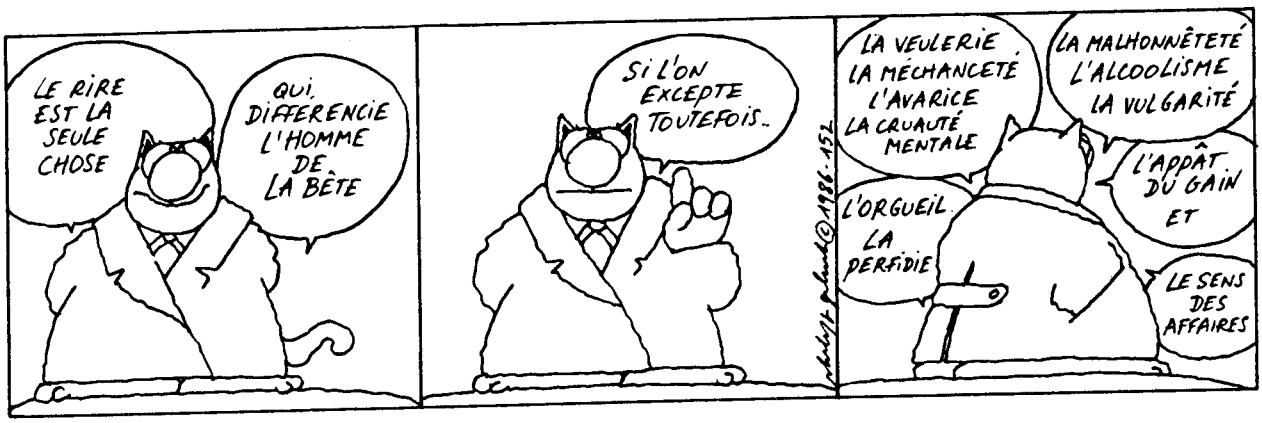
NOMBRES ALGEBRIQUES ET APPROXIMATIONS


LE RIRE EST LA SEULE CHOSE
QUI DIFFERENCIE L'HOMME DE LA BÊTE
SI L'ON EXCEPTE TOUTEFOIS..
LA VEULERIE LA MÉCHANCETÉ L'AVARICE LA CRUAUTÉ MENTALE
L'ORGUEIL LA PERFIDIE
LA MALHONNÊTETÉ L'ALCOOLISME LA VULGARITÉ
L'APPÂT DU GAIN ET
LE SENS DES AFFAIRES
Mark Alan © 1986 - 152

NOMBRES ALGEBRIQUES ET APPROXIMATIONS

Résumé

Nous donnons tout d'abord une description très simple des nombres algébriques réels et disutons la calculabilité en temps polynomial des opérations arithmétiques et de la recherche des zéros relativement à cette description.

Nous discutons ensuite le même problème pour une description beaucoup plus sophistiquée des nombres algébriques, réels ou complexes. Cette description est basée sur le système D5 . Nous donnons dans ce cadre des majorations de temps de calcul uniformément polynomiales par rapport à la taille des entrées **et au** "degré a priori" des nombres algébriques décrits.

Nous discutons l'efficacité des méthodes approximatives, et leur necessité lorsqu'on manipule de "vrais" nombres réels ou complexes, en particulier pour ce qui concerne la recherche des racines.

Ceci nous conduit à étudier la classe des fonctions définies sur un intervalle compact et approchables en temps polynomial par des polynômes à coefficients rationnels, au sens de la norme uniforme. Cette classe de fonctions est en fait la classe des fonctions Gevrey qui sont calculables en temps polynomial au sens de Ko-Friedman. Nous obtenons dans ce cadre des théorèmes agréables et de démonstration très simple, qui améliorent les résultats précédents de Ko-Fridman et de Müller sur les fonctions analytiques calculables en temps polynomial.

Mots clé

Nombres algébriques, codage, calcul formel, système D5, calculabilité en temps polynomial, fonctions calculables en temps polynomial, classe de Gevrey.

ALGEBRAIC NUMBERS AND APPROXIMATIONS

Abstract

We give a very simple description of real algebraic numbers and discuss the polynomial time computability of arithmetic operations and searching roots (relatively to this description).

We discuss then the same problem for a much more sophisticated description of real or complex algebraic numbers. This description is based upon the D5 system. We give systematic uniformly polynomial majorations (for the computing time) relatively to the input size **and** the "a priori degree" of the described algebraic numbers.

We discuss the efficiency or approximative methods, and their necessity when we have to handle Cauchy real or complex numbers, in particular for the root searching problem.

This leads us to study the class of functions on a compact interval that are polynomial time approximable (in the uniform norm) by rational polynomials. This class of functions is in fact the class of Gevrey functions that are polynomial time computable (in the sense of Ko-Friedman). We obtain good and simple theorems concerning this class, improving previous results of Ko-Friedman and Müller on polynomial time computable analytic functions.

Introduction

Ce travail se situe dans la lignée directe de **[Lom1]** . Nous reprenons ici en partie la terminologie développée dans cet article, en l'explicitant autant que possible à chaque fois. Après l'algèbre linéaire en temps polynomial dans les corps les plus usuels (extensions de type fini de $\mathbb{Q}$ ou d'un corps fini), nous étudions maintenant dans quelle mesure les calculs dans la clôture algébrique de $\mathbb{Q}$ peuvent être présentés de manière à être en temps polynomial. Comme on peut s'y attendre, les résultats sont moins agréables et une explosion exponentielle de la taille des objets manipulés et du temps de calcul semble à peu près inévitable.

L'importance d'avoir une bonne description des nombres algébriques réels ou complexes dans les systèmes de calcul formel n'est plus à démontrer. Nous développons dans cette étude quelques considérations à ce sujet. Le point de vue qui nous guide est de montrer l'efficacité des méthodes approximatives dans la solution de ce problème et de problèmes connexes.

Dans le chapitre A , nous explicitons la présentation des nombres algébriques réels la plus naïve qu'on puisse imaginer : un nombre algébrique réel est donné par un polynôme P de $\mathbb{Z}[X]$ qui l'annule et par un intervalle où ce polynôme change de signe tout en étant strictement monotone. Pour que cette dernière condition soit tout à fait simple à constater, nous demandons que la dérivée de P reste de signe constant de manière évidente, en donnant un sens précis à ceci. Autrement dit, aucun recours au théorème de Sturm, et aux calculs de polynômes sous-résultants qu'il implique, n'est utilisé dans cette description. La recherche des racines réelles d'un polynôme est également faite de la manière naïve (celle du lycée) : on cherche les racines de sa dérivée et on dresse un tableau de variation. Il s'avère que, tant qu'on ne se préoccupe que de complexité en temps polynomial, cette présentation des réels algébriques et les calculs qu'elle induit sont *aussi bons* que ceux relevant de méthodes nettement plus sophistiquées. Pour résumer: les lois de corps et la recherche des racines d'un polynôme de $\mathbb{Z}[X]$ sont en temps polynomial, mais ces opérations enchaînées conduisent à une explosion de la taille des objets manipulés. A la fin du chapitre, nous donnons des conditions suffisantes pour remplacer $\mathbb{Q}$ par un autre sous-corps de $\mathbb{R}$ et obtenir néanmoins les mêmes majorations de temps de calculs.

