RAIRO. INFORMATIQUE THÉORIQUE ET APPLICATIONS

ANNETTE PAUGAM

Résolution du problème de l'ellipse et du cercle par l'algorithme de Hörmander

RAIRO. Informatique théorique et applications, tome 24, n° 2 (1990), p. 161-188

<http://www.numdam.org/item?id=ITA_1990__24_2_161_0>

Informatique théorique et Applications/Theoretical Informatics and Applications
(vol. 24, n° 2, 1990, p. 161 à 188)

RÉSOLUTION DU PROBLÈME DE L'ELLIPSE ET DU CERCLE PAR L'ALGORITHME DE HÖRMANDER (*)

par Annette PAUGAM (1)

Communiqué par J. BERSTEL

Résumé. – *Nous donnons une solution algorithmique du problème de Kahan : trouver des conditions sur les nombres réels a, b, x_0, y_0 pour que l'ellipse $((x-x_0)^2/a^2)+((y-y_0)^2/b^2)-1=0$ soit intérieure au cercle unité $x^2+y^2-1=0$. Notre solution est obtenue à partir d'une démonstration de Hörmander du principe de l'élimination des quantificateurs, basée sur l'étude de tableau de signes de polynômes. La comparaison des différents résultats obtenus sur le problème de Kahan algorithmique ou non (travaux de Collins, Arnon et Mignotte, Lazard) montre, sur cet exemple simple en apparence, la difficulté pratique d'obtenir une « bonne » méthode de décision pour l'algèbre élémentaire et la géométrie.*

Abstract. – *We give an algorithmic solution to Kahan's problem: find conditions on real numbers a, b, x_0 and y_0 so that the ellipse $((x-x_0)^2/a^2)+((y-y_0)^2/b^2)-1=0$ lies inside the unit circle $x^2+y^2-1=0$. Our solution is obtained by following Hormander's proof of quantifier elimination, using the study of sign's table of polynomial. The comparison of differents results on the Kahan's problem, algorithmic or no (works by Collins, Arnon and Mignotte, Lazard) illustrates, on this apparently simple example, the practical difficult to obtain a "good" decision method for elementary algebra and geometry.*

INTRODUCTION

Le problème de l'ellipse et du cercle posé par Kahan en 1975 [11], était le suivant :

Trouver une condition nécessaire et suffisante, sous forme d'une combinaison booléenne d'équations et d'inégalités polynomiales en x_0, y_0, $1/a$ et $1/b$ pour que l'ellipse

$$E(x, y)=(x-x_0)^2/a^2+(y-y_0)^2/b^2-1=0$$

(*) Reçu janvier 1988, révisé en juin 1988.
(1) I.R.M.A.R., Université de Rennes-I, Campus de Beaulieu, 35042 Rennes Cedex.

Informatique théorique et Applications/Theoretical Informatics and Applications
0988-3754 90/02 161 28/$4.80/

soit incluse dans le cercle

$$C(x, y) = x^2 + y^2 - 1 = 0,$$

ce qui peut s'écrire

$$(\forall x) \quad (\forall y) \qquad E(x, y) = 0 \quad \Rightarrow \quad C(x, y) < 0 \tag{1}$$

où C et E sont des polynômes de $\mathbb{Q}[x_0, y_0, 1/a, 1/b]\,[x, y]$.

La formule (1) est une formule de l'algèbre élémentaire, c'est-à-dire une formule construite à partir d'égalités et d'inégalités polynomiales en utilisant les connecteurs logiques et les quantificateurs universels et existentiels, portant sur des variables réelles. Le problème de Kahan consiste à calculer une formule équivalente de l'algèbre élémentaire, mais sans quantificateurs. C'est donc une application algorithmique du fameux principe de l'élimination des quantificateurs dont les premières démonstrations ont été écrites par Tarski et Seidenberg.

THÉORÈME (Principe de Tarski-Seibenberg [18, 19]) : *Pour toute formule de l'algèbre élémentaire, on peut calculer une formule équivalente sans quantificateurs.*

Les applications potentielles de ce principe sont nombreuses :

– Pour le problème de Kahan, le principe démontre que l'on peut calculer effectivement des conditions polynomiales en x_0, y_0, $1/a$ et $1/b$, équivalentes à l'inclusion de l'ellipse dans le cercle. C'est l'élimination de deux quantificateurs universels.

– En géométrie réelle, on considère les ensembles semi-algébriques, sous-ensembles de $\mathbb{R}^n$ définis un nombre fini d'égalités et d'inégalités polynomiales. Le principe de Tarski-Seidenberg démontre que la projection d'un semi-algébrique est un semi-algébrique et donne un algorithme pour calculer les conditions qui définissent cette projection. C'est l'élimination d'un quantificateur existentiel. Ce calcul de la projection est un résultat fondamental de géométrie algébrique réelle. On en déduit, par exemple, la calculabilité de l'adhérence et de l'intérieur d'un ensemble semi-algébrique.

Dans cet article, nous nous intéresserons uniquement au premier problème. Sur cet exemple, simple en apparence, on constate que la démonstration de Seidenberg [18] donne un algorithme, difficile à appliquer directement [17]. Dans les années 70-80, Collins [1] a mis au point un programme : « cylindrical algebraic decomposition », qui en théorie permet d'éliminer les quantificateurs. Cet algorithme utilise différents algorithmes notamment sur la représentation des réels algébriques [8] et sur les résultants et sous-résultants [6]. Une

utilisation intelligente de l'algorithme de Collins a permis à Arnon de résoudre un cas particulier facile du problème de Kahan ($y_0 = 0$) [2, 3]. Puis le problème a été résolu par Laüer [12], Mignotte [16] et Lazard [13] hors algorithme. Ensuite Arnon et Mignotte ont traité le problème général par différentes méthodes, utilisant partiellement l'algorithme de Collins [4].

Nous exposerons ici une autre méthode pour résoudre ce problème, basée sur une démonstration de Hörmander du principe de Tarski-Seidenberg [5, 9]. Elle donne un résultat rapide pour $y_0 = 0$ (seul cas résolu par Arnon de manière totalement algorithmique). Pour $y_0 \neq 0$, les résultats, plus complexes que ceux de Lazard et Mignotte, sont obtenus de manière systématique.

Le plan de l'article est le suivant :

1. La démonstration de Hörmander du principe de Tarski-Seidenberg.
2. Application au problème de Kahan.
3. Comparaison des différentes méthodes.

I. DÉMONSTRATION DE HÖRMANDER DU PRINCIPE DE TARSKI-SEIDENBERG

La démonstration de ce principe se réduit par récurrence sur le nombre de quantificateurs et par des calculs de logique élémentaire à l'élimination d'un quantificateur existentiel dans une conjonction de conditions polynomiales. Pour plus de généralité, nous supposerons les variables à valeur dans un corps réel clos, c'est-à-dire un corps ordonnable dans lequel tout élément positif admet une racine carrée et tout polynôme de degré impair a au moins une racine. Le principe de l'élimination des quantificateurs est alors ramené à la démonstration du théorème suivant :

THÉORÈME I.1 : *Soient* $P_i(\underline{X}, Y) = a_{0,i}(\underline{X}) + \ldots + a_{m_i,i}(\underline{X})\, Y^{m_i}$ *pour* $i = 1, \ldots, s$ $(\underline{X} = X_1, \ldots, X_n)$, *une suite de s polynômes en* $n+1$ *variables* $(\underline{X}, Y)$ *à coefficients dans* $\mathbb{Z}$ *et soit* $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_s)$ *des signes* $(<, >, =)$ *pour ces polynômes.*

Alors on peut calculer une combinaison booléenne $S(\underline{X})$ *d'équations et d'inégalités polynomiales en* $\underline{X}$ *à coefficients dans* $\mathbb{Z}$ *telle que pour tout corps réel clos R et tout* $\underline{x}$ *dans* R^n, *il existe y appartenant à R tel que*

$$P_1(\underline{x}, y)\, \varepsilon_1\, 0 \qquad et \ \ldots \ et \qquad P_s(\underline{x}, y)\, \varepsilon_s\, 0 \tag{*}$$

si et seulement si $S(\underline{x})$ *est vrai dans* R^n.

On remarque dans cet énoncé que le calcul de $S(\underline{X})$ est indépendant du corps réel clos R et du $\underline{x}$ choisi dans R^n.

L'outil de base de la démonstration algorithmique de Hörmander est la notion de tableau de signes d'une suite de polynômes $P_1, \ldots, P_s$ en Y à coefficients dans un corps réel clos R.

Le tableau de signes des P_i: $T_R(P_1, \ldots, P_s)$, sera la donnée du nombre N de zéros : $y_1 \ldots y_N$ des différents polynômes P_i, et d'un tableau à s lignes et $2N+1$ colonnes donnant le signe $(+, -, 0)$ de chaque polynôme P_i en chaque zéro y_j et sur chaque intervalle $]-\infty, y_1[$, $]y_i, y_{i+1}[$, $]y_N, +\infty[$. Il est à noter que ce tableau ne donne pas la valeur des zéros. Si les degrés des P_i sont inférieurs à m, alors $N \leqq sm$ et la tableau de signes des P_i est un élément de l'ensemble fini $\mathscr{T}_{s,m} = \coprod_{N \leqq sm} \{-, 0, +\}^{s(2N+1)}$.

