle $\sum_{f \in X^*} (r,f)f$. L'ensemble des f dans X^* tels que $(r,f) \neq 0$ est le support de la série r. Muni de la somme $(r+r',f) = (r,f)+(r',f)$, du produit $(rr',f) = \sum_{f_i f_j = f} (r,f_i)(r',f_j)$ et des équivalences modulo n $r \equiv r'$ modulo n si et seulement si $(r,f) = (r',f)$ pour tout f de longueur inférieure où égale à n, l'ensemble des séries formelles est un anneau topologique.

A la grammaire propre G on associe le système d'équations en les séries formelles inconnues v_i pour $1 \leq i \leq n$

$$v_i = \sum_{p_{i_k} \in p_i} p_{i_k} \qquad (1)$$

la solution unique de (1) est obtenue par le procédé classique d'approximation : c'est la limite dans le demi anneau topologique des séries formelles positives de la suite de n-uples de polynômes formels $(r_{j_1},\ldots r_{j_n})$ définie par

$$r_{o1},\ldots,r_{or} = (0\ 000\ \ 0)$$

$$r_{j+1,i} = \Psi_i\ (r_{j1},\ldots,r_{jr}) \qquad \text{pour } 1 \leq i \leq n$$

Ψ_i consistant pour tout i à substituer à toute variable v_k dans $\sum_{p_{i_k} \in p_i} p_{i_k}$ le polynôme r_{j_k}

La première composante ρ_o de la solution de (1) a pour support le langage L(G) et (ρ_o,f) est le degré d'ambiguité de f relativement à G
["le nombre de descriptions structurales de f dans G"]
G est dite non ambigue si ρ_o est la série caractéristique de son support (ide $(\rho_o,f) = 1$ ou 0 pour tout $f \in X^*$).
Alors L langage context free sur X est non ambigu s'il existe une grammaire G non ambigue telle que L = L(G)

Un sous-ensemble particulier des context free sur X est formé des langages réguliers (ou langages de Kleene ou langages d'états finis) c'est la classe engendrée par les grammaires linéaires droites ou gauche c'est-à-dire possédant l'une des deux propriétés :

pour tout $i \subset p_i$ $\quad X^* V \cup X^*$

pour tout $i \subset p_i$ $\quad V X^* \cup X^*$

un caractère algébrique connu en est le suivant :

Soit $L \subset X^*$, la congruence de Teissier : $m \sim m'$ ssi $fmg \in L \Leftrightarrow fm'g \in L$ est la congruence la plus grossière sur X^* pour laquelle L est une partie saturée c'est-à-dire une réunion de classes, on l'appelle congruence syntaxique de L (notée (L)) le monoïde quotient $X^*/(L)$ est le monoïde syntaxique de L et les langages réguliers sur X sont les parties de X^* dont le monoïde syntaxique est fini.

Ces généralités posées revenons à notre objet : les langages bornés.
Le fait qu'un langage borné soit contenu dans un produit de n sous monoïdes monogènes conduit à considérer comme l'a fait S. Ginsburg le sous ensemble de $\mathbb{N}^n$, dont il est canoniquement l'image.

Plus précisément si on note w_i^* le monoïde libre construit sur w_i mot de X^* Ψ étant l'application canonique de $\mathbb{N}^n$ dans $w_1^* \ldots w_n^*$, $w_1, \ldots, w_n$ étant des mots de X^* alors l'ensemble $\Psi^{-1}(L)$ où L est un langage context free dans $w_1^* \ldots w_n^*$ à des propriétés remarquables.
D'après le théorème de S. Ginsburg $\Psi^{-1}(L)$ est réunion finie d'ensemble de la forme c+B ou $c \in \mathbb{N}^n$ et B est un sous monoïde stratifié de $\mathbb{N}^n$ considéré comme monoïde additif
[on dira qu'un sous monoïde B est stratifié s'il admet un système fini de générateurs A satisfaisant aux deux conditions :
- pour tout $a \in A$ $\quad a_i$ est nul pour toutes les valeurs de $i = 1 \ldots n$ sauf 2
- il n'existe pas a et a' dans A tels que :

$$a_i \neq 0 \quad a_i \neq 0 \quad a'_j \neq 0 \quad a'_j \neq 0 \quad \text{et} \quad i < j < i' < j']$$

si L est non ambigu alors les ensembles c+B sont disjoints et le sous monoïde stratifié B est libre.

Il semblait intéressant d'opérer, en sens inverse et à partir de sous ensembles de $\mathbb{N}^n$ arithmétiquement définis, d'étudier leur image dans $w_1^*...w_n^*$ et relier par conséquent certaines propriétés, arithmétiques à la théorie des langages context free.

L'objet du présent travail effectué sous la direction de M. NIVAT est l'étude d'une famille de langages bornés particuliers appelés langage bornés arithmétiques. Ils se partagent en deux familles :

premièrement les langages R définis par $R = \Psi[f^{-1}(\bar{d}) \cap \mathbb{N}^n]$ où $\bar{d}$ est une classe dans $\mathbb{Z}$ d'une congruence arithmétique sur Z et f une forme linéaire du $\mathbb{Z}$ module $\mathbb{Z}^n$

deuxièmement les langages L définis par $L = \Psi[f^{-1}(\bar{d}) \cap \mathbb{N}^n]$ où $d \in \mathbb{Z}$ et f est une forme linéaire de $\mathbb{Z}^n$

<u>Résumé du notre travail</u>

On justifie de l'introduction des langages R nommés réguliers arithmétiques dans le chapitre I, en se plaçant tout d'abord dans le cas où $w_1^*...w_n^*$ est un monoïde libre, ce qui entraîne que Ψ^{-1} est une application définie; la définition entraîne qu'ils sont l'intersection avec $w_1^*...w_n^*$, lui-même un langage régulier, de leur fermeture commutative dont le monoïde syntaxique est isomorphe au quotient de $\mathbb{Z}$ par la congruence arithmétique ρ c'est-à-dire un groupe commutatif fini. [§1] ce qui justifie l'appellation de régulier arithmétique.
Ceci nous incite à introduire la classe des réguliers bornés totaux définis comme intersection avec $w_1^*...w_n^*$ d'un langage régulier sur X dont le monoïde syntaxique est un groupe fini commutatif. [§3] en application du résultat de Ginsburg et Spanier (1966)

[tout langage régulier dans $w_1^*...w_n^*$ est réunion finie de produits de n langages réguliers contenus dans $w_1^*, w_2^*, ..., w_n^*$]

dont on donne une nouvelle démonstration en précisant la nature des langages composants [§2] , on obtient le résultat que tout régulier borné total est réunion finie d'intersections de réguliers arithmétiques. [§3]
Pour caractériser tous les réguliers bornés il fallait une classe plus forte : ce sont les réguliers pseudo arithmétiques P[§4] définis par
$P = \Psi(f^{-1}(\bar{d}) \cap N^n)$ où f est une forme linéaire à coefficients tous positifs d une classe d'une congruence ρ d'indice fini non nécessairement arithmétique sur $\mathbb{N}$.
Le résultat déjà cité de Ginsburg et Spanier permet d'obtenir les réguliers bornés quelconques comme réunion finie d'intersection de réguliers pseudo arithmétiques ; avec le corollaire : tout régulier dans $w_1^*,\ldots,w_n^*$, sous monoïde libre, est l'intersection d'un langage régulier abélien avec $w_1^*\ldots w_n^*$
[les réguliers abéliens étant les parties de X^* dont le monoïde syntaxique est commutatif.]

Ces résultats sont naturellement valables dans le cas où $w_1,\ldots,w_n$ sont des lettres a_i distinctes de X. On les étend [§5] au cas où $w_1,\ldots,w_n$ sont des mots quelconques par l'homomorphisme défini par $h(a_i) = w_i$ avec les restrictions suivantes :

La fermeture des réguliers arithmétiques n'est plus la classe des réguliers bornés totaux mais celle de leurs images homomorphes qui ne coïncide pas avec la précédente. Le dernier corollaire cité est faux.

Le chapitre II justifie de l'introduction de la deuxième famille de langages bornés arithmétiques nommés context free arithmétiques.
Comme précédemment on en place tout d'abord dans le cas où $w_1^*\ldots w_n^*$ est un sous-monoïde libre. Alors les context free arithmétiques et leur fermeture commutative sont des langages bornés non ambigus pour la raison qu'ils peuvent être donnés comme image inverse dans une application injective de $w_1^*\ldots w_n^*$ dans un monoïde libre Y^* de l'intersection d'un langage de Dyck et d'un langage régulier tous deux sur Y [§12].

[Le langage de Dyck considéré est appelé langage de Dyck par abus de langage il a les propriétés de factorisation qui font du langage de Dyck un langage non ambigu [§11].]

Ceci répond à une question de S. Ginsburg trouver une famille de langages bornés non ambigus.

On décrit dans [§2] des automates à pile qui appartiennent à la famille des "déterministic Push down automata" de Ginsburg qui reconnaissent les context free arithmétiques et leur fermeture commutative. Il en résulte par application de théorèmes généraux relatifs aux automates de cette famille que les complémentaires dans X^* des context free arithmétiques et de leur fermeture commutatives sont des langages non ambigus.

Les paragraphes suivant étudient des propriétés en relation plus directe avec l'arithmétique, ainsi [§3] le problème de décision du langage vide qui est celui de l'existence des solutions de l'équation diophantrienne $f(x) = d$ se ramène à celui de l'existence des solutions de la relation $f(x)\rho\, d$ où ρ est une congruence arithmétique par conséquent c'est le problème de décision du langage vide d'un régulier arithmétique. De même le problème de la dégénérescence d'un context free arithmétique, en un langage régulier est tout aussi soluble : un context free arithmétique non vide n'est régulier que si les coefficients de la forme f sont tous de même signe.

Il résulte du théorème de Ginsburg cité dans les généralités et du [§1] que l'ensemble D des solutions de l'équation diophantrienne $f(x) = d$ est réunion finie disjointe de sous ensembles de la forme $c+B$ où $c \in \mathbb{N}^n$ et B est un sous monoïde libre stratifié de $\mathbb{N}^n$.

La donnée explicite d'une telle décomposition de D (D est appelée alors partie non ambigue de $\mathbb{N}^n$ [41] permet de construire une grammaire explicite non ambigue d'un context free arithmétique [§42]. On essaye donc l'existence étant prouvée de construire une telle décomposition de D.

On donne [§43] des conditions suffisantes sur la distribution des signes des coefficients de f pour que la construction par récurrence soit possible. Cette condition est directement liée à la forte contrainte de la stratification, elle n'est plus réalisée à partir de n = 8.

Dans [§5] et dernier on donne comme corollaire du théorème de Ginsburg une décomposition des langages context free bornés généraux sur les context free arithmétiques.

Des context free arithmétiques dans $w_1^* \ldots w_n^*$ où $w_1, \ldots w_n$ sont des mots quelconques de X^* on ne sait pas dire s'ils sont en général ambigus où non par suite aucune propriété particulière ne s'y attache.

CHAPITRE I

LANGAGES REGULIERS ARITHMETIQUES

ET PSEUDO-ARITHMETIQUES

INTRODUCTION

Les langages bornés ont été introduits par GINSBURG soit X^* le monoïde libre construit sur l'alphabet X $L \subset X^*$ est borné s'il existe $W_1,\ldots,W_n$ mots de X^* tels que $L \subset W_1^* \ldots W_n^*$.

La notion de langage régulier est bien connue on se servira de la caractérisation suivante :

$L \subset X^*$ est régulier si et seulement s'il existe φ homomorphisme de X^* dans un monoïde fini tel que $\varphi^{-1} \circ \varphi(L) = L$.

On se propose de donner une caractérisation des langages réguliers bornés à l'aide des langages particuliers "réguliers arithmétiques et pseudo-arithmétiques bornés" qu'on définira plus loin.

On établira ces résultats tout d'abord pour les réguliers contenus dans $a_1^* \ldots a_n^*$ où $a_1,\ldots,a_n$ sont des lettres distinctes (tout mot de $a_1^* \ldots a_n^*$ a une factorisation unique) [paragraphes 1, 2, 3, 4] et on les étendra [paragraphe 5] aux réguliers contenus dans $W_1^* \ldots W_n^*$ $W_1,\ldots,W_n$ étant des mots quelconques sur X^*.

Ces résultats précisent ceux de [GINSBURG et SPANIER (1966)] sur les "Bounded Regular Sets".

§ 1 - REGULIERS ARITHMETIQUES BORNES

Définition :

Soit $a_1,\ldots,a_n$ lettres distinctes

On appelle régulier arithmétique borné (notation : $L \in RA(a_1,\ldots,a_n)$
Les parties L de $a_1^* \ldots a_n^*$ telles que :
soit Ψ l'injection de $\mathbb{N}^n$ dans $a_1^* \ldots a_n^*$ (1)

Il existe f forme linéaire à coefficients dans $\mathbb{Z}$

d et $\gamma \in \mathbb{N}$ $\quad o \leq d < \gamma$

$\underset{\gamma}{\equiv}$ une congruence arithmétique sur $\mathbb{Z}$

$$\left(x \underset{\gamma}{\equiv} y \quad \text{ssi } x = y (\text{mod } \gamma)\right)$$

et $\quad L = \{x \in a_1^* \ldots\ldots a_n^*,\ f \circ \Psi^{-1}(x) \underset{\gamma}{\equiv} d\}$

Justification :

Soit $R(a_1,\ldots,a_n)$ l'ensemble des langages réguliers contenus dans $a_1^* \ldots\ldots a_n^*$

alors $RA(a_1,\ldots,a_n) \subset R(a_1,\ldots,a_n)$

On pose $\quad X = \{a_1,\ldots,a_n\}$

Ψ étant l'injection de $\mathbb{N}^n$ dans $a_1^* \ldots a_n^*$

On a $\Psi^{-1}(a_i) = (ooo \underset{\substack{\uparrow \\ i}}{1} ooo)$ pour tout i $\quad o < i < n$

Soit $L = RA(a_1,\ldots,a_n)$ on considère le quotient de $\mathbb{Z}$ par la congruence $\underset{\gamma}{\equiv}$ dont on note les éléments $\bar{\alpha}$ pour $\alpha \in \mathbb{Z}$

Soit $f = (\alpha_1,\ldots\ldots,\alpha_n)$, $\alpha_i \in \mathbb{Z}$ pour tout i

On pose $\varphi(a_i) = \overline{f \circ \Psi^{-1}(a_i)} = \bar{\alpha}_i$

$$\varphi(a_i\, a_\gamma) = \overline{f \circ \Psi^{-1}(a_i) + f \circ \Psi^{-1}(a_\gamma)} = \bar{\alpha}_i + \bar{\alpha}_\gamma$$

en tenant compte de la loi sur $\mathbb{Z}_\gamma$ $\quad \bar{\alpha} + \bar{\beta} = \alpha + \beta$

φ s'étend à un homomorphisme de X^* dans $\mathbb{Z}_\gamma$

alors $\varphi^{-1}(d) \subset X^*$ est régulier

en outre $a_1^* \ldots\ldots a_n^*$ est régulier

On a $L = \varphi^{-1}(\bar{d}) \cap a_1^* \ldots\ldots a_n^*$ donc L est régulier

<u>Remarque</u> :

$\varphi^{-1}(\bar{d})$ coïncide avec sa fermeture commutative

$\mathbb{Z}_\gamma$ étant un groupe commutatif ; c'est un langage régulier abélien.

---

(1) soit $(x_1,\ldots\ldots,x_n) \in M^n$ $\qquad \Psi(x_1 \ldots\ldots x_n) = a_1^{x_1} \ldots\ldots a_n^{x_n}$

§ 2 - STRUCTURE DES LANGAGES REGULIERS BORNES

Lemme 21 :

Pour toute congruence ρ d'indice fini sur $\mathbb{N}$ il existe un entier p tel que $\{o\}$, $\{1\},\ldots,\{p-1\}$ forment p classes distinctes de la congruence ρ les autres classes étant celles d'une congruence arithmétique sur $\mathbb{N} + p$.

Pour la démonstration de ce lemme classique, se reporter à CLIFFORD et PRESTON [ALGEBRAIC THEORY OF SEMI GROUPS].

Corollaire :

Si ϕ est un homomorphisme de $\mathbb{N}$ sur M monoïde fini alors il existe G groupe cyclique fini contenu dans M il existe p et λ appartenant à $\mathbb{N}$ et :

(i) la restriction de φ à $\{o,1,2,\ldots,p-1\}$ est une bijection sur $M - G$

(ii) $\varphi^{-1}(G) = \mathbb{N} + p$ est la congruence sur $\mathbb{N} + p$ définie par $x \; u \; y$ ssi $\varphi(x) = \varphi(y)$ est la congruence arithmétique $\underset{\gamma}{\equiv}$

Preuve : Il suffit d'appliquer le lemme précédent à la congruence d'indice fini ρ sur $\mathbb{N}$ $x \rho y$ ssi $\varphi(x) = \varphi(y)$. On pose $G = \phi(\mathbb{N}+p)$ il est clair que φ est une bijection de $\{o,1,2,\ldots\ldots,p-1\}$ sur $M - G$ et que la congruence sur $\mathbb{N} + p$ associée à φ est une congruence arithmétique soit H le quotient de $\mathbb{N} + p$ par cette congruence.