Dans le chapitre B , nous étudions le problème de savoir si les défauts constatés dans la présentation naïve peuvent être tournés en utilisant une présentation plus sophistiquée. Chaque fois qu'un calcul conduit à manipuler des objets trop gros (par rapport à la taille des entrées), il est a priori possible de tourner la difficulté en *n'effectuant pas le calcul et en indiquant seulement qu'il devrait être fait* . C'est par exemple le secret de la présentation des entiers en base 10 par rapport aux entiers "batons". Cette méthode universelle souffre cependant de quelques inconvénients. Si on l'applique par exemple pour la représentation des nombres algébriques réels, on obtient certes une représentation toujours compacte des nombres manipulés, mais le test de comparaison est, très probablement, en temps exponentiel ou pire. Nous discutons cette question dans le § a) et démontrons qu'en tout état de cause, il faut a priori accepter de céder du terrain d'un coté ou de l'autre. Récemment, D. Duval et C. Dicrecenzo ont développé et implanté un sytème de représentation appelé D5 , dans lequel les nombres algébriques sont donnés comme solutions d'équations algébriques emboitées. Dans le § b) , nous étudions le comportement de représentations des nombres algébriques dans le cadre D5 et nous vérifions que les calculs qui peuvent être qualifiés d'élémentaires (y compris certains calculs de déterminants, donc l'algèbre linéaire) sont *presque* en temps polynomial. En fait, le temps est polynomial, non par rapport à la taille des entrées, mais par rapport à la

taille qu'occuperaient a priori ces entrées si elles étaient traduites dans une présentation naïve comme celle développée au chapitre A . Le gain peut apparaître assez mince. La souplesse de D5 ou de systèmes analogues, relativement à la présentation naïve (ou une représentation analogue) est néanmoins bien certaine. D'autre part, le fait de raisonner dans D5 pour les calculs de majoration est actuellement la meilleure manière de comprendre clairement ce qui se passe avec les nombres algébriques et où se situent les difficultés. Par exemple, le fait que les calculs de déterminants ne peuvent pas, a priori, être traités par la méthode de Bareiss. Ou encore, les majorations que nous obtenons dans le cadre D5 , par leur caractère uniforme, sont meilleures que celles qui pouvaient résulter de la simple application des résultats obtenus dans $\mathbb{R}_{alg}$.
Dans le § c) nous nous situons dans un cadre directement hérité de D5 , mais en abandonnant ce qui fait une bonne partie de la philosophie de D5 , c.-à-d. que nous levons *a priori* l'ambiguité sur la solution considérée d'un système d'équations algébriques emboitées en le caractérisant par une approximation convenable. Nous obtenons les résultats qui pouvaient être espérés a priori, du même genre que ceux obtenus au § b). Notons enfin que les résultats obtenus s'appliquent, via les mêmes méthodes, dans différents cadres voisins: nombres algébriques réels, nombres algébriques complexes, nombres algébriques p-adiques, clôture algébrique d'un corps de fonctions $\mathbf{F}(X)$ où $\mathbf{F}$ est un corps fini.

L'étude faite en B c) a montré l'efficacité assez bonne des méthodes approximatives pour calculer avec des nombres algébriques réels ou complexes.

Nous examinons dans le chapitre C deux théorèmes "en temps polynomial" qui relèvent par leur nature même de méthodes approximatives. Ces méthodes sont indispensables chaque fois qu'on a à résoudre un problème dont les variables sont dans $\mathbb{R}$, $\mathbb{C}$, ou un espace de fonctions.
Le théorème fondamental de l'algèbre est de ceux-là. Il peut être traité soit par une méthode approximative en tant que telle (algorithme de Schönage ou de Victor Pan), soit en utilisant une théorème effectif de perturbation des racines.
Quand on passe à la recherche des racines réelles d'un polynôme à coefficients réels, une méthode classique comme la méthode de Sturm devient impraticable dans un contexte constructif pour la simple raison qu'il n'y a pas de test d'égalité à 0 pour un nombre réel "en général". L'affirmation classique selon laquelle on peut situer les racines réelles d'un polynôme à coefficients réels devient *fausse* d'un point de vue constructif. Il y a néanmoins un substitut constructif à cette affirmation: la possibilité de dresser un tableau de signes "approché" pour un tel polynôme (cf § C b) pour plus de précision).

Ceci nous amène à faire une brève étude, au § C c) , de la classe des fonctions "approchables en temps polynomial par des polynômes, pour la norme uniforme, sur un intervalle compact". Cette classe est en fait celle des fonctions Gevrey $\mathcal{P}$-calculables. Tous les calculs élémentaires dans cette classe de fonction s'avèrent être en temps polynomial, pour des raisons tout à fait immédiates. Nous obtenons ainsi une amélioration des théorèmes de Ko-Friedman ([**KF1**] et [**KF2**]) et Müller ([**Mü2**]) concernant les fonctions analytiques et $\mathcal{P}$-calculables, et une simplification de leurs preuves.
Nous concluons par quelques perspectives de travail dans le cadre ainsi tracé : la géométrie algébrique réelle exacte dans la clôture réelle de $\mathbb{Q}$ pourrait, selon nous, être avantageusement remplacée par une géométrie algébrique réelle approximative dans tous les problèmes appliqués.

Quelques points de terminologie :

Nous reprenons, surtout dans le A) , la terminologie de **[Lom1]**.

Nous parlons d'une $\mathcal{P}$-fonction ou d'une $\mathcal{P}$-opération pour signifier "opération calculable en temps polynomial par rapport à la taille des entrées".

Nous parlons d'un $\mathcal{P}$-ensemble E ou d'un ensemble $\mathcal{P}$-présenté E pour parler d'un ensemble dénombrable codé (présenté) dans un langage A^* sur un alphabet fini A lorsque les conditions suivantes sont réalisées: 1) les mots qui codent les éléments de E forment une $\mathcal{P}$-partie de A^* (c.-à-d. une partie $\mathcal{P}$-testable de A^*), et 2) le test d'égalité dans E (pour deux mots de A^* qui codent des éléments de E) est un $\mathcal{P}$-test.