Pour un tableau de signes des P_i comme polynôme en Y, donné, ($\underline{x}$ fixé) l'existence d'un y répondant à (*) se détermine en regardant chaque colonne du tableau. Supposons $\deg_Y P_i \leqq m$. Pour $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_s)$ donné, les tableaux de signes répondant à (*) forment un sous-ensemble $\mathscr{T}(\varepsilon)$ de $\mathscr{T}_{s,m}$. *A priori* chaque tableau de signes dépend de $\underline{x}$. En fait, Hörmander [5, 9] montre que l'appartenance d'un tableau de signe à $\mathscr{T}(\varepsilon)$ ne dépend que d'un nombre fini de conditions polynômiales en $\underline{x}$.

Avant de montrer ce résultat, nous allons voir que le tableau de signes de $P_1, \ldots, P_s$ est entièrement déterminé par la connaissance du tableau de signes d'une suite de polynômes, déduite de $P_1, \ldots, P_s$, et comportant un polynôme de moins de degré maximum. Cet algorithme sur les tableaux de signes est donné dans la preuve du lemme suivant

LEMME 1.2 : *Il existe un algorithme permettant de déterminer une partie de* $\mathscr{T}_{2s,m}$ *et de définir une application* φ_s *de cette partie de* $\mathscr{T}_{2s,m}$ *dans* $\mathscr{T}_{s,m}$ *tel que pour tout corps réel clos* R *et pour toute suite* $P_1, \ldots, P_s$ *de polynômes de* $R[Y]$ *de degré* $\leqq m$, *telle que* $\deg P_1 = m$, *on a*

$$T_R(P_1, \ldots, P_s) = \varphi_s(T_R(P_1', P_2, \ldots, P_s, R_1, \ldots, R_s))$$

où P_1' *est le polynôme dérivé de* P_1 *et*

R_1 *le reste de la division de* P_1 *par* P_1'

R_i *le reste de la division de* P_1 *par* P_i *pour* $2 \leqq i \leqq s$.

Remarque : Dans ce lemme, φ_s est indépendant du corps réel clos R et des polynômes P_i. C'est une fonction portant uniquement sur les tableaux de signes.

Indication de démonstration : Le principe du calcul de φ_s est le suivant : Étant donné un tableau de signes de $P_1', P_2, \ldots, P_s, R_1, \ldots, R_s$ on en

déduit celui de $P_1, \ldots, P_s$ en utilisant les propriétés élémentaires des polynômes sur les corps réels clos :

(1) le signe de P'_1 donne les variations de P_1;

(2) le signe de R_1 donne le signe de P_1 au zéro de P'_1 (extrêma de P_1);

(3) le signe de $R_i (i > 1)$ donne le signe de P_1 au zéro de P_i;

(4) le signe de P_1 à $(-\infty)$ est l'opposé du signe de P'_1 à $(-\infty)$;

(5) le signe de P_1 à $(+\infty)$ est le signe de P'_1 à $(+\infty)$

et les zéros s'intercalent par le théorème des valeurs intermédiaires.

L'algorithme correspondant à ce lemme à partir du tableau de signes de $P'_1, P_2, \ldots, P_s, R_1, \ldots, R_s$, en entrée, consiste à :

(1) Introduire dans le tableau, le signe de P_1 aux zéros de $P'_1, P_2, \ldots, P_s$. (On ajoute une ligne au tableau.)

(2) D'après le signe de P_1, introduire les zéros de P_1, en contrôlant qu'il n'y a pas de contradiction avec le signe de P'_1. (On ajoute des colonnes.)

(3) Effacer $P'_1, R_1, \ldots, R_s$ et leurs zéros du tableau (des lignes et des colonnes).

Les détails de cet algorithme : « remontée des tableaux de signes », figurent dans [17]. A partir de ce lemme, on obtient une démonstration algorithmique du principe de Tarski-Seidenberg.

Démonstration du théorème I.1 : Le calcul de $S(\underline{x})$ commence par le calcul des suites successives de polynômes déduites de $P_1, \ldots, P_s$ selon le lemme 1.2, jusqu'à l'obtention de « constantes », polynômes de degré 0 en Y. A ce niveau, on rencontre deux difficultés.

1. Pour les divisions, les polynômes sont à coefficients dans $\mathbb{Z}[\underline{X}]$ et non dans un corps. On fait donc des pseudo-divisions : divisions usuelles multipliées par une puissance bien choisie du coefficient dominant du diviseur. Pour cela, j'ai utilisé les travaux de Brown, Traub [6] et Loos [14], sur la pseudo-division, l'algorithme d'Euclide et les sous-résultants.

2. La construction d'une suite déduite de $P_1, \ldots, P_s$, dépend du degré exact des P_i en Y. Chaque division se fait sous conditions polynômiales en $\underline{X}$, traduisant que le coefficient dominant des P_i en Y soit non nul. Les suites déduites forment donc un arbre, chaque branche correspondant à des conditions en $\underline{X}$ sur les coefficients dominants des polynômes en Y.

Ensuite à partir de tous les signes possibles de ces « constantes », on calcule, par l'algorithme de remontée donné par le lemme 1.2, les tableaux de signes successifs, jusqu'à celui de $P_1, \ldots, P_s$.

La dernière étape consiste à tester parmi les tableaux obtenus pour $P_1, \ldots, P_s$, lesquels correspondent aux conditions demandées. Ceci se fait en regardant les colonnes du tableau de $P_1, \ldots, P_s$.

Le principal défaut de cet algorithme est la croissance exponentielle en fonction du degré du nombre des polynômes déduits de $P_1, \ldots, P_s$ et donc une croissance doublement exponentielle des nombres de conditions de signe à considérer.

Dans l'exemple que j'ai traité, j'ai essayé de caractériser les tableaux de signes favorables pour $P_1, \ldots, P_s$, et d'en déduire les caractéristiques des tableaux de signes favorables pour les $2s$ polynômes déduits. De cette manière, je n'ai pas à parcourir entièrement l'arbre des tableaux de signes, mais seulement la partie nécessaire pour obtenir les conditions (*).

II. LES RÉSULTATS DE HÖRMANDER APPLIQUÉS AU PROBLÈME DE KAHAN

Soit (E) l'ellipse d'équation

$$E=(x-x_0)^2/a^2+(y-y_0)^2/b^2-1=0$$

et le cercle unité

$$C=x^2+y^2-1=0.$$

On cherche des conditions sur x_0, y_0, a et b pour que l'ellipse soit incluse dans le cercle, soit

$$(\forall x) \quad (\forall y) \qquad (E=0 \Rightarrow C<0) \tag{1}$$

il s'agit ici d'éliminer les quantificateurs universels dans une relation polynomiale en x_0, y_0 $1/a$ et $1/b$ avec a et b strictement positifs. L'algorithme nous permet d'établir la liste de tous les tableaux de signes possibles pour E et C comme polynôme en y. Les tableaux de signes vérifiant (1) correspondront à un sous-ensemble de $\mathscr{T}_{2,2}$, donc à une combinaison booléenne d'égalités et d'inégalités polynomiales en x_0, y_0, $1/a$ et $1/b$ à coefficients entiers.

J'ai abordé le problème de Kahan avec les techniques de Hörmander pour l'étude des tableaux de signes. Ces techniques amènent à distinguer le cas $y_0=0$ du cas $y_0 \neq 0$.

Pour $y_0=0$, j'ai calculé toutes les suites de polynômes déduites par application directe de l'algorithme. J'ai raccourci l'étude des tableaux de signes en caractérisant les tableaux de signes vérifiant (1) sans les écrire explicitement.

Pour $y_0 \neq 0$, certaines remarques sur le signe des polynômes obtenus me permettent de réduire le calcul des suites déduites par l'algorithme de Hörmander. Lorsque le signe de certains polynômes est apparu comme « évident », je n'ai pas continué à les faire intervenir dans la suite de l'algorithme. La démarche pour l'étude des tableaux de signes est identique à celle du cas précédent.

Je distinguerai à chaque étape les choix ou les astuces de calcul, des parties d'application systématique de l'algorithme.

II.1. Élimination de Y

Choix préliminaires : Avant d'utiliser l'algorithme, remarquons que l'on s'intéresse seulement au signe de C quand E est nul, donc la première démarche naturelle pour éliminer y est de faire la pseudo-division de C par E, comme polynôme en Y

$$1/b^2\, C = E + P$$

où P est une parabole

$$P(x, y) = 2y_0(y - y_0)/b^2 + F(x)$$

avec

$$F(x) = (x^2 + y_0^2 - 1)/b^2 - [(x - x_0)^2/a^2 - 1].$$

On a alors à étudier le signe de P, de degré 1 en Y, quand E est nul, c'est-à-dire :

$$(\forall x) \quad (\forall y) \qquad (E(x, y) = 0 \Rightarrow P(x, y) < 0).$$

Cette première division fait partie d'ailleurs des constructions de liste de Hörmander si l'on choisit C (et non E) comme polynôme de plus haut degré, mais sans cette remarque on a déjà quatre polynômes à cette étape au lieu de deux.

Déroulement de l'algorithme : Écrivons les suites de polynômes successives obtenues lorsqu'au cours des pseudo-divisions, aucun coefficient dominant ne s'annule. On remarque que, sur cet exemple, les coefficients dominants en Y ne font jamais intervenir x, ce qui simplifie l'utilisation de l'algorithme.