H est un groupe cyclique fini.

Soit η la projection de $\mathbb{N} + p$ sur H

si $\bar{\varphi}$ est défini par $\bar{\varphi} \circ \eta(x) = \varphi(x)$

φ est un homomorphisme de H sur G et on a le diagramme commutatif suivant :

$$\begin{tikzcd} \mathbb{N} + p \arrow[r, "\varphi"] \arrow[d, "\eta"'] & G \\ H \arrow[ur, "\bar{\varphi}"'] & \end{tikzcd}$$

alors $G = \varphi(\mathbb{N}+p)$ est un groupe cyclique fini.

<u>Lemme 22</u> :

Soit Ψ l'injection de $\mathbb{N}^n$ dans $a_1^* \ldots\ldots a_n^*$, si $L \in R(a_1,\ldots,a_n)$ alors $\Psi^{-1}(L)$ est réunion finie d'ensembles disjoints

$\{B_i\}_{i \in I \text{ fini}}$ et pour tout $i \in I$, $B_i = C_{i_1} \times \ldots \times C_{i_n}$ (Produit Cartésien)

et $C_{i_\gamma} = \varphi_\gamma^{-1}(m_{i_\gamma})$ $\quad m_{i_\gamma} \in M_\gamma$ monoïde fini

φ_γ est un homomorphisme de $\mathbb{N}$ sur M_γ, pour tout $1 \leqslant \gamma \leqslant n$

<u>Preuve</u> : $L \in R(a_1,\ldots,a_n)$ alors il existe φ homomorphisme de X^* dans M monoïde fini et $\varphi^{-1} \circ \varphi(L) = L$

On pose $\varphi_\gamma = \varphi \circ \Psi_\gamma$ $\quad \Psi_\gamma$ étant l'injection de $\mathbb{N}$ dans a_γ^*

pour $1 \leqslant \gamma \leqslant n$ $\quad M_\gamma = \varphi \circ \Psi_\gamma(\mathbb{N})$

soit $M' = \phi(a_1^* \ldots\ldots a_n^*) \subset M$

$$\xi : a_1^* \ldots\ldots a_n^* \longrightarrow M_1 \times \ldots \times M_n$$

défini par

$$\xi \circ \Psi\,(p_1,\ldots,p_n) = \varphi_1\,(p_1),\ldots,\varphi_n\,(p_n)$$

$\eta : M_1 \times \ldots \times M_n \longrightarrow M$

défini par $\eta(m_1,\ldots,m_n) = m_1 \ldots m_n$ (produit dans M)

alors on a le diagramme suivant :

$$\begin{tikzcd}
a_1^* \ldots\ldots a_n^* \arrow[r, "\xi"] \arrow[rrr, bend left] & M_1 \times \ldots \times M_n \arrow[r, "\eta"] & M' \\
 & \mathbb{N}^n \arrow[u, "\Psi"] \arrow[ur, "\xi \circ \Psi"] \arrow[urr, "\varphi \circ \Psi"] & \\
 & & \phi
\end{tikzcd}$$

$L \in R(a_1,\ldots,a_n)$ alors $\varphi\,(L) \subset M'$

et $L = \bigcup_{x \in \varphi(L)} \varphi^{-1}(x)$

d'où $\Psi^{-1}(L) = \bigcup_{x \in \varphi(L)} (\varphi \circ \Psi)^{-1} x$

et $\Psi^{-1}(L) \bigcup_{x \in \varphi(L)} \bigcup_{y \in \eta^{-1}(x)} (\xi \circ \Psi)\, y$

les $\left\{ (\xi \circ \Psi)^{-1}\, y \right\}_{\eta(y) \in \varphi(L)}$ sont évidemment disjoints et en nombre fini

d'autre part

$$(\xi \circ \Psi)^{-1} y = \varphi_1^{-1}(y_1) \times \ldots\ldots \times \varphi_n^{-1}(y_n) \quad \text{où} \quad y = y_1,\ldots,y_n$$

ce qui démontre ce lemme.

Corollaire immédiat :

Si $L \in R(a_1,\ldots,a_n)$ alors L est réunion finie d'ensembles $\{B_i\}_{i \in I}$

et pour tout i $\quad B_i = A_{i_1} \ldots A_{i_n}$

pour tout γ $1 \leqslant \gamma \leqslant n$ $\quad A_{i_\gamma}$ est un langage régulier dans a_γ^*

En effet $A_{i_\gamma} = \Psi_\gamma \circ \varphi_\gamma (m_{i_\gamma})$

et $\Psi_\gamma \circ \varphi_\gamma^{-1} (m_{i_\gamma}) = \left[\varphi_\gamma \circ \Psi_\gamma^{-1}\right]^{-1} m_{i_\gamma}$

$\varphi_\gamma \circ \Psi_\gamma^{-1}$ est un homomorphisme de a_γ^* sur M_γ monoïde fini donc A_{i_γ} est régulier.

Ce résultat est exactement celui du lemme 12 de [GINSBURG et SPANIER] (1966).

Corollaire 22 :

Si $L \in R(a_1,\ldots,a_n)$ alors $\Psi^{-1}(L) = \bigcup_{i \in I} (B_i)_{\text{fini}}$

et pour tout $i \in I$

$$B_i = C_{i_1} \times \ldots \times C_{i_n}$$

et il existe $\lambda_1,\ldots,\lambda_n$ tels que C_{i_γ} est réduit à un élément ou bien est une progression arithmétique de raison λ_γ.

NB : Ce corollaire précise la nature des langages A_{i_γ} du corollaire précédent.

Preuve : On applique le résultat du corollaire du lemme 21 aux résultats du lemme 22.

On a $C_{i_\gamma} = \varphi_\gamma^{-1} \quad m_{i_\gamma}$ où φ_γ est un homomorphisme de $\mathbb{N}$ sur un monoïde M_γ alors soit G_γ le groupe cyclique contenu dans M_γ.

Si $m_{i_\gamma} \in G_\gamma$ alors $\varphi_\gamma^{-1} m_{i_\gamma}$ est une classe d'une congruence $\equiv_{\lambda_\gamma}$ sur $\mathbb{N} + p$ ide une progression arithmétique de raison λ_γ. Si $m_{i_\gamma} \notin G_\gamma$ alors $\varphi_\gamma^{-1} (m_{i_\gamma})$ est réduit à un élément.

§ 3 - REGULIERS TOTAUX BORNES

On a vu dans § 1 que tout régulier arithmétique borné dans $a_1^* \ldots\ldots a_n^*$ était égal à $\varphi^{-1}(d) \cap a_1^* \ldots\ldots a_n^*$ où φ est un homomorphisme de X^* dans $\mathbb{Z}_\gamma$ groupe commutatif fini.

On considère le sous-ensemble des langages abéliens sur X suivant :

Définition 31 :

$L \subset X^*$ est dit régulier total s'il existe φ homomorphisme de X^* dans un groupe commutatif fini tel que $^{-1} \circ \varphi(L) = L$

Exemple :

Si $a_{i_{|x|}}$ est le nombre d'occurence de a_i dans $x \in X^*$

On pose $\eta(x) = (a_{1_{|x|}}, \ldots\ldots, a_{n_{|x|}})$

soit f forme linéaire à coefficients dans $\mathbb{Z}$

alors $L = \{x \in X^* \mid f \circ \eta(x) \equiv d\}$ est régulier total

en effet L n'est autre que $\varphi^{-1}(\bar{d})$ où φ est l'homomorphisme dans $\mathbb{Z}_\gamma$ du § 1.

<u>Propriété 31</u> :

Les réguliers totaux forment une classe fermée pour le passage au complément les intersections et les réunions finies.

<u>Preuve</u> : Il suffira de montrer que si L_1 et L_2 sont réguliers totaux il en est de même pour $L_1 \cup L_2$ et $L_1 \cap L_2$ (pour le complément c'est évident) soit φ_1, G_1 resp(φ_2, G_2) relatifs à L_1 resp L_2.

Le produit direct $G_1 \times G_2$ est un groupe commutatif fini l'application $\varphi : X^* \longrightarrow G_1 \times G_2$ défini par :

$\varphi(x) = \varphi_1(x), \varphi_2(x)$ est un homomorphisme

$$\varphi(xy) = \varphi_1(xy), \varphi_2(xy) = (\varphi_1(x), \varphi_2(x)) \cdot (\varphi_1(y), \varphi_2(y)) = \varphi(x) \quad \varphi(y)$$

alors $\varphi^{-1} \circ \varphi(L_1 \cup L_2) = L_1 \cup L_2$

en effet $\varphi(x) = \varphi(y)$, $y \in L_1 \cup L_2$ entraîne $\varphi_1(x) \in L_1$ ou $\varphi_2(x) \in L_2$

donc $x \in L_1 \cup L_2$

de même $\varphi^{-1} \circ \varphi(L_1 \cap L_2) = L_1 \cap L_2$

Remarque :

$a_1^* \ldots a_n^*$ n'est pas régulier total

En effet il existe p $\varphi(a_1^p) = e$ élément neutre de G

si $a_1^* \ldots a_n^*$ est régulier total

alors pour $x \in a_1^* \ldots a_n^*$ $x a_1^p$ appartiendrait aussi à $a_1^* \ldots a_n^*$ ce qui est absurde.

On est conduit à introduire les réguliers totaux bornés.

Définition 32 :

$L \subset X^*$ est dit régulier total borné $(L \in RT(a_1,\ldots,a_n))$ s'il existe L_1 régulier total dans X^* et $L = L_1 \cap a_1^* \ldots a_n^*$

Proposition 31 :

$RA(a_1,\ldots,a_n) \subset RT(a_1,\ldots,a_n)$

ceci découle directement du § 1

$L \in RA(a_1,\ldots,a_n)$ alors $L = \varphi^{-1}(\bar{d}) \cap a_1^* \ldots a_n^*$

et $\varphi^{-1}\bar{d}$ est régulier total

Propriété 32 :

$RT(a_1,\ldots,a_n)$ est fermé pour les intersections et les réunions finies ceci découle de la propriété 31.

Corollaire 32 :

Soit $\overline{RA}(a_1,\ldots,a_n)$ la fermeture de $RA(a_1 \ldots a_n)$ par réunions et intersections finies alors $\overline{RA}(a_1,\ldots,a_n) \subset RT(a_1 \ldots a_n)$.

Lemme 31 :

Si $L \in RT(a_1,\ldots,a_n)$ et Ψ_γ l'injection de $\mathbb{N}$ sur a_γ^* il existe $\varphi^{-1} \circ \varphi L = L$ et $\varphi \circ \Psi_\gamma$ est un homomorphisme de $\mathbb{N}$ sur M_γ groupe cyclique contenu dans M.

Preuve : On considère l'automate $A = \{S,X,\eta,q_o,F\}$ (1)

où $S = \{q_o,q_1,\ldots,q_n,q_d\}$

$F = S - \{q_d\}$, $\eta : S \times X \longrightarrow S$ définie par

$$\eta(a_i a_\gamma) = \begin{cases} q_\gamma & \text{si } i \leqslant \gamma \\ q_d & \text{si } i > \gamma \end{cases}$$

$\eta(q_d\, a_i) = q_d$

A reconnait $a_1^* \ldots a_n^*$. On construit canoniquement à partir de A l'homomorphisme $\varphi_1 : X^* \longrightarrow S^S$, $\varphi_1(x) = f_x$

$f_x(s) = \bar{\eta}(s,x)$ $\bar{\eta}$ étant l'extension de η à $S \times X^*$

On a $\varphi_1^{-1} \circ \varphi_1 (a_1^* \ldots a_n^*) = a_1^* \ldots a_n^*$

$\varphi_1(a_\gamma^2) = f_{a_\gamma} \circ f_{a_\gamma} = f_{a_\gamma}$ donc $\varphi_1 (a_\gamma^*)$ est un groupe réduit à un seul élément.

Si φ_2 est l'homomorphisme dans G groupe fini associé à L_2 tel que :

$$L_2 \cap a_1^* \ldots a_n^* = L$$

---

(1) Notation de Ginsburg pour finite state acceptor

de même que dans la démonstration de la propriété 31

$\varphi : X^* \longrightarrow G \times S^s$, $\varphi(x) = \varphi_2(x)$, $\varphi_1(x)$ est tel que $\varphi^{-1} \circ \varphi(L) = L$ en outre $\varphi \circ \Psi_\gamma(|N) = \varphi_2 \circ \Psi_\gamma(|N) \times \varphi_1 \circ \Psi_\gamma(|N)$ car $\varphi_1 \circ \Psi_\gamma(|N)$ est réduit à un élément $\varphi \circ \Psi_\gamma(|N)$ est un groupe cyclique fini contenu dans $G \times S^s$ ce qui termine la démonstration.

<u>Proposition 32</u> :

$$RT(a_1,\ldots,a_n) = \overline{RA}(a_1,\ldots,a_n)$$

<u>Preuve</u> : On se borne à montrer que $RT(a_1,\ldots,a_n) \subset \overline{RA}(a_1,\ldots,a_n)$; soit $L \in R(a_1,\ldots,a_n)$ et d'après le lemme 22

$$\Psi^{-1}(L) = \bigcup_{i \in I \text{ fini}} B_i$$

$$B_i = C_{i_1} \times \ldots \times C_{i_n} \qquad C_{i_\gamma} = (\varphi \circ \Psi_\gamma)\, m_{i_\gamma},\ m_{i_\gamma} \in M_\gamma$$

monoïde fini ; en raison du lemme 31 il existe ϕ tel que $\varphi^{-1} \circ \varphi(L) = L$ et $\varphi \circ \Psi_\gamma(|N)$ est un groupe cyclique, alors, pour ce φ là, M_γ est un groupe cyclique.

La congruence associée à $\varphi \circ \Psi_\gamma$ sur $|N$ est une congruence arithmétique $\underset{\lambda_\gamma}{\equiv}$ on pose d_{i_γ} un élément de $(\varphi \circ \Psi_\gamma^{-1})\, m_{i_\gamma}$

alors si f_γ est la forme linéaire $(\circ\circ\circ \underset{\gamma}{\uparrow} \circ\circ\circ)$

$$B_i = \bigcap_{1 \leq \gamma \leq n} \{x \in \mathbb{N}^n \mid f_\gamma(x) \underset{\lambda_\gamma}{\equiv} d_{i_\gamma}\}$$

d'où :

$$\Psi(B_i) = \bigcap_{1 \leq \gamma \leq n} \{x \in a_1^* \ldots a_n^* \mid f_\gamma \circ \Psi^{-1}(x) \underset{\lambda_\gamma}{\equiv} d_{i_\gamma}\}$$

ce qui montre que $\Psi(B_i) \in \overline{RA}(a_1,\ldots,a_n)$

par suite $L = \bigcup_{i \in I} \Psi(B_i) \in \overline{RA}(a_1,\ldots,a_n)$

§ 4 - REGULIERS PSEUDO ARITHMETIQUES BORNES $PA(a_1,\ldots,a_n)$

Il existe des langages réguliers bornés qui ne sont pas réguliers totaux bornés, ceux-ci ayant toujours un nombre infini d'éléments ou aucun, par exemple les parties finies de $a_1^* \ldots a_n^*$; on va ici introduire une classe analogue à celle des réguliers arithmétiques pour caractériser tous les réguliers bornés.