Nous notons $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ les $\mathcal{P}$-ensembles correspondants présentés en binaire. Nous notons $\mathbb{N}_1$ pour le $\mathcal{P}$-ensemble des entiers naturels présenté en unaire.

De manière générale **lg**(x) désignera la longueur d'un mot représentant l'objet x (élément de X) dans la présentation choisie de l'ensemble X . Pour des éléments de $\mathbb{Z}$, ce sera donc la taille pour l'écriture en binaire.

Les polynômes de $\mathbb{Q}[X]$, $\mathbb{Q}[X,Y]$, $\mathbb{Q}[X_1,X_2,...,X_n]$ sont supposés donnés en présentation dense. Si $P \in \mathbb{Q}[X_1,X_2,...,X_n]$ et si $d_{X_j} = d_j$, si l_{creux} et l_{dense} représentent la longueur de P dans une présentation creuse et une présentation dense (les coefficients étant toujours écrits en binaire), on a : $l_{creux} \leq l_{dense} \leq d_1...d_n.l_{creux}$. Les résultats de complexité qui font intervenir l_{dense} sont alors facilement traduisibles en résultats qui font intervenir l_{creux} .

A) LE CORPS DES NOMBRES REELS ALGEBRIQUES : PRESENTATION NAIVE

Introduction

Nous étudions dans ce chapitre la présentation des nombres réels algébriques la plus naïve qui soit. Selon ce point de vue, un nombre algébrique est donné par un polynôme P à coefficients entiers qui l'annule et un intervalle sur lequel P change de signe et P' est évidemment de signe constant. Cela suffit à rendre de complexité $\mathcal{P}$ les calculs élémentaires concernant les nombres algébriques. Mais les calculs "en cascade" ont un comportement exponentiel. Nous verrons dans le chapitre B qu'il est difficile d'espérer beaucoup mieux.
On notera qu'on se passe entièrement des algorithmes de décomposition en facteurs premiers dans $\mathbb{Q}[X]$. Autrement dit, on ne sait jamais a priori si le polynôme donné qui annule un réel algébrique ξ est le polynôme minimum de ξ ou non.

a) Présentation de $\mathbb{R}_{alg}$

Evidence du signe constant d'un polynôme sur un intervalle donné

Définition A.a1 :

- Soient $P \in \mathbb{Q}[X]$, a et $b \in \mathbb{Q}$ avec $a.b \geq 0$, $a < b$. On écrit $P = P_1 + P_2$, où P_1 est la somme des monômes strictement croissants sur l'intervalle $[a , b]$ et P_2 est la somme des monômes décroissants.
- on dira que le nombre $P_1(a) + P_2(b)$ est le *minorant-évident de P sur l'intervalle* $[a , b]$ et que $P_1(b) + P_2(a)$ est le *majorant-évident de P sur l'intervalle* $[a , b]$
- si maintenant a et $b \in \mathbb{Q}$ avec $a < 0 < b$, on appelera minorant-évident (resp. majorant-évident) de P sur $[a , b]$ le plus petit (resp. le plus grand) des minorants-évidents (resp. majorants-évidents) de P sur $[a , 0]$ et sur $[0 , b]$
- pour $a < b$ dans $\mathbb{Q}$, on dira que P est *évidemment-de-signe-constant sur l'intervalle* $[a , b]$ lorsque le majorant-évident et le minorant-évident de P sur $[a , b]$ ont même signe, non nul.

Il est clair qu'un majorant-évident est un majorant, et que si un polynôme P est évidemment-de-signe-constant sur un intervalle $[a , b]$, alors il est de signe constant sur cet intervalle.

De plus le majorant-évident d'un polynôme sur un intervalle plus petit est inférieur au majorant-évident sur l'intervalle initial. De même l'évidence du signe constant sur un intervalle implique l'évidence du signe constant sur tout intervalle plus petit.

Lemme 1 : Soient P et Q dans $\mathbb{Q}[X]$, avec pgcd(P,Q) = 1 , et [r', r] un intervalle rationnel.
On peut $\mathcal{P}$-calculer (à partir des polynômes P et Q et des rationnels r' , $r \in \mathbb{Q}$) un entier $n \in \mathbb{N}_1$ tel que :
si $r \le a \le b \le r'$ et $| b - a | \le 1/2^n$, alors P ou Q est évidemment-de-signe-constant sur [a , b]

preuve> Nous allons traiter le cas où $r' < 0 < r$, qui est le plus compliqué (si $r.r' \ge 0$ l'adaptation est immédiate). Nous notons p et q les degrés de P et Q .
Supposons tout d'abord $0 \le a < b$. Ecrivons $P = P_1 + P_2$ et $Q = Q_1 + Q_2$ en vue de tester "l'évidence du signe constant". Soit M un majorant de $|P_1'|, |P_2'|, |Q_1'|, |Q_2'|$ sur [0 , r] .
On a alors :

$|P(a) - (P_1(a) + P_2(b))| = |P_2(a) - P_2(b)| \le M.(b - a)$, et
$|P(a) - (P_1(b) + P_2(a))| = |P_1(a) - P_1(b)| \le M.(b - a)$.

De sorte que :

$|P(a)| > M.(b - a) \Rightarrow$ P est évidemment-de-signe-constant sur [a , b] .

De même :

$|Q(a)| > M.(b - a) \Rightarrow$ Q est évidemment-de-signe-constant sur [a , b] .

Par ailleurs on sait $\mathcal{P}$-calculer U(X) et V(X) tels que :

$P(X).U(X) + Q(X).V(X) = \mathbf{Res}(P,Q)$ (le résultant de P et Q) .

Les coefficients de U et V sont des cofacteurs de la matrice de Sylvester de P et Q , d'après l'inégalité de Hadamard sur les déterminants on a donc la majoration :

$$| \text{coeff de U ou V} | \le || P ||_2^{q-1} \, || Q ||_2^{p-1} \, \mathbf{sup}(|| P ||_2, || Q ||_2) \quad (1)$$

Soit N un majorant de |U(x)| et |V(x)| sur [r', r] . On a alors les implications :
$|P(a)| < |\mathbf{Res}(P,Q)|/2N \Rightarrow |Q(a)|.|V(a)| > |\mathbf{Res}(P,Q)|/2 \Rightarrow |Q(a)| > |\mathbf{Res}(P,Q)|/2N$.
Donc : $|P(a)| \ge |\mathbf{Res}(P,Q)|/2N$ **ou** $|Q(a)| \ge |\mathbf{Res}(P,Q)|/2N$.
Donc : **si** $b - a < |\mathbf{Res}(P,Q)|/2NM$, **alors** P ou Q est évidemment-de-signe-constant sur [a , b] .

