SUR LA COMPLEXITE DU PRINCIPE DE TARSKI-SEIDENBERG

Joos HEINTZ
IAM
Viamonte 1636
(1055) Buenos Aires ARGENTINE

Marie-Françoise ROY
IRMAR
Université de Rennes I CEDEX
35042 Rennes CEDEX

Pablo SOLERNO
IAM
Viamonte 1636
(1055) Buenos Aires ARGENTINE

Résumé

Cet article est consacré à un algorithme d'élimination des quantificateurs dans les corps réels clos dont la complexité est simplement exponentielle en séquentiel (et polynomiale en parallèle) en le nombre de variables dès lors que le nombre d'alternances de quantificateurs est fixé.

Abstract

This paper is devoted to an algorithm for quantifier elimination in the real closed case which is of complexity single exponential in the number of variables in the sequential model (and polynomial in the parallel model) as soon as the number of alternations of quantifiers is fixed.

1. Introduction

Nous remercions Teresa Krick et Henri Lombardi pour l'aide qu'ils nous ont apportée et leurs nombreuses suggestions utiles concernant ce travail.

Avant d'énoncer plus précisément notre résultat, il est nécessaire de faire quelques rappels sur la géométrie algébrique réelle et la complexité d'algorithmes.

1.1. Rappels sur la géométrie algébrique réelle.

Pour les notions et preuves résumés dans ce paragraphe on peut voir [3].

Definition : Un *corps réel clos* est un corps ordonné où tout élément positif a une racine carrée et où tout polynôme de degré impair a une racine.

Dans tout l'article $\mathbf{A}$ désigne un anneau intègre et $\mathbf{R}$ un corps réel clos contenant $\mathbf{A}$.

Définition : Un *ensemble semi-algébrique* S de $\mathbf{R}^n$ défini sur $\mathbf{A}$ est un ensemble défini par une combinaison booléenne d'inégalités polynomiales à coefficients dans $\mathbf{A}$.

Beaucoup de constructions restent à l'intérieur du domaine semi-algébrique et on a les résultats suivants:

Théorème (théorème de projection): La projection d'un ensemble semi-algébrique de $\mathbf{R}^{n+1}$ défini sur $\mathbf{A}$ sur $\mathbf{R}^n$ est un ensemble semi-algébrique défini sur $\mathbf{A}$.

Théorème (composantes semi-algébriquement connexes) : Les composantes (semi-algébriquement) connexes d'un ensemble semi- algébrique défini sur $\mathbf{A}$ sont en nombre fini. Ce sont des ensembles semi-algébriques définis sur $\mathbf{A}$.

Le théorème de projection peut se reformuler en termes logiques. On doit pour celà introduire la définition suivante.

Définition : Une *formule du langage des corps ordonnés à paramètres dans* $\mathbf{A}$ est construite en un nombre fini d'étapes à partir des formules atomiques qui sont des égalités et inégalités portant sur des polynômes à coefficients dans $\mathbf{A}$, à l'aide des connecteurs logiques (ou, et, non) et de quantificateurs ($\exists$, $\forall$) portant sur les éléments du corps. Une *formule prénexe* est une formule où tous les quantificateurs apparaissent au début de la formule. Toute formule est équivalente à une formule prénexe.

Exemple : Si S est un ensemble semi-algébrique de $\mathbf{R}^n$, son adhérence peut-être décrite par la formule suivante

$$\{(x_1,...,x_n) \in \mathbf{R}^n \mid \forall \epsilon \geq 0 \exists (y_1,...,y_n) \in S \sum_{i=1,...,n} (x_i - y_i)^2 \leq \epsilon^2\}$$

Le théorème de projection admet l'importante conséquence suivante qu'on démontre par induction sur le nombre de quantificateurs:

Théorème (principe de Tarski-Seidenberg ou élimination des quantificateurs) : Un sous-ensemble de $\mathbf{R}^n$ défini par une formule du langage des corps ordonnés à paramètres dans $\mathbf{A}$ est un ensemble semi-algébrique défini sur A.

On en déduit le corollaire suivant.

Théorème (principe de transfert) : Soit $\mathbf{R}'$ un corps réel clos contenant $\mathbf{R}$ et Φ une formule du langage des corps ordonnés à paramètres dans $\mathbf{A}$. La formule Φ est vraie dans $\mathbf{R}$ si et seulement si elle est vraie dans $\mathbf{R}'$.

Notation : Si $\mathbf{R}'$ est un corps réel clos contenant $\mathbf{R}$ et S est un ensemble semi-algébrique de $\mathbf{R}^n$ on note $S(\mathbf{R}')$ le sous-ensemble de $\mathbf{R}'^n$ défini par la même formule sans quantificateurs que S. Le fait que cette notation ait un sens est une conséquence facile du principe de transfert.

1.2. Notions de complexité d'algorithme.

Les polynômes que nous considérons sont à coefficients dans $\mathbf{A}$. Les opérations arithmétiques considérées sur $\mathbf{A}$ sont les additions, multiplications, divisions exactes (lorsqu'on sait d'avance que le rapport est encore dans $\mathbf{A}$) et déterminations du signe d'un élément de $\mathbf{A}$ (dans $\mathbf{R}$).

On évaluera la complexité des algorithmes en considérant les paramètres suivants:

d le degré total des polynômes

t la taille des coefficients (dans le cas où $\mathbf{A} = \mathbf{Z}$)

s le nombre des polynômes

n le nombre de variables

m le nombre d'alternances de quantificateurs dans une formule prénexe

Les algorithmes sont décrits par des réseaux arithmétiques sur $\mathbf{A}$ ([12]).

Définition : La *complexité séquentielle* est la taille du réseau arithmétique donc le nombre d'opérations arithmétiques sur $\mathbf{A}$ nécessaires pour l'algorithme. La *complexité séquentielle binaire* (lorsque $\mathbf{A} = \mathbf{Z}$) est le nombre d'opérations binaires nécessaires pour l'algorithme. La *complexité parallèle* est la profondeur du réseau arithmétique.

Nous nous intéresserons aussi à la parallélisation des algorithmes.

Définition : On dit qu'un algorithme de complexité séquentielle polynomiale (resp. simplement exponentielle, resp. doublement exponentielle) est *bien parallélisable* si sa complexité paralèle est polylog (i. e. polynomiale dans le log) (resp. polynomiale, resp. simplement exponentielle). La largeur du réseau arithmétique qui définit l'algorithme est alors polynomiale (resp. simplement exponentielle , resp. doublement exponentielle).

Exemple : Les calculs de déterminants sont bien parallélisables: on connaît un algorithme pour le le calcul des déterminants d'une matrice à coefficients dans $\mathbf{A}$ de taille x par un réseau arithmétique de profondeur $\log^2(x)$ et de taille polynomiale en x ([2]).

Ce résultat permet de bien paralléliser tous les calculs à base d'algèbre linéaire.

1.3. Nos résultats

Le résultat essentiel de cet article est le suivant:

La réponse au problème de l'élimination des quantificateurs (principe de Tarski-Seidenberg) est donnée par un algorithme de complexité doublement exponentielle en m bien parallélisable.

Ce résultat a déjà été annoncé dans [19](voir aussi [28]et [24]).

La preuve repose sur une nouvelle démonstration du résultat suivant :

On peut décider si S est vide en complexité séquentielle simplement exponentielle, en complexité parallèle polynomiale en n ([4]; voir aussi [23]et [18]).

Cet énoncé d'élimination des quantificateurs est similaire à un résultat récent dans le cas algébriquement clos ([7],[11]).

Les problèmes mathématiquement significatifs ont un petit nombre d'alternances de quantificateurs. Les résultats ci-dessus signifient qu'ils sont résolubles en complexité simplement exponentielle (polynomiale en parallèle). C'est le cas par exemple du calcul de l'adhérence ou de l'intérieur d'un ensemble semi-algébrique, de la distance entre entre deux ensembles semi-algébriques, etc...

1.4. Historique du sujet.

Le théorème d'élimination des quantificateurs est dû à Tarski et Seidenberg et a été publié au début des années 50 ([31],[26]). La complexité des algorithmes qu'on peut déduire de leurs démonstrations est hyper exponentielle.