Définition :

$L \in PA(a_1,\ldots,a_n)$ s'il existe ρ congruence d'indice fini sur $\mathbb{N}$

$d \in \mathbb{N}$, f forme linéaire à coefficients positifs

et $\qquad L = \{x \in a_1^* \ldots a_n^* \mid f \circ \Psi^{-1}(x) \; \rho \; d\}$

Propriété 41 :

$PA(a_1,\ldots,a_n) \subset R(a_1,\ldots,a_n)$

On considère l'homomorphisme $\varphi : X^* \longrightarrow \mathbb{N}/_\rho$ (quotient de $\mathbb{N}$ par ρ)

défini par $\varphi(a_i) = f \circ \Psi^{-1}(a_i)$

$$\varphi(a_i\, a_\gamma) = \varphi(a_i) + \varphi(a_\gamma) \qquad \left[\mathbb{N}/_\rho \text{ a une structure de monoïde fini} \right]$$

alors $\varphi^{-1}(d)$ est régulier de même que $\varphi^{-1}(\bar{d}) \cap a_1^* \ldots\ldots a_n^* = L$

<u>Propriété 42</u> :

$RA(a_1,\ldots\ldots,a_n) \subset PA(a_1,\ldots,a_n)$

<u>Preuve</u> : soit $L \in RA(a_1,\ldots\ldots,a_n)$

$$L = \{x \in a_1^* \ldots\ldots a_n^*\} \mid f \circ \Psi^{-1}(x) \underset{\gamma}{\equiv} d\}$$

avec $\quad f = (\alpha_1,\ldots\ldots,\alpha_n) \qquad \alpha_i \in \mathbb{Z}$ pour $0 < i \leqslant n$

On considère $f' = (\alpha_1',\ldots\ldots,\alpha_n')$ et $\alpha_i' \; E \; \mathbb{N}$

$$\alpha_i' \underset{\gamma}{\equiv} \alpha_i$$

Si φ' est l'homomorphisme de X^* dans $\mathbb{N}$ associé à f'

On a $\varphi'(a_j) = f' \circ \varphi^{-1}(a_j) = i(f\circ\Psi^{-1}a_j^\gamma) = i\big(\varphi(a_j)\big)$

où i est l'injection de $\mathbb{Z}_\gamma$ dans $\mathbb{N}_\gamma$

par un abus de langage raisonnable on peut écrire $\varphi' = \varphi$ sur X^*

Il s'ensuit que

$$L' = \{x \in a_1^* \ldots\ldots a_n^* \mid f' \circ \Psi^{-1}(x) \underset{\gamma}{\equiv} d\} = L$$

et il est clair que $L' \in PA(a_1,\ldots,a_n)$

Exemple 41 :

Soit f_1 la forme (1 0000)

et ρ la congruence dont les classes sont $\{0\}$, $\{1\}$, $\{2\}$, $\bar{3}$, $\bar{4}$

où $\bar{3} = \{x > 2 \mid x \underset{2}{\equiv} 3\}$

$\bar{4} = \{x > 2 \mid x \underset{2}{\equiv} 4\}$ (cf lemme 21)

alors $L = \{x \in a_1^* \ldots\ldots a_n^* \mid f_1 \circ \Psi^{-1}(x) \rho 1\} = a_1 a_2^* \ldots\ldots a_n^*$

L n'est pas régulier arithmétique sinon il existerait f et γ, soit f_1 la première coordonnée de f

et $a_1 a_2^* \ldots\ldots a_n^* \subset L \implies a_1 (a_1^{f_1 \gamma})^* a_2^* \ldots\ldots a_n^* \subset L$

ce qui est absurde.

Proposition 41 :

$\overline{PA}(a_1, \ldots\ldots, a_n)$ étant la fermeture par réunion et intersection finie de $PA(a_1, \ldots\ldots, a_n)$

alors $PA(a_1, \ldots\ldots, a_n) = R(a_1, \ldots, a_n)$

Preuve : D'une manière analogue que pour la proposition 32, les C_{i_γ} étant les ensembles de la décomposition de $L \in R(a_1, \ldots, a_n)$,

$C_{i_\gamma} = (\varphi \circ \Psi_\gamma)^{-1} m_{i_\gamma}$ $\quad m_{i_\gamma} \in M_\gamma$ monoïde fini

soit ρ_γ la congruence associée à $\varphi \circ \Psi_\gamma$

alors $\Psi(B_i) = \bigcap_{1 \leqslant \gamma \leqslant n} \{x \in a_1^* \ldots\ldots a_n^* \mid f_\gamma \circ \Psi^{-1}(x) \underset{\gamma}{\rho} d_{i_\gamma}\}$

où d_{i_γ} est tel que $\varphi \circ \Psi_\gamma (d_{i_\gamma}) = m_{i_\gamma}$

il est clair que $\{x \in a_1^* \ldots\ldots a_n^* \mid f_\gamma \circ \Psi^{-1}(x) \rho_\gamma d_{i_\gamma}\} \in PA(a_1,\ldots,a_n)$

par suite

$$L = \bigcup_{i \in I} \Psi B_i \in \overline{PA}(a_1,\ldots\ldots,a_n)$$

comme $\overline{PA}(a_1,\ldots\ldots,a_n) \subset R(a_1,\ldots a_n)$ la démonstration est achevée.

<u>Proposition 42</u> :

Tout langage de $R(a_1,\ldots,a_n)$ peut s'écrire comme intersection d'un langage régulier abélien avec $a_1^* \ldots\ldots a_n^*$

si on se reporte à la propriété 41

on voit que si $L \in PA(a_1,\ldots\ldots,a_n)$ alors $L = \varphi^{-1}(\bar{d}) \cap a_1^* \ldots\ldots a_n^*$

il est clair que $\varphi^{-1}(\bar{d})$ coïncide avec sa fermeture commutative $\left[\mathbb{N}/_\rho\right.$ est un monoïde abélien$\left.\right]$

il s'ensuit que si $L \in \overline{PA}(a_1,\ldots\ldots,a_n)$

$L = L' \cap a_1^* \ldots\ldots a_n^*$ ou L' est un langage abélien régulier, comme résultat d'opérations d'intersection et de réunion finie sur des langages abéliens réguliers il en est de même pour $R(a_1,\ldots\ldots,a_n)$.

§ 5 - <u>EXTENSION AUX LANGAGES BORNES REGULIERS DANS $W_1^* \ldots\ldots W_n^*$ où $W_1,\ldots\ldots,W_n$ SONT DES MOTS SUR X</u>

Un mot de $W_1^* \ldots\ldots W_n^*$ ayant en général plusieurs factorisations. L'application Ψ de $\mathbb{N}^n$ dans $W_1^* \ldots\ldots W_n^*$ n'est pas injective en général et on ne peut définir directement les réguliers arithmétiques et pseudo-arithmétiques comme dans les paragraphes précédents.

On considère alors $Y = \{a_1, \ldots\ldots, a_n\}$ les a_i étant des lettres distinctes et l'homomorphisme $h : Y^* \longrightarrow X^*$ défini par $h(a_i) = W_i$ relativement à cet homomorphisme on a les propriétés suivantes connues.

Propriété 51 :

Pour tout $L \in R(W_1, \ldots, W_n)$ il existe $L' \in R(a_1, \ldots, a_n)$ et $h(L') = L$

Propriété 52 :

Pour tout $L \in R(a_1, \ldots, a_n)$ alors $h(L) \in R(W_1, \ldots, W_n)$

Cette dernière propriété résulte du lemme 22 (lemme 12 de GINSBURG et SPANIER).

Définition 51 :

$RA(W_1, \ldots, W_n)$ $\big($resp $PA(W_1, \ldots\ldots, W_n)\big)$ est l'image par h de $RA(a_1, \ldots, a_n)$ $\big($resp $PA(a_1 \ldots\ldots a_n)\big)$

Remarque :

L'image par h d'un langage abélien n'est pas nécessairement un langage abélien. Ceci entraîne que l'image d'un régulier total n'est pas forcément un régulier total d'où la proposition :

Proposition 51 :

$$h\big(R^T(a_1, \ldots, a_n)\big) = \overline{RA}(W_1, \ldots, W_n)$$

Proposition 52 :

$$R(W_1, \ldots, W_n) = \overline{PA}(W_1, \ldots, W_n)$$

La démonstration est immédiate en utilisant la propriété 51.
Les propriétés de fermeture sont conservées par h

Remarque :

On n'a pas de propriété analogue à la proposition 42. (Voir le contre exemple ci-dessous).

Exemple de langage dans $R(W_1,\ldots,W_n)$ qui n'est pas intersection d'un régulier abélien avec $W_1^* \ldots W_n^*$

soit $X = \{ab\}$ $\quad W_1 = a \quad W_2 = b \quad W_3 = a \quad L = (a^2)^* b^*(a^3)^*$

alors $(a^2)^* b^*(a^3)^* \in R(W_1,W_2,W_3)$

s'il existait $\mathcal{L}$ abélien sur X tel que $\mathcal{L} \cap W_1^* W_2^* W_3^* = L$

alors $a^2 b a^3 \in \mathcal{L}$ de même que $a^3 b a^2$

$a^3 b a^2 \in W_1^* W_2^* W_3^* \cap \mathcal{L}$ mais n'appartient pas à L d'où la contradiction donc il n'existe pas de langage abélien $\mathcal{L}$ sur X tel que $L = \mathcal{L} \cap W_1^* \ldots W_3^*$ ce qui justifie la remarque précédente.

Lemme 51 (Critère) :

Soit a, u, b, v, c mots sur X

$L \in a u^* b v^* c$ est régulier si et seulement s'il existe $m, \pi, \lambda_1, \lambda_2$ dans $\mathbb{N}$ tels que la condition suivante soit réalisée :

(C) si $a u^k b v^\ell c \in L$ pour $k \geqslant m$ (resp $\ell \geqslant \pi$)
alors $a u^k (u^{\lambda_1})^* b v^\ell c \in L$ (resp $a u^k b v^\ell (v^{\lambda_2})^* c \in L$)

<u>Preuve</u> : Soit h l'homomorphisme de $Y = \{\alpha, u, \beta, v, \gamma\}$ dans X^* tel que

$h(\alpha) = a$, $h(u) = u$, $h(\beta) = b$, $h(v) = v$, $h(\gamma) = o$

en adaptant légèrement le lemme 22 on a pour L_1 régulier dans $\alpha\, u^* \beta\, v^* \gamma$ le résultat suivant :

Ψ étant l'injection de $\mathbb{N}^2$ dans $\alpha\, u^* \beta\, v^* \gamma$

$$\Psi^{-1}(L_1) = \bigcup_{i \in I \text{ fini}} \alpha\, C_{i_1}\, \beta\, C_{i_2}\, \gamma$$

pour tout $i \in I$ C_{i_1} (resp C_{i_2}) est une classe d'une congruence d'indice fini ρ_1 (resp ρ_2) sur $\mathbb{N}$

En utilisant le lemme 21 il existe m, π, λ_1, λ_2 tels que

$$(C_1)\begin{cases} \text{si } k \geqslant m \text{ resp } (\ell \geqslant \pi) \text{ est dans } C_{i_1} \text{ (resp } C_{i_2}) \text{ pour un } i \in I, \text{ alors} \\ \text{pour tout } p \text{ entier, } k + p\lambda_1 \text{ (resp } \ell + p\lambda_2) \text{ est dans } C_{i_1} \text{ (resp } C_{i_2}) \end{cases}$$

Comme corollaire de la propriété 51 on a :

Si L est régulier dans $a\, u^* b\, v^k c$ alors il existe L_1 régulier dans $\alpha\, u^* \beta\, v^* \gamma$ et $h(L_1) = L$

alors par $h \circ \Psi$ la propriété précédente (C_1) donne la condition (C)
La réciproque est immédiate

pour tout $(k\ell) < m + \lambda_1, \pi + \lambda_2$

On pose :

$$U_{k\ell} = \{x = a\, u^{k + p\lambda_1}\, b\, v^{\ell + q\lambda_2} \mid x \in L,\ pq \in \mathbb{N}^2\}$$

Il est clair que $U_{k\ell}$ s'il n'est pas vide est régulier

En effet, avec la condition (C)

$\bigcup_{k\ell}$ est fini ou bien est l'image par un homomorphisme de $\alpha\ u^{*}\ \beta\ \gamma$ ou de $\alpha\ \beta\ \overset{+*}{v}\gamma$ ou de $\alpha\ u^{*}\ \beta\ v^{*}\ \gamma$ et avec un corollaire de la propriété 52 il est régulier.

Par suite L est régulier car :

$$L = \bigcup_{k,\ell<m+\lambda_1,\pi+\lambda_2} \bigcup_{k\ell}$$

CHAPITRE II

LANGAGES CONTEXT FREE ARITHMETIQUES

INTRODUCTION

En utilisant les notations du chapitre I on appelera context free arithmétique tout sous-ensemble L de $a_1^* \ldots a_n^*$ tel qu'il existe $d \in \mathbb{Z}$ f forme linéaire sur $\mathbb{Z}^n$ et $L = \{x \in a_1^* \ldots a_n^*, f \circ \Psi^{-1}(x) = d\}$.

On montrera par deux voies différentes, la deuxième faisant fortement appel à des résultats connus, que les context free arithmétiques sont des langages context free non ambigus (1) (§1 et §2) de même que leur fermeture commutative.

Dans le §3 on donne une caractérisation des context free arithmétiques relative à la forme f.

Dans le §4 on construit effectivement pour $n \leqslant 7$ une grammaire non ambiguë en se servant de propriétés de factorisation de l'ensemble des solutions de l'équation diophantienne $f(x) = d$ où $x \in \mathbb{N}^n$.

---

(1) on prend pour critère de non ambiguïté celui de CHOMSKY et SCHUTZEMBERGER (1963) : il existe une grammaire telle que la série formelle qu'elle engendre soit caractéristique de son support.

Dans le 5ème et dernier paragraphe on donne une caractérisation des langages bornés de GINSBURG à l'aide des réguliers arithmétiques et des context free arithmétiques.

§ 1 - PREUVE DE LA NON AMBIGUITE DES CONTEXT FREE ARITHMETIQUES

1.1. Pseudo langage de Dyck.

Définition :

Soit $\Sigma = \{\varepsilon_0\ \varepsilon_1\ \ldots\ \varepsilon_n\}$

$I = \{0,1,2,\ldots,n\}$ $I = J \cup K$, $J \cap K = \phi$

ϕ homomorphisme de Σ^* dans $\mathbb{Z}$ défini par $\phi(\varepsilon_i) = 1$ si $i \in J$
-1 si $i \in K$

$\varphi^{-1}(o)$ est un pseudo langage de Dyck.

Remarque :

Si μ est l'homomorphisme conservant la longueur de Σ^* dans $\{\varepsilon,\eta\}^*$

tel que : $\mu(\varepsilon_i) = \varepsilon$ pour $i \in J$

$\mu(\varepsilon_i) = \eta$ pour $i \in K$

$\mu(\varphi^{-1}(o))$ est le langage de Dyck sur $\{\varepsilon,\eta\}$

La propriété dont la démonstration suit est une propriété connue des langages de Dyck.

Définition :

Notons D^* le pseudo langage de Dyck $\varphi^{-1}(o)$ et D l'ensemble des mots premiers de D^*. $x \in D$ s'il n'a pas de facteur gauche (ou droit) propre dans D^*.

Propriété 111 :

Tout mot de D^* se factorise d'une manière unique en mots de D.

Preuve : On raisonne par récurrence sur la longueur de $x \in D^*$ $|x| = 2$ alors $x = \varepsilon_i \varepsilon_j$, $ij \in (J \times K) \cup (K \times J)$, x est premier.

Hypothèse de récurrence :

Si $|x| < n$ alors x se décompose d'une manière unique en produit de mots premiers.

Soit $x \in D^*$ et $|x| = n$ si x est premier il satisfait à la propriété. si x n'est pas premier alors il a un facteur gauche propre dans D^* y et z satisfont à l'hypothèse de récurrence ; alors x est un produit de mots premiers. La décomposition est unique sinon il existerait des mots premiers ayant un facteur gauche propre dans D^*.

Propriété 112 :

D est la réunion de 2 ensembles disjoints D_1 et D_2 et tout mot de D_1 (resp D_2) est de la forme $\varepsilon_i \times \varepsilon_j$ où $ij \in J \times K$ (resp $K \times J$) et $x \in D_1^*$ (resp D_2^*)

Preuve : Soit $y \in D$ $|y| > 2$ alors $y = \varepsilon_i \times \varepsilon_j$ et $x = \varepsilon_{i_1} \ldots \varepsilon_{i_p}$. On considère la suite $u_k = \phi(\varepsilon_i \varepsilon_{i_1} \ldots \varepsilon_{i_k})$; pour $k \leqslant p$ on a $|u_k - u_{k+1}| = 1$; si $\phi\, \varepsilon_i \varepsilon_j \neq o$ alors $u_k = o$ pour $k < p$ ce qui entraîne une contradiction. Donc $\phi(\varepsilon_i \varepsilon_j) = o$ ce qui entraîne $\phi(x) = o$ et $x \in D^*$. Si $ij \in J \times K$ (resp $K \times J$) alors $x \in D_1^*$ (resp D_2^*) sinon $y = \varepsilon_i x_1 \varepsilon_k + \varepsilon_\ell x_2 \varepsilon_j$ où $k, \ell \in K \times J$ (resp $J \times K$) et $\varepsilon_i x_1 \varepsilon_k$ est facteur gauche propre de y, ce qui est absurde.

Remarque :

La propriété : « tout mot de D_1 (resp D_2) est de la forme $\varepsilon_i \times \varepsilon_j$ où $ij \in (J \times K)$ (resp $K \times J$) et $x \in D_1^*$ (resp D_2^*)» est une définition récursive de D_1 (resp D_2).