Dans le cas où $a < b \le 0$, on a une conclusion analogue avec une valeur M' à la place de M . On pose donc M" = **sup**(M,M') pour obtenir la même minoration dans les 2 cas.

Enfin, dans le cas où $a < 0 < b$, on a pareillement :

$|P(0)| \ge |\mathbf{Res}(P,Q)|/2N$ **ou** $|Q(0)| \ge |\mathbf{Res}(P,Q)|/2N$

$|P(0)| > M".(b - a) \Rightarrow$ P est évidemment-de-signe-constant sur [a , 0] et sur [0 , b] .

$|Q(0)| > M".(b - a) \Rightarrow$ Q est évidemment-de-signe-constant sur [a , 0] et sur [0 , b] .

Conclusion : dans tous les cas,
si $b - a < |\mathbf{Res}(P,Q)|/2NM"$, **alors** P ou Q est évidemment-de-signe-constant sur [a , b] .

---

[1] $|| P ||_2 := (\sum \mathbf{cf}_i(P)^2)^{1/2}$

La minoration a été $\mathcal{P}$-calculée dans $\mathbb{Q}$ à partir des polynômes P et Q et des rationnels r' , r . Il ne reste qu'à en déduire $n \in \mathbb{N}_1$ tel que :

$$1/2^n < |\mathbf{Res}(P,Q)| \,/2NM'' \,. \qquad \square$$

Appliquons ce lemme avec $Q = P'$: lorsque P ou P' est de signe constant sur [a , b] , P admet au plus une racine sur l'intervalle et on sait déterminer s'il en admet une ou non.

Notons que, pour $P = \sum_{0 \le i \le p} a_i.X^i$, l'intervalle $[\,-r\,,\,r\,]$ contient toutes les racines réelles de P dès que r est égal à

$$1 + \sup_{0 \le i < p} (|a_i|/|a_p|) \text{ ou encore à } \sup_{0 \le i < p} \left(\sqrt[p-i]{p|a_i|/|a_p|} \right).$$

Le lemme 1 justifie donc une méthode de calcul **PSPACE** pour situer les racines réelles d'un polynôme sans facteur carré de $\mathbb{Q}[X]$: découper l'intervalle [- r , r] en intervalles de longueur assez petite pour que le test du signe évidemment constant marche sur tous ces intervalles, soit pour P , soit pour P' .
De plus, dans le lemme 1 , la dépendance de n par rapport aux bornes de l'intervalle rationnel pourrait être supprimée (parce que P(x) est évidemment-de-signe-constant pour | x | assez grand), avec pour prix à payer une valeur de n trop grande si l'intervalle rationnel est petit .

Par ailleurs, ce lemme justifie la présentation suivante de l'ensemble $\mathbb{R}_{alg}$ des réels algébriques :

Présentation naïve de $\mathbb{R}_{alg}$

Définition A.a2 : Nous désignerons par $\mathbb{R}_{alg}$ l'ensemble des réels algébriques présenté de la manière décrite ci-dessous. (pour le moment on ne sait pas si c'est une $\mathcal{P}$-présentation).

Un nombre réel algébrique u est présenté par un triplet

(P , a , b) $\in \mathbb{Z}[X] \times \mathbb{Q} \times \mathbb{Q}$ vérifiant les conditions suivantes :

- $a < b$
- P est un polynôme sans facteur carré (i.e.: vérifiant $\mathbf{Res}(P,P') \neq 0$)
- $P(a).P(b) < 0$, et P' est évidemment-de-signe-constant sur [a , b]
- $P(u) = 0$ et $u \in [\,a\,,\,b\,]$

On remarquera qu'un réel algébrique rationnel c/d peut être représenté par (d.X − c , a , b) , avec $a < c/d < b$, et que l'injection canonique $\mathbb{Q} \to \mathbb{R}_{alg}$ est une $\mathcal{P}$-opération. On verra plus précisément que $\mathbb{Q}$ s'identifie à une $\mathcal{P}$-partie de $\mathbb{R}_{alg}$. (corollaire de la prop A.a6)

On remarquera également qu'on n'exige pas que le polynôme P soit irréductible.

On sait que les réels algébriques forment un ensemble discret, mais on ne peut affirmer d'emblée que la séparation de 2 réels algébriques peut être testée en temps polynomial. On a néanmoins tout de suite :

Proposition A.a3 :

$\mathbb{R}_{alg}$ ainsi présenté est une $\mathcal{P}$-partie du $\mathcal{P}$-ensemble $\mathbb{Z}[X] \times \mathbb{Q} \times \mathbb{Q}$

preuve> La 4ème condition doit être prise pour une définition de u lorsque les 3 autres conditions sont vérifiées. Ces conditions sont vérifiées au moyen d'un calcul de résultant, et d'évaluations du polynôme P et de polynômes "extraits" de P' en vue de tester l'évidence du signe constant. Tous ces calculs sont en temps polynomial. ❑

Opérations élémentaires avec des rationnels

Proposition A.a4 : On peut construire :

- une $\mathcal{P}$-opération Pr : $\mathbb{R}_{alg} \times \mathbb{Q} \to \mathbb{R}_{alg}$ vérifiant Pr(u,r) = u.r
- une $\mathcal{P}$-opération Sm : $\mathbb{R}_{alg} \times \mathbb{Q} \to \mathbb{R}_{alg}$ vérifiant Sm(u,r) = u + r
- une $\mathcal{P}$-opération Comp : $\mathbb{R}_{alg} \times \mathbb{Q} \to (<, =, >)$ qui donne le résultat de la comparaison du réel algébrique u au rationnel r
- une $\mathcal{P}$-opération Min : $\mathbb{R}_{alg} \to \mathbb{Q}$ qui donne, pour un réel algébrique non nul, une minoration rationnelle de sa valeur absolue

preuve> Soit u le réel algébrique représenté par (P , a , b) , avec d = deg(P) , $P = \sum_i c_i.X^i$ et r le rationnel.
Pour la somme et le produit : on fait le changement de variable Y = X + r ou Y = X.r : cela n'affecte pas le changement de signe pour P ni l'évidence-du-signe constant pour P' .
Pour la comparaison : si r est extérieur à l'intervalle] a , b [, la comparaison est immédiate, sinon on calcule le signe de P(r) : si P(r) = 0 alors r = u , sinon on compare avec le signe de P(a) pour situer r , soit sur l'intervalle] a , u [soit sur l'intervalle] u , b [.
Pour la minoration de la valeur absolue (u étant supposé non nul) : on majore 1/u en considérant le polynôme aux inverses, ce qui donne : $|u| > |c_0| / (|c_0| + \sup_{i>0}(|c_i|))$. ❑