Dans les travaux de Collins datant des années 70 on utilise les progrès du calcul formel pour obtenir un algorithme séquentiel (non parallélisable) polynomial en d, s, t doublement exponentiel en n ([6]).

Les résultats de Ben-Or Kozen Reif ([1]) complétés par Fitchas Galligo Morgenstern ([10]) donnent un algorithme doublement exponentiel en n, bien parallélisable.

Les travaux de Weispfenning ([34]) et Davenport-Heintz ([9]) fournissent une borne inférieure doublement exponentielle en n en séquentiel, (simplement exponentielle en n en parallèle [10]).

Les résultats de Grigor'ev Vorobjov ([16]) permettent de décider si S est vide en complexité simplement exponentielle. Leurs algorithmes ne sont pas bien parallélisables. Enfin Grigor'ev ([15]) montre que le problème de la décision peut se résoudre en complexité doublement exponentielle dans le nombre d'alternances de quantificateurs m. Dans cet article, Grigor'ev conjecturait que le théorème est vrai aussi pour l'élimination des quantificateurs.

Nous répondons à cette question de Grigor'ev par des algorithmes bien parallélisables. Un résultat de même nature que le notre a été obtenu indépendamment par J. Renegar (voir [24]).

1.5. Plan de l'article

Dans la section 2 nous étudierons des algorithmes de base sur les polynômes en une variable. Il s'agit de déterminer par des algorithmes en temps polynomial bien parallélisables les signes d'une liste de polynômes aux racines rélles d'un polynôme. On utilise la méthode des sous-résultants et la suite de Stum-Habicht, dont les méthodes de calcul sont basées sur l'algèbre linéaire.

La section 3 est le coeur de l'article. On y étudie la complexité des ensembles semi-algébriques. Après avoir énoncé les résultats et fait une première réduction dans le paragraphe 1, on opère dans le paragraphe 2 une réduction géométrique au cas d'une hypersurface qui s'inspire largement de [16]. On étudie alors les extremas d'une fonction bien choisie sur cette hypersurface et on est ramené à considérer un nombre fini de points critiques. On peut alors projeter ce nombre fini de points sur une droite bien choisie, et on a à résoudre un problème en une variable qu'on traite par les méthodes de la section 2. On fait appel à quelques notions de géométrie différentielle et à la géométrie semi-algébrique sur des corps réels clos non archimédiens. L'algorithme et sa complexité sont précisés dans le paragraphe 3. On utilise pour celà des conséquences récentes du Nullstellensatz effectif (élimination des quantificateurs dans le cas algébriquement clos ([11]). Nos algorithmes sont essentiellement différents de ceux de [16]. Une conséquence de notre algorithme est une nouvelle démonstration donnée dans la section 4 d'un résultat de Vorobjov sur le diamètre d'une boule rencontrant toutes les composantes d'un ensemble semi-algébrique.

La section 4 est consacrée à l'élimination d'un seul quantificateur. Il s'agit de paramétriser les résultats de la section 2, et l'obtention de la complexité souhaitée passe par les résultats de la section 3.

Enfin le résultat principal est démontré dans la section 5. Il s'agit essentiellement de réinterpéter l'algorithme de la section 3 dans une situation avec paramétres. Les résultats de la section 4 sont utilisés dans la dernière étape de l'algorithme. Il reste ensuite à faire une induction sur le nombre de blocs de quantificateurs de la formule.

2. Les polynômes en une variable

2.1. Le problème de base

L'appproche classique semi-numérique où on approche les racines par dichotomie ([6]) n'est pas bien parallélisable. Des recherches d'éléments primitifs et des algorithmes de factorisation y sont nécessaires. De plus elle ne fonctionne pas dans le cas non archimédien.

Ici nous adoptons une approche purement formelle, basée sur une généralisation du théorème de Sturm et des algorithmes d'algèbre linéaire.

Tout est ramené au problème de base suivant:

$P, Q_1, ..., Q_s$ étant des polynômes en une variable à coefficients dans **A**

(*) "déterminer les signes des polynômes $Q_1, ..., Q_s$ aux racines de P dans **R**".

On note d un entier supérieur aux degrés de P et des $Q_1, ..., Q_s$. Nous allons voir qu'on a une solution à (*) en complexité polynomiale en s et d, bien parallélisable.

Pour cela on fait appel à la théorie des sous-résultants.

2.2. Les sous-résultants

Définition : Le *déterminant polynomial* d'une matrice M à m lignes et n colonnes avec $n \geq m$ est le polynôme de degré inférieur ou égal à $n - m$ dont le coefficient de X^j est le déterminant de la matrice carrée obtenue en rajoutant aux $m - 1$ premières colonnes de M celle de numéro $n - j$.

Soit P un polynôme de degré p et Q un polynôme de degré $\leq q$. Le *polynôme sous-résultant* $S_j(P, p, Q, q)$, $j < \inf(p, q)$, est le déterminant polynomial de la matrice dont les lignes sont les polynômes

$$X^{(q-1-j)}P, ..., XP, P, X^{(p-1-j)}Q, ..., XQ, Q$$

dans la base $X^{(p+q-1-j)}, .., X, 1$. Les sous-résultants sont nuls ou égaux (à une constante multiplicative près) aux restes successifs de P et Q (voir [22]ou [13]ou [14]).

2.3. La suite de Sturm-Habicht

La suite de Sturm-Habicht ([13],[14]) se déduit par des changements de signe systématiques de la suite des sous-résultants de $P'Q$ et de P.

Définition : Soit P un polynôme de degré p à coefficients dans **A**. On notera $\mathrm{cd}(P)$ le coefficient dominant de P. La *suite de Sturm-Habicht de* P est définie par

$$\mathrm{StHa}_p(P) = \mathrm{cd}(P)P$$

$$\mathrm{StHa}_{p-1}(P) = \mathrm{cd}(P)P'$$

$$\mathrm{StHa}_j(P) = \frac{1}{\mathrm{cd}(P)}(-1)^{\frac{(p-j)(p-j-1)}{2}} S_j(P, p, P', p-1)$$

pour $j \leq p - 2$.

Soit Q un autre polynôme à coefficients dans **A** de degré q supérieur ou égal à 1. La *suite de Sturm-Habicht de* P *et* Q est définie par

$$\mathrm{StHa}_p(P, Q) = \mathrm{cd}(P)P$$

$$\mathrm{StHa}_j(P, Q) = \frac{1}{\mathrm{cd}(P)}(-1)^{\frac{(p-j)(p-j-1)}{2}} S_j(P, p, P'Q, p+q-1).$$

pour $j \leq p - 1$

La suite de Sturm-Habicht est à coefficients dans **A**. Les coefficients de Sturm-Habicht $\mathrm{stha}_j(P, Q)$ sont les coefficients de X^j dans $\mathrm{StHa}_j(P, Q)$.

On a la variante suivante du théorème de Sturm-Sylvester ([29],[30]).

Définition : Soit l une liste de polynômes à coefficients dans A. On définit $v(l)$ comme la différence entre le nombre de variations de signes dans l en $-\infty$ et en $+\infty$.

Proposition 1 : Soient P et Q deux polynômes à coefficients dans **A**.

La quantité $v(StHa(P,Q))$ est la différence entre le nombre $c_{>0}(P,Q)$ des racines de P dans **R** rendant $Q > 0$ et le nombre $c_{<0}(P,Q)$ des racines de P dans **R** rendant $Q < 0$. D'autre part $v(StHa(P,Q))$ ne dépend que des signes des $\text{stha}_j(P,Q)$.

Démonstration: voir [13]ou [14].

Corollaire : On a l'égalité

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} c_{=0}(P;Q) \\ c_{>0}(P;Q) \\ c_{<0}(P;Q) \end{pmatrix} = \begin{pmatrix} s(P,1) \\ s(P,Q^2) \\ s(P,Q) \end{pmatrix}$$

qui permet de calculer $c_{>0}(P,Q), c_{<0}(P,Q)$ et $c_{=0}(P,Q)$.

Remarque : On peut remplacer si on le souhaite les calculs de suite de Sturm-Habicht par des calculs de signature de formes quadratiques (méthode d'Hermite) comme dans [10]. Il est en fait montré dans la dernière section de [14]que les déterminants à calculer dans les deux cas sont identiques.