1. $\{E, P\}$

(coef. dominants : $1/b^2$, $2y_0/b^2$) (degré en Y: 2, 1).

2. $\{\partial E/\partial y,\ P,\ I = E \bmod_y (\partial E/\partial y),\ R = (4\, y_0^2/b^2)\, E \bmod_y P\}$

(coef. dominants : $2/b^2$, $2\, y_0/b^2$, I, R) (degré en Y: 1, 1, 0, 0).

3. $\{\partial P/\partial y,\ \partial E/\partial y,\ I,\ R,\ F = [P \bmod_y \partial E/\partial y]\}$

(coef. dominants : $2\, y_0/b^2$, $2/b^2$, I, R, F) (degré en Y: 0, 1, 0, 0, 0).

4. $\{\partial^2 E/\partial y^2,\ \partial P/\partial y,\ I,\ R,\ F\}$

(coef. dominants : $2/b^2$, $2\, y_0/b^2$, I, R, F) (degré en Y: 0, 0, 0, 0, 0)

avec

$$\mathrm{F}(x) = (x^2 + y_0^2 - 1)/b^2 - [(x - x_0)^2/a^2 - 1]$$

$$I(x) = (x - x_0)^2/a^2 - 1,$$

$$R(x) = 4\, y_0^2\, [(x - x_0)^2/a^2 - 1]/b^2 + F(x)^2 = \text{résultant } (E, C).$$

On remarque que, en général, F est de degré 2 en x et R de degré 4.

Avec nos hypothèses a et b non nuls, le seul coefficient dominant pouvant s'annuler est $2\, y_0/b^2$.

Étudions les différents tableaux de signes en y pour $y_0 \neq 0$. Le signe de $\partial^2 E/\partial y$ donne celui de $\partial E/\partial y$ et montre que E passe par un minimum. En ce minimum, E est du signe de $I = E \bmod_Y \partial E/\partial y$ et P du signe de $F = P \bmod_Y \partial E/\partial y$. On veut obtenir : P négatif dès que E s'annule. D'abord, pour que E s'annule, il faut et il suffit que son minimum I soit négatif ou nul. Pour que P soit négatif dès que E s'annule, comme P (degré 1) est monotone en y, il faut et il suffit que P soit négatif sur l'intervalle fermé entre les deux racines de E en y (distinctes ou confondues), et pour cela il faut et il suffit que

(1) P soit négatif en un point entre les racines de E en Y, au minimum de E et

(2) P s'annule à l'extérieur des racines de E en y, c'est-à-dire pour $E > 0$.

Hörmander remplace ces deux conditions par $F(x) < 0$ et $R(x) > 0$.

La conclusion de

$$(\forall x) \quad (\forall y) \quad (E(x, y) = 0 \Rightarrow F(x) < 0 \text{ et } R(x) > 0)$$

étant indépendante de y on peut écrire l'énoncé sous la forme

$$(\forall x) \quad ((\exists y\, \mathrm{E}(x, y) = 0) \Rightarrow F(x) < 0 \text{ et } R(x) > 0).$$

Pour x fixé, le signe de $\partial^2 E/\partial y$ montre que E admet un minimum. Pour que E (degré 2) ait des racines en y, il faut et il suffit que ce minimum soit négatif ou nul et le minimum de E est donné par I.

L'assertion (1) de départ est alors remplacée par

$$(\forall x) \quad I(x) \leqq 0 \quad \Rightarrow \quad F(x) < 0 \qquad \text{et} \qquad R(x) > 0.$$

On a bien éliminé le quantificateur $(\forall y)$ pour $y_0 \neq 0$.

Remarque : Cette étude de tableaux de signes est automatisable. Elle utilise simplement le fait que $\deg_y(P) = 1$ et $\deg_y(E) = 2$.

Pour $y_0 = 0$, les polynômes sont tronqués et les suites de polynômes sont alors réduites à

1. $\{E,\ F_0 = (1/b^2\, C) \bmod_Y E\}$

(coef. dominants : $1/b^2$, F_0) (degré en Y: 2, 0).

2. $\{\partial E/\partial y,\ F_0,\ I = E \bmod_y (\partial E/\partial y)\}$

(coef. dominants : $2/b^2$, F_0, I) (degré en Y: 1, 0, 0).

3. $\{\partial^2 E/\partial y^2 = 2/b^2,\ F_0,\ I\}$

(coef. dominants : $2/b^2$, F_0, I) (degré en Y: 0, 0, 0),

avec

$$F_0(x) = (x^2 - 1)/b^2 - [(x - x_0)^2/a^2 - 1].$$
$$I(x) = (x - x_0)^2/a^2 - 1.$$

On retrouve la valeur de $F(x)$ pour $y_0 = 0$, et cette fois $F_0(x)$ est le résultant de E et C en y.

On remarque qu'aucun coefficient dominant ne s'annule, donc que l'élimination de y sera terminée après l'étude des tableaux de signes de ces polynômes.

La conclusion de

$$(\forall x) \quad (\forall y) \qquad (E(x, y) = 0 \Rightarrow F_0(x) < 0)$$

étant indépendante de y on peut écrire l'énoncé sous la forme

$$(\forall x) \quad ((\exists y\, E(x, y) = 0 \Rightarrow F_0(x) < 0).$$

Pour x fixé, le signe de $\partial^2 E/\partial y$ montre que E admet un minimum. Pour que E ait des racines en y, il faut et il suffit que ce minimum soit négatif ou nul et le minimum de E est donné par I. On est ainsi ramené pour $y_0 = 0$ à l'énoncé équivalent à (1)

$$(\forall x) \quad [([x - x_0]^2/a^2 - 1) \leqq 0 \Rightarrow F_0(x) < 0]$$

la variable y est éliminée.

Pour résumer, nous avons obtenu un énoncé en x équivalent à (1) sous la forme

$$\{y_0 \neq 0 \text{ et } [(\forall x)\ I(x) \leqq 0 \Rightarrow F(x) < 0 \text{ et } R(x) > 0]\}$$

ou

$$\{y_0 = 0 \text{ et } [(\forall x)\ I(x) \leqq 0 \Rightarrow F_0(x) < 0]\}.$$

Dans la suite, nous appellerons I l'intervalle $I(x) \leqq 0$.

II.2. Élimination de x

II.2.1. $y_0 = 0$

Remarques préliminaires : Le polynôme $I(x)$ se factorise de manière évidente en

$$I(x) = ((x - x_0)/a - 1)\,((x - x_0)/a + 1).$$

Pour étudier le signe de F_0 et I, nous remplacerons I par ces deux facteurs

$$I_{x_0 - a} = (x - x_0)/a + 1 \qquad \text{et} \qquad I_{x_0 + a} = (x - x_0)/a - 1.$$

Déroulement de l'algorithme : Voici alors les différentes étapes obtenues. Posons $\alpha = 1/b^2 - 1/a^2$

$1 - \{F_0, I_{x_0 - a}, I_{x_0 + a}\}$

(coef. dominants : α, $1/a$, $1/a$) (degré en X : 2, 1, 1).

2. $\{F_0', I_{x_0 - a}, I_{x_0 + a}, (4\alpha^2 F_0) \bmod_X F_0' = \operatorname{dis} F_0,$

$$((1/a^2) F_0) \bmod_X I_{x_0 - a} = (1/a^2) F_0(x_0 - a),$$

$$((1/a^2) F_0) \bmod_X I_{x_0 + a} = (1/a^2) F_0(x_0 + a)\}$$

(coef. dominants : 2α, $1/a$, $1/a$, $\operatorname{dis} F_0$, $(1/a^2) F_0(x_0 - a)$, $(1/a^2) F_0(x_0 + a)$)

(degré en X : 1, 1, 1, 0, 0, 0)

3. $\{F_0'', I_{x_0-a}, I_{x_0+a}, \operatorname{dis} F_0, (1/a^2)F_0(x_0-a), (1/a^2)F_0(x_0+a),$

$(1/a^2)F_0' \bmod_X I_{x_0-a} = (1/a^2)F_0'(x_0-a),$

$(1/a^2)F_0 \bmod_X I_{x_0+a} = (1/a^2)F_0'(x_0+a)\}$

(coef. dominants : 2α, $1/a$, $1/a$, $\operatorname{dis} F_0$, $(1/a^2)F_0(x_0-a)$, $(1/a^2)F_0(x_0+a)$),

$(1/a^2)F_0'(x_0+a)$, $(1/a^2)F_0'(x_0+a))$ (degré en X : 0, 1, 1, 0, 0, 0, 0).

Les étapes suivantes serviraient à situer l'un par rapport à l'autre x_0-a et x_0+a. Je ne les écris donc pas, a étant supposé positif.

Si le coefficient dominant $\alpha = 1/b^2 - 1/a^2 = 0$, le coefficient dominant de F_0 devient $2x_0/a^2 = F_0'$, il ne reste à l'étape 2 que F_0', I_{x_0-a}, I_{x_0+a}, $F_0(x_0-a)$ et $F_0(x_0+a)$. Le signe de F_0' est celui de x_0. Donc

Si $a^2 = b^2$ *et* $x_0 > 0$

F_0 est monotone croissante. Donc F_0 est négatif sur I si et seulement si $F_0(x_0+a) = ((x_0+a)^2 - 1)/b^2 < 0$. Le numérateur se factorise, et compte tenu des signes de a, b et x_0, ceci équivaut à $x_0 + a < 1$.