Proposition 112 :

D^* possède une grammaire context free non ambiguë.

Preuve :

Soit la grammaire G_1 d'axiome S_1

$$S_1 \longrightarrow \sum_{ij \in J \times K} \varepsilon_i \, \varepsilon_j + \varepsilon_i \, S'_1 \, \varepsilon_\gamma$$

$$S'_1 \longrightarrow S'_1 \, S_1 + S_1$$

alors tout mot de $L(G_1)$ est de la forme $\varepsilon_i \times \varepsilon_j$ où $x \in \left[L(G_1)\right]^*$ par suite $L(G_1) = D_1$

G_1 est une grammaire non ambiguë : en suivant [Chomsky et Schutzenberger] (1963) on construit la série formelle de G_1 à partir des équations :

$$\rho = \sum_{i,j \in J \times K} \varepsilon_i \, \varepsilon_j + \varepsilon_i \, \rho' \, \varepsilon_j = f_1 \, (\rho, \rho')$$

$$\rho' = \rho' \rho + \rho \qquad = f_2 \, (\rho, \rho')$$

en construisant les polynômes successifs :

$$\rho^0 = o \; ; \; \rho'^0 = o \qquad \rho^i = f_1(\rho^{i-1}, \rho'^{i-1}) \; ; \; \rho'^i = f_2(\rho^{i-1}, \rho'^{i-1})$$

ρ^{∞} est la semie formelle engendrée par G_1 ; on se propose de montrer qu'elle est caractéristique de son support à savoir que G_1 est non ambiguë.

<u>Hypothèse de récurrence</u> :

pour $i < n$ ρ^i et ρ'^i sont caractéristiques de leurs supports. Le support de ρ^i est contenu dans D_1 celui de ρ'^i dans D_1^* pour $i = n$ on a :

$$\rho^n = \sum_{ij \in J \times K} \varepsilon_i \, \varepsilon_j + \varepsilon_i \, \rho'^{n-1} \varepsilon_j$$

ρ'^{n-1} étant caractéristique de son support il en est de même de

$\sum_{ij \in J \times K} \varepsilon_i \, \rho'^{n-1} \varepsilon_j$ dont le support est constitué de mots de D_1 de longueur supérieure à 2 ; $\sum_{ij \in J \times K} \varepsilon_i \, \varepsilon_j$ est caractéristique de son support dont les mots sont de longueur 2 et appartiennent à D_1 ; les supports des deux polynômes sont disjoints et dans D_1. Donc ρ^n est caractéristique de son support qui est dans D_1.

$$\rho'^n = \rho'^{n-1} \, \rho^{n-1} + \rho^{n-1}$$

les supports de $\rho'^{n-1} \, \rho^{n-1}$ et de ρ^{n-1} sont disjoints car $D_1 \, D_1^* \cap D_1 = \emptyset$ (le mot vide n'est pas dans D_1^*)

$\rho'^{n-1} \, \rho^{n-1}$ est caractéristique de son support sinon il existe $x \neq z$ dans D_1^* et $t \neq y$ dans D_1 tels que $xy = zt$ ce qui contredit la propriété III.

Par suite ρ'^n est caractéristique de son support contenu dans D_1^*

Conclusion :

ρ^{∞} est caractéristique de son support donc G_1 est non ambiguë.

On a la même propriété pour la grammaire G_2 qui génère D_2

$$S_2 \longrightarrow \sum_{ij \in K \times J} \varepsilon_i \, \varepsilon_j + \varepsilon_i \, S_2' \, \varepsilon_j$$

$$S_2' \longrightarrow S_2' \, S_2 + S_2$$

La grammaire G constituée des règles de G_1 et de G_2 et de $S \longrightarrow S_1 + S_2 + SS_1 + SS_2$ génère D^*, et est non ambiguë du fait de la propriété 111, donc D^* pseudo langage de Dyck est non ambigu.

1.2. Image inverse d'un ensemble de messages context free sur un code préfixe affine.

Définition :

(i) Rappel : $A \subset Y^*$ est un code si et seulement si A^* est un sous monoïde libre sur A

A^* est l'ensemble des messages du code

(ii) Code préfixe affine :

$M \subset Y^*$ est un code affine préfixe si et seulement si $M = \{W_o\} \cup A$ où $W_o \subset Y^*$ et $A \subset Y^*$ est un code possédant en outre la propriété que un mot de A n'a pas de facteur gauche propre dans A^* (A est une partie finie de Y^*).

N.B. Les précédentes définitions concernant Code et Code préfixe sont tirées de l'article de M. NIVAT dans "Automata theory" Ed Caianiello Academic Press.

(iii) <u>image inverse d'un ensemble de messages</u> :

Soit Σ un alphabet tel qu'il existe h une bijection de Σ sur A h s'étend canoniquement en un monomorphisme de Σ^* dans Y^* soit $L \subset W_o A^*$ L est un ensemble de messages sur M et $\phi^{-1}(L) \subset \Sigma^*$ en est l'image inverse, où $\phi(x) = w_o h(x)$, $x \in \Sigma^*$.

Propriété 121 :

Si M est un code affine préfixe il existe un automate d'états finis déterministe $\mathcal{A} = \{K, Y, \delta, s_o, F\}$ tel que : $F = \{q_o, q_1, \ldots, q_n\}$ ∘ si $A = \{w_1, \ldots, w_n\}$; $\mathcal{A}$ reconnait les messages de M et pour tout i $\quad o \leqslant i \leqslant n \quad \delta(s_o, f) = q_i$ équivaut à $f = x\, w_i$, x étant un message du code ou bien vide, $\left[\text{si } i = o\right]$

Preuve : soit $\mathcal{A}_i$ l'automate $(K_i, Y, \delta_i, s_i, q_i)$ qui reconnait w_i on suppose que K_i contient d_i tel que $\delta(d_i\ a) = d_i$ pour tout $a \in Y$ et que $\delta_i(s_i, f) = q_i$ équivaut à $f = w_i$.

On pose :

$$K = K_o - \{q_o\} \cup \left[(K_1 - \{q_1\}) \times \ldots\ldots \times (K_n - \{q_n\})\right] \cup \{q_o, q_1, \ldots, q_n\}$$

δ est défini par les égalités suivantes :

(o) $\quad \delta(p_o\ a) = \delta_o(p_o\ a)$ pour tout $a \in Y$, tout $p_o \in K_o - \{q_o\}$

$\delta\ (q_o\ a) = \delta(s_1, \ldots, s_n,\ a)$ pour tout $a \in Y$

(1) $\quad \delta(p_1, \ldots, p_n, a) = \delta_1(p_1, a), \ldots, \delta_n(p_n, a)$ pour tout a tel que $\delta_i(p_i\ a) \neq q_i$ pour tout i

$\delta(p_1, \ldots, p_n, a) = q_i$ si il existe i unique tel que $\delta_i(p_i\ a) = q_i$

$\delta(p_1 \ldots p_n, a) = d_o$ dans les autres cas

(2) $\delta(q_i,a) = \delta(s_1,\ldots s_n,a)$ pour tout i et tout $a \in Y$ soit donc $\mathcal{A} = \{K,Y,\delta,s_o,F\}$ où $F = \{q_o,q_1,\ldots,q_n\}$ alors w_o est facteur gauche de tout mot accepté par $\mathcal{A}$ sinon pour un x accepté par $\mathcal{A}$ et pour toute partie gauche z de x on a $\delta(s_o\ z) \neq q_o$ et $\delta(s_o\ z) \in K_o$ alors $\delta(s_o,x) \in K_o - \{q_o\}$ ce qui est absurde.

<u>Hypothèse de récurrence</u> :

Pour tout $f \in Y^*$ tel que $|f| \leqslant n$ $\left[\delta(q_o,f) = q_i\right]$ équivaut à $\left[f \in A^* \text{ et } w_i \text{ est facteur droit de } f\right]$. soit f et $|f| = n + 1$ tel que $\delta(q_o,f) = q_i$ alors f a un facteur gauche dans A sinon grâce aux règles (1) $\delta(q_o,f) \neq q_i$ pour tout i soit w_j tel que $f = w_j\ x$; on a $\delta(q_o\ w_j) = q_j$ et $\delta(q_j\ x) = \delta(q_o\ x)$ car w_j n'a pas de facteur gauche dans A^*. L'hypothèse de récurrence s'applique à x : $x \in A^*$ et w_i est facteur droit de x ; par suite $f \in A^*$ et w_i est facteur droit de f. <u>réciproquement</u> soit f et $|f| = n + 1$ tel que $f \in A^*$ et w_i facteur droit de f f se factorise d'une manière unique en mots de A $f = y\ w_i$, $y \in A^*$; $y = z\ w_j$. Par hypothèse de récurrence on a $\delta(q_o\ y) = q_j$ et par suite $\delta(q_o\ f) = \delta(q_j\ w_i) = q_i$. w_o étant facteur gauche de tout mot accepté par $\mathcal{A}$

On a :

$\delta(s_o\ f) = q_i$ entraîne $f = w_o\ g$ et $\delta(q_o,g) = q_i$ alors $g \in A^*$ et w_i est facteur droit de g ; donc f est un message du code M. Réciproquement, si f est un message du code M, alors $f = w_o\ g$ et $g \in A^*$ donc il existe j et w_j est facteur droit de g, alors $\delta(s_o\ f) = q_j$.

Lemme 122 :

Soit M langage régulier reconnu par un automate $\mathcal{A} = \{K, Y, \delta, s_o, q\}$ n'ayant qu'un état final et D un langage CF non ambigu sur Y alors $M \cap D$ possède une grammaire non ambiguë la démonstration de ce lemme est empruntée à [Guisburg] [théorème 321 p 88]

soit $G = (V, Y, P, \sigma)$ une grammaire non ambiguë de D à toute règle $\xi \longrightarrow \xi_1 \ldots \xi_k$ $(\xi, \xi_1, \ldots, \xi_k \in V)$ on associe les règles de $\bar{G}$

(a) $(p\xi p') \longrightarrow (p\xi_1 p_1)\ (p_1 \xi_2 p_2) \ldots\ldots (p_{k-1} \xi_k p')$ pour toute suite $p, p_1, p_2 \ldots p_{k-1}, p'$ d'éléments de K.

Pour tout $\alpha \in Y$

(b) $(p\alpha p') \longrightarrow \alpha$ est une règle de $\bar{G}$ si $\delta(p\alpha) = p'$ $\bar{G}$ est la grammaire $(\bar{V}, Y, \bar{P}, \bar{\sigma})$

où $\bar{\sigma} = (s_o, \sigma, q)$ $\bar{P}$ l'ensemble des règles (a) et (b)

$\bar{V} = Y \cup (K \times V \times K)$

si $a_1 \ldots\ldots a_t$ est généré par $\bar{G}$

$(s_o a_1 p_1) \ldots\ldots (p_{t-1}, a_t, q)$ tel que $\delta(p_i a_{i+1}) = p_{i+1}$ pour $o \leqslant i \leqslant t-1$

avec $p_o = s_o$

$p_t = q$

est l'unique chaîne sur $(K \times Y \times K)^*$ qui se dérive en $a_1 \ldots a_t$ il s'ensuit qu'à toute dérivation terminée de f dans $\bar{G}$ on peut associer une dérivation terminée de f dans G et $\delta(s_o\ f) = q$ d'une manière unique. La réciproque est vraie ; G étant non ambiguë il en est de même de $\bar{G}$ et $\bar{G}$ génère $D \cap M$.

Proposition 121 :

Si D est un langage non ambigu sur Y, $\{w_o\} \cup A$ un code affine préfixe dans Y^*, h le monomorphisme de Σ^* dans Y^* tel que $h(\Sigma^*) = A^*$, l'image inverse de $D \cap w_o A^*$ est un langage non ambigu sur Σ.

Preuve : Pour tout i on considère l'ensemble des messages du code qui sont dans D et qui admettent w_i comme facteur droit soit U_i si $\mathcal{A}_{q_i}$ est l'automate de la propriété 121 dans lequel on a réduit l'ensemble des états finaux à $\{q_i\}$ il est clair que $U_i = D \cap T(\mathcal{A}_{q_i})$ $\left[T(\mathcal{A}_i) \text{ est l'ensemble reconnu par } \mathcal{A}_{q_i}\right]$

et $D \cap w_o A^* = \bigcup_{o \leqslant i \leqslant n} U_i$ les U_i étant disjoints

et $\varphi = w_o h$ de Σ^* dans Y^* on a :

$\bar{\varphi}^1(D \cap w_o A^*) = \bigcup_{o \leqslant i \leqslant n} \bar{\varphi}^1(U_i)$ et les $\bar{\varphi}^1(U_i)$ sont disjoints il suffira de montrer que :

$\bar{\varphi}^1(U_i)$ est un langage non ambigu sur Σ

en utilisant le lemme 122 soit $\bar{G}_i$ la grammaire non ambiguë qui génère U_i on substitue aux règles terminales de $\bar{G}_i$ les règles suivantes :

(c) $(p\alpha p') \longrightarrow \varepsilon$ (mot vide) si $p\ \alpha\ p' \longrightarrow \alpha$ est une règle de $\bar{G}_i$ et si $p' \notin \{q_1 \ldots\ldots q_n\}$

(d) $(p\alpha p') \longrightarrow a_j \in \Sigma$ si $p\ \alpha\ p' \longrightarrow \alpha$ est une règle de $\bar{G}_i$ et si $p' = q_j$

on obtient une nouvelle grammaire $\bar{G}'_i$ sur le vocabulaire terminal Σ

on se propose de montrer que $\bar{G}'_i$ est une grammaire non ambiguë de $\bar{\varphi}^{1}(\cup_i)$

on considère l'ensemble des productions de $\bar{G}'_i$ susceptible de se dériver en Mots de Σ^* par application des règles (c) et (d) ces productions sont identiques à celles de $\bar{G}_i$ qui sont susceptibles de se dériver en mots de Y^* par application des règles terminales de $\bar{G}_i$ (productions presque terminales) si on note (q_j, w_i, q_i) la séquence

$$(q_j \alpha_1 p_1)\ (p_1 \alpha_2 p_2)\ \ldots\ldots\ (p_{k-1} \alpha_k q_i) \quad \text{telle que} \quad \delta(q_j w_i) = q_i$$
$$\text{et} \quad \alpha_1\ \alpha_2 \ldots \alpha_k = w_i$$

toute production presque terminale de $\bar{G}'_i$ se dérivant en un message du code dans $\bar{G}_i$ se factorise d'une manière unique en (1) $(s_o w_o q_o)\ (q_o w_{i_1} q_{i_1})\ \ldots\ldots\ (q_{i_{n-1}} w_{i_n} q_i)$ et se dérive en $a_{i_1}\ \ldots\ldots\ a_{i_n}$ dans $\bar{G}'_i$ alors $\bar{G}'_i$ est non ambiguë puisque $\bar{G}_i$ l'est il est clair que le langage généré par $\bar{G}'_i$ est $\bar{\varphi}^{1}(\cup_i)$ donc $\bar{\varphi}^{1}(\cup_i)$ est un langage non ambigu sur Σ par suite $\bar{\varphi}^{1}(D \cap w_o A^*)$ est un langage non ambigu sur Σ

<u>Proposition 122</u> :

Application aux context free arithmétiques et à leur fermeture commutative.

Les context free arithmétiques et leur fermeture commutative sont langages non ambigus.

Remarque préalable :

La proposition 121 se particularise à $D \cap w_0 w_1^* \dots w_n^*$ où $\{w_1, \dots, w_n\}$ est un code préfixe A dans Y^* à savoir :

soit $\Sigma = \{a_1 \dots a_n\}$ et h le monomorphisme de Σ^* dans Y^* tel que $h(a_i) = w_i$ soit $\varphi = w_0 h$ alors $\varphi^{-1} (D \cap w_0 w_1^* \dots w_n^*)$ est un langage non ambigu sur Σ

Il suffit de calquer la démonstration de la proposition 121 en utilisant l'automate $\mathcal{R}$ déduit de $\mathcal{A}$ en remplaçant les règles (2) de $\mathcal{A}$ par $\delta(q_i, a) = \delta(d_1, \dots, d_{i-1}, s_i, s_{i+1}, \dots, s_n, a)$ pour tout i et tout $a \in Y$ $\mathcal{B}$ reconnaît alors les messages du code affine préfixe $\{w_0\}$ A qui sont contenus dans $w_0 w_1^* \dots w_n^*$.