Les bonnes propriétés de spécialisation de la suite de Sturm-Habicht énoncées dans la proposition suivante nous serons utiles dans la section 4.

Proposition 2 : Soient $P(Y_1,...,Y_k)(X)$ et $Q(Y_1,...,Y_k)(X)$ des polynômes de $\mathbf{A}[Y_1,...,Y_k][X]$ et p le degré de P. Soient $y_1,...,y_k$ des points de $\mathbf{R}^k$ tels que le degré de $P_1 = P(y_1,...,y_k)(X)$ soit encore p. Notons $Q_1(X) = Q(y_1,...,y_k)(X)$. La suite l obtenue en substituant $y_1,...,y_k$ à $Y_1,...,Y_k$ dans la suite de Sturm-Habicht de P est telle que $v(l)$ est la différence entre le nombre $c_{>0}(P_1,Q_1)$ des racines de P_1 dans **R** rendant $Q_1 > 0$ et le nombre $c_{<0}(P_1,Q_1)$ des racines de P_1 dans **R** rendant $Q_1 < 0$.

Démonstration: voir [13]ou [14].

2.4. La solution au problème de base

Revenons à notre problème (∗).

Proposition 3 : Soient $P, Q_1,...,Q_s$ des polynômes en une variable de degrés inférieur ou égal à d à coefficients dans **A**. Il existe un réseau arithmétique sur **A** de taille polynomiale en d et s et de profondeur polylog en d et s qui résoud le problème

(∗) "déterminer aux racines de P dans **R** les signes des polynômes $Q_1,...,Q_s$ ".

Démonstration: En utilisant la proposition 1 et en calculant les 3^s suites de Sturm correspondant à tous les produits possibles des Q_i et Q_i^2 on peut résoudre le problème (∗) grâce à une matrice de taille 3^s en complexité polynomiale en d, mais exponentielle en s. En utilisant le fait que, parmi les 3^s conditions de signes possibles, au plus d sont réalisées on peut éviter la croissance exponentielle du nombre de suites de Sturm à calculer en limitant la taille de la matrice à $3d$ à chaque étape (voir [1]ou [25]). Ceci se parallélise bien car on peut séparer les polynômes en paquets de taille égale et faire log$s$ étapes ([1]). On obtient la complexité parallèle et séquentielle sont correctes car les calculs de polynômes de la suite de Sturm-Habicht sont des calculs de déterminants extraits de matrices de Sylvester (cf 1.2.).

Corollaire 1 : Soient $P, Q_1,...,Q_s$ des polynômes en une variable de degrés inférieur ou égal à d à coefficients entiers de taille inférieure ou égale à t. On peut résoudre le problème

(∗) "déterminer aux racines de P dans **R** les signes des polynômes $Q_1,...,Q_s$ " par un algorithme de complexité séquentielle binaire polynomiale en d ,s et t.

Démonstration: La taille des déterminants considérés qui sont tous extraits de matrices de Sylvester est polynomiale en d, s et t.

2.4. Le problème de base pour le corps des séries de Puiseux

Définition : Si ϵ est une variable on note $\mathbf{R}\langle\epsilon\rangle$ le *corps des séries de Puiseux en ϵ à coefficients dans* $\mathbf{R}$ dont les éléments sont les séries

$$\sum_{i\geq i_0, i\in\mathbf{Z}} a_i \epsilon^{\frac{i}{q}}$$

avec $i_0 \in \mathbf{Z}$, $a_i \in \mathbf{R}$, $a_{i_0} \neq 0$ et $q \in \mathbf{N}$ ([33]). On définit $a_k = 0$ pour $k < i_0$.

Le corps $\mathbf{R}\langle\epsilon\rangle$ est réel clos ([3]). L'élément ϵ y est infiniment petit positif (positif et plus petit que toutes les éléments de $\mathbf{R}$ positifs). Les éléments de $\mathbf{R}\langle\epsilon\rangle$ bornés sur $\mathbf{A}$ forment un *anneau de valuation* noté $\mathbf{V}(\epsilon)$. On note eval l'application de $\mathbf{V}(\epsilon)$ dans $\mathbf{R}$ qui à

$$\sum_{i\geq i_0, i\in\mathbf{Z}} a_i \epsilon^{\frac{i}{q}}$$

avec $i_0 \in \mathbf{N}$, $a_{i_0} \neq 0$ associe a_0.

Dans le corps réel clos $\mathbf{R}' = \mathbf{R}\langle\frac{1}{\Omega}\rangle$ l'élément Ω est infiniment grand positif (positif et plus grand que les éléments positifs de $\mathbf{R}$). L'ordre induit sur le corps $\mathbf{R}(\Omega)$ par $\mathbf{R}'$ est noté $+\infty$. Les éléments de $\mathbf{R}'$ bornés sur $\mathbf{R}$ sont les éléments de la forme

$$\sum_{i\geq i_0, i\in\mathbf{Z}} a_i (\frac{1}{\Omega})^{\frac{i}{q}}$$

avec $i_0 \in \mathbf{N}$, $a_i \in \mathbf{R}$, $a_{i_0} \neq 0$. Les éléments de $\mathbf{R}'$ bornés sur $\mathbf{A}$ forment un anneau de valuation noté $\mathbf{V}'$. On note eval' l'application de $\mathbf{V}'$ dans $\mathbf{R}$ qui à

$$\sum_{i\geq i_0, i\in\mathbf{Z}} a_i (\frac{1}{\Omega})^{\frac{i}{q}}$$

avec $i_0 \in \mathbf{N}$, $a_{i_0} \neq 0$ associe a_0.

Corollaire 2 : Soient $P, Q_1, ..., Q_s$ des polynômes de $\mathbf{A}[\Omega][X]$ de degrés inférieur ou égal à d. Il existe un réseau arithmétique sur $\mathbf{A}$ de taille polynomiale en d et s et de profondeur polylog en d et s qui résoud le problème

(*) "déterminer aux racines de P dans $\mathbf{R}'$ les signes des polynômes $Q_1, ..., Q_s$ "

Démonstration: On utilise le fait que tous les calculs de suite de Sturm-Habicht sont des calculs de déterminants extraits de matrices de Sylvester et se font dans l'anneau $\mathbf{A}[\Omega]$.

La morale de ce corollaire est la suivante: la réponse au problème (*) en $\Omega = +\infty$ ou en $\Omega = 0$ s'obtient par des algorithmes dont les complexités sont de même nature (ramenés en opérations arithmétiques sur $\mathbf{A}$).

On aura aussi besoin des résultats suivants :

Corollaire 3 : Soient $P, Q_1, ..., Q_s$ des polynômes de $\mathbf{A}[\Omega][X]$ de degrés inférieurs ou égaux à d. Il existe un réseau arithmétique sur $\mathbf{A}$ de taille polynomiale en d et s et de profondeur polylog en d et s qui calcule un polynôme en X de degré polynomial en d et s tel que pour tout élément M de $\mathbf{R}$ supérieur à la plus grande racine de P dans $\mathbf{R}$ et pour tout s-uple de conditions de signe $(\sigma_1, ...\sigma_s)$

il existe une racine de $P(\Omega)(X)$ dans $\mathbf{R}'$ donnant à $Q_1(\Omega)(X), ..., Q_s(\Omega)(X)$ les signes $(\sigma_1, ...\sigma_s)$

si et seulement si

il existe une racine de $P(M)(X)$ dans $\mathbf{R}'$ donnant à $Q_1(M)(X), ..., Q_s(M)(X)$ les signes $(\sigma_1, ...\sigma_s)$.

Démonstration : Le polynôme P s'obtient en faisant le produit de tous les coefficients principaux des polynômes en Ω qui apparaissent dans les calculs de suite de Sturm-Habicht nécessaires dans le Corollaire 2. Il est alors clair que le signe de tous ces coefficients principaux est le même dans $\mathbf{R}$ en M que dans $\mathbf{R}'$.