Si $a^2 = b^2$ *et* $x_0 < 0$

On obtient de même la condition $F_0(x_0-a) < 0$, c'est-à-dire $x_0 - a > -1$.

Si $a^2 = b^2$ *et* $x_0 = 0$

$F_0(x) = 1 - 1/b^2$. La condition peut s'écrire $b^2 - 1 < 0$, c'est-à-dire, l'ellipse est un cercle ($a=b$) de rayon inférieur à 1, centrée à l'origine.

Si $a^2 < b^2$

F_0'' est négatif, F_0' décroît et F_0 admet un maximum sur $\mathbb{R}$ en x_1. Le signe de F_0' en x_0-a et x_0+a situe x_1 par rapport à I.

– *Si* $F_0'(x_0-a) \geqq 0$ et $F_0'(x_0+a) \leqq 0$, x_1 appartient à l'intervalle I, le maximum de F_0 est atteint sur I. Sa valeur $F_0(x_1) = F_0 \bmod_X F_0'$ doit être strictement négative. Dans l'algorithme, on écrit cette condition $(4a^2 F_0) \bmod_X F_0' < 0$.

– Si $F_0'(x_0-a) < 0$, F_0 décroît sur l'intervalle I car $F_0' < 0$, donc F_0 négative sur I équivaut à $F_0(x_0-a) < 0$.

– Si $F_0'(x_0+a) > 0$, F_0 croît sur l'intervalle I car $F_0' > 0$, donc F_0 négative sur I équivaut à $F_0(x_0+a) < 0$.

Si $a^2 > b^2$

F_0'' est positif, F_0' croît et F_0 admet un minimum sur $\mathbb{R}$. Son maximum sur I est donc atteint en $x_0 - a$ ou $x_0 + a$, les conditions s'écrivent alors

$$F_0(x_0+a) < 0 \qquad \text{et} \qquad F_0(x_0-a) < 0.$$

Résultats simplifiés : Résumons les résultats obtenus en remplaçant les inégalités

$$F_0(x_0+a) = (x_0+a)^2 - 1/b^2 < 0 \qquad \text{et} \qquad F_0(x_0-a) = (x_0-a)^2 - 1/b^2 < 0$$

par des inégalités portant sur leurs deux facteurs, et

$$F_0'(x_0-a) = (2/ab^2)[x_0 a + (b^2-a^2)] < 0$$

par

$$x_0 a/(b^2-a^2) < -1 \quad (x_0 < 0)$$
$$F_0'(x_0+a) = (2/ab^2)[x_0 a - (b^2-a^2)] > 0$$

par

$$x_0 a/(b^2-a^2) > 1 \quad (x_0 > 0).$$

On obtient pour $y_0 = 0$

$$(a^2 = b^2 \text{ et } x_0 > 0 \text{ et } x_0 + a < 1)$$

ou

$$(a^2 = b^2 \text{ et } x_0 < 0 \text{ et } x_0 - a > -1)$$

ou

$$(a^2 = b^2 \text{ et } x_0 = 0 \text{ et } b^2 < 1)$$

ou

$$(a^2 < b^2 \text{ et } -1 \leqq x_0 a/(b^2-a^2) \leqq 1 \text{ et (dis } F_0 < 0))$$

ou

$$(a^2 < b^2 \text{ et } x_0 a/(b^2-a^2) < -1 \text{ et } -1 < x_0 - a)$$

ou

$$(a^2 < b^2 \text{ et } x_0 a/(b^2-a^2) > 1 \text{ et } x_0 + a < 1)$$

ou

$$(a^2 > b^2 \text{ et } -1 < x_0 - a \text{ et } x_0 + a < 1).$$

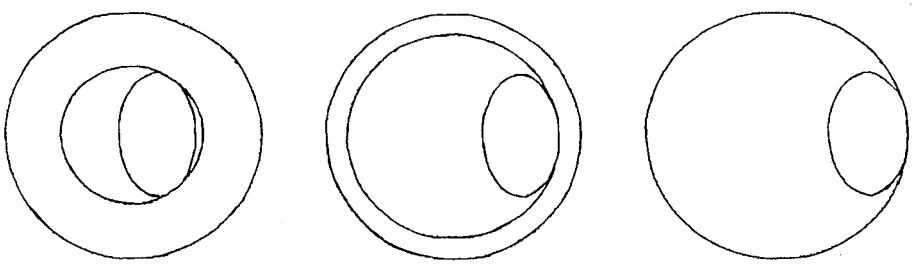
J'ai simplifié les inégalités polynômiales en $1/a$ et $1/b$ obtenues par des multiplications par $1/a^2$ $1/b^2$ ou des polynômes positifs en $1/a^2$ et $1/b^2$, a et b étant supposés positifs.

Remarques géométriques sur les résultats obtenus pour $y_0 = 0$: Lorsqu'un point (x, y) est sur l'ellipse on a

$$C(x, y) = x^2 + y^2 - 1 = F_0(x).$$

Donc $1 + F_0(x)$ nous donne la distance à l'origine d'un point (x, y) de l'ellipse. Dans le cas où l'ellipse est horizontale ($a^2 > b^2$), on voit que la distance extrêmale est atteinte en $x_0 - a$ ou $x_0 + a$.

Par contre, lorsque l'ellipse est verticale, la distance extrêmale est atteinte en $x_0 - a$ ou $x_0 + a$ seulement pour $|x_0 a/(b^2 - a^2)| > 1$. Dans l'autre cas, $-1 \leqq x_0 a/(b^2 - a^2) \leqq +1$, l'ellipse a deux points à distance extrêmale de l'origine. C'est le seul cas où l'on ait à calculer le discriminant du résultant de C et E qui nous permet d'évaluer cette distance extrêmale. Les conditions obtenues sont donc très simples du point de vue géométrique.



$x_0 = 0{,}2$
$y_0 = 0$
$a = 0{,}3$
$b = 0{,}5$
$\dfrac{x_0 a}{b^2 - a^2} = 0{,}375$

$x_0 = 0{,}533$
$y_0 = 0$
$a = 0{,}3$
$b = 0{,}5$
$\dfrac{x_0 a}{b^2 - a^2} = 1$

$x_0 = 0{,}7$
$y_0 = 0$
$a = 0{,}3$
$b = 0{,}5$
$\dfrac{x_0 a}{b^2 - a^2} = 1{,}312$

Figure 1.

II.2.2. $y_0 \neq 0$

Nous avons à éliminer x dans l'énoncé

$$(\forall x) \quad (I(x) \leqq 0 \Rightarrow (F(x) < 0 \text{ et } R(x) > 0)) \qquad (**)$$

avec

$$I(x) = (x - x_0)^2/a^2 - 1$$

$$F(x) = -I(x) + (x^2 + y_0^2 - 1)/b^2 \qquad \text{et} \qquad \deg_x F(x) = 2$$

$$R(x) = (4y_0^2/b^2)\, I(x) + F(x)^2 = \text{résultant}(E, C) \qquad \text{et} \qquad \deg_x R(x) = 4.$$

Remarques géométriques : Avant de procéder à l'élimination de x par l'algorithme, regardons la signification géométrique de la condition (**).

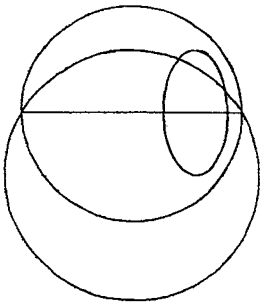
Il est clair que le résultant $R(x)$ aura un rôle à jouer dans tout algorithme d'élimination de y puisque ces zéros nous donnent l'abscisse des points d'intersection du cercle et de l'ellipse. Pour ce qui est du signe de F, regardons d'abord la signification de $F(x) < 0$ sur I.

Si l'on prend (x, y) point de l'ellipse au-dessus d'un point x de I, on a

$$C = b^2 P = 2y_0 (y - y_0) + b^2 F(x),$$

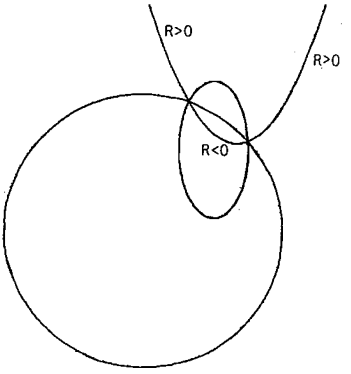
ce qui s'écrit aussi

$$x^2 + (y - y_0)^2 = 1 - y_0^2 + b^2 F(x).$$



$x_0 = 0{,}5$
$y_0 = 0{,}5$
$a = 0{,}25$
$b = 0{,}5$

Figure 2.



$x_0 = 5$
$y_0 = 6$
$a = 2{,}5$
$b = 5$

Figure 3.