Calcul des valeurs approchées d'un réel algébrique à partir de sa présentation

Définition A.a5 : Soit A un $\mathcal{P}$-ensemble et f une fonction de A vers $\mathbb{R}$. On dira que f est une $\mathcal{P}$-fonction, ou une $\mathcal{P}$-suite, ou encore que f est une fonction $\mathcal{P}$-calculable, si il existe une $\mathcal{P}$-opération F : $A \times \mathbb{N}_1 \to \mathbb{Q}$ telle que F(z,n) est une approximation de f(z) avec la précision 2^{-n} [1]

En particulier, lorsque A est réduit à un point, on obtient la notion de $\mathcal{P}$-nombre réel (ou réel de complexité $\mathcal{P}$). Des définitions analogues vaudraient d'ailleurs pour toute autre classe de complexité.

Proposition A.a6 :

Il existe une $\mathcal{P}$-opération F : $\mathbb{R}_{alg} \times \mathbb{N}_1 \to \mathbb{D}$ telle que $F(v,n) = r/2^n$ (avec $r \in \mathbb{Z}$) est une approximation de v avec la précision 2^{-n} .
C'est-à-dire: l'injection naturelle $\mathbb{R}_{alg} \to \mathbb{R}$ est une $\mathcal{P}$-fonction

On notera que cette opération F n'est pas une fonction puisque F(v,n) va dépendre de la présentation de v . Par ailleurs, il est clair qu'on aurait le même résultat en remplaçant 2^{-n}

---

[1] On peut exiger de plus que F(z,n) soit de la forme $r/2^n$ ($r \in \mathbb{Z}$) de sorte que $f(z) \in [(r-1)/2^n , (r+1)/2^n]$

(représenté par l'entrée n) par un élément $\varepsilon > 0$ arbitraire de $\mathbb{D}$ ou $\mathbb{Q}$. Cette proposition A.a6 est un cas particulier du lemme suivant :

Lemme 2 :

Soit DICHOT le $\mathcal{P}$-ensemble formé des triplets $(P , a , b) \in \mathbb{Q}[X] \times \mathbb{Q} \times \mathbb{Q}$ vérifiant $P(a).P(b) < 0$.

Il existe une $\mathcal{P}$-opération $G : DICHOT \times \mathbb{N}_1 \to \mathbb{D}$ telle que $G(P , a , b , n) = r/2^n$ (avec $r \in \mathbb{Z}$) est une approximation avec la précision 2^{-n} d'une racine u de P située sur l'intervalle [a , b] (la racine u ne dépendant pas de n).

preuve> L'opération G est obtenue par une dichotomie classique, avec pour points de départ a et b . Posons $k = \mathbf{sup}(0 , n + [\mathbf{log}_2(b - a)]^1)$. Après avoir divisé k fois l'intervalle [a , b] en 2 , on obtient un intervalle de longueur $< 2^{-n}$. Les bornes successives des intervalles sont de la forme $x_i = (m_i.a + (2^i - m_i).b)/2^i$; avec $2^i \geq m_i \geq 0$, $i \leq k$. Il y a k évaluations $P(x_i)$ nécessaires. Si l'intervalle obtenu est [a ' , b '] , on prend $r = \mathbf{Ent}(b' .2^n)$.
Le tout est un $\mathcal{P}$-calcul à partir de l'entrée (P , a , b , n) . ❑

Un corollaire de la proposition A.a6 est le suivant :

Corollaire : $\mathbb{Q}$ s'identifie à une $\mathcal{P}$-partie de $\mathbb{R}_{alg}$

preuve> Soit u = (P , a , b) et soit c la valeur absolue du coefficient dominant de P et d son degré. Il s'agit de tester en temps polynomial si u est rationnel : il suffit de calculer l'entier algébrique $u.c = (c^{d-1}.P(X/c) , a.c , b.c)$ avec une précision meilleure que 1/2 : si l'intervalle obtenu contient un entier k , on teste si $P(k/c) = 0$. ❑

Une autre conséquence immédiate est la :

Proposition A.a7 : (réduction de la taille d'un réel algébrique)

Il existe un polynôme Q et une $\mathcal{P}$-opération $Rd : \mathbb{R}_{alg} \to \mathbb{R}_{alg}$ telle que : si $u = (P , a , b)$, alors $Rd(u) = v = (P , a' , b')$ et :

- $u = v$ au sens de $\mathbb{R}_{alg}$
- $\mathbf{lg}(a') + \mathbf{lg}(b') < Q(\mathbf{lg}(P))^2$

preuve> Cela résulte clairement des lemmes 1 et 2 et de la $\mathcal{P}$-majoration des valeurs absolues des racines réelles de P . ❑

Recherche d'une racine par dichotomie sur un intervalle rationnel

Proposition A.a8 : Le lemme 2 peut être précisé de la manière suivante :

Il existe un polynôme Q et une $\mathcal{P}$-opération $Rac : \mathbb{Q}[X] \times \mathbb{Q} \times \mathbb{Q} \to \mathbb{R}_{alg}$ telle que :

Si $P(a).P(b) < 0$, alors $Rac(P , a , b) = u$ avec : u est une racine de P sur] a , b [, et $\mathbf{lg}(u) < Q(\mathbf{lg}(P))$.

---

[1] [x] est un entier majorant le réel x

[2] De manière générale lg(x) désignera la longueur d'un mot représentant l'objet x (élément de X) dans la présentation choisie de l'ensemble X . Pour des éléments de $\mathbb{Z}$, ce sera la taille pour l'écriture en binaire.

preuve> On calcule $R = P/pgcd(P,P')$, qui admet les mêmes racines que P , puis une approximation suffisante d'une racine de P sur $[a , b]$ pour que R' soit de signe-évidemment-constant sur l'intervalle obtenu. ❑

Test d'égalité. Séparation dans le cas de non égalité

Théorème A.a9 : Il existe une $\mathfrak{P}$-opération $V : \mathbb{R}_{alg} \times \mathbb{R}_{alg} \to \mathbb{D} \times \mathbb{D} \times (< , = , >)$ telle que :