Corollaire 4 : Soient $P, Q_1, ..., Q_s$ des polynômes de $\mathbf{Z}[\Omega][X]$ de degrés inférieurs ou égaux à d avec des coefficients de taille inférieure ou égale à t. Il existe algorithme de complexité séquentielle binaire polynomiale en d et s et t qui calcule un polynôme en X de degré polynomial en d et s dont les coefficients sont de taille

polynomiale en d et s et linéaire en t tel que pour tout réel M supérieur à la plus grande racine de P et pour tout s-uple de conditions de signe $(\sigma_1, ...\sigma_s)$

il existe une racine de $P(\Omega)(X)$ dans le corps des séries de Puiseux à coefficients réels donnant à $Q_1(\Omega)(X), ..., Q_s(\Omega)(X)$ les signes $(\sigma_1, ...\sigma_s)$

si et seulement si

il existe une racine réelle de $P(M)(X)$ donnant à $Q_1(M)(X), ..., Q_s(M)(X)$ les signes $(\sigma_1, ...\sigma_s)$.

3. Complexité des ensembles semi-algébriques

3.1. Les résultats annoncés et la réduction au cas des conjonctions

Précisons l'énoncé du théorème que nous allons démontrer:

Théorème 1 : Soit S un ensemble semi-algébrique de $\mathbf{R}^n$ défini par une combinaison booléenne de longueur L portant sur une famille F de s polynômes $F_1, ..., F_s$ à coefficients dans $\mathbf{A}$ de degrés totaux $\deg(F_1), ..., \deg(F_s)$ et $D = \sum_{i=1,...,s} \deg(F_i)$. Il existe un réseau arithmétique sur $\mathbf{A}$ de taille $O(L)D^{n^{O(1)}}$ et de profondeur $O(\log L) + (n \log D)^{O(1)}$ qui décide si S est vide ou non. De plus si S est non vide, le réseau arithmétique construit au moins un point par composante (semi-algébriquement) connexe de S. Les coordonnées de ces points vérifient des équations polynomiales de degré $D^{O(n)}$.

En réexaminant l'algorithme obtenu on obtiendra immédiatement le résultat suivant.

Corollaire : Soit S un ensemble semi-algébrique de $\mathbf{R}^n$ défini par une combinaison booléenne de longueur L portant sur une famille F de s polynômes $F_1, ..., F_s$ à coefficients dans $\mathbf{Z}$ de degrés totaux $\deg(F_1), ..., \deg(F_s)$ et $D = \sum_{i=1,...,s} \deg(F_i)$. Il existe un algorithme permettant de décider si S est vide ou non en complexité séquentielle binaire $O(L)D'^{n^{O(1)}}$ où D' est une borne supérieure à D et à la longueur binaire des coefficients de $F_1, ..., F_s$.

La réduction au cas des conjonctions que nous allons maintenant opérer est semblable à [16].

Définition : Un *F-semi-algébrique de base* de $\mathbf{R}^n$ est un ensemble semi-algébrique non vide défini par une conjonction de conditions de signes $(> 0, < 0, = 0)$ portant sur des polynômes extraits de la famille.

Si $n = 1$ il est clair que le nombre de toutes les composantes semi-algébriquement connexes de tous les F-semi-algébriques de base est en $O(D)$, c'est à dire que parmi les 3^D conjonctions de signes possibles la plupart sont vides. Ceci est un phénomène général.

On a en effet le résultat suivant.

Proposition 4 : Le nombre de toutes les composantes semi-algébriquement connexes de tous les F-semi-algébriques de base est en $D^{O(n)}$.

Démonstration : Similaire à celle donnée dans [16], lemme 1. La technique est la suivante: on se ramène au cas algébrique, on utilise les bornes de Thom et Milnor sur le nombre de composantes connexes d'un ensemble algébrique sur un corps réel clos ([3]) et on fait appel à un raffinnement du théorème de Bezout [17].

On va démontrer tout d'abord le théorème 2.

Définition : Un réseau arithmétique sur $\mathbf{A}$ est *admissible* s'il a une profondeur $(n \log D)^{O(1)}$ et une taille $D^{n^{O(1)}}$.

Théorème 2 : Soient D et n des entiers naturels fixés. Il existe un réseau arithmétique admissible $N = N_{D,n}$ qui pour toute famille F de s polynômes $F_1, ..., F_s$ à coefficients dans $\mathbf{A}$ en n variables avec $D = \sum_{i=1,...,s} \deg(F_i)$ et pour tout s-uple de conditions de signe σ , $\sigma_i \in \{> 0, < 0, = 0, \geq 0, \leq 0\}$ décide si

$$S_\sigma = \{x \in \mathbf{R}^n \mid F_1(x)\sigma_1, ..., F_s(x)\sigma_s\}$$

est non vide. De plus si S_σ est non vide ce réseau arithmétique construit au moins un point par composante (semi-algébriquement connexe) de S_σ.

Le théorème 2 a pour conséquence le corollaire suivant:

Corollaire (énumération des conditions consistantes) : Soient D et n des entiers naturels fixés. Pour toute famille F de s polynômes $F_1, ..., F_s$ à coefficients dans $\mathbf{A}$ en n variables avec $D = \sum_{i=1,...,s} \deg(F_i)$ il existe un réseau arithmétique admissible $N = N_{D,n}$ qui détermine les s-uples de conditions de signe σ , $\sigma_i \in \{> 0, < 0, = 0, \geq 0, \leq 0\}$ tels que

$$S_\sigma = \{x \in \mathbf{R}^n \mid F_1(x)\sigma_1, ..., F_s(x)\sigma_s\}$$

est non vide. De plus si S_σ est non vide ce réseau arithmétique construit au moins un point par composante (semi-algébriquement connexe) de S_σ.

Démonstration: On construit les conditions de signes consistantes grâce au théorème 2 en rajoutant progressivement les polynômes $F_1, ..., F_s$ par une stratégie analogue à celle suivie dans la proposition 3 (il y aura donc $\log s$ étapes). La proposition 4 permet de borner à chaque étape le nombre de cas considérés.

Le théorème 1 se déduit alors facilement du corollaire. Le reste du paragraphe 3 est consacré à la démonstration du théorème 2.

3.2. La démarche géométrique

Il est facile en rajoutant une variable et en changeant éventuellement le signe de certains polynômes de supposer que toutes les σ_i sont ≥ 0. Supposons que

$$S = \{x \in \mathbf{R}^n \mid F_1(x) \geq 0, ..., F_s(x) \geq 0\}$$

est non vide.

3.2.1. Réduction au cas algébrique lisse

Notation : Si ϵ, δ et Ω sont des variables on notera

$$\mathbf{R}' = \mathbf{R}(\frac{1}{\Omega})^\sim$$

$$\mathbf{R}'_1 = \mathbf{R}'(\epsilon)^\sim, \mathbf{R}'_2 = \mathbf{R}'_1(\delta)^\sim$$

$$\mathbf{C} = \mathbf{R}[i], \mathbf{C}' = \mathbf{R}'[i]$$

(avec $i^2 = -1$).

$$\mathbf{C}'_j = \mathbf{R}'_j[i](j = 1, 2)$$

On note enfin $\mathbf{V}'$ (resp. $\mathbf{V}'_1$, $\mathbf{V}'_2$) l'anneau de valuation formé des éléments de $\mathbf{R}'$ bornés sur $\mathbf{R}$ (resp. de $\mathbf{R}'_1$ bornés sur $\mathbf{R}'$, $\mathbf{R}'_2$ bornés sur $\mathbf{R}'_1$) et eval' (resp eval'$_1$, eval'$_2$) les applications correspondantes de $\mathbf{V}'$ dans $\mathbf{R}$ (resp. $\mathbf{V}'_1$ dans $\mathbf{R}'$, $\mathbf{V}'_2$ dans $\mathbf{R}'_1$).

Posons

$$F_0 = \Omega - ({X_1}^2 + ... + {X_n}^2)$$

$$S' = \{x \in \mathbf{R}'^n \mid F_0(x) \geq 0, F_1(x) \geq 0, ..., F_s(x) \geq 0\}$$

On a les inclusions

$$S \subset S' \subset S(\mathbf{R}')$$

Puisque S est supposé non vide, S' est donc non vide.
Il est clair que S' est borné sur $\mathbf{R}'$ par Ω.