Le signe de $F(x)$ sur I donne donc la position d'un point correspondant (x, y) de l'ellipse par rapport à un cercle C_{y_0} de centre $(0, y_0)$ et de rayon $1-y_0^2$. C'est le cercle coupant le cercle unité pour $y = y_0$.
Si $F(x) < 0$, $x^2 + (y-y_0)^2 < 1-y_0^2$, donc $|y_0| < 1$, et l'ellipse est intérieure au cercle C_{y_0} de rayon inférieur à 1. Cette condition de degré 2 en x implique donc que l'ellipse ne soit pas « trop grande » ni « trop éloignée » de l'origine. Mais de manière plus précise pour $a < b$ et $F(x) < 0$, l'ellipse ne coupe le cercle unité qu'en deux points au plus, ce que l'on retrouvera dans l'étude du signe de $R(x)$ pour $a < b$. Par contre, si $a > b$, on peut toujours avoir quatre points d'intersection, donc quatre zéros pour R. Cette condition apporte beaucoup moins dans ce cas. Cette remarque géométrique confirme ce que j'avais en fait trouvé d'abord par les calculs : il est préférable pour $a < b$ d'éliminer y d'abord, et d'éliminer x d'abord, pour $a > b$.

La condition, $R(x) > 0$ sur I, est moins intéressante. Avec les divisions de l'algorithme, on peut voir en prenant un point (x, y) sur la parabole $P(x, y) = 0$, que le signe de $R(x)$ est positif sur la parabole, si (x, y) est extérieur au cercle et à l'ellipse, et négatif, s'il est intérieur aux deux, la parabole étant la seule courbe du faisceau de degré 1 en y. La condition $R(x) > 0$ pour $I(x) \leqq 0$ équivaut en fait à $R(x) \neq 0$ pour $I(x) \leqq 0$, c'est-à-dire l'ellipse ne rencontre pas le cercle. La position (intérieure ou extérieure) va être précisée par le signe de $F(x)$.

Revenons à l'algorithme de Hörmander.

Remarques préliminaires : Pour étudier les polynômes $F(x)$ et $R(x)$ sur l'intervalle I, on est conduit à calculer le polynôme R''

$$R''(X) = 4y_0^2/b^2 I''(x) + 2(F'^2 + FF'') = 4y_0^2/a^2 b^2 + 2F'^2 + 2F(1/b^2 - 1/a^2).$$

Lorsque $F < 0$ sur I, on remarque alors que $R'' > 0$ pour $a^2 \leqq b^2$, c'est-à-dire $a \leqq b$. Cette remarque raccourcit considérablement l'étude du signe de R (polynôme de degré 4) par l'algorithme. Nous nous placerons donc par la suite dans l'hypothèse $a \leqq b$. Dans le cas $a > b$, on éliminera la variable x en premier, et on obtiendra les conditions sur x_0, y_0, a et b en intervertissant le rôle des variables. Cette remarque nous conduit aussi à écrire d'abord les conditions pour que $F(x)$ soit négatif sur I.

En supposant ces conditions sur F vérifiées, nous chercherons à obtenir $R > 0$ sur I. Nous arrêterons le calcul des dérivées de R, à R'', dont le signe est connu, positif sur I. Les restes modulo R'' à la troisième étape ne seront pas calculés car R'' n'est jamais nul sur I.

On peut remarquer, sur les résultats obtenus dans le cas $y_0 = 0$, que l'étude par l'algorithme du signe d'un polynôme A sur l'intervalle I qui doit se faire en établissant un tableau de signes pour $A(x)$ et $I(x)$, revient en fait à établir le tableau de signes de A, en y ajoutant la valeur des différents polynômes en $x_0 - a$ et $x_0 + a$. Nous ajouterons donc à chaque étape les valeurs des polynômes en $x_0 - a$ et $x_0 + a$.

Nous laisserons de côté le cas trivial où le coefficient dominant de R: $(1/b^2 - 1/a^2)^2$ est nul (cas d'un cercle).

Étudions le cas général $y_0 \neq 0$ et $a < b$.

Déroulement de l'algorithme : Écrivons les suites successives de polynômes pour F en y ajoutant à chaque étape la valeur en $x_0 - a$ et $x_0 + a$. Si l'on note $\alpha = 1/b^2 - 1/a^2$, on remarque que dans ce cas α est négatif.

1. $\{F\}$

(coef. dominants: α) (degré en X: 2).

2. $\{F(x_0 - a), F(x_0 + a), F', (4\alpha^2 F) \bmod_X F' = \operatorname{dis} F\}$

(coef. dominants: $F(x_0 - a)$, $F(x_0 + a)$, α, $\operatorname{dis} F$) (degré en X: 0, 0, 1, 0).

3. $\{F(x_0-a),\ F(x_0+a),\ F'(x_0-a),\ F'(x_0+a),$
$F''=2\alpha,\ (4\alpha^2 F) \bmod_X F' = \operatorname{dis} F\}$

(degré en X: 0, 0, 0, 0, 0, 0).

$F'' < 0$, donc F' est strictement décroissante et F admet un maximum au zéro de F'.

– Si $F'(x_0-a) < 0$ alors $F' < 0$ sur I et F décroît sur I. Le maximum de F est atteint en x_0-a. $F < 0$ sur I se traduit par $F(x_0-a) < 0$.

– Si $F'(x_0+a) > 0$, $F' > 0$ sur I et F croît sur I. De même F négatif sur I se traduit par $F(x_0+a) < 0$.

– Si $F'(x_0-a) \geqq 0$ et $F'(x_0+a) \leqq 0$, le maximum de F est sur I et son signe est celui de $(4\alpha^2 F) \bmod_X F'$. La condition s'écrit alors

$$\operatorname{dis} F < 0.$$

Remarques géométriques : On remarque que $F(x_0+a) < 0$ et $F(x_0-a) < 0$ sont des conditions qui traduisent simplement (x_0-a, y_0) et (x_0+a, y_0) intérieurs au cercle unité. Les conditions $F'(x_0-a) < 0$ et $F'(x_0-a) > 0$ traduisent respectivement

$$(x_0-a) < -b^2/a < 0 \qquad \text{et} \qquad (x_0+a) > b^2/a > 0$$

ce qui minore $|x_0-a|$ (resp. x_0+a) par le rayon de courbure de l'ellipse au point (x_0-a, y_0) [resp. (x_0+a, y_0)]: b^2/a. Elle signifie que l'ellipse est intérieure au cercle de centre $(0, y_0)$ de rayon $|x_0-a|$ ou $|x_0+a|$; dans le cas contraire l'ellipse a deux points de contact avec le cercle centré en $(0, y_0)$ qui l'enveloppe. Nous retrouvons des résultats analogues au cas $y_0=0$.

Reprenons l'algorithme. Pour le signe de R, on obtient les suites successives de polynômes auxquelles s'ajouteront donc à chaque étape les valeurs des polynômes en x_0-a et x_0+a.

1. $\{R\}$

(coef. dominants: α^2) (degré en X: 4).

2. $\{R',\ P_1=(16\alpha^4 R) \bmod_X R'\}$

(coef. dominants: $4\alpha^2$, P_1'') (degré en X: 3, 2),

3. $\{R'',\ P_1,\ P_2=(P_1'')^2 R' \bmod_X P_1\}$

(coef. dominants: $12\alpha^2$, P_1'', P_2') (degré en X: 2, 2, 1).

4. $\{P_1', R'', P_2, P_3 = (P_2''^2 P_1) \bmod_X P_1', P_4 = (P_2'^2 P_1) \bmod_X P_2\}$

(coef. dominants : P_1'', $12\alpha^2$, P_2', P_3, P_4) (degré en X: 1, 2, 1, 0, 0)

5. $\{P_1'', R'', P_2, P_3, P_4, P_5 = (P_2'^2 P_1') \bmod_X P_2\}$

(coef. dominants : P_1'', $12\alpha^2$, P_2', P_3, P_4, P_5) (degré en X: 0, 2, 1, 0, 0, 0)

6. $\{P_2', R'', P_1'', P_3, P_4, P_5\}$ (degré en X: 0, 2, 0, 0, 0, 0).

On remarque que, parmi les constantes, on retrouve les coefficients dominants des polynômes de la suite de Sturm de R: P_1'', P_2'. Et P_4 est, à un coefficient multiplicatif près, le discriminant du résultant R de C et E.

En général, on a 15 conditions à tester, mais en fait nous allons voir que chaque branche de l'arbre des conditions ne comporte que 9 tests au plus, car on teste les polynômes soit en $x_0 - a$ et $x_0 + a$, soit en un point de l'intervalle, zéro d'un autre polynôme.

Avant d'étudier le signe de P_1 qui nous donnera le signe des extrêmums de R, regardons si R admet ou non un extrêmum sur I.

Comme $R'' > 0$ sur I, R' croît. Distinguons selon que R' admet ou non un zéro sur I.

– *Si* $R'(x_0 - a) > 0$, $R' > 0$ sur I donc R est croissant sur I et est positif sur I si et seulement si $R(x_0 - a) > 0$.

– *Si* $R'(x_0 + a) < 0$, $R' < 0$ sur I, R est décroissant sur I et est positif sur I si et seulement si $R(x_0 + a) > 0$.

– *Si* $R'(x_0 - a) \leqq 0$ *et* $R'(x_0 + a) \geqq 0$, R' s'annule sur I une seule fois puisque $R'' > 0$. $R > 0$ sur I équivaut alors à $R > 0$ en son minimum sur I pour $R' = 0$. Ce qui équivaut à $P_1 > 0$ en l'unique zéro de R' sur I.