Preuve de la proposition 122 : Soit L context free arithmétique sur $\Sigma = \{a_1 \dots a_n\}$ alors $L = \{x \in a_1^* \dots a_n^* \mid f \circ \bar{\Psi}^1(x) = d\}$ si $\mu(x) = |a_1|_x, |a_2|_x \dots, |a_n|_x$ où $|a_i|_x$ est le nombre d'occurences de a_i dans x

$\bar{L} = \{x \in \Sigma^* \mid f \circ \mu(x) = d\}$ est la fermeture commutative de L

$Y = \{\varepsilon_0, \varepsilon_1, \dots, \varepsilon_n\}$

L'application h de Σ^* dans Y^* définie par $h(a_i) = \varepsilon_i^{\alpha_i}$ où $\alpha_i = |f_i|$ s'étend en un monomorphisme de Σ^* dans Y^* $h(\Sigma)$ est un code préfixe si $\varphi = \varepsilon_0^{|d|} h$, λ homomorphisme de Y^* dans $\mathbb{Z}$ défini par $\lambda(\varepsilon_i) = \frac{f_i}{|f_i|}$ pour $i \neq o$ $\lambda(\varepsilon_0) = \frac{d}{|d|}$, D^* le pseudo langage de Dyck $\{x \in Y^* \mid \lambda(x) = o\}$ non ambigu [voir 1.1]

alors on a $\bar{L} = \bar{\varphi}^{1}\left(D^{*} \cap \varepsilon_{o}^{|d|} \cdot h(\Sigma^{*})\right)$

et $L = \bar{\varphi}^{1}\left(D^{*} \cap \varepsilon_{o}^{|d|}\left(h(a_{1}^{*} \ldots . a_{n}^{*})\right)\right)$

et d'après la proposition 121 et la remarque précédente, L et $\bar{L}$ sont tous deux des langages context free non ambigus.

§ 2 - RECONNAISSANCE

Etant donné f forme linéaire sur $\mathbb{N}^{n}$ et $d \in \mathbb{Z}$ on se propose de décrire un automate à pile qui reconnait le langage

$$L = \{x \in a_{1}^{*} \ldots . a_{n}^{*} \mid f \circ \bar{\psi}^{1}(x) = d\}$$

Cet automate appartient à la famille des automates Push Down décrits par GINSBURG.

Notations :

Soit Σ alphabet d'entrée, Σ' alphabet de la pile, K ensemble des états, $F \subset K$ ensemble des états finaux, Z_{o} symbole distingué dans Σ' $s_{o} \in K$ état initial, δ une application de $(K \times \Sigma \times \Sigma') \cup (K \times \Sigma')$ dans les parties finies de $K \times (\Sigma')^{*}$;

$A = \{K, \Sigma, \Sigma', \delta, Z_{o}, s_{o}, F\}$ est un automate Push doun

(q,w,α) où $q \in K$, $w \in \Sigma^{*}$, $\alpha \in \Sigma'^{*}$ est une configuration de l'automate A : A est dans l'état q, w est un mot d'entrée α est le contenu de la pile.

Une transition de A est une relation ($\vdash$) entre 2 configurations de A elles sont de deux sortes :

(i) $(q, xw, \alpha Z) \vdash (q', w, \alpha f)$

$w \in \Sigma^*$ $x \in \Sigma$, $Z \in \Sigma'$, $\alpha \in \Sigma'^*$

si et seulement si $(q', f) \in \delta(q, x, Z)$ $f \in \Sigma'^*$

(ii) $(q, w, \alpha Z) \vdash (q', w, \alpha f)$

$w \in \Sigma^*$

si et seulement si $(q', f) \in \delta(q, Z)$

Une chaîne $y \in \Sigma^*$ est acceptée par A $\left[y \in T(A)\right]$ si et seulement si il existe une suite finie $c_1, \ldots\ldots, c_n$ de configurations telle que pour $1 \leqslant i < n$ $c_i \vdash c_{i+1}$ et $c_1 = (s_o, y, Z_o)$ $c_n = (q, \lambda, \alpha)$ où $q \in F$ et λ est le mot vide. $\left[\text{notation } c_1 \overset{*}{\vdash} c_n\right]$

On appelle Push doun arithmétique l'automate Push doun dont la définition suit $\left[\text{notation PD}\mathcal{A}\right]$

2.1. <u>Définition du PD$\mathcal{A}$</u>.

Soit f forme linéaire sur $\mathbb{N}^n$, ψ l'injection de $\mathbb{N}^n$ dans $a_1^* \ldots\ldots a_n^*$, $d \in \mathbb{Z}$ en adoptant les notations précédentes on considère

$$A = \{K, \Sigma, \Sigma', \delta, Z_o, s_o, F\}$$

dans lequel $K = \{q_1, \ldots\ldots, q_n, \bar{q}_1, \ldots\ldots \bar{q}_n\} \cup \{s_o, q_d\}$

$$\cup \{q_{ij}, 1 \leqslant i \leqslant n, 0 \leqslant j \leqslant |f \circ \bar{\psi}^1(a_i)|\}$$

$\Sigma = \{a_1, \ldots, a_n\}$; $\Sigma' = \{Z_o, \varepsilon, \eta\}$; $F = \{\bar{q}_1 \ldots\ldots \bar{q}_n\}$

δ est défini par les égalités suivantes :

(i) $\delta(s_o, Z_o) = (q_1, Z_o \varepsilon^{|d|})$ si $d > o$

$(q_1, Z_o \eta^{|d|})$ si $d < o$

$(\overline{q_1}, Z_o)$ si $d = o$

(ii) $\delta(q_i, a_j, u) = (q_{j, |f o \bar{\psi}^1(a_j)|}, u)$ pour $u \in \{\varepsilon, \eta\}$ et $1 \leqslant i \leqslant j \leqslant n$

$\delta(\overline{q_i}, a_j, Z_o) = (q_{j, |f o \bar{\psi}^1(a_j)|}, Z_o)$

(iii) $\delta(q_{i,o}, u) = (q_i, u)$ pour $u \in \{\varepsilon, \eta\}$

$\delta(q_{i,o}, Z_o) = (\overline{q_i}, Z_o)$ et $1 \leqslant i \leqslant n$

(iv) $\delta(q_i, a_j, u) = (q_d, u)$ pour tout $u \in \Sigma'$

$\delta(\bar{q}_i, a_j, Z_o) = (q_d, Z_o)$ et tout ij tels que $i > j$

$\delta(q_d, a_i, u) = (q_d\ u)$ pour tout i et tout $u \in \Sigma'$

(v) $\delta(q_{i,j}, u) = (q_{i,j-1}, g_i(u))$ pour tout $u \in \Sigma'$ et tout i,j tels que $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant |f o \bar{\psi}^1(a_i)|$

$g_i(u)$ est défini par le tableau suivant

	u	ε	η	Z_o
$g_i(u)$	$f o \bar{\psi}^{1}(a_i) > o$	λ	$\eta\eta$	$Z_o\eta$
	$f o \bar{\psi}^{1}(a_i) < o$	$\varepsilon\varepsilon$	λ	$Z_o\varepsilon$

(vi) pour tout $(q,(a_i),u)$ de $(K\times\Sigma\times\Sigma) \cup (K\times\Sigma')$ qui n'est pas partie gauche d'une des égalités précédentes

On a $\delta(q,(a_i),u) = \emptyset$ $\left[\text{la partie vide de } K\times\Sigma'^{*}\right]$

2.2. <u>Propriétés du PD $\mathcal{A}$</u>.

Il résulte de la définition de δ que le PD $\mathcal{A}$ est déterministe : pour toute configuration c s'il existe c' telle que $c \vdash c'$ alors c' est unique de plus pour toute chaîne $y \in \Sigma^*$ il existe une suite finie et unique de transitions qui conduit de la configuration initiale (s_o,y,Z_o) à une configuration finale qui peut être (q_i,λ,α), $(\bar{q}_i,\lambda,Z_o)$ où (q_d,λ,α) $\alpha \in \Sigma'^{*}$
[cette dernière propriété est appelée Loop Free par GINSBURG]
en effet soit $y = a_{i_1} \ldots\ldots a_{i_p}$ et il existe j tel que $i_1 \leqslant i_2 \leqslant \ldots \leqslant i_j$ et $i_j > i_{j+1}$ c'est le cas le plus général pour un mot de Σ^* ; si $y \in a_1^*\ldots a_n^*$ alors $j = p$

On a :

$$(s_o, a_{i_1} \dots a_{i_p}, Z_o) \vdash (q_1, a_{i_1} \dots a_{i_p}, Z_o\alpha_o) \overset{*}{\vdash} (q_{i_1,o}, a_{i_2} \dots a_{i_p}, Z_o\alpha_{i_1})$$

et pour tout $2 \leqslant k \leqslant j$.

$$(q_{i_{k-1},o}, a_{i_k} \dots a_{i_p}, Z_o\alpha_{i_{k-1}}) \overset{*}{\vdash} (q_{i_k,o}, a_{i_{k+1}} \dots a_{i_p}, Z_o\alpha_{i_k})$$

par une suite de transitions finie et unique

$$(q_{i_j,o}, a_{i_{j+1}} \dots a_{i_p}, Z_o\ {}_{i_j}) \overset{*}{\vdash} q_d,\ a_{i_{j+2}} \dots a_{i_p},\ Z_o\ \alpha_{i_j} \overset{*}{\vdash} (q_d, \lambda, Z_o\alpha_{i_j})$$

$(q_d, \lambda, \alpha_{i_j})$ est une configuration finale

si $y \in a_1^* \dots a_n^*$

alors $(q_{i_p,o}, \lambda, Z_o\ {}_{i_p}) \vdash (q_{i_p}, \lambda, Z_o\alpha_{i_p})$ si $\alpha_{i_p} \neq \lambda$

$\vdash (\bar{q}_{i_p}, \lambda, Z_o)$ si $\alpha_{i_p} = \lambda$

<u>Calcul sur la pile</u> :

Les seules transitions qui modifient la pile empruntent les égalités (v) et (i) il est clair d'après la définition de $g_i(u)$ que si αu est un mot du groupe libre (1) construit sur $\{\varepsilon, \eta\}$ il en est de même pour $\alpha\ g_i(u)$ de plus si φ est l'isomorphisme du groupe libre sur $\mathbb{Z}$ on a $\varphi\ \alpha\ g_i(u) = \varphi(\alpha u) \pm 1$ alors pour toute chaîne y sur Σ^* on a pour $1 \leqslant k \leqslant j$ α_{i_k} appartient au groupe libre et $\varphi(\alpha_{i_k}) = \varphi(\alpha_{i_{k-1}}) - f \circ \bar{\psi}^1 (a_{i_k})$; d'après les égalités (i)

$\varphi(\alpha_o) = d$

(1) pour une définition du groupe libre voir MARSHALL Hall Jr. Theory of groups Mac Millan Company N.Y. 1959

par suite

si $y \in a_1^* \ldots a_n^*$

on a $\varphi(\alpha_{i_p}) = d - f \circ \bar{\psi}^1(y)$

2.3. $T(A) = \{x \in a_1^* \ldots a_n^* \mid f \circ \bar{\psi}^1(x) = d\} = L$

On a vu que si $x \in a_1^* \ldots a_n^*$ alors

$(s_o, x, Z_o) \overset{*}{\vdash} (q_{i,o}, \lambda, Z_o\alpha)$ et $\varphi(\alpha) = d - f \circ \bar{\psi}^1(x)$

si $x \in L$ $d - f \circ \bar{\psi}^1(x) = o$ et $\varphi(\alpha) = o$ comme α appartient au groupe libre

on a $\alpha = \lambda$

donc

$(s_o, x, Z_o) \overset{*}{\vdash} (q_{i,o}, \lambda, Z_o) \vdash (\bar{q}_i, \lambda, Z_o)$

$\bar{q}_i \in F$ donc $x \in T(A)$

réciproquement si $x \in T(A)$

alors $x \in a_1^* \ldots a_n^*$ sinon la configuration finale est $(q_d, \lambda, Z_o\alpha)$ et x serait rejetée

alors $(s_o, x, Z_o) \overset{*}{\vdash} (q_{i,o}, \lambda, Z_o) \vdash (\bar{q}_i, \lambda, Z_o)$

donc $\alpha = \lambda$ comme $\varphi(\alpha) = d - f \circ \bar{\psi}^1(x)$ $x \in L$

2.4. Conséquences.

Il résulte des propriétés du PD$\mathcal{R}$ que si A' est l'automate déduit de A en remplaçant F par K-F A' reconnait $\Sigma^* - T(A)$ en effet pour tout $x \in \Sigma^* - T(A)$ il existe une suite finie et unique de transitions qui conduit de

(s_o, x, Z_o) à une configuration finale du type (q_f, λ, α) ou (q_d, λ, α) alors x est acceptée par A' la réciproque est évidente. A' est déterministe. En utilisant le résultat suivant de GINSBURB.

[Le langage reconnu par un automate Push doun déterministe est un langage context free non ambigu.]

On a T(A) et T(A') sont des langages context free non ambigus.

Ce qui entraîne :

Les context free arithmétiques sont des langages context free non ambigus [c'est exactement le résultat du paragraphe 1].

[Le complémentaire d'un langage context free arithmétique est un langage context free non ambigu.]

<u>Application à la fermeture commutative</u> :

Soit A_c l'automate déduit de A

en supprimant les égalités (iv) et en ajoutant aux égalités (ii)

[pour tout i,j $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant n$] à la place de [pour tout i,j tels que $1 \leqslant i \leqslant j \leqslant n$]

alors A_c a les propriétés suivantes

pour toute chaîne y sur Σ^* $\quad y = a_{i_1} \ldots a_{i_p}$

$$(s_o, a_{i_1} \ldots a_{i_p}, Z_o) \vdash (q_1, a_{i_1} \ldots a_{i_p}, Z_o\alpha_o) \overset{*}{\vdash} (q_{i_1,o}, a_{i_2} \ldots a_{i_p}, Z_o\alpha_{i_1})$$

$$\overset{*}{\vdash} (q_{i_{p-1}o}, a_{i_p}, Z_o\alpha_{i_{p-1}}) \overset{*}{\vdash} (q_{i_p,o}, \lambda, Z_o\alpha_{i_p}) \vdash (q_{i_p}, \lambda, Z_o\alpha_{i_p}) \text{ si } \alpha_{i_p} \neq \lambda$$

$$\vdash (\bar{q}_{i_p}, \lambda, Z_o) \text{ si } \alpha_{i_p} = \lambda$$

ceci d'une manière unique

de plus $\alpha_o\ \alpha_{i_1}, \ldots, \alpha_{i_p}$ appartiennent au groupe libre et

$\varphi(\alpha_{i_k}) = \varphi(\alpha_{i_{k-1}}) - f \circ \bar{\psi}^1(a_{i_k})$ pour $k \geqslant 1$, $\varphi(\alpha_o) = d$; ceci entraîne,

si $\mu(x) = |a_1|_x, \ldots, |a_n|_x$ (voir § 1), que $\varphi(\alpha_{i_p}) = d - f \circ \mu(x)$.

Il s'ensuit que $T(A_c) = \{x \in \Sigma^* \mid f \circ \mu(x) = d\}$.

Il est clair que $T(A_c)$ est la fermeture commutative de $T(A)$; on retrouve de la même manière que pour $T(A)$ le résultat du paragraphe 1.

[Si L est context free arithmétique, sa fermeture commutative est un langage context free non ambigu.]

en outre de même que $T(A') = \Sigma^* - T(A)$

on a $T(A'_c) = \Sigma^* - T(A_c)$

et A'_c est un automate déterministe

d'où

[Le complémentaire de la fermeture commutative d'un context free arithmétique est un langage context free non ambigu.]

§ 3 - CARACTERISATION DES CONTEXT FREE ARITHMETIQUE PAR LA FORME LINEAIRE QUI LES DEFINIT

3.1. On note $(\alpha_1, \ldots\ldots, \alpha_n)$ les coefficients de f

$L = \{x \in a_1^* \ldots a_n^* \mid f \circ \bar{\psi}^1(x) = d\}$; si les coefficients de f ne sont pas tous de même signe alors L n'est pas régulier s'il n'est pas vide.

Preuve : On n'a pas : $\alpha_i \leqslant o$ pour tout i ou $\alpha_i \geqslant o$ pour tout i

alors $\{1\ 2\ \ldots\ n\} = J \cup K$, $J \cap K = \emptyset$ et $i \in J \Longrightarrow \alpha_i > o$.

Supposons L non vide

alors il existe $p_1\ p_2\ \ldots\ p_n$ tels que $\Sigma\, \alpha_i\, p_i = d$

pour tout $t \in \mathbb{N}$ on considère $p_i(t) = p_i - \alpha_k\, t$ pour tout $i \in J$

$(\alpha_k < o)$, $p_k(t) = p_k + t \sum_J \alpha_i$, $p_j(t) = p_j$ pour tout $j \in k - \{k\}$.

il est clair que :

$$\Sigma\, \alpha_i\, p_i(t) = d \quad \text{et} \quad a_1^{p_1(t)} \ldots a_n^{p_n(t)} \in L \text{ pour tout } t \in \mathbb{N}$$

soit $\ell = \sup\ (J \cup \{k\})$ on considère pour tout $t \in \mathbb{N}$ les mots

$x(t) = \prod_{i<\ell} a_i^{p_i(t)}$; et la relation d'équivalence invariante à gauche sur

$$\Sigma^* \left[x \cup y \quad \text{ssi} \quad zx \in L \Longleftrightarrow zy \in L\right]$$

pour tout $z \in \Sigma^*$

si L était régulier cette relation serait d'indice fini.