On se ramène au cas d'une hypersurface algébrique sur $\mathbf{R}'$ en considérant l'équation

$$G = ((F_0 + \epsilon)...(F_s + \epsilon) - \epsilon^{s+1})$$

à coefficients dans $\mathbf{R}[\Omega, \epsilon]$ où ϵ est une nouvelle variable indépendante.

On a le résultat suivant dû à Grigor'ev et Vorobjov:

Lemme 1 : Notons $\mathrm{B}(S')$ le bord de S' (c'est -à-dire les point x de S' tels que pour tout r positif de $\mathbf{R}'$ la boule de centre x et de rayon r n'est pas tout entière contenue dans S'). Ce bord est non vide puisque S' est borné sur $\mathbf{R}'$.

Définissons

$$S'(\epsilon) = \{x \in {\mathbf{R}'_1}^n \mid F_0(x_1, ..., x_n) + \epsilon > 0, ..., F_s(x_1, ..., x_n) + \epsilon > 0\}$$

On a les propriétés suivantes:

a) Sur chaque composante semi-algébriquement connexe de

$$\{x \in {\mathbf{R}'_1}^n \mid G(x_1, ..., x_n) \geq 0\}$$

le signe des polynômes $F_i + \epsilon$ est fixé.

b) Soit $x \in \mathrm{B}(S')$. Il existe un point $z \in S'(\epsilon)$ qui annule G et tel que $\mathrm{eval'}_1(z) = x$.

Démonstration: ([16]Section1, lemme 2).

Corollaire : Soit $x \in \mathrm{B}(S')$ et z comme dans le point b) du lemme. La composante semi-algébriquement connexe C de

$$V_1 = \{x \in {\mathbf{R}'_1}^n \mid G(x_1, ..., x_n) = 0\}$$

contenant z est toute entière contenue dans $S'(\epsilon)$ et est bornée sur $\mathbf{R}'$ par $\Omega + 1$.

Démonstration: d'après le point a) du lemme.

Notons ${V'}_1$ l'intersection de V_1 avec la boule ouverte de centre 0 et de rayon $\Omega + 1$.

On se ramène au cas d'une hypersurface algébrique lisse en rajoutant une deuxième variable indépendante δ et en considérant l'équation $G^2 - \delta$ à coefficients dans $\mathbf{R}[\Omega, \epsilon, \delta]$.

L'ensemble algébrique V_2 de ${\mathbf{R}'_2}^n$ défini par l'équation $G(X_1, ..., X_n)^2 = \delta$ est non singulier. Pour tout point $x \in {V_1}'$ il existe un point $z \in V_2$ tel que $\mathrm{eval'}_2(z) = x$ d'après le lemme 3 de [16]Section 1. Soit ${V'}_2$ l'intersection de V_2 avec la boule ouverte de centre 0 et de rayon $\Omega + 1$. On a $\mathrm{eval'}_2({V_2}') = {V_1}'$.

On a alors le lemme suivant dû à Grigor'ev et Vorobjov.

Lemme 2 : Les composantes semi-algébriquement connexes de ${V_1}'$ sont des unions finies d'images par $\mathrm{eval'}_2$ des composantes semi-algébriquement connexes de ${V_2}'$.

Démonstration: ([16]Section1, lemme 1).

A part quelques petits points techniques, la modification essentielle apportée à [16]est que l'on a travaillé avec une boule de rayon infiniment grand Ω ce qui ne posera pas de problème dans la suite grâce au corollaire 2 de la proposition 3.

3.2.2. Réduction à la dimension 0

L' ensemble V_2 est non-singulier, donc la fonction $\Delta = \sum_{i=1,...,n} \frac{\partial G}{\partial X_i}^2$,est strictement positive sur V_2. Notons W_2 le sous-ensemble de ${C'_2}^n$ défini par $G^2 = \delta$ et $\Delta \neq 0$. Considérons maintenant les expressions

$$T = \frac{1}{\Delta}(\frac{\partial G}{\partial X_1}^2, ..., \frac{\partial G}{\partial X_n}^2)$$

qui définissent des applications T' de W_2 dans ${C'_2}^n$ (resp. T de V_2 dans ${\mathbf{R}'_2}^n$).

Proposition 5 : Pour tout λ à coordonnées rationnelles positives appartenant à un ouvert de Zariski de l'hyperplan H d'équation $\sum_{i=0,..,n} X_i = 1$ l'idéal I_λ de ${C_2}'[X_1, ..., X_n, Y]$ défini par les équations

$$G(X_1, ..., X_n)^2 = \delta, T(X_1, ..., X_n) = \lambda, Y\Delta = 1$$

est de dimension 0.

Démonstration: Par des arguments classiques sur la dimension des fibres ([27]Chapitre 1, Corollaire du Théorème 7) on montre que l'image de W_2 dans H est de dimension $n-1$ et on trouve un ouvert de Zariski de H tel que pour tout λ l'idéal I_λ est de dimension ≤ 0. On conclut que la dimension est bien 0 sur tout un ouvert de Zariski de H en regardant les maxima des fonctions $\sum_{i=1,...,n} \mu_i X_i$ (avec $\mu_i{}^2 = \lambda_i$, $\sum_{i=1,..,n} \lambda_i = 1$, λ_i réel) sur les composantes connexes bornées de V_2 (voir [18]).

Pour un λ à coordonnées rationnelles positives tel que I_λ soit de dimension 0 l'ensemble $T^{-1}(\lambda)$ contient au moins un point par composante semi-algébriquement connexe bornée de V_2.

3.2.3. Projection sur une droite

Soit λ tel que la dimension de l'idéal I_λ est zéro. Puisque l'idéal I_λ est de dimension 0, on peut trouver des polynômes $P_i(\Omega, \epsilon, \delta, X_i)$ de degrés $D^{0(n)}$ dans I_λ à coefficients dans $\mathbf{A}[\Omega, \epsilon, \delta]$.

On divise ces polynômes par la plus grande puissance possible de δ et on fait $\delta = 0$ ce qui donne des polynômes $P_{1i}(\Omega, \epsilon, X_i)$ à coefficients dans $\mathbf{A}[\Omega, \epsilon]$. Le polynôme $P_1 i$ n'est pas de degré zéro en X_1 car un des zéros de I_λ a toutes ses coordonnées dans $\mathbf{V'_2}$. On note $I_{1,\lambda}$ l'idéal engendré par les équations P_{1i}. D'après le lemme 2 toute composante semi-algébriquement connexe de V'_1 intersecte l'ensemble des zéros de $I_{1,\lambda}$ dans $\mathbf{R}'_1$: $I_{1,\lambda}$ est donc de dimension zéro.

On divise les P_{1i} par la plus grande puissance possible de ϵ et on fait $\epsilon = 0$ ce qui définit des polynômes $p_i(\Omega, X_1)$. On note I'_λ l'idéal engendré par les $p_i(\Omega, X_1)$. D'après le corollaire du lemme 1 toute composante connexe de S' intersecte l'ensemble des zéros de I'_λ et I'_λ est donc de dimension zéro.

On note J_λ le radical de I'_λ.

On considère une direction de projection vérifiant le propriété (P):

"deux zéros (dans $C_2{}^{'n}$) distincts de J_λ ont leurs coordonnées sur l'axe des X_1 distinctes".

La base standard correspondant au nouveau système de coordonnées associé à cette projection a un polynôme $p(\Omega, X_1)$, les autres polynômes de la base sont de la forme $X_i - h_i(\Omega, X_1)$ ([20]). Les coefficients de q et des h_i sont dans $\mathbf{K}(\Omega)$ où $\mathbf{K}$ désigne le corps de fractions de $\mathbf{A}$. Toute composante semi-algébriquement connexe de S' intersecte l'ensemble des zéros de I'_λ.

3.2.4. Retour aux inégalités

Substituons maintenant les h_i aux X_i $(i = 2, ..., n)$ dans les inégalites de départ $F_1, ..., F_s$ ce qui donne, en se débarrassant des dénominateurs des conditions $q_1, ..., q_s$ ne dépendant plus que de Ω et X_1.

Pour toute composante semi-algébriquement connexe S'' de S' il y a alors une solution x_1 dans $\mathbf{R}'$ de l'égalité $p(\Omega, X_1) = 0$ rendant les $q_1(\Omega, X_1), ...q_s(\Omega, X_1)$ positifs ou nuls et telle que $x_1, h_2(\Omega, x_1), ..., h_n(\Omega, x_1)$ appartienne à S''.