Il reste alors à regarder selon les signes des constantes de la sixième étape les tableaux de signes qui conviennent à la deuxième étape. Nous commençons par établir les variations de P_1 en utilisant P_1'', puis le signe de P_1 en ses extrêmums donné par P_3. Ensuite il faut situer les zéros de P_1 par rapport à I, puis par rapport au zéro de R' sur I. On utilise alors P_2 et P_1'.

Si $P_1'' > 0$

P_1' est croissante et P_1 admet un minimum sur $\mathbb{R}$, la valeur de ce minimum est P_3.

Si $P_3 > 0$

P_1 est partout positif, donc positif en l'unique zéro de R' sur I.

Si $P_3 = 0$

$P_1 > 0$ sauf au zéro commun avec l'unique zéro de P_1', il faut et suffit que ce point ne soit pas l'unique zéro de R' dans I.

Pour cela, soit ce zéro est extérieur à I $(P_1'(x_0+a)\,P_1'(x_0-a) > 0)$, soit il est dans I et $R' \neq 0$ quand $P_1 = P_1' = 0$, c'est-à-dire $P_2 \neq 0$ quand $P_1 = P_1' = 0$, ou encore, puisque P_2 a un unique zéro car de degré 1, il faut et il suffit que ce zéro de P_2 ne soit pas zéro de P_1, ce qui donne la condition $P_4 \neq 0$. En résumé pour $P_3 = 0$, il faut

$$P_1'(x_0+a)\,P_1'(x_0-a) > 0$$

ou

$$(P_1'(x_0+a)\,P_1'(x_0-a) \leqq 0 \text{ et } P_4 \neq 0).$$

Si $P_3 < 0$

Dans ce cas, P_1 a deux racines et il est positif à l'extérieur des racines. Il faut donc que lorsque $R' = 0$ sur I, on se trouve à l'extérieur des racines de P_1. Regardons toutes les situations possibles.

(1) Si les zéros de P_1 sont tous les deux avant I ou tous les deux après I,

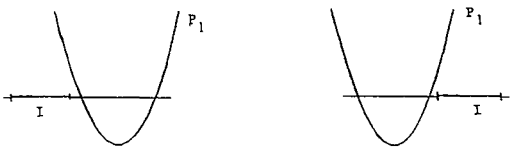


Figure 4.

P_1 est positif sur I. On peut l'écrire $P_1(x_0-a) > 0$ et $P_1(x_0+a) > 0$ et P_1' ne s'annule pas sur I, c'est-à-dire

$$P_1(x_0-a) > 0 \qquad \text{et} \qquad P_1(x_0+a) > 0$$

et

$$P_1'(x_0-a)\,P_1'(x_0+a) > 0.$$

(2) Si les zéros de P_1 sont de chaque côté de I, $P_1 \leqq 0$ sur I, donc ne répond pas à la question.

La condition

$$P_1(x_0-a) \leqq 0 \qquad \text{et} \qquad P_1(x_0+a) \leqq 0$$

donne une réponse négative.

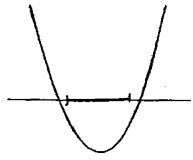


Figure 5.

(3) Si les zéros de P_1 sont tous les deux dans I, ce qui peut se traduire par $P_1(x_0-a) > 0$ et $P_1(x_0+a) > 0$ et $P_1'(x_0-a)\,P_1'(x_0+a) \leqq 0$ (minimum de P_1' sur I), alors le zéro de R' est extérieur aux racines de P_1

– si et seulement si R' a le même signe non nul en les deux racines de P_1,

– si et seulement si P_2 a le même signe non nul en les deux racines de P_1. Comme $\deg P_2 \leqq 1$, on peut traduire cela en [$P_2' \neq 0$ et P_2 s'annule à l'extérieur des racines de P_1 ($P_1 > 0$)] ou ($P_2' = 0$ et $P_2 \neq 0$).

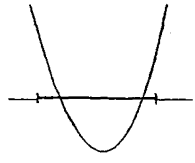


Figure 6.

On obtient ainsi la condition

$$(P_1(x_0-a) > 0 \qquad \text{et} \qquad P_1(x_0+a) > 0$$

et

$$P_1'(x_0-a)\,P_1'(x_0+a) \leqq 0$$

et

$$[(P_2' \neq 0 \text{ et } P_4 > 0) \text{ ou } (P_2' = 0 \text{ et } P_2 > 0)].$$

(4) Si $P_1(x_0-a) \leqq 0$ et $P_1(x_0+a) > 0$ alors si x_1 et x_2 sont les deux racines de P_1 on a $x_1 \leqq x_0-a \leqq x_2 < x_0+a$ on veut que $R'(x)=0$ pour $x \in]x_2, x_0+a]$.

Comme R' croît, ceci équivaut à $R'(x_2) < 0$ ou encore $P_2(x_2) < 0$. Mais P_2 est de degré 1, donc monotone en x, donc

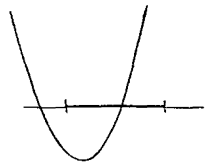


Figure 7.

$$P_2(x_2) < 0 \quad \Leftrightarrow \quad \begin{array}{l} P_2 \text{ décroît et } P_2(x)=0 \text{ pour un } x \text{ de }]-\infty, x_2[\\ \text{ou} \\ P_2 \text{ croît et } P_2(x)=0 \text{ pour un } x \text{ de }]x_2+\infty[\\ \text{ou} \\ P_2 \text{ constant et } P_2 < 0. \end{array}$$

L'algorithme nous donne le signe de P_1 : (P_4) et de P_1' : (P_5), lorsque P_2 s'annule.

L'intervalle $]-\infty, x_2[$ est la réunion de $]-\infty, x_1[\cup]x_1, x_2[$ caractérisé respectivement par $(P_1 \geqq 0$ et $P_1' < 0)$ et par $(P_1 < 0)$. L'intervalle $]x_2+\infty[$ est caractérisé par $P_1 > 0$ et $P_1' > 0$.

Les trois conditions précédentes peuvent donc s'écrire

$$P_2(x_2) < 0 \quad \Leftrightarrow \quad \begin{array}{l} P_2' < 0 \quad \text{et} \quad P_4 \geqq 0 \quad \text{et} \quad P_5 < 0 \\ \text{ou} \\ P_2' < 0 \quad \text{et} \quad P_4 < 0 \\ \text{ou} \\ P_2' > 0 \quad \text{et} \quad P_4 > 0 \quad \text{et} \quad P_5 > 0 \\ \text{ou} \\ P_2' = 0 \quad \text{et} \quad P_2 < 0 \end{array}$$

(5) Si $P_1(x_0-a) > 0$ et $P_1(x_0+a) \leqq 0$ alors si x_1 et x_2 sont les deux racines de P_1 on a $x_0-a < x_1 \leqq x_0+a \leqq x_2$ et on veut $R'(x)=0$ pour $x \in [x_0-a, x_1[$.

Comme R' croît, ceci équivaut à $R'(x_1) > 0$ ou encore $P_2(x_1) > 0$. Comme P_2 est de degré 1, donc monotone (pour $P_2' \neq 0$), on aboutit à

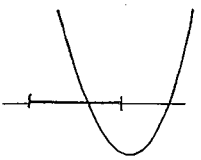


Figure 8.

$$P_2(x_2) > 0 \Leftrightarrow \begin{array}{l} P_2 \text{ décroît et } P_2(x) = 0 \text{ pour un } x \text{ de }]x_1, +\infty[\\ \text{ou} \\ P_2 \text{ croît et } P_2(x) = 0 \text{ pour un } x \text{ de }]-\infty, x_1[\\ \text{ou} \\ P_2' = 0 \quad \text{et} \quad P_2 > 0. \end{array}$$

En utilisant de nouveau les caractérisations des intervalles $]-\infty, x_1[$, $]x_1, x_2[$, $]x_2, +\infty[$ par les signes de P_1 et P_1', on obtient

$$P_2(x_1) > 0 \Leftrightarrow \begin{array}{l} (P_2' < 0 \text{ et } (P_4 < 0 \text{ ou } (P_4 \geqq 0 \text{ et } P_5 < 0))) \\ \text{ou} \\ (P_2' > 0 \text{ et } P_4 > 0 \text{ et } P_5 < 0) \\ \text{ou} \\ (P_2' = 0 \text{ et } P_2 > 0). \end{array}$$