- $V(u,v) = (c,d,"=") \Rightarrow u = v$
- $V(u,v) = (c,d,"<") \Rightarrow u < c < d < v$ et $(d - c) > (v - u)/2$
- $V(u,v) = (c,d,">") \Rightarrow u > d > c > v$ et $(d - c) > (u - v)/2$

Proposition A.a10 : Avec l'opération V définie ci-dessus pour tester l'égalité de 2 réels algébriques, la présentation $\mathbb{R}_{alg}$ est une $\mathfrak{P}$-présentation de l'ensemble des réels algébriques.

preuve> La proposition A.a10 découle immédiatement de la précédente. Voyons celle-ci.
Soit $u = (P_1 , a_1 , b_1) , v = (P_2 , a_2 , b_2)$. On calcule $R = pgcd(P_1 , P_2)$, $R_1 = P_1/R$, $R_2 = P_2/R$.
Un seul des 2 polynômes R et R_1 change de signe sur l'intervalle $[a_1 , b_1]$; appelons le S pour un instant. En appliquant les lemmes 1 et 2 , on peut $\mathfrak{P}$-calculer un intervalle $[a'_1 , b'_1]$ contenu dans $[a_1 , b_1]$ et tel que , d'une part S change de signe , d'autre part S' soit évidemment-de-signe-constant sur l'intervalle. On peut alors affirmer $u = (S , a'_1 , b'_1)$.
On procède de même pour v .
Deux cas se présentent alors.
– *Le 1er cas* est celui où u et v sont tous deux racines de R . Si les intervalles $[a'_1 , b'_1]$ et $[a'_2 , b'_2]$ se coupent, on a $u = v$; sinon, il suffit de calculer u et v avec une précision meilleure que $|u - v|/4$ pour obtenir c et d : or ceci ne prendra pas trop de temps puisqu'on a :

$$\mathbf{log}(|\mathbf{Res}(R,R')|) = k'.\mathbf{log}(r) + \sum_{i \neq j} \mathbf{log}(|x_i - x_j|) ,$$

où les x_i sont les racines de R , r son coefficient dominant positif, $k' = 2.deg(R) - 1$. Soit M un majorant des $| x_i |$. On obtient donc :

$$2.\mathbf{log}(|u - v|) > \mathbf{log}(|\mathbf{Res}(R,R')|) - \sum_{i \neq j, \{x_i,x_j\} \neq \{u,v\}} \mathbf{log}(|x_i - x_j|) - k'.\mathbf{log}(r) >$$

$$\mathbf{log}(|\mathbf{Res}(R,R')|) - k.\mathbf{log}(2.M) - k'.\mathbf{log}(r)$$

où $k = deg(R).(deg(R) - 1) - 2$
On peut donc terminer en appliquant le lemme 2 .
– *Le 2ème cas* est celui où u et v sont racines de 2 polynômes qui n'ont pas de racine commune. Appelons les S et T . Il s'agit de nouveau de calculer u et v avec une précision meilleure que $|u - v|/4$ pour obtenir c et d : ceci ne prendra pas trop de temps puisqu'on a de même :

$\mathbf{log}(|\mathbf{Res}(S,T)|) = k_1.\mathbf{log}(r_2) + k_2.\mathbf{log}(r_1) + \mathbf{log}(|x_i - y_j|)$; x_i racines de S , y_j racines de T , $k_1 = deg(S)$, $k_2 = deg(T)$, r_1 coeff dominant positif de S , r_2 coeff dominant positif de T; et donc :

$\log(|u - v|) > \log(|\mathrm{Res}(S,T)|) - k_1.\log(r_2) - k_2.\log(r_1) - (k_1.k_2 - 1).\log(M + N)$; M majore les $|x_i|$, N majore les $|y_j|$. ❑

La preuve du théorème A.a9 fait intervenir des calculs de résultants et PGCD .
Mais en pratique, si on sait, par un argument quelconque, que $u \neq v$, il suffit, pour séparer u et v , de les calculer chacun avec une précision de plus en plus grande, jusqu'à ce qu'ils soient séparés. Ce calcul est une dichotomie ne faisant intervenir que des évaluations de polynômes, et le nombre d'étapes est raisonnable.

Dans le cas où on ne sait pas si $u = v$ ou non , on commence par calculer le résultant des 2 polynômes. Si le résultant est non nul on est ramené au cas précédent. S'il est nul, le pgcd R est donné au cours du calcul et il faut calculer R_1 , R_2 : ce surcroît de calculs est néanmoins compensé par le fait que désormais, u et v seront plus simples à manipuler puisque racines de polynômes de degrés moindres.

Nous pouvons cependant remarquer que la preuve du théorème A.a9 nous fournit explicitement, un écart en deça duquel deux réels algébriques sont nécessairement confondus. C'est ce que nous précisons dans la proposition suivante. Il en découle que les calculs de PGCD ne sont jamais indispensables pour la comparaison des réels algébriques.

Théorème A.a11 :

a) Soient P et Q 2 polynômes à coefficients réels, premiers entre eux, u une racine de P et v une racine de Q . Posons $p := d(P)$, $q := d(Q)$, M := un majorant des modules des racines de P , N := un majorant des modules des racines de Q ,
$n := p \log_2(|cd(Q)|) + q \log_2(|cd(P)|) + (p.q-1).\log_2(M+N) - \log_2(|Res(P,Q)|)$
Si $|u - v| < 1/2^n$, **alors** $u = v$

b) Soient $u = (P, a, b)$, $v = (Q, c, d)$ 2 éléments de $\mathbb{R}_{alg}$.
Mêmes notations qu'en a) pour p , q , N et M ;
$n := p.\mathbf{lg}(cd(Q)) + q.\mathbf{lg}(cd(P)) + (p.q-1).\mathbf{lg}(M+N)$
Si $|u - v| < 1/2^n$, **alors** $u = v$
Si P et Q sont unitaires on peut prendre $n = (p.q-1).\mathbf{lg}(M+N)$

preuve> On utilise les majorations établies dans la preuve du théorème A.a9 . Pour le b) on remarque que le résultant de 2 polynômes à coefficients entiers est un entier, donc de logarithme ≥ 0 . Les majorations établies dans la preuve du théorème A.a9 couvrent alors également le cas où P et Q ne sont pas premiers entre eux (et en particulier le cas $P = Q$).❑

Remarque : La majoration ci-dessus, obtenue par un calcul grossier, peut sans doute être améliorée. Selon cette majoration, pour connaître le polynôme minimum P d'un nombre algébrique α , sachant que $d(P) \leq p$ et **sup**(| coeffs de P |) $\leq H$, $\mathbf{lg}(H) = h$, il suffit de connaître α avec une précision de $1/2^n$ où $n = 2\,p\,h + p^2\,(1+h) + 1$.
Dans [**KLL**] les auteurs, en utilisant l'algorithme LLL (cf [**LLL**] ou [**Val**]), retrouvent les coefficients de P en temps polynomial dès que sont connus s bits du développement binaire de α , où $s \geq p^2/2 + ((3p+4)\,\mathbf{lg}(p+1))/2 + 2\,p\,h$.