3.3. L'algorithme et sa complexité

3.3.1. Calculs algébriques

On part de l'ensemble

$$S = \{x \in \mathbf{R}^n \mid F_1(X_1, ..., X_n) \geq 0, ..., F_s(X_1, ..., X_n) \geq 0\}$$

(dont on ne sait pas s'il est non vide).

On pose $F_0(\Omega, X_0, ..., X_n) = \Omega - ({X_0}^2 + ... + {X_n}^2)$ et $G = ((F_0 + \epsilon)...(F_s + \epsilon) - \epsilon^{s+1})$.

Pour des λ à coordonnées rationnelles on considère les équations

$$G(X_0, ..., X_n)^2 = \delta, T(X_0, ..., X_n) = \lambda, Y\Delta = 1$$

qui définissent un idéal I_λ.

Précisons comment choisir λ pour que I_λ soit de dimension 0..

Proposition 6 : Il existe une fonction $M(D, n)$ d'ordre $D^{n^{O(1)}}$ avec les propriétés suivantes: si S est non vide il existe un vecteur λ de la forme

$$(\frac{1}{n^2{\gamma_1}^2}, ...\frac{1}{n^2{\gamma_{n-1}}^2}, 1 - \sum_{i=0,...,n-1} \frac{1}{n^2{\gamma_i}^2})$$

avec les γ_i entiers positifs et plus petits que $M(D,n)$ telque I_λ soit de dimension 0.

Démonstration: Le fait que l'idéal I_λ soit de dimension 0 se traduit par une formule du langage des corps : on exprime le fait que les zéros de I_λ sont non vides et qu'il n'existe existe aucun vecteur tel que la projection des zéros de I_λ sur ce vecteur soit une droite. On obtient ainsi (voir [18]) une formule avec quatre blocs de quantificateurs en $4n+1$ variables où les polynômes sont de degrés $D^{O(n)}$. L'élimination des quantificateurs dans le cas algébriquement clos permet de déterminer avec une complexité admissible l'ensemble des "mauvais λ" pour lesquels I_λ n'est pas de dimension 0. Comme l'ensemble des "mauvais λ" n'est pas Zariski dense d'après la Proposition 5, c'est qu'on a construit un polynôme A de degré $D"$ d'ordre $D^{n^{O(1)}}$ qui s'annule sur l'ensemble des mauvais λ. Il suffit alors de remarquer que si on prend $M(D,n) = D" + 1$ il y a toujours un λ qui n'annule pas A dans l'ensemble des vecteurs

$$(\frac{1}{n^2{\gamma_1}^2}, \ldots . \frac{1}{n^2{\gamma_{n-1}}^2}, 1 - \sum_{i=0,\ldots,n-1} \frac{1}{n^2{\gamma_i}^2})$$

avec les γ_i entiers positifs et plus petits que $M(D,n)$.

Pour déterminer un "bon λ " on les essaye l'un après l'autre et on fait un test de dimension ≤ 0 de complexité admissible en utilisant l'élimination des quantificateurs dans le cas algébriquement clos (voir [11]).

Si pour aucune valeur on ne trouve la dimension zéro c'est que S' est vide, donc que S est vide.

Sinon on a obtenu un λ tel que I_λ est de dimension zéro et on détermine en résolvant un système linéaire des polynômes $P_i(X_i)$ de I_λ de degrés $D^{O(n)}$ en chacune des variables X_i.

De la construction précédente on peut déduire le thórème suivant:

Théorème 3 : Soit $V \subset \mathbf{R}^n$ une hypersurface lisse compacte donnée par une équation f de degré d telle qu'en tout point de V le vecteur $(\frac{\partial g}{\partial X_1}, \ldots, \frac{\partial g}{\partial X_n})$ soit en tout point non nul. On peut déterminer en temps admissible un vecteur à coordonnées rationnelles λ et un idéal de dimension 0 I_λ d'équations

$$f(X_1, \ldots, X_n) = 0, T(X_1, \ldots, X_n) = \lambda, Y\Delta = 1$$

avec

$$\Delta = \sum_{i=1,\ldots,n} \frac{\partial G}{\partial X_i}^2,$$

$$T = \frac{1}{\Delta}(\frac{\partial G}{\partial X_1}^2, \ldots, \frac{\partial G}{\partial X_n}^2)$$

dont les zéros réels sont les points critiques de la fonction $\sum_{i=1,\ldots,n} \lambda_i X_i$ sur V.

Après s'être débarrassés des dénominateurs et avoir fait $\delta = 0$ puis $\epsilon = 0$ comme dans 3.2.4. on obtient les équations d'un idéal I'_λ.

Si I'_λ n'est pas de dimension zéro, S' est vide, donc aussi S. Sinon I'_λ est engendré par des polynômes $p_i(\Omega, X_1)$ de degrés $D^{O(n)}$ et on calcule son radical J_λ par la méthode indiquée dans [21]:

le radical de I'_λ est alors engendré par les équations de I'_λ et les polynômes sans facteurs carrés $\frac{1}{PGCD(p(X_i),p'(X_i))}p(X_i)$.

On fait un nombre convenable (borné par $D^{O(n)}$) d'essais de directions rationnelles de projections pour réaliser la propriété (P): on fait le changement de variable correspondant et on calcule la base standard (pour l'ordre où X_1 est plus petit que les autres variables) en complexité admissible ([7]). Si J_λ a la propriété (P) pour cette direction de projection on s'en aperçoit en regardant la forme de la base standard : elle a un polynôme $p(\Omega, X_1)$ de degre $D^{0(n)}$ les autres polynômes de la base sont de la forme $X_i - h_i(\Omega, X_1)$, avec $h_i(\Omega, X_1)$ de degré $D^{0(n)}$.

On substitue les h_i aux X_i dans $F_1, \ldots, F_s$ et on se débarrase des dénominateurs ce qui définit les $q_1, \ldots, q_s$.

La construction précédente a la conséquence suivante (voir aussi [4]), qu'on peut voir comme un "théorème de l'élément primitif" sans factorisation.

Proposition 7 : Soit I un idéal de dimension 0 défini par des équations de degré d à n variables à coefficients dans un $\mathbf{K}$ contenu dans un corps réel clos $\mathbf{R}$ et $Z(I)$ l'ensemble des zéros de I dans $\mathbf{R}$. On peut définir avec une complexité admissible un polynôme P et des polynômes P_i à coefficients dans $\mathbf{K}$ tels que tout point de $Z(I)$ est de la forme $(P_1(t), ..., P_n(t))$ avec t une racine réelle de P.

3.3.2. Calculs semi-algébriques

La réponse au problème (**)
"a-t-on un zéro de $p(\Omega, X_1)$ dans $\mathbf{R}'$ qui rende

$$q_1(\Omega, X_1) \geq 0, ..., q_s(\Omega, X_1) \geq 0"$$

est alors de complexité admissible d'après le corollaire 2 de la proposition 4.

Si la réponse est non c'est que S' est vide donc que S est vide. Si elle est oui, c'est que S' est non vide dans $\mathbf{R}'^n$, donc que $S(\mathbf{R}')$ est non vide, et on est assuré d'avoir dans l'ensemble des zéros de J_λ un point au moins par composante connexe de $S(\mathbf{R}')$.

On revient alors à des points à coordonnées dans S en utilisant le Corollaire 3 de la Proposition 3.

Si c'est nécessaire, on peut individualiser les points obtenus en effectuant leur codage "à la Thom" ; on obtient ainsi au moins un représentant par composante connexe de S avec une complexité admissible.

3.4. Un résultat de Vorobjov

On cherche à donner une borne M qui pour un ensemble semi-algébrique S de $\mathbf{R}^n$ défini par une conjonction booléenne de longueur L portant sur une famille F de s polynômes $F_1, ..., F_s$ à coefficients entiers de taille inférieure à t avec $D = \sum_{i=1,...,s} \deg(F_i)$ permette d'assurer que si l'ensemble est non vide il contient un point dans la boule de rayon M. Une telle borne était obtenue dans [32]par des méthodes assez compliquées.