Il n'en est rien car les mots $z(t)$ ne sont pas équivalents 2 à 2 (c'est-à-dire il y a une infinité de classes)

en effet $x(t)\ z(t) \in L$ et pour tout $u \neq t$ $x(t)\ z(t) \notin L$

$$z(t) = \prod_{i \geqslant \ell} a_i^{p_i(t)}$$

On montre que pour tout $u \neq t$

$$f\big(x(t)\ z(t)\big) - f\big(x(t)\ z(u)\big) \neq o$$

On a :

$$\sum_{i<\ell} \alpha_i\ p_i(t) + \alpha_\ell\ p_\ell(t) + \sum_{i>\ell} \alpha_i\ p_i - \sum_{i<\ell} \alpha_i\ p_i(t) - \alpha_\ell\ p_\ell(u) - \sum_{i>\ell} \alpha_i\ p_i$$

$= \alpha_\ell\big(p_\ell(t) - p_\ell(u)\big)$; α_ℓ est non nul et $p_\ell(t) - p_\ell(u)$ vaut $(u-t)\alpha_k$ ou bien $(t-u) \sum_J \alpha_i$ donc $f\big(x(t)\ z(t)\big) - f\big(x(t)\ z(u)\big) \neq o$ pour $u \neq t$ la relation $\sim$ n'est pas d'indice fini L n'est pas régulier.

3.2. Si les coefficients de f sont tous de même signe alors L est vide ou régulier.

<u>Preuve</u> :

(i) les α_i sont tous nuls

alors si $d = o$ $\quad L = a_1^* a_2^* \ldots a_n^*$ qui est régulier

si $d \neq o$ $\quad L = \emptyset$

(ii) les α_i ne sont pas tous nuls et L n'est pas vide ceci entraîne que d est du signe des α_i

soit $J = \{i,\ \alpha_i \neq o\}$

l'équation $\sum_J \alpha_i\ p_i = d$ a un nombre fini de solutions

à toute solution on associe le langage régulier

$\prod_i a_i^{f(p_i)}$ ou $f(p_i) = p_i$ si $i \in J$

$f(p_i) = x$ si $i \notin J$

alors $L = \bigcup_H \left(\prod_i a_i^{f(p_i)}\right)$ ou H est l'ensemble des solutions H étant fini L est régulier.

<u>Corollaire</u> :

L context free arithmétique est context free si et seulement si les coefficients de f ne sont pas tous de même signe (quand il n'est pas vide) autrement il est régulier.

3.3. Pour tout vrai context free arithmétique L il existe un régulier arithmétique borné L' tel que :

$$L = \emptyset \iff L' = \emptyset$$

<u>Corollaire</u> :

Pour tout vrai context free arithmétique L il existe un algorithme tel qu'après un nombre fini de pas on puisse décider si L est vide ou non.

<u>Preuve</u> : On ne s'intéresse qu'au cas où les α_i ne sont pas tous de même signe.

il existe k tel que $\alpha_k < o$ on pose $\gamma = -\alpha_k$

soit $J \cup K$ tels que $\alpha_i < o$ pour tout $i \in J$

$\alpha_i \leqslant o$ pour tout $i \in K$

et $J \cup K = \{1\ 2\ \ldots\ n\} - \{k\}$

soit d_γ le reste de la division de d par γ

on pose $f' = (\alpha'_1 \ldots\ldots \alpha'_n)$ où $\alpha'_i = \alpha_i$ pour $i \neq k$

$\alpha'_k = o$

$L' = \{x \in a_1^* \ldots a_n^* \mid f' \circ \bar{\psi}^1 x \underset{\gamma}{\equiv} d_\gamma\}$; $L' \in RA(a_1 \ldots a_n)$

[voir chapitre I]

(i) $L \neq \emptyset \Longrightarrow L' \neq \emptyset$

si $L \neq \emptyset$ il existe $p_1 \ldots\ldots p_n$ et

$\sum_{J \cup K} \alpha_i p_i = \gamma p_k + d = \gamma(\gamma + p_k) + d_\gamma$

donc $\sum_{J \cup K} \alpha_i p_i \underset{\gamma}{\equiv} d_\gamma$ d'où $f'(p_1 \ldots\ldots p_n) \underset{\gamma}{\equiv} d_\gamma$

donc $L' \neq \emptyset$

(ii) $L' \neq \emptyset \Longrightarrow L \neq \emptyset$

si $L' \neq \emptyset$ soit $p_1 \ldots\ldots p_n$ tel que $f'(p_1 \ldots\ldots p_n) \underset{\gamma}{\equiv} d_\gamma$

alors $\sum_{J \cup K} \alpha_i p_i = x\gamma + d_\gamma$ $x \in \mathbb{Z}$

soit $p'_i = p_i + \mu\gamma$ $\mu > o$ pour tout $i \in J$

$p'_i = p_i + \gamma$ pour tout $i \in K$

on a :

$$\sum_{J\cup K} \alpha_i p'_i = \sum_{J\cup K} \alpha_i p_i + \left(\mu \sum_J \alpha_i + \sum_K \alpha_i\right) \gamma$$

on peut choisir μ pour que $\mu \sum_J \alpha_i + \sum \alpha_i + x - y > 0$

on pose $p'_k = \mu \sum_J \alpha_i + \sum_K \alpha_i + x - y$

alors on a $\sum_{J\cup K} \alpha_i p'_i = p'_K \gamma + d$ donc $L \neq \emptyset$

§ 4 - CONSTRUCTION D'UNE GRAMMAIRE NON AMBIGUE D'UN CONTEXT FREE ARITHMETIQUE POUR $n < 7$

4.1. Notations définitions.

On considère $\mathbb{N}^n$ en tant que monoïde commutatif pour la somme soit A une famille finie dans $\mathbb{N}^n$ et $c \in \mathbb{N}^n$ l'ensemble $c + A^*$ est dit linéaire [A^* est l'ensemble des combinaisons linéaires à coefficients dans $\mathbb{N}$ des éléments de A.]

A est dit stratifiée (1) si tout élément de A a au plus 2 coordonnées non nulles,

[pour $u \in A$ on note $k(u)$ et $\ell(u)$ les rangs successifs de ses coordonnées]

et si pour tout couple u,v d'éléments de A on n'a pas croisement des coordonnées :

---

(1) la notion de stratification est empruntée à Ginsburg p 159.

On n'a pas $k(u) < k(v) < \ell(u) < \ell(v)$ ou $k(v) < k(u) < \ell(v) < \ell(u)$

A est <u>libre</u> si A^* est un monoïde libre sur A

$B \subset \mathbb{N}^n$ est une partie <u>non ambiguë</u> si elle est réunion finie d'ensembles linéaires disjoints chacune étant engendré par une famille libre et stratifiée.

Ordre sur une famille stratifiée A.

On considère $B = \{k(u), \ell(u) \mid u \in A\}$

et sur B l'ordre partiel : $i,\gamma \leqslant k,\ell$ ssi $k \leqslant i \leqslant \gamma \leqslant \ell$

et $k < \ell$

en ordonnant arbitrairement les sous-ensembles de A du type :

$\{u \in A \mid k(u), \ell(u) = k,\ell\}$ on a un ordre total sur A :

$u < v \iff k(u), \ell(u) < k(v), \ell(v)$ ou $k(u), \ell(u) = k(v), \ell(v)$ et $u < v$

dans $\{w \in A \mid k(w), \ell(w) = k(u), \ell(u)\}$

4.2. Soit B une partie non ambiguë dans $\mathbb{N}^n$ et ψ l'injection de $\mathbb{N}^n$ dans $a_1^* \ldots a_n^*$ alors $\psi(B)$ est un langage context free non ambigu dans $a_1^* \ldots a_n^*$.

On se bornera a en faire la preuve pour B linéaire engendré par une famille libre et stratifiée A_i: on exhibe une grammaire non ambiguë de $\psi(B)$.

Notations :

On considère la famille $A = \{u_1, \ldots, u_p\}$ munie de l'ordre partiel défini plus haut. $A_i = \{u_k \in A \mid u_k < u_i\}$; $\{u_{i_1}, \ldots, u_{i_r}\}$ est l'ensemble des maximaux dans A_i ; A étant stratifiée on a en notant $k_{i_1} = k(u_{i_1})$ $k_{i_1} \leqslant \ell_{i_1} \leqslant k_{i_2} \leqslant \ell_{i_2} \ldots \leqslant k_{i_r} \leqslant \ell_{i_r}$ soit $c = \alpha_1, \ldots, \alpha_n$ et $B = c + A^*$ on a $\psi(c) = a_1^{\alpha_1}, \ldots, a_n^{\alpha_n}$.

on pose :

$$x_{i,\gamma} = \prod_{i<k\leqslant\gamma} a_k^{\alpha_k} \qquad 1 \leqslant i \leqslant \gamma \leqslant n$$

$$y_{i\gamma} = \prod_{i\leqslant k\leqslant\gamma} a_k^{\alpha_k}$$

$$z_{i\gamma} = \prod_{i\leqslant k<\gamma} a_k^{\alpha_k}$$

A tout $u_i \in A$ on associe la variable S_i ; p_i et q_i sont les valeurs des coordonnées non nulles de u_i.

Preuve de 4.2 : Pour tout i $\quad 1 \leqslant i \leqslant p$ $\quad \{u_{i_1}, \ldots, u_{i_r}\}$ étant les maximaux dans A_i

On considère l'équation :

$$(i) \quad S_i = a_{k_i}^{p_i} S_i a_{\ell_i}^{q_i} + x_{k_i,k_{i_1}} S_{i_1} y_{\ell_{i_1},k_{i_2}} S_{i_2} y_{\ell_{i_2},k_{i_3}} \cdots\cdots S_{i_r} z_{\ell_{i_r},\ell_i}$$

si A_i est vide alors :

$$(i) \quad S_i = a_{k_i}^{p_i} S_i a_{\ell_i}^{q_i} + \prod_{k_i<m<\ell_i} a_m^{\alpha_m}$$

de plus $\{u_{o_1},\ldots, u_{o_r}\}$ étant la famille des maximaux dans A on considère l'équation

$$(o) \quad S = x_{o,k_{o_1}} S_{o_1} y_{\ell_{o_1},o_2} S_{o_2} \cdots\cdots S_{o_r} z_{\ell_{o_r},n+1}$$

on se propose de montrer que la première composante du système $\mathcal{S}$ d'équation (i) pour $o \leqslant i \leqslant p$ est caractéristique de son support $\psi(B)$

On raisonne par récurrence sur le nombre d'éléments de A

$|A| = 1$: le système $\mathcal{S}$ se réduit à

$$(o) \quad S = x_{o_{k_1}} S_1 z_{\ell_1,n+1}$$

$$(1) \quad S_1 = a_{k_1}^{p_1} S_1 a_{\ell_1}^{q_1} + \prod_{k_1<m<\ell_1} a_m^{\alpha_m}$$

Il est clair que la 1ère composante de la solution 5 est caractéristique de son support $\psi(c+u_1^*)$

Hypothèse de récurrence :

Pour $p < k$ $A = \{u_1, \ldots, u_p\}$ étant une famille libre et stratifiée munie de l'ordre défini plus haut, $c \in \mathbb{N}^n$ la première composante de la solution de $\mathcal{S}$ relatif à A et c est caractéristique de $\psi(c+A^*)$.

Démontrons qu'il en est de même pour $p = k$

soit $A = \{u_1, \ldots, u_k\}$ et $c \in \mathbb{N}^n$

① Il y a plus d'un maximal dans A

soit u_i le premier de la liste on pose $A_1 = A_i \cup \{u_i\}$, $A_2 = A - A_1$ soit $c_1 = (\alpha_1 \ldots \alpha_{\ell_i} \; o \; o \; o)$ et $c_2 = (o, o, o, \alpha_{\ell_{i+1}} \alpha \ldots \alpha_n)$ en vertu de l'hypothèse de récurrence.

La première composante de la solution de $\mathcal{S}_1$ relatif à A_1 et c_1 est caractéristique de $\psi(c_1+A_1^*)$, de même pour $\mathcal{S}_2$ et $\psi(c_2+A_2^*)$ φ_1 et φ_2 étant les parties droites des équations (o) de $\mathcal{S}_1$ et $\mathcal{S}_2$ le système $\mathcal{S}'$ formé de la réunion $\mathcal{S}_1 - (o) \cup \mathcal{S}_2 - (o)$ et de l'équation (o) $S = \varphi_1 \; \varphi_2$ n'est autre que le système $\mathcal{S}$ relatif à A et c φ_1 et φ_2 étant caractéristique de leurs supports il en est de même pour $\varphi_1 \; \varphi_2$ dont le support est $\psi(c_1+A_1^*)\,\psi(c_2+A_2^*)$ soit $\psi(c+A^*)$ (si $\varphi_1 \; \varphi_2$ n'était pas caractéristique alors A^* ne serait pas libre sur A).

② Il n'y a qu'un maximal dans A

soit u_i alors $A' = A - \{u_i\}$ et $\psi(c') = \prod_{k_i<m<\ell_i} a_m^{\alpha_m}$ satisfont à l'hypothèse de récurrence.

En ajoutant à $\mathcal{S}'$ l'équation $S = x_{o\,k_i}\, S_i\, y_{\ell_i n}$ et en substituant à l'équation (o) de $\mathcal{S}$ l'équation $S_i = a_{k_i}^{p_i}\, S_i\, a_{\ell_i}^{q_i} + \varphi_o$

où φ_i était le $2^{\text{ème}}$ membre de l'équation (o) dans $\mathcal{S}'$ on obtient le système $\mathcal{S}$ relatif à A et c

on a $\psi(c+A^*) = \{x_{o\,k_i}\, (a_{k_i}^{p_i})^r\, \psi(c'+A'^*)\, (a_{\ell_i}^{q_i})^r\, y_{\ell_i r} \mid r \in \mathbb{N}\}$ le membre de droite n'est autre que le support de la solution de $\mathcal{S}$ pour les mêmes raisons que précédemment (liberté de A^*). La solution de $\mathcal{S}$ est caractéristique.

On se propose dans ce qui suit de montrer pour $n \leqslant 7$ que les solutions de l'équation diophantienne $\sum_i^n \alpha_i\, p_i = d$ forment une partie non ambiguë de $\mathbb{N}^n$ en donnant une décomposition explicite.

[Ginsburg (1) a montré que pour tout langage context free borné L non ambigu dans $a_1^* \ldots a_n^*$ $\bar{\psi}^1(L)$ était une partie non ambiguë de $\mathbb{N}^n$ les context free arithmétiques étant des langages context free bornés non ambigus comme il a été montré dans les paragraphes précédents on a une preuve indirecte du fait que les solutions de $\sum_i \alpha_i\, p_i = d$ forment une partie non ambiguë de $\mathbb{N}^n$]

---

(1) lemme 612 p 188

4.3. Structure de l'ensemble des solutions de l'équation Diophantienne.

$$\sum_{i=1}^{n} \alpha_i p_i = d \quad , d, \alpha_i \in \mathbb{Z} \quad , 1 \leqslant i \leqslant n \ , \ n \leqslant 7$$

Notations :

D est l'ensemble des solutions de $\sum_{i=1}^{n} \alpha_i p_i = d$ dans $\mathbb{N}^n$

f est la forme $(\alpha_1, \ldots\ldots, \alpha_n)$

Π_j est la projection de $\mathbb{N}^n$ sur lui-même parallèlement à la $j^{\text{ème}}$ composante.

Définition :

Soit A une partie de la base canonique de $\mathbb{Z}^n$ considéré comme module sur $\mathbb{Z}$. A est stratifiable relativement à (f,j,B) $\left[1 \leqslant j \leqslant n,\ B \subset \Pi_j(\mathbb{N}^n)\right]$ si $A \subset \Pi_j(\mathbb{N}^n)$ et si $A' \cup B'$ est une famille libre et stratifiée A' étant définie par :

$$A' = \{|\alpha_j| \ a + |\alpha_{k(a)} \ | \ (o\ o\ o \underset{j}{|} o\ o) \ , \ a \in A\}$$

$k(a)$ est le rang de la coordonnée non nulle de $a \in A$ et B' par

$B' = \{|\alpha_j \ | \ b, \ b \in B\}$

Lemme 431 :

soit $u_K = \{x \in \Pi_j(\mathbb{N}^n), \ f(x) \geqslant K\}$

<u>si</u> il existe j tel que $\alpha_j < o$ et pour tout $K \in \mathbb{Z}$ u_k est réunion finie d'ensemble disjoint de la forme $c + A_1^* + A_2^*$ où A_2 est une partie de la base canonique dans $\Pi_j(|N^n)$ et A_2 est stratifiable relativement à (f, j, A_1).