Vue le théorème A.a11, on note l'importance particulière dans la pratique d'une méthode de calcul particulièrement rapide des approximations de nombres réels algébriques. C'est par exemple le cas de la méthode de Newton (cf [**Mü1**]).

Proposition A.a12 :

Soient $u = (P, a, b) \in \mathbb{R}_{alg}$, $x_0 \in]a, b[$, $r_0 = \inf(|x_0 - a|, |x_0 - b|)$, M un majorant de $|P^{(2)}(x)|$ sur $[a, b]$

a) Si $|P(x_0)| \leq \inf(|P'(x_0)|/2M, r_0/2)$, la méthode de Newton peut être appliquée pour le calcul d'approximations de u en démarrant avec x_0

b) Si cette condition est réalisée et si on utilise les techniques de multiplication rapide, le calcul d'une approximation avec la précision 2^{-n} est alors en temps O(n log(n) loglog(n)) (l'unique entrée est n en unaire)

c) Un point x_0 de $\mathbb{D} \cap]a, b[$ vérifiant a) peut être calculé en temps polynomial (pour l'entrée u).

preuve> *Le a)* résulte des majorations classiques pour la convergence de la méthode de Newton. Cf par exemple, **[DM]** p 164 .

Le b) résulte essentiellement du fait que l'itération dans la méthode de Newton est quadratique. D'autre part, à chaque itération, le calcul n'est fait qu'approximativement: on arrête dès que le nombre significatif de décimales est obtenu. Détails dans **[Mü1]** .

Le c) : on peut calculer en temps polynomial :
– un majorant $m \geq 1$ de $|P'(x)|$ sur $[a, b]$ – un majorant M de $|P^{(2)}(x)|$ sur $[a, b]$ – un minorant > 0 s de $|P'(u)|$ – un minorant r de $|u - a|/2$ et $|u - b|/2$.

Si x_0 vérifie $|x_0 - u| < \inf(s/4mM, r/3m)$, alors on a $r_0 > 2r/3$, $P(x_0) < r_0/2$, $|P'(x_0)| > s/2$ et $|P(x_0)| < s/4M < |P'(x_0)|/2M$. ❑

Remarque : Ceci montre que tout réel algébrique est "individuellement" un réel de complexité en temps O(n log(n) loglog(n)), mais c'est une appréciation très individualiste dans la mesure où P, a, b ne sont pas considérées comme des entrées. Si on fait précéder la méthode de Newton d'une méthode par dichotomie cela relativise le résultat obtenu. Notons cependant que si $P^{(2)}$ est de signe constant sur $[a, b]$ la méthode de Newton fonctionne sans phase préparatoire, en démarrant de l'extrémité de l'intervalle où P et $P^{(2)}$ sont de même signe. Une solution "définitive" (au problème de calculer très rapidement les approximations rationnelles des nombres réels algébriques qu'on manipule) consisterait à donner toujours un réel algébrique sous la forme (P, a, b, x_0) sur un intervalle tel que la condition a) de la proposition A.a12 soit vérifiée. C'est grosso modo la solution que nous développons dans le B , dans le cadre des systèmes d'équations emboitées.
Il serait intéressant d'étudier la complexité du calcul si on utilise le processus itératif dit "méthode regula falsi" (ou méthode de la sécante), qui a également une bonne vitesse de convergence. Cf par exemple **[Ost]** p 55 pour une comparaison des mérites respectifs des méthodes de Newton et "regula falsi".

b) $\mathbb{R}_{alg}$ comme $\mathcal{P}$-structure

Calcul d'un réel algébrique à partir de ses valeurs approchées

Nous établissons maintenant un théorème général utile pour montrer qu'une fonction à valeur dans $\mathbb{R}_{alg}$ est $\mathcal{P}$-calculable.

Théorème A.b1 : Soit A un $\mathcal{P}$-ensemble et f une fonction de A vers $\mathbb{R}_{alg}$.
Pour que f soit $\mathcal{P}$-calculable, il faut et suffit que les 2 conditions suivantes soient vérifiées :

- la fonction $f : A \to \mathbb{R}$ est une $\mathcal{P}$-fonction
- il existe une $\mathcal{P}$-opération $G : A \to \mathbb{Z}[X] - \{0\}$ telle que :
  si $G(z) = S$, alors $f(z)$ est racine de S.

preuve> Les conditions sont clairement necessaires.
Montrons qu'elles sont suffisantes : soit $z \in A$ et voyons comment calculer $f(z)$.
Tout d'abord on remplace $S = G(z)$ par $T = S/pgcd(S,S')$.
Par le lemme 1 , on détermine un entier n tel que, sur tout intervalle de longueur $\leq 2^{-n}$, T ou T' est évidemment-de-signe-constant sur l'intervalle.
Ensuite on calcule une approximation dyadique de f(z) avec la précision 2^{-n-1} , ce qui permet de situer f(z) sur un intervalle rationnel [a , b] de longueur $\leq 2^{-n}$, sur lequel T s'annule et T' est évidemment-de-signe-constant.
Le réel algébrique f(z) est donc correctement représenté par (T , a , b).
Le tout est un $\mathcal{P}$-calcul à partir de l'entrée z ❑

Remarque : on peut éviter de calculer n (par utilisation du lemme 1) : on calculerait des intervalles successifs (pour i = 0,1,2,...) de longueur $< 2^{-i}$, sur lesquels se trouve f(z) , en testant à chaque fois l'évidence du signe constant pour le polynôme T' . On est sûr d'aboutir en un temps raisonnable.