Le principe que nous considérons est simple: on peut considérer les polynômes précédemment obtenus $p(\Omega, X_1), q_1(\Omega, X_1), ...q_s(\Omega, X_1)$. On leur applique le Corollaire 4 du Théorème 3 ce qui définit M. L'ensemble S est donc non vide si et seulement si son intersection avec la boule de centre O et de rayon M, est non vide.

De plus, si S est non vide chaque composante connexe de S rencontre la boule de centre 0 et de rayon M.

On a des polynômes en Ω de degré $D^{O(n)}$ avec des coefficients de taille $t(D^{O(n)})$.

D'où le théorème:

Théorème 4(Vorobjov): Il existe une borne M en $2^{t(D^{O(n)})}$ telle que pour tout ensemble semi-algébrique non vide S de $\mathbf{R}^n$ défini par une combinaison booléenne de polynômes de degré borné par D avec des coefficients entiers de taille bornée par t la boule de centre O et de rayon M rencontre chaque composante connexe de S.

Démonstration: On utilise les résultats précédents et les majorations classiques des racines d'un polynôme en fonction des coefficients.

4. L'élimination d'un quantificateur

On va montrer le résultat suivant.

Théorème 5 : Soient $P, Q_1, ..., Q_s$ des polynômes en X à coefficients dans l'anneau $\mathbf{A}[Y_1, ..., Y_k]$, $k \leq n$. SoitΦ la formule

$$\exists X\, P(Y_1, ..., Y_k, X) = 0, Q_1(Y_1, ..., Y_k, X) \geq 0, ..., Q_s(Y_1, ..., Y_k, X) \geq 0$$

et soit U l'ensemble semi-algébrique

$$\{(Y, X) \in \mathbf{R}^{k+1} \mid P(Y, X) = 0, Q_1(Y, X) \geq 0, ..., Q_s(Y, X) \geq 0\}$$

On pose $D = \deg(P) + \sum_{i=1,...,s} \deg(Q_i)$. Il existe un réseau arithmétique admissible sur $\mathbf{A}$ construisant

- une partition de $\mathbf{R}^k$ en ensembles semi-algébriques T_m

- pour tout m une famille finie de fonctions semi-algébriques continues de T_m dans R $\xi_{m,j\,j=1,...,n_m}$ telles que le graphe de $\xi_{m,j}$ vérifie

$$P(Y, \xi_{m,j}(Y)) \in U,$$

et telle que $U \cap (T_m \mathrm{x} \mathbf{R})$ est l'union des graphes des $\xi_{m,j}$'s si ni P ni Q_i ne sont identiquement nuls sur T_m. On obtient en particulier une formule sans quantificateurs Ψ équivalente à Φ dans $\mathbf{R}$.

Démonstration: Soit

$$P = c_d(Y_1, ..., Y_k)X^d + ... + c_i(Y_1, ..., Y_k)X^i + ... + c_0(Y_1, ..., Y_k).$$

On teste tout d'abord si les d ensembles

$$S_i = \{(y_1, ..., y_k) \in \mathbf{R}^k \mid c_d(y_1, ..., y_k) = 0, ..., c_{i+1}(y_1, ..., y_k) = 0, c_i(y_1, ..., y_k) \neq 0\}$$

sont non vides en utilisant le théorème 2. On pourra donc supposer en se restreignant aux différents S_i non vides supposer que P est de degré fixé. Traitons déjà le cas $s = 1$. On calcule la suite de Sturm-Habicht de P et de Q_1 par rapport à la variables X. Elle a des coefficients dans l'anneau $\mathbf{A}[Y_1, ..., Y_k]$. Les coefficients de Sturm- Habicht forment donc une liste de polynômes de $\mathbf{A}[Y_1, ..., Y_k]$. D'après la proposition 4 et le corollaire du théorème 2 le nombre de conditions de signes consistantes non vides portant sur cette liste de polynômes est en $D^{O(n)}$ (et non en 3^D comme il semblerait a priori) et peut être calculée en temps admissible. Pour chacune de ces situations consistantes on continue le calcul par la méthode de [1]donnée dans la proposition 3 en utilisant la proposition 2. A chaque étape on a en utilisant de nouveau le corollaire du théorème 2 une liste de situations consistantes de longueur en $D^{O(n)}$ portant sur des listes de $O(D)$ polynômes au plus ce qui donne la bonne complexité. On a ainsi construit la partition semi-algébrique de $\mathbf{R}^k$ cherchée et les fonctions $\xi_{m,j}$ cherchées consistent à suivre continûment les racines de P. La formule sans quantificateurs annoncée s'obtient en prenant l'union de tous les T_m pour lesquels le nombre des fonctions $\xi_{m,j}$ est non nul.

Corollaire : Soient $P, Q_1, ..., Q_s$ des polynômes en X à coefficients dans l'anneau $\mathbf{A}[\Omega][Y_1, ..., Y_k]$ de degrés totaux d. SoitΦ la formule

$$\exists X P(Y_1, ..., Y_k, \Omega, X) = 0, Q_1(Y_1, ..., Y_k, \Omega, X) \geq 0, ..., Q_s(Y_1, ..., Y_k, \Omega, X) \geq 0.$$

On pose $D = \deg(P) + \sum_{i=1,...,s} \deg(Q_i)$.

Il existe un réseau arithmétique admissible sur $\mathbf{A}$ construisant

- une partition de $\mathbf{R'}^k$ en ensembles semi-algébriques T_m

- pour tout m une famille finie de fonctions semi-algébriques coontinues de T_m dans R' $\xi_{m,j\,j=1,...,n_m}$ telles que le graphe de $\xi_{l,j}$ vérifie

$$P(Y, \Omega, \xi_{m,j}(Y)) \in U,$$

et telle que $U \cap (T_m \mathrm{x} \mathbf{R'})$ est l'union des graphes des $\xi_{m,j}$ si ni P ni Q_i ne sont identiquement nuls sur T_m. On obtient en particulier une formule sans quantificateurs Ψ équivalente à Φ dans $\mathbf{R'}$.

Commentaire :

Le théorème 4 est la clé pour obtenir la complexité souhaitée de l'élimination des quantificateurs comme l'avait déjà remarqué Grigor'ev ([15]).

5. Complexité du principe de Tarski-Seidenberg

On va démontrer le théorème suivant:

Théorème 6 : Soit $\mathbf{A}$ un anneau intègre contenu dans un corps réel clos $\mathbf{R}$. Soit Φ une formule du langage des corps ordonnés à paramètres dans $\mathbf{A}$ supposée sous forme prénexe avec m blocs de quantificateurs où apparaissent s polynômes en n variables à coefficients dans $\mathbf{A}$ dont la somme des degrés totaux est inférieure ou égale D. On peut construire un réseau arithmétique de taille $O(L)D^{n^{O(m)}}$ et de profondeur $O(\log L)n^{O(m)}\log D^{O(1)}$ qui calcule une formule sans quantificateurs équivalente à Φ .

Corollaire : Soit Φ une formule du langage des corps ordonnés à paramètres dans $\mathbf{Z}$ supposée sous forme prénexe avec m blocs de quantificateurs où apparaissent s polynômes en n variables à coefficients dans

A dont la somme des degrés totaux est inférieure ou égale D. On peut construire un algorithme qui calcule une formule sans quantificateurs équivalente à Φ en complexité séquentielle binaire $O(L)D'^{n^{O(m)}}$ où D' est un entier supérieur à t et D.

Le principe de la démonstration est clair: on travaille par induction sur le nombre de blocs. A chaque étape on est amené à éliminer un bloc de quantificateurs de même nature (disons des quantificateurs existentiels).

On utilise pour celà une version paramétrée de l'algorithme de test pour connaitre si un ensemble semi-algébrique est vide.

Théorème 7 : Soit S un ensemble semi-algébrique de $\mathbf{R}^{(k+n)}$ défini par s polynômes en $k+n$ variables tels que la somme de leurs degrés est inférieure ou égale à D.