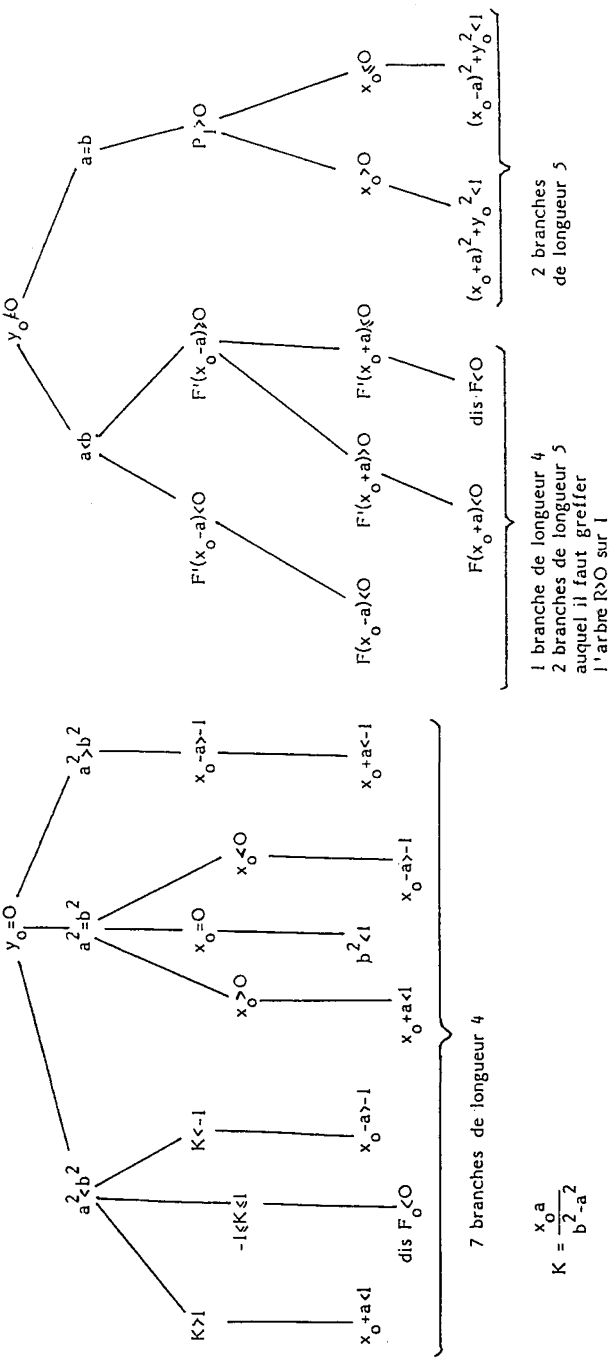
Ceci termine l'étude du cas $P_3 < 0$ avec $P_1'' > 0$. On étudie de même les cas $P_1'' < 0$ et $P_1'' = 0$ (voir les détails dans ma thèse).

Nous obtenons ainsi toutes les conditions possibles en x_0, y_0 a et b pour que l'ellipse soit incluse dans le cercle. On peut les résumer sous forme d'arbre.

III. COMPARAISON DES DIFFÉRENTES MÉTHODES

La conception de l'algorithme de Hörmander [5, 9] est totalement différente de celle de l'algorithme de Collins. Elle est semi-algébrique, portant entièrement sur des inégalités et des résultats purement réels : théorème des valeurs

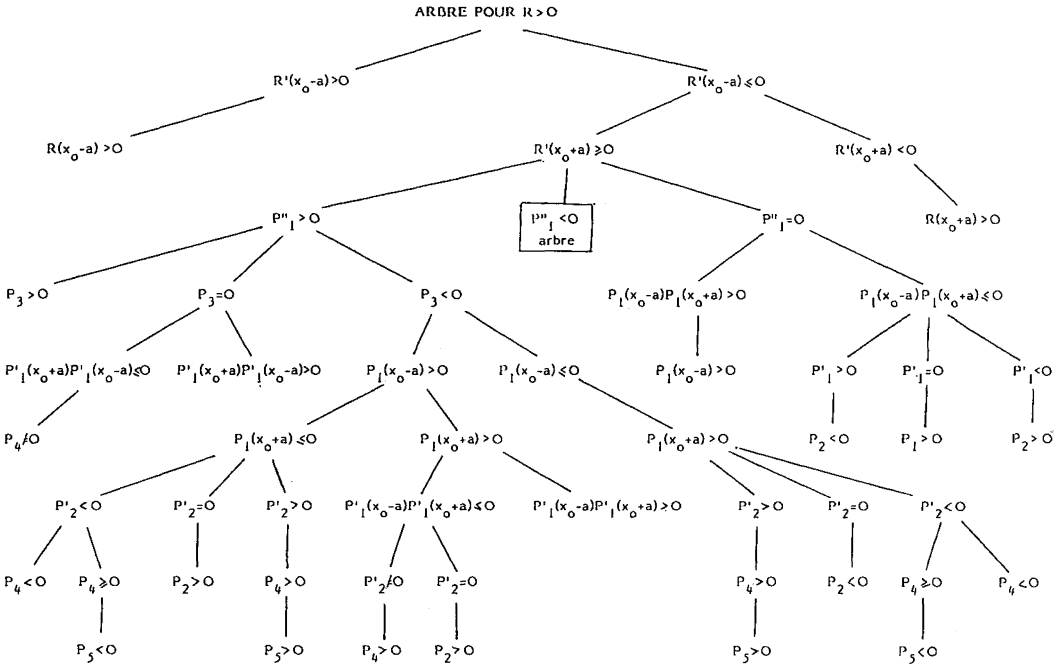
ARBRE DEDUIT DE HORMANDER

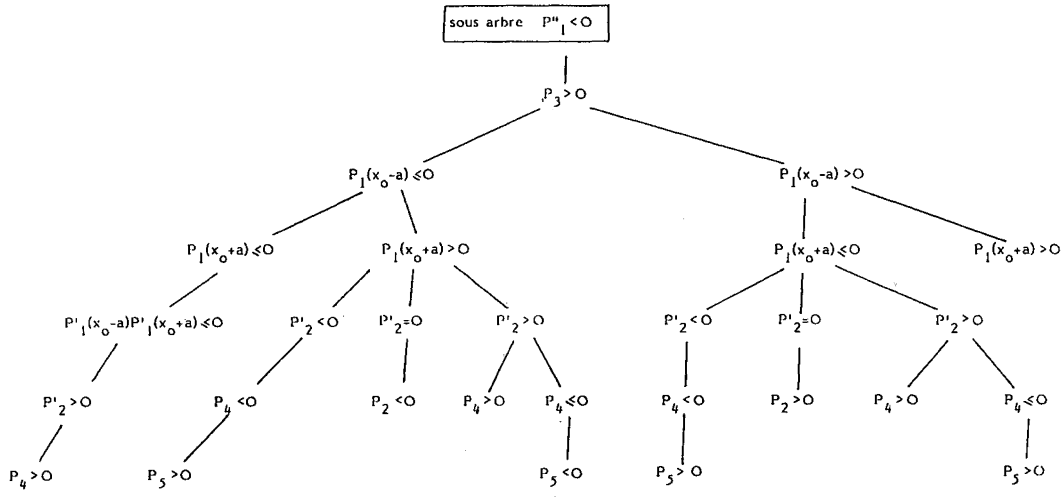


$$K = \frac{x_o a}{b^2 - a^2}$$

Si $x_o = 0$, éliminer x d'abord (arbre similaire à $y_o = 0$)

Si $a > b$ et $x_o \neq 0$, éliminer x d'abord (arbre similaire à $y_o \neq 0$ et $a < b$)


ARBRE POUR R>O
R'(x0-a)>O
R'(x0-a)≤O
R(x0-a)>O
R'(x0+a)≥O
R'(x0+a)<O
P''1>O
P''1<O arbre
P''1=O
R(x0+a)>O
P3>O
P3=O
P3<O
P1(x0-a)P1(x0+a)>O
P1(x0-a)P1(x0+a)≤O
P'1(x0+a)P'1(x0-a)≤O
P'1(x0+a)P'1(x0-a)>O
P1(x0-a)>O
P1(x0-a)≤O
P1(x0-a)>O
P'1>O
P'1=O
P'1<O
P4≠O
P1(x0+a)≤O
P1(x0+a)>O
P1(x0+a)>O
P2<O
P1>O
P2>O
P'2<O
P'2=O
P'2>O
P'1(x0-a)P'1(x0+a)≤O
P'1(x0-a)P'1(x0+a)>O
P'2>O
P'2=O
P'2<O
P4<O
P4≥O
P2>O
P4>O
P'2≠O
P'2=O
P4>O
P2<O
P4≥O
P4<O
P5<O
P5>O
P4>O
P2>O
P5>O
P5<O


sous arbre P''1<O
P3>O
P1(x0-a)≤O
P1(x0-a)>O
P1(x0+a)≤O
P1(x0+a)>O
P1(x0+a)≤O
P1(x0+a)>O
P'1(x0-a)P'1(x0+a)≤O
P'2<O
P'2=O
P'2>O
P'2<O
P'2=O
P'2>O
P'2>O
P4<O
P5>O
P4<O
P2<O
P4>O
P4≤O
P4<O
P2>O
P4>O
P4≤O
P5<O
P5>O
P5>O

intermédiaires, variation d'un polynôme en fonction du signe de sa dérivée, signe d'un polynôme à l'infini.

Par contre, l'algorithme de Collins [1] est plutôt algébrique, utilisant dans sa récurrence des conditions ($=0$) ou ($\neq 0$) et des arguments portant sur la clôture algébrique du corps réel clos : propriétés des résultants et des sous-résultants. Ces calculs lui permettent de déterminer successivement pour chaque variable, le nombre de zéros dans la clôture algébrique. Les méthodes réelles et les inégalités s'introduisent pour calculer les zéros réels [8] de la dernière variable en appliquant le théorème de Sturm [10]. La suite de l'algorithme, pour déterminer le signe des polynômes en 1 puis 2, ..., n variables utilise les calculs sur les nombres algébriques [15].