<u>Alors</u> D est une partie non ambiguë de $|N^n$

<u>Preuve du lemme 431</u> : On suppose donc qu'il existe $j \mid \alpha_j < o$ on pose $\Gamma = |\alpha_j|$ et on considère la partition de $|N^n$ par la relation ρ :

$$p_1, \ldots\ldots, p_n \; \rho \; p'_1, \ldots\ldots, p'_n \quad \text{ssi} \quad p_i - p'_i \underset{\Gamma}{\equiv} o \quad \text{pour tout} \quad i \neq j$$

ρ est d'indice fini ; toute classe $\bar{v}_\ell$ est de la forme :

$$\bar{v}_\ell = v_\ell + \Gamma\left[\Pi_\gamma(|N^n)\right] + (o\ o\ o\ \underset{\substack{\uparrow \\ j}}{|}\ o\ o\ o)^* \quad \text{avec} \quad v_\ell = (u_\ell^1, \ldots\ldots \underset{\substack{\uparrow \\ j}}{o} \ldots\ldots u_\ell^n)$$

$$o \leqslant u_\ell^i < \Gamma \quad \text{pour tout} \quad i$$

on a évidemment $D = \underset{\ell}{u}(\bar{v}_\ell \cap D)$

<u>Structure de $\bar{v}_\ell \cap D$:</u>

N.B. Dans le cas qui nous intéresse $\psi(D)$ est context free alors il existe toujours j tel que $\alpha_j < o$ sinon $\psi(D)$ est régulier (cf paragraphe précédent).

Si $y \in \bar{v}_\ell \cap D$ alors $f(v_\ell) \underset{\Gamma}{\equiv} d_\Gamma$ où $d = z\Gamma + d_\Gamma$, $o \leqslant d_\Gamma < \Gamma$ soit t_ℓ le quotient de $f(v_\ell)$ par Γ on a $f(v_\ell) = t_\ell\,\Gamma + d_\Gamma$ l'ensemble $u_{z-t_\ell} = \{x \in \Pi_\gamma(|N^n),\ f(x) \geqslant z - t_\ell\}$ a la propriété suivante :

$$\bar{v}_\ell \cap D = \{v_\ell + \Gamma x + (o, o \ldots, f(x) - \underset{\substack{\uparrow \\ \gamma}}{(z-t_\ell)}, o \ldots o)\ ,\ x \in u_{z-t_\ell}.$$

en effet soit $y \in \bar{v}_\ell \cap D$

on a $y = v_\ell + \Gamma x + w(o\ o\ o \underset{\substack{\uparrow \\ \gamma}}{|}\ o\ o)$ et $f(y) = d$ $\quad w \in |N$

d'autre part

$$f(y) = f(v_\ell) + \Gamma f(x) - \Gamma w$$

$$= t_\ell\,\Gamma + d_\Gamma + \Gamma(z-t_\ell) + \Gamma\Big[f(x) - (z-t_\ell)\Big] - \Gamma(w) = d$$

$$= d + \Gamma\Big[f(x) - (z-t_\ell)\Big] - \Gamma w = d$$

alors $w = f(x) - (z-t_\ell)$ et $x \in u_{z-t_\ell}$

la réciproque est immédiate.

u_{z-t_ℓ} vérifiant les hypothèses du lemme 31

On a $u_{z-t_\ell} = {}_k\, v_k$ chaque v_k étant de la forme $c + A_1^* + A_2^*$ où A_2 est stratifiable relativement à (f, j, A_1).

En posant $D_k = \{v_\ell + \Gamma x + \left[f(x) - (z-t_\ell)\right] (o\ o\ o \underset{j}{|} o\ o\ o),\ x \in v_k\}$

on a $\bar{v}_\ell \cap D = \bigcup_k D_k$; les D_k étant disjoints puisque les v_k le sont.

$A_1' \cup A_2'$ étant la famille libre et stratifiée formée à partir de A_1 et A_2, soient φ_1, φ_2, φ_2' les éléments de A_1^*, A_2^*, $A_2'^*$ φ_2' est l'élément obtenu à partir de φ_2 par la substitution

$$a \longrightarrow |\alpha_\gamma|\ a + |\alpha_{k(a)}|\ (o\ o\ o \underset{j}{|} o\ o)$$

Montrons que D_k est engendrée par $A_1' \cup A_2'$

soit $y \in D_k$ on a $y = v_\ell + \Gamma x + \left[f(x) - (z-t_\ell)\right] (o\ o\ o \underset{j}{|} o\ o\ o)$

et $x \in v_k$

On pose $\xi = v_\ell + \Gamma c + \left[f(c) - (z-t_\ell)\right] (o\ o\ o \underset{\uparrow j}{|} o\ o\ o)$

$\xi \in D_k$ puisque $c \in v_k$

$(x = c + \varphi_1 + \varphi_2)$

par suite $y = \xi + \Gamma(\varphi_1 + \varphi_2) + f(\varphi_1 + \varphi_2)\ (o\ o\ o \underset{j}{|} o\ o\ o)$

φ_1' est l'élément de $A_1'^*$ obtenu par la substitution

$$a \longrightarrow |\alpha_\gamma| a$$

alors $y = \xi + \varphi'_1 + \varphi'_2$ car $f(\varphi_1) = o$

la réciproque est immédiate.

Les $\bar{v}_\ell \cap D$ étants disjoints D est donc réunion disjointe d'ensembles non ambigus D est non ambigu. φ .E.D.

Définition :

On appellera générateur tout élément u de $\mathbb{N}^n$ ayant au plus deux coordonnées non nulles et tel que $f(u) = o$
on note u_{ij} un générateur u tel que $k(u) = i$
$\ell(u) = j$

Remarque :

D étant l'ensemble des solutions de $\sum_{i=1}^{n} \alpha_i p_i = d$ si D est non ambiguë les éléments des familles libres et stratifiées qui l'engendrent sont des générateurs.

Lemme 432 :

$D = \{x \in \mathbb{N}^n, f(x) = d\}$ étant non ambiguë pour tout $n < r$ et tout g pour $n = r$ si f ou $-f$ satisfait à une des deux conditions suivantes :

1. les coefficients non nuls de f sont tous de même signe

2. il existe j, $1 \leqslant j \leqslant n$ tel que $\alpha_j < o$ et on peut ordonner l'ensemble des coefficients strictement positifs de f en une suite $\alpha_{i_1}, \ldots\ldots, \alpha_{i_p}$ telle que :

(i) $i_\ell < i_k < j \Longrightarrow \ell > k \qquad j < i_\ell < i_k \Longrightarrow \ell < k$

(ii) si u est un générateur dans $\Pi_j(|N^n)$, u et u_{ji_1} (ou u_{i_1j})

forment une famille stratifiée

si u est un générateur dans $\Pi_{i_1}\ \Pi_j(|N^n)$ u et u_{ji_2}

(ou $u_{i_{2j}}$) forment une famille stratifiée

si u est un générateur dans $\Pi_{i_{p-1}} \ldots \Pi_{i_1}\ \Pi_j(|N^n)$ alors u

et $u(j,i_p)$ forment une famille stratifiée.

<u>Alors</u> $D = \{x \in |N^n,\ f(x) = d\}$ est une partie non ambiguë

<u>Preuve du lemme 432</u> : si f vérifie ① alors comme on l'a vu dans le paragraphe précédent $D = \bigcup_{i \in I \text{ fini}} D_i$ les D_i étant disjoints,

$D_i = c_i + A_i^*$ ou A_i est une famille de générateurs ayant une seule coordonnée non nulle alors D est non ambiguë

si f vérifie ②

soit $j \leqslant r$ tel que $\alpha_j < o$ et $\alpha_{i_1}, \ldots, \alpha_{i_p}$ la suite ordonnéedes coefficients positifs de f possédant les propriétés (i) et (ii).

On se propose de montrer que les ensembles

$u_z = \{x \in \Pi_j(|N^r),\ f(x) \geqslant z\}$ satisfont aux hypothèses du lemme 31.

On notera dans ce qui suit $u_{i_\ell} = (o\ o\ o\ \underset{i_\ell}{1}\ o\ o\ o)$ élément de la base canonique de $\mathbb{Z}^r$; A_{i_ℓ} la famille $\{u_{i_1}, \ldots, u_{i_\ell}\}$ pour $1 \leq \ell \leq p$

pour $o \leq m \leq p$ on considère les ensembles C_m

$$C_o = B_o + A^*_{i_1}$$

$$C_m = B_m + A^*_{i_{(m+1)}} \quad \text{pour } o < m < p$$

$$C_p = B_p + A^*_{i_p}$$

avec $B_o = \{x \in \Pi_j(|N^r),\ z \leq f(x) < z + \alpha_{i_1}\}$

$$B_m = \{x \in \Pi_{i_m} \ldots\ldots \Pi_{i_1}\ \Pi_j(|N^r),\ z + \sum_{\ell=1}^{m} a_{i_\ell} \leq f(x) < z + \sum_{\ell=1}^{m+1} \alpha_{i_\ell}\}$$

$$B_p = \{x \in \Pi_{i_p} \ldots\ldots \Pi_{i_1}\ \Pi_j(|N^r),\ f(x) \geq z + \sum_{\ell=1}^{p} \alpha_{i_p}\}$$

(α) $$u_z = \{x \in \Pi_j(|N^r),\ f(x) \geq z = \bigcup_{m=o}^{p} C_m$$

il est évident que $\bigcup_{m=o}^{p} C_m \subset u_z$

démontrons l'inclusion inverse

pour tout $y \in u_z$ on considère la suite finie :

$y_o = y$

$y_\ell = \Pi_{i_\ell}(y_{\ell-1})$ si $f\left(\Pi_{i_\ell}(y_{\ell-1})\right) \geqslant z + \sum_{k=1}^{\ell} \alpha_{i_k}$

$y_\ell = x_\ell = y_{\ell-1} - \lambda_\ell u_{i_\ell}$ dans le cas contraire où λ_ℓ est tel que :

$$\begin{cases} z + \sum_{k=1}^{\ell-1} \alpha_{i_1} \leqslant f(x\ ) < z + \sum_{k=1}^{\ell} \alpha_{i_k} \\ \lambda_\ell \geqslant o \end{cases}$$

La suite s'interromp dès que $y_\ell = x_\ell$ pour $\ell \leqslant p$

si $y_\ell = x_\ell$ pour $\ell < p$ | alors $x_\ell \in \Pi_{i_{\ell-1}} \ldots\ldots \Pi_{i_1} \Pi_j\ (|N^r)$

et $z + \sum_{k=1}^{\ell-1} \alpha_{i_k} \leqslant f(x_\ell) < z + \sum_{k=1}^{\ell} \alpha_{i_k}$

donc $x_\ell \in B_{\ell-1}$ comme on a $y = x_\ell + \sum_{k=1}^{\ell} \lambda_k u_{i_k}$ $y \in C_{\ell-1}$

par suite $y \in \bigcup_{m=o}^{p} C_m$ [le raisonnement est identique si $\ell = p$]

β Les C_m sont disjoints

soit $y_1 \in C_{m_1}$ $y_2 \in C_{m_2}$ avec pour fixer les idées

$p \geqslant m_2 > m_1 \geqslant o$

On a $\quad z + \sum_{\ell=1}^{m_1} \alpha_{i_\ell} \leqslant f\, \Pi_{i_{m_1}} \ldots\ldots \Pi_{i_1}\, \Pi_j\, (y_1) < z + \sum_{\ell=1}^{m_1+1} \alpha_{i_\ell}$

et $\quad z + \sum_{\ell=1}^{m_2} \alpha_{i_\ell} \leqslant f\, \Pi_{i_{m_2}} \ldots\ldots \Pi_{i_1}\, \Pi_j (y_2)$

Comme $\quad m_2 \geqslant m_1 + 1 \quad$ on a

$$f\, \Pi_{i_{m_2}} \ldots\ldots \Pi_{i_1}\, \Pi_j (y_2) > f\, \Pi_{i_{m_2}} \ldots\ldots \Pi_{i_1}\, \Pi_j\, (y_1)$$

ce qui ne saurait être réalisé si $y_2 = y_1$ donc $C_{m_1} \cap C_{m_2} = \emptyset$

pour tout $m \quad o < m < p$ vérifie les hypothèses du lemme 31

pour tout $m \quad o < m < p$

On considère la forme $\bar{f}$ sur $\mathbb{Z}^{r-(m+1)}$ définie par

$\bar{f}\left[i(x)\right] = f(x)$ où i est l'injection de $\Pi_{i_m} \ldots\ldots \Pi_{i_1}\, \Pi_j\, (\mathbb{N})$

dans $\mathbb{N}^{r-(m+1)}$ alors $\bar{B}_m = \{x \in \mathbb{N}^{r-m+1},\ z + \sum_{\ell=1}^{m} \alpha_{i_\ell} \leqslant \bar{f}\, x < z + \sum_{\ell=1}^{m+1} \alpha_{i_\ell}\}$

$$\bar{B}_m = \bigcup_{o \leqslant t < \alpha_{i_{(m+1)}}} \bar{D}_t^{\,m}$$

avec $\quad \bar{D}_t^m = \{x \in \mathbb{N}^{r-(m+1)},\ \bar{f}\, x = z + \sum_{\ell=1}^{m} \alpha_{i_\ell} + t\}$

par hypothèse $\bar{D}_t^m$ est une partie non ambiguë dans $|N^{r-(m+1)}$

il est clair que l'application inverse de i conserve cette propriété

on pose $D_t^m = i^{-1}\,\bar{D}_t^m$

et on a $B_m = \bigcup_{0 \leqslant t < \alpha_{i(m+1)}} D_t^m$ les D_t^m étant disjoints puisque les $\bar{D}_t^m$ le sont.

Donc B_m est non ambiguë

tout générateur u de B_m étant tel que $f(u) = o$ il satisfait à la condition "u et $u_{i_{m+1},j}$ forment une famille stratifiée" $(u \in \Pi_{i_m} \ldots \Pi_{i_1} \Pi_j(|N^r))$

soit donc $c + A_1^*$ une composante de B_m, $c + A_1^* + A_{i_{m+1}}^*$ est une composante de C_m et on montre en prenant les notations du lemme 31 que $A_1' \cup A_{i_{m+1}}'$ est une famille libre et stratifiée.

Les éléments de A_1 sont des générateurs dans $\Pi_{i_\ell} \ldots \Pi_{i_1} \Pi_j |N^r$ pour $1 \leqslant \ell \leqslant m$ ils forment une famille stratifiée avec $u_{i_\ell,j}$ pour $1 \leqslant \ell \leqslant m+1$ en raison des hypothèses (ii) et du fait que A_1 est une famille stratifiée.

De plus $A_1 \cup \{u_{i_\ell j}, 1 \leqslant \ell \leqslant m+1\}$ est libre :

en effet $A_1 \cup \{u_{i_\ell,j},\ 1 \leq \ell \leq m\}$ est libre pour des raisons évidentes et on ne peut avoir $\lambda u^*_{i_{\ell+1},j} = x + \sum_{1\leq\ell\leq m} \lambda_\ell u_{i_\ell,j}$ avec $x \in A_1^*$

sinon $\sum_{1\leq\ell\leq m} \lambda_\ell u_{i_\ell,j} = 0$ et $\lambda u_{i_{\ell+1},j} = x$ ce qui est absurde x n'ayant pas de coordonnée de rang j

il en résulte que $A_1 \cup A'_{i_{(m+1)}}$ est libre et stratifiée ; par suite $A'_1 \cup A'_{i_{(m+1)}}$ est libre et stratifiée donc $A_{i_{m+1}}$ est stratifiable par rapport à (f,j,A_1) et C_m vérifie les hypothèses du lemme 431.

γ2 $m = p$.

$B_p = \{x \in \Pi_{i_p} \ldots \Pi_{i_1} \Pi_j(|N^r),\ f(x) \geq z + \sum_{\ell=1}^{p} \alpha_{i_\ell}\}$ est fini pour la raison que $\bar{f}$ a tous ses coefficients négatifs

alors C_p vérifie trivialement les hypothèses du lemme 431.

<u>Conclusion</u> :

$u_z = \{x \in \Pi_j(|N^r),\ f(x) \geq z\}$ vérifie les hypothèses du lemme 431 comme réunion finie d'ensembles disjoints qui les vérifient et par suite (lemme 431)

$D = \{x \in |N^r,\ f(x) = d\}$ est une partie non ambiguë.