Il existe un réseau arithmétique de taille $D^{(k+n)^{O(1)}}$ et de profondeur $((k+n)\log D)^{O(1)}$ construisant

- une partition de $\mathbf{R}^k$ en ensembles semi-algébriques T_m
- pour chaque m une famile finie de fonctions semi-algébriques continues de T_m dans R^n $\xi_{m,j\,j=1,...,l_m}$ telles que le graphe de $\xi_{m,j}$ appartient à S et telle que pour tout $y \in T_m$ les graphes des $\xi_{m,j}$ intersectent chaque composante (semi-algébriquement) connnexe de $S \cap (\{y\}\mathrm{x}\mathbf{R}^n)$.

Démonstration : On reprend dans le cas paramétré l'algorithme du paragraphe 3.3. précédent. Les calculs à faire en 3.3.1. sont des calculs algébriques et la complexité annoncée résulte de l'élimination des quantificateurs dans le cas algébriquement clos [11]. A chaque étape du calcul on introduit une subdivision des paramètres en ensembles constructibles (définis par des combinaisons booléennes de $=$ et $\neq$). Remarquons que les calculs de base standard "paramétrée" se font par l'algèbre linéaire puisque on peut borner par avance les degrés des polynômes.

La seule partie "réelle" (où on introduira des ensembles semi-algébriques et non des ensembles constructibles) est 3.3.2. : on résoud le problème $(**_k)$

"trouver la condition booléenne T sur $Y_1, ..., Y_k$ telle que pour tout choix de $y_1, ..., y_k$ vérifiant T un des zéros (dans $\mathbf{R}'$) de $p(y_1, ..., y_k, \Omega, X_1)$ est tel que les signes de

$$q_1(y_1, ..., y_k, \Omega, X_1), ..., q_s(y_1, ..., y_k, \Omega, X_1)$$

sont positifs ou nuls ".

On utilise pour celà le corollaire du théorème 5.

Corollaire : On peut trouver une combinaisons booléenne T sur les Y_j qui soit équivalente à

$$\exists X_1 ... \exists X_n S(Y_1, ..., Y_k, X_1, ..., X_n)$$

grâce à un réseau arithmétique de taille $D^{(k+n)^{O(1)}}$ et de profondeur $((k+n)\log D)^{O(1)}$.

Les polynômes décrivant la conditions T sont de degré $D^{O(n+k)}$.

Démonstration: Il suffit de prendre l'union des T_m donnés par le théorème 7 tels que le nombre des fonctions $\xi_{m,j}$ construites dans le théorème soit non nul.

Démonstration du Théorème 6 : L'algorithme d'élimination des quantificateurs s'obtient maintenant simplement par induction sur le nombre de blocs et on obtient aisément le théorème 6 et son corollaire.

Remarque : Cet algorithme ne produit pas de démonstration du principe de Tarski-Seidenberg: on a utilisé pour établir sa correction des résultats de géométrie semi-algébrique sur un corps réels clos qui utilisent le théorème de Tarski-Seidenberg.

Références

[1]Ben-Or M. , Kozen D. , Reif J. : The complexity of elementary algebra and geometry. J. of Computation and Systems Sciences 32 251-264 (1986).

[2]Berkowitz S. J.: On computing the determinant in small parallel time using a small number of processors. Information Processing Letter 18 (1984), 147-150.

[3]Bochnak J., Coste M., Roy M.-F.: Géométrie algébrique réelle. Springer Verlag (1987).

[4]Canny J.: Some algebraic and geometric computations in PSPACE. ACM Symposium on the theory of computation, 460-467 (1988).

[5]Canny J.: The complexity of robot motion planning. MIT Press (1989).

[6]Collins G. : Quantifier elimination for real closed fields by cylindric algebraic decomposition. In Second GI Conference on Automata Theory and Formal Languages. Lecture Notes in Computer Sciences, vol. 33, pp. 134-183, Springer-Verlag, Berlin (1975).

[7]Caniglia L., Galligo A., Heintz J.: Some new effectivity bounds in computational geometry. Proc. AAECC-6 (Rome 1988) - Best Paper Award AAECC-6. Springer Lecture Notes in Computer Science 357 131-151.

[8]Coste M. , Roy M.-F. : Thom's lemma, the coding of real algebraic numbers and the topology of semi-algebraic sets. J. of Symbolic Computation 5 121-129 (1988).

[9]Davenport J., Heintz J.: Real quantifier elimination is doubly exponential. J. of Symbolic Computation 5 29-35 (1988).

[10]Fitchas N., Galligo A., Morgenstern J.: Algorithmes rapides en séquentiel et en parallle pour l'élimina- tion des quantificateurs en géométrie élémentaire. Séminaire Structures Ordonnées, U.E.R. de Math. Univ. Paris VII (1987).

[11]Fitchas N., Galligo A., Morgenstern J.: Precise sequential and parallel complexity bounds for the quantifier elimination of algebraically closed fields. A paraître dans Journal of Pure and Applied Algebra.

[12]von zur Gathen J.: Parallel arithmetic computations: a survey. Proc 13th Conf. MFCS (1986).

[13]Gonzalez L., Lombardi H., Recio T., Roy M.-F.: Sturm-Habicht sequences. Proceedings ISSAC 1989.

[14]Gonzalez L., Lombardi H., Recio T., Roy M.-F.: Sous-résultants et spécialisation de la suite de Sturm. A paraître au RAIRO Informatique théorique.

[15]Grigor'ev D.: Complexity of deciding Tarski algebra. J. Symbolic Computation 5 (1988) 65-108.

[16]Grigor'ev D., Vorobjov N.: Solving systems of polynomial inequalities in subexponential time. J. Symbolic Computation 5 (1988) 37-64.

[17]Heintz J.: Definability and fast quantifier elimination in algebraically closed fields. Theor. Comput. Sci. 24 (1983) 239-277.

[18]Heintz J., Roy M.F., Solernó P. : On the complexity of semialgebraic sets. Proc. IFIP (San Francisco 1989).

[19]Heintz J., Roy M.-F., Solerno P.: Complexité du principe de Tarski- Seidenberg. Compte-Rendus de l'Académie des Sciences Paris. 309 825-830 (1989).

[20]Kobayashi H., Moritsugu S., Hogan R.W.: On solving systems of algebraic equations. Soumis au Journal of Symbolic Computation.

[21]Krick T. : Thèse. Université de Buenos Aires (en préparation).

[22]Loos R.: Generalized poynomial reaminder sequences. Dans Computer Algebra, Symbolic and Algebraic Computation 115-138. Edit par Buchberger, Collins, Loos. Springer Verlag 1982.

[23]Renegar J.: A faster PSPACE algorithm for deciding the existential theory of the reals. Technical Report 792, Cornell University Ithaca (1988).

[24]Renegar J.: On the computational complexity and geometry of the first order theory of the reals. Technical Report 856, Cornell University Ithaca (1989).

[25]Roy M.-F., Szpirglas A.: Complexity of computations with real algebraic numbers. A paraître au Journal of Symbolic Computation.

[26]Seidenberg A.: A new decision method for elementary algebra and geometry. Ann. Math. 60 365-374 (1954).

[27]Shafarevitch I.S. Algebraic geometry. Springer Verlag (1974)

[28]Solernó P.: Complejidad de conjuntos semialgebraicos. Thèse. Université de Buenos Aires 1989.

[29]Sturm C.: Mémoire sur la résolution des équations numériques. Ins. France Sc. Math. Phys. 6 (1835).

[30]Sylvester J. T. : On a theory of syzygetic relations of two rational integral functions,comprising an application to the theory of Sturm's function. Trans. Roy. Soc. London (1853). Reprint in: Sylvester : Collected Math Papers. Chelsea Pub. Comp. NY 1983 vol 1 429-586.

[31]Tarski A.: A decision method for elementary algebra and geometry. Berkeley (1951).

[32]Vorobjov N.: Bounds of real roots of a system of algebraic equations. Notes of Sci. Seminars of Leningrad Department of Math. Steklov Inst. 13 7-19.

[33]Walker R.: Algebraic curves. Princeton University Press (1950).

[34]Weipsfenning V.: The complexity of lieanr problems in fields. J. Symbolic Computation 5 3-27 (1988).

Joos Heintz et Pablo Solernó
Instituto Argentino de Matematica
CONICET
Viamonte 1636
(1055) Buenos Aires
ARGENTINA

Marie-Françoise Roy
IRMAR
Université de Rennes I
35 042 Rennes CEDEX