Une autre différence essentielle entre ces deux algorithmes est que Hörmander n'élimine que les variables quantifiées, alors que Collins élimine tous les paramètres. De ce fait il y a des cas où la méthode de Hörmander me paraît clairement plus efficace que celle de Collins : C'est lorsque l'on a beaucoup de paramètres et peu de quantificateurs. Pour le montrer prenons un exemple beaucoup plus simple que celui de Kahan. Trouver une combinaison booléenne de conditions polynomiales en x_0, y_0, a, b, x pour que $(\forall y)$ $(E=0 \Rightarrow C \leqq 0)$. C'est l'élimination d'une seule variable y avec cinq variables non quantifiées. On a vu que Hörmander nous donnait un résultat rapide avec des conditions polynomiales portant sur cinq « constantes » y_0, R, I, F, F_0 (cas $y_0=0$). Avec Collins, on aboutit toute de suite à l'étude de $R(x)$ pour $y_0 \neq 0$ et $(x^2-1)/b^2 F_0(x)$ pour $y_0=0$. Mais on ne sait pas quelle condition de signes exiger sur ces « constantes », on est donc ramené à l'étude du polynôme de degré 4 : $R(x)$, qui ne peut se traiter directement à la machine sans simplification préalable du problème [4].

Dans la disjonction obtenue par l'algorithme de Hörmander, certains cas peuvent être vides, ceux qui correspondent à des conditions incompatibles sur les paramètres. Un moyen de déterminer lesquelles de ces conditions sont incompatibles serait d'éliminer toutes les variables.

Cependant un avantage de l'algorithme de Hörmander, est qu'il permet d'obtenir une procédure de décision souvent rapide, calculant pour chaque valeur des paramètres en entrée, les seuls pseudo-restes utiles, évitant dans de nombreux cas le calcul du résultant. On peut le constater sur l'exemple de l'ellipse et du cercle : de nombreuses branches de l'arbre de décision ne font pas intervenir P_4 (discriminant/X du résultant/Y de l'ellipse et du cercle). Certaines branches sont très courtes.

Par contre il faut reconnaître que la croissance exponentielle du nombre de polynômes en fonction du degré rend difficile l'application directe de cet algorithme pour un polynôme général de degré 4.

Comparons les différents résultats obtenus dans le cas de l'ellipse et du cercle. Dans le cas $y_0 = 0$, l'algorithme de Hörmander est très performant puisque l'élimination de Y est automatique, aboutissant directement à des polynômes de degré 2 en X. L'élimination de X est également facile compte tenu du degré. Tandis que par l'algorithme de Collins, pour le seul calcul de la décomposition cellulaire, Arnon et Mignotte mettent 75 minutes pour trouver 2 291 cellules. Il faut ensuite sélectionner par des tests dans chaque cellule, les bonnes cellules. Puis rassembler les cellules définies par les mêmes inégalités, en visualisant géométriquement la décomposition obtenue.

De plus il est intéressant de comparer les résultats de la méthode de Hörmander avec les résultats obtenus géométriquement. On peut remarquer que Mignotte [16] obtenait exactement les mêmes polynômes en utilisant les multiplicateurs de Lagrange, traduisant que les gradients du cercle et de l'ellipse sont colinéaires pour les extrêmums de C lorsque $E = 0$. Rappelons que la condition, dis $F_0 < 0$, n'était utile que dans le cas où l'ellipse pour $a < b$ avait deux points à distance extrêmale de l'origine, points de tangence avec un cercle centré à l'origine (II.2.1). Cette condition, après simplification par des constantes strictement positives, peut s'écrire $(b^2 x_0^2 - (b^2 - 1)(a^2 - b^2)) < 0$.

Lazard [13] a trouvé une amélioration de ces conditions pour $0 < a < b$ par une remarque géométrique : si le rayon de courbure maximal de l'ellipse : b^2/a est plus petit que 1, les deux points de tangence éventuels de l'ellipse avec un cercle centré à l'origine ne peuvent se trouver à l'extérieur du cercle unité. La condition $(b^2 x_0^2 - (b^2 - 1)(a^2 - b^2) < 0)$, disant que ces points sont intérieurs au cercle unité (II.2.1), ne reste plus à tester que dans le cas $b^2/a \geqq 1$. Effectivement, on vérifie bien par le calcul que si $b^2 < a < b$ et $-1 \leqq x_0 a/(b^2 - a^2) \leqq 1$ alors on a toujours $b^2 x_0^2 - (b^2 - 1)(a^2 - b^2) < 0$.

Pour $y_0 \neq 0$, on peut constater que toutes les solutions comportent une partie non algorithmique. Le nombre de polynômes obtenus avec Hörmander est trop important. La difficulté rencontrée par Arnon porte sur le nombre de cellules à tester. Pour éliminer Y, il fait une paramétrisation de l'ellipse $X = 2t/(1+t^2)$, $Y = (1-t^2)/(1+t^2)$, qui le ramène facilement à l'étude d'un polynôme de degré 4 en t. Seule l'élimination de X est partiellement automatisable, avec un travail préliminaire simplifiant l'étude du signe d'un polynôme de degré 4.

Un autre problème se pose pour tous les algorithmes : c'est la taille des conditions polynomiales obtenues par un travail direct sur l'équation du cercle et de l'ellipse. Sur ce point on obtient une légère amélioration en appliquant l'algorithme de Collins exposé dans l'article de Loos [14], pour le calcul des suites de pseudo-restes. Mais les résultats obtenus restent encore énormes : le polynôme P_4 (discriminant/X du résultant/Y de l'ellipse et du cercle) fait 11 205 lignes de listing ! Et ce polynôme apparaît dans les deux algorithmes.

Signalons qu'un programme en Pascal mis au point par B. Bougaut tourne sur quelques étapes de la remontée des tableaux de signes, selon la méthode de Hörmander, pour l'étude du signe d'un polynôme.

CONCLUSION

Pour conclure ce travail, je pense que l'on devra encore bien améliorer les algorithmes existants d'élimination des quantificateurs pour qu'ils deviennent vraiment utilisables de manière entièrement automatique, les solutions proposées au problème de Kahan utilisant toutes des astuces diverses et peu algorithmiques.

BIBLIOGRAPHIE

1. D. S. Arnon, G. E. Collins et S. McCallum, *Cylindrical Algebraic Decomposition I and II: the Basic Algorithm*, Siam J. Comput., vol. 13, n° 4, nov. 84, p. 865-889.
2. D. S. Arnon, *Towards Mechanical Solution of Kahan Ellipse Problem I*, Computer Algebra, Lectures Notes, 162, Springer-Verlag, 1983.
3. D. S. Arnon, *On Mechanical Quantifier Elimination For Elementary*. Algebra and Geometry: Solution of a non Trivial Problem, Eurocal 85, Lectures Notes 204, p. 270-271, Springer-Verlag, 1985.
4. D. S. Arnon et M. Mignotte, *On Mechanical Quantifier Elimination For Elementary Algebra and Geometry*, J. Symbolic Computation, Vol. 5, 1988, p. 237-259.
5. J. Bochnak et M. Coste, M.-F. Roy, *Géométrie Algébrique Réelle*, Ergebnisse der Mathematik, Springer-Verlag, 1987.
6. W. S. Brown et J.-F. Traub, *On Euclid's Algorithm and the Theory of Subresultants*, J. Assoc. Comput. Math., vol. 18, n° 4, 1971, p. 505-514.
7. G. E. Collins, *Quantifier Elimination for Real Closed Fields: a Guide to the Litterature*, Computer Algebra Symbolic and Algebraic Computation, Springer-Verlag, 1982-1983.
8. G. E. Collins et R. Loos, *Real Zeros of Polynomials*, Computer Algebra Symbolic and Algebraic Computation, Springer-Verlag, 1982-1983.

9. HÖRMANDER, *The Analysis of Linear Partiel Differential Operators*, tome 2, Springer-Verlag, 1983.
10. N. JACOBSON, *Basic Algebra I*, San Francisco, Freeman, 1974.
11. W. KAHAN, *« Problem=9: an Ellipse Problem »*, SIGSAM Bulletin of the Assoc. Comp. Math., vol. 9, 1975, p. 11.
12. M. LAUER, *A solution to Kahan's problem* (SIGSAM problem n° 9); SIGSAM Bulletin of the Ass. Com. Math., vol. 11, 1977, p. 16-20.
13. D. LAZARD, *Quantifier Elimination: Optimal Solution for 2 Classical Examples*, J. Symbolic Computation, vol. 5, 1988, p. 261-266.
14. R. LOOS, *Generalized Polynomial Remainder Sequences*, Computer Algebra Symbolic and Algebraic Computation, Springer-Verlag, 1982-1983.
15. R. LOOS, *Computing in Algebraic Extensions*, Computer Algebra Symbolic and Algebraic Computation, Springer-Verlag, 1982-1983.
16. M. MIGNOTTE, *Solution au problème de Kahan* (non publié).
17. A. PAUGAM, *Comparaison entre 3 algorithmes d'élimination des quantificateurs sur les corps réels clos*, Thèse, 1986.
18. A. SEIDENBERG, *A New Decision Method for Elementary Algebra*, Ann. of Math. 60, 1954, p. 365-374.
19. A. TARSKI, *A Decision Method for Elementary Algebra and Geometry*, Prepared for publication by J. C. C. MacKinsey, Berkeley, 1951.