φ.E.D.

<u>Remarques</u> :

- La démonstration précédente, pour autant que soient vérifiées les hypothèses (ii) pour f et pour $\bar{f}$ fournit un procédé de construction de D.

- pour qu'un élément u de $|N^r$ ayant deux coordonnées non nulles puisse être un générateur il faut et il suffit qu'on ait $\alpha_{k(u)}\,\alpha_{\ell(u)} < 0$

il suffit donc de connaître le signe des coefficients de f pour savoir si celle ci satisfait à la condition ② ii du lemme 432 on introduit donc la <u>configuration</u> de f qui est le n-uple des signes de ses coefficients notée cf. Comme on se propose de démontrer que f satisfait à la condition ② ii du lemme 432 par une récurrence finie on ne considérera pas le cas où f à des coefficients nuls ceci nous ramenant à la vérification de la condition ② ii pour des valeurs de n inférieures.

<u>Exemple</u> :

$$f = \alpha_1\,\alpha_2 \circ \alpha_4\,\alpha_5$$

alors si $\bar{f} = \alpha_1\,\alpha_2\,\alpha_4\,\alpha_5$ satisfait à la condition ② ii il en sera de même pour f.

- La condition ② i qui n'a pas servie à la démonstration au lemme 432 trouve son utilité dans la démonstration de la proposition qui suit.

<u>Proposition 433</u> :

Pour $\pi \leqslant 7$ si f ne vérifie pas la condition ① alors f ou $-f$ vérifie la condition ② .

433-1 :

Pour tout $n \geqslant 2$ si $p = 1$ tout f satisfait à la condition (2) i étant le rang du symbole $+$ dans c_f tout générateur a une coordonnée de rang i alors toute famille de générateurs est stratifiée.

La condition i est trivialement vérifiée.

433-2 :

Pour tout $n \geqslant 4$ si $p = 2$ tout f satisfait à la condition (2) dans c_f figure nécessairement la succession $s = - +$ ou $s = + -$ sinon les coefficients de f seraient tous de même signe.

soit j le rang du symbole $-$ de s

i_1 $+$ de s

i_2 le rang du symbole $+$ qui n'est pas dans s $(p=2)$

il est clair que si u générateur $\in \Pi_j(|N^n)$ alors u et $u_{i_1 j}$ forme une famille stratifiée

de même si $u \in \Pi_{i_1} \Pi_j(|N^n)$ alors u a nécessairement une de ses coordonnées de rang i_2 alors u et $u_{i_2,j}$ forment une famille stratifiée la condition (i) est trivialement vérifiée.

433-3 :

Pour $n = 5$ tout f satisfait à la condition (2)

en effet on a forcément $p = 1$ ou $p = 2$ en prenant c_f ou $\underline{c}_f$ suivant les cas c'est donc une conséquence de 433-1 et 433-2.

433-4 :

Pour $n = 5$ $p = 3$ et tout f tel que $p = p(c_f)$ il existe une suite régulière des rangs des symboles $+$ telle que f satisfasse à la condition ② .

Dans c_f il existe nécessairement une succession s de 3 symboles dont deux sont $+$ soit k le rang du 1er symbole de s

On a les configurations possibles

(i) $- + +$

(ii) $+ - +$

(iii) $+ + -$

en prenant

(i)	$i_1 = k + 1$	$i_2 = k + 2$	$j = k$
(ii)	$i_1 = k$	$i_2 = k + 2$	$j = k + 1$
(iii)	$i_1 = k + 1$	$i_2 = k$	$j = k + 2$

et pour i_3 le rang du symbole $+$ qui n'est pas dans s

on vérifie dans chaque cas que $i_1\ i_2\ i_3$ est régulière et que la condition ② est satisfaite.

433-5 :

Pour $n = 6$ $p = 3$ f satisfait à la condition ②

ⓐ α_1 et $\alpha_6 > o$

alors pour $\alpha_2, \alpha_3, \alpha_4, \alpha_5$ on a les configurations possibles

(i) $+ - - -$
(ii) $- + - -$
(iii) $- - + -$
(iv) $- - - +$

on vérifie que si

(i) $j = 3,\ i_1 = 2,\ i_2 = 1,\ i_3 = 6$
(ii) $j = 2,\ i_1 = 3,\ i_2 = 1,\ i_3 = 6$
(iii) $j = 5,\ i_1 = 4,\ i_2 = 6,\ i_3 = 1$
(iv) $j = 4,\ i_1 = 5,\ i_2 = 6,\ i_3 = 1$

La condition ② est satisfaite dans chaque cas

ⓑ α_1 ou $\alpha_6 < o$

on suppose $\alpha_1 < o$ ce qui ne restreint pas la généralité d'après 433-4 pour $\alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6$ on a $p = 3$ $n = 5$ et il existe j, i_1, i_2, i_3 tels que la condition ② soit satisfaite.

Les seuls nouveaux générateurs qu'on introduit sont $u_{1,i_1}, u_{1,i_2}, u_{1,i_3}$ ils sont dans $\Pi_j(|N^6)$ et forment avec $u_{j\ i_1}$ une famille stratifiée du fait que i_1, i_2, i_3 est régulière.

de même u_{1,i_2} et u_{1,i_3} sont dans $\Pi_{i_1} \Pi_j |N^6$ et forment avec u_{ji_2} une famille stratifiée.

Enfin $u(1,i_3) \in \Pi_{i_2} \Pi_{i_1} \Pi_j (|N^6)$ et forme avec $u(i_{3,j})$ une famille stratifiée donc f satisfait à la condition ②.

<u>433-6</u> :

Pour $n = 6$ tout f satisfait à la condition ②

Les seuls cas à considérer sont $p = 1, 2, 3$ et sont couverts par 331, 332 et 335'.

<u>433-7</u> :

Pour $n = 7$ $p = 3$ tout f satisfait à la condition ②

ⓐ α_1 ou $\alpha_7 < 0$

on suppose $\alpha_1 < 0$ ce qui ne restreint pas la généralité alors d'après 433-5' $\alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7$ satisfait à la condition ② et par le même raisonnement que 433-5' ⓑ on montre qu'il en est de même pour $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7$

ⓑ α_1 et $\alpha_7 > 0$

$\alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6$ présente les configurations possibles suivantes :

(i)	+ - - - -	on peut voir facilement que la condition ② R est
(ii)	- + - - -	satisfaite dans les cas (i) (ii) (iv) et (v)
(iii)	- - + - -	il n'en est pas de même dans le cas (iii)
(iv)	- - - + -	(+ - - + - - +)
(v)	- - - - +	

Pour tout choix de j et tout choix de i_1 il existe un générateur dans $\Pi_j(|N^7)$ qui n'est pas stratifié avec u_{ji_1}

La configuration opposée a de bonnes propriétés (- + + - + + -) la condition ② est satisfaite pour $j = 4$, $i_1 = 3$, $i_2 = 2$, $i_3 = 5$, $i_4 = 6$ par conséquent f satisfait à la condition ② .

<u>433-8</u> :

Pour $n = 7$ tout f satisfait à la condition ②
Les cas $p = 1, 2, 5, 6$ sont couverts par 433-1 et 433-2
Le cas $p = 3$ est couvert par 433-7 de même que le cas $p = 4$

Donc pour $n \leqslant 7$ f ou $-f$ satisfait à la condition ② φ.E.D.

<u>Proposition 434</u> :

Pour $n \leqslant 7$ $D = \{x \in |N^n, f(x)\ d\}$ est une partie non ambiguë et on sait la construire.

<u>Preuve</u> : Pour $n = 2$ D est une partie non ambiguë on le démontre dans le cas non trivial par exemple $f = (\alpha_1, \alpha_2)$ avec $\alpha_1 > o$ $\alpha_2 < o$ ce qui ne nuit pas à la généralité. On a $j = 2$; soit c le plus petit entier tel que $\alpha_1 c \geqslant z$ alors $u_z = c, o + (1, o)^*$ il est clair que les hypothèses du lemme 31 sont satisfaites, le procédé de la démonstration du lemme 31 permet de construire D.

Pour $n \leqslant 7$ on décompose u_z suivant le procédé du lemme 32 en ensembles c_i desquels on extrait les ensembles $u_z i$ relatifs à c_i auquels on applique le même traitement jusqu'à obtenir des ensembles u_z pour lesquels $\bar{f}$ est une forme sur $\mathbb{N}^2$ c'est possible puisque les $\bar{f}$ sont des formes sur $\mathbb{N}^k$ avec $k \leqslant 7$ et qu'elles satisfont toutes à la condition ② ou ① il ne reste plus qu'à appliquer le lemme 31 a chaque élément de la décomposition pour obtenir une décomposition explicite de D en ensembles linéaires disjoints engendrés par des familles libres et stratifiées.

Remarque :

Pour $n \geqslant 8$ et $2 < p < n - 2$ il existe des configurations c telles que la condition ② n'est satisfaite ni par c ni par son opposée.

Exemple :

$n = 8$	$p = 4$	$+ - + - + - + -$
$n = 9$	$p = 6$	$+ - ++ - ++ - +$

§ 5 - FACTORISATION DES LANGAGES CONTEXT FREE BORNES

Définition :

On étend la notion de context free arithmétique aux ensembles contenus dans $w_1^* w_2^* \ldots w_n^*$ où $w_1, w_2, \ldots, w_n$ sont des mots sur Y alphabet fini de la même manière que pour les réguliers arithmétiques du chapitre I.

$L \subset w_1^* \ldots w_n^*$ est context free arithmétique s'il est l'image par l'homomorphisme alphabétique $\varphi(a_i) = w_i$ pour $1 \leqslant i \leqslant n$ d'un context free arithmétique dans $a_1^* a_2^* \ldots a_n^*$.

Ce qui revient à dire que si ψ est l'application de $\mathbb{N}^n$ dans $w_1^* \ldots w_n^*$ telle que $\psi(p_1 \ldots p_n) = w_1^{p_1} \ldots w_n^{p_n}$ alors pour tout context free arithmétique L dans $w_1^* \ldots w_n^*$ il existe f forme linéaire sur $\mathbb{N}^n$ et $d \in \mathbb{Z}$ tels que $L = \psi \circ \bar{f}^1(d)$.

Remarque :

Les résultats des § 1 2 3 4 s'étendent aisément aux context free arithmétiques dans $w_1^* \ldots w_n^*$ où $\{w_1, \ldots, w_n\}$ est un code.

Context free pseudo-arithmétique :

L est context free pseudo-arithmétique dans $a_1^* \ldots a_n^*$ si $L = K \cap L'$ où K est régulier pseudo-arithmétique dans $a_1^* \ldots a_n^*$

L est context free arithmétique dans $a_1^* \ldots a_n^*$

de même que précédemment les context free pseudo-arithmétiques dans $w_1^* \ldots w_n^*$ sont les images homomorphes des context free pseudo-arithmétiques dans $a_1^* \ldots a_n^*$.

Remarque :

Si $\{w_1, \ldots, w_n\}$ est un code alors φ est un homomorphisme injectif et si L est context free pseudo-arithmétique dans $w_1^* \ldots w_n^*$ alors $L = K \cap L'$ où $K = \emptyset(K_1)$, $L' = \emptyset(L'_1)$ K_1 étant pseudo-arithmétique L'_1 étant context free arithmétique dans $a_1^* \ldots a_n^*$.

$$L = \varphi(L_1) = \varphi(K_1 \cap L'_1) = \varphi(K_1) \cap \varphi(L'_1)$$

<u>Lemme 51</u> :

L'image par ψ d'un sous monoïde affine de $\mathbb{N}^2$ n'ayant qu'un seul générateur est un context free pseudo-arithmétique dans $w_1^* \, w_2^*$

soit $A = (c_1 \; c_2) + (u_1, u_2)^*$ $\quad c_1, c_2, u_1, u_2 \in \mathbb{N}$

soit $f = (u_2, u_1)$ et $d = u_2 \, c_1 - u_1 \, c_2$

alors $A \subset \bar{f}^1(d)$

soit $B = \{x_1 \; x_2 \mid x_1 \geqslant c_1 \text{ et } x_1 \underset{u_1}{\equiv} c_1\}$

on a : $\bar{f}^1(d) \cap B = A$

ψ se factorise en $\varphi \circ \psi'$ est l'injertion dans $a_1^* \, a_2^*$ alors

$\psi'(A) = \psi'(\bar{f}^1(d)) \cap \psi'(B)$

$\psi'(B)$ est un régulier pseudo arithmétique dans $a_1^* \, a_2^*$

$\psi'(\bar{f}^1(d))$ est context free arithmétique dans $a_1^* \, a_2^*$ par suite $\psi(A) = \varphi \circ \psi'(A)$ est context free pseudo arithmétique dans $w_1^* \, w_2^*$

<u>Remarque</u> :

On n'a pas en général $\psi(\bar{f}^1(d)) \cap \psi(B) = \psi(\bar{f}^1(d)) \cap \psi(B) = \psi(\bar{f}^1(d) \cap B)$

prenons par exemple $w_1 = w_2$

et $(u_1 \; u_2)$ ayant un pgcd $t \neq 1$

soit (α, B) tel que $u_1 \, u_2 = t(\alpha, B)$ on pose :

$$p_1 = c_1 + (\lambda t+\Gamma)\alpha$$
$$p_2 = c_2 + (\lambda t+\Gamma)\beta \qquad \text{où} \quad 1 \leqslant \Gamma < t \qquad \lambda > 0$$

Il est clair que $w_1^{p_1} w_2^{p_2} \in \psi(\bar{f}^1(d))$ d'autre part en posant :

$$p'_1 = c_1 + \lambda t\ \alpha$$
$$p'_2 = c_2 + (\lambda t+\Gamma)\ \beta + \Gamma(\alpha) \qquad \text{on a} \quad p'_1\ p'_2 \in B$$

Or

$$w_1^{p'_1} w_2^{p'_2} = w_1^{p_1} w_2^{p_2}$$

alors

$w_1^{p_1} w_2^{p_2} \in \psi\ \bar{f}^1(d) \cap \psi(B)$ il n'existe pas $p''_1\ p''_2$ tels que

$w_1^{p_1} w_2^{p_2} = w_1^{p''_1} w_2^{p''_2}$ et tels que $p''_1\ p''_2 \in \bar{f}^1(d) \cap B$ donc

$$\psi\ \bar{f}^1(d) \cap \psi(B) \neq \psi(\bar{f}^1(d) \cap B)$$

<u>Produit étoile</u> :

On définit le produit étoile d'un sous-ensemble borné de Y^* par un sous-ensemble borné dans $w^*\ z^*$ w et z étant des mots sur Y de la façon suivante :

L est borné dans Y^*

$H \subset w^*\ z^*$

$$L * H = \{f\ g\ h \mid g \in L \quad f \in w^*,\ h \in z^*\}$$

et $fh, \in H$

N.B. Le produit étoile de Ginsburg est un cas particulier du précédent.

Proposition 51 :

Tout langage context free borné s'obtient comme réunion finie de produits de langages context free bornés possédant la propriété suivante :

Il existe $L_1, \ldots\ldots, L_p$ langages context free pseudo arithmétiques dans $w_1^{1*} w_2^{1*}, \ldots\ldots\ldots, w_1^{p*} w_2^{p*}$ où w_j^i est un mot dans Y^*

et $L = L_1 * L_2 * L_3 * \ldots * L_p$ (1)

(1) cette écriture doit s'entendre $\left(\left(\left(\left(L_1 * L_2\right) * L_3\right) * \ldots\right) * L_p\right)$

Preuve : On se sert du résultat de PARIKH amélioré par GINSBURG pour tout langage context free borné L dans $w_1^* \ldots w_n^*$ $\bar{\psi}^1(L)$ est réunion finie de sous monoïdes affines stratifiés (cf § 4).

La construction s'opère d'une manière analogue à celle du § 4 (démonstration de : si A est une partie non ambiguë alors $\psi(A)$ est non ambiguë) en utilisant le lemme 51.

BIBLIOGRAPHIE

GINSBURG : The Mathematical theory of context free langages
Mac Graw Hill 1966.

CHOMSKY and CHUTZENBERGER : Algebraic theory of context free langages dans computer Programming and Formal Systems North Holland 1963.

GINSBURG
SPANIER : Bounded régular Sets 1966.

CLIFFORD A.H.
PRESTON G.B. : The Algébraic theory of Semi groups Vol 1 (1961).
MARSHALL Hall, Jr. Theory of Groups
Mac Millan Company 1959

NIVAT Thèse 1967

NIVAT AUTOMATA THEORY Ed Caianiello
Academic Press.

VU

Grenoble, le

Le Président de la Thèse

VU

Grenoble, le

Le Doyen de la Faculté des Sciences

VU, et permis d'imprimer,

Le Recteur de l'Académie de GRENOBLE