Evaluation $\mathbb{Q}[X] \times \mathbb{R}_{alg} \to \mathbb{R}_{alg}$

Proposition A.b2 :
Il existe une $\mathcal{P}$-opération $Ev : \mathbb{Q}[X] \times \mathbb{R}_{alg} \to \mathbb{R}_{alg}$ telle que : $Ev(Q,u) = Q(u)$
(égalité au sens de $\mathbb{R}_{alg}$)

preuve> Soit $u = (P , a , b)$ et M_P la matrice standard admettant P comme polynôme minimum et caractéristique. Le réel algébrique Q(u) est racine du polynôme caractéristique S de la matrice $Q(M_P)$. Le calcul de S à partir de P et Q est un $\mathcal{P}$-calcul.
Ensuite, on majore $|Q'|$ sur l'intervalle [a , b] par un entier m , ce qui permet, via le théorème des accroissements finis et le lemme 2 , de calculer un rationnel approchant Q(u) avec une précision meilleure que 2^{-n} , comme $\mathcal{P}$-calcul à partir des entrées Q , a , b , n (n dans $\mathbb{N}_1$).
On conclut par le théorème A.b1 . ❑

Evaluation $\mathbb{Q}[X,Y] \times \mathbb{R}_{alg} \times \mathbb{R}_{alg} \to \mathbb{R}_{alg}$

Théorème A.b3 :

Il existe une $\mathcal{P}$-opération Ev2 : $\mathbb{Q}[X,Y] \times \mathbb{R}_{alg} \times \mathbb{R}_{alg} \to \mathbb{R}_{alg}$ telle que :

$Ev2(R,u,v) = R(u,v)$ (égalité au sens de $\mathbb{R}_{alg}$)

Théorème A.b4 : : $\mathbb{R}_{alg}$ est un $\mathcal{P}$-corps-ordonné

preuve> Démontrons d'abord le théorème A.b3:
Supposons $u = (P_1 , a_1 , b_1)$, $k_1 = \deg(P_1)$, $v = (P_2 , a_2 , b_2)$, $k_2 = \deg(P_2)$. On va calculer R(u,v) en utilisant le théorème A.b1 .
Tout d'abord, à partir d'une majoration de $|\partial R/\partial X|$ et $|\partial R/\partial Y|$ sur le rectangle $[a_1 , b_1] \times [a_2 , b_2]$ on obtient un module de Lipschitz pour R , ce qui permet de calculer un rationnel approchant R(u,v) avec une précision meilleure que 2^{-n} , comme $\mathcal{P}$-calcul à partir des entrées R , u , v , n (n dans $\mathbb{N}_1$) .
Il reste à $\mathcal{P}$-calculer un polynôme S de $\mathbb{Z}[X]$ annulant R(u,v) .
Pour cela nous considérons l'idéal $\mathcal{J}$ de $\mathbb{Q}[X,Y]$ engendré par $P_1(X)$ et $P_2(Y)$. Le corps $\mathbb{Q}(u,v)$ est un quotient de l'algèbre $\mathcal{A} = \mathbb{Q}[X,Y]/\mathcal{J}$ par l'homomorphisme qui envoie X et Y en u et v . Il nous suffit donc de déterminer un polynôme S de $\mathbb{Z}[X]$ tel que S(R) = 0 dans $\mathcal{A}$.
Or $\mathcal{A}$ possède comme base "canonique" les monômes $X^i.Y^j$ où $i < k_1$ et $j < k_2$.
Nous savons que nous pouvons $\mathcal{P}$-calculer dans $\mathbb{Q}[X,Y]$ les polynômes $1 , R , R^2 , R^3 ,\ldots ,$ R^h ($h = k_1.k_2 - 1$) à partir des entrées R , k_1 , k_2 , puisque $\mathbb{Q}[X,Y]$ est un anneau c-$\mathcal{P}$-c .
Comme par ailleurs les relations de dépendance linéaire sont $\mathcal{P}$-calculables dans $\mathbb{Q}$, nous aurons terminé la preuve du théorème après avoir démontré le lemme suivant :

Lemme : L'application de $\mathbb{Q}[X,Y] \times \mathbb{Z}[X] \times \mathbb{Z}[Y]$ vers $\mathbb{Q}[X,Y]$ qui, au triplet (T , P_1 , P_2) associe le polynôme " T(X,Y) réduit modulo $P_1(X)$ et $P_2(Y)$ " est une $\mathcal{P}$-fonction.

(cette application consiste à exprimer l'élément T sur la base "canonique" dans l'algèbre $\mathcal{A} = \mathbb{Q}[X,Y]/\mathcal{J}$ où $\mathcal{J}$ est l'idéal ($P_1(X)$, $P_2(Y)$)) .

preuve du lemme : Comme l'addition est c-$\mathcal{P}$-c, il suffit de le montrer lorsque le polynôme Q est un monôme $X^n.Y^m$. On réduit séparément X^n modulo $P_1(X)$ et Y^m modulo $P_2(Y)$ puis on fait le produit, et on sait que ce sont des $\mathcal{P}$-opérations.
prouvons maintenant le théorème A.b4:
On sait déjà que la relation d'ordre est $\mathcal{P}$-décidable. Pour l'addition et la multiplication, on applique le Théorème précédent avec X + Y et X.Y . Il reste à voir le calcul de l'inverse d'un élément non nul. Ce qui se démontre sans problème avec le théorème A.b1 . ❑

Remarque : Ces théorèmes ne signifient pas vraiment qu'on peut calculer dans $\mathbb{R}_{alg}$, en effet l'addition et le produit ne sont pas complètement $\mathcal{P}$-calculables : par exemple la somme de n nombres quadratiques est en général de degré 2^n , et donc les calculs en série explosent du fait de la taille des objets utilisés. Nous discutons ce problème plus en détail dans B a) . Notons enfin qu'il existe des extensions infinies de $\mathbb{Q}$ contenues dans $\mathbb{R}_{alg}$ et où cependant les additions et produits en chaîne n'explosent pas (restent polynomialement majorés en taille, et donc en temps de calcul), par exemple :

$$K := \bigcup_n \mathbb{Q}[\sqrt[r^n]{2}] \quad \text{où r est un entier fixé}$$

On peut en effet se convaincre qu'il s'agit ici de la même présentation (à $\mathcal{P}$-équivalence près) que celle donné dans l'exemple de **[Lom1]** p 32 .

Evaluation $\mathbb{Q}[X_1,X_2,...,X_n] \times \mathbb{R}_{alg}^{\ n} \to \mathbb{R}_{alg}$

Proposition A.b5 :

L'application $E : \mathbb{Q}[X_1,X_2,...,X_n] \times \mathbb{R}_{alg}^{\ n} \to \mathbb{R}$ définie par :

$E(R , (\xi_1,\xi_2,...,\xi_n)) = R(\xi_1,\xi_2,...,\xi_n)$ est une $\mathcal{P}$-fonction

preuve> Supposons $\xi_i = (P_i , a_i , b_i)$. Comme l'évaluation $\mathbb{R}_{alg} \to \mathbb{R}$ est une $\mathcal{P}$-fonction, il suffit de savoir $\mathcal{P}$-calculer un module de Lipschitz pour R sur le pavé $B := \prod_i [a_i , b_i]$, donc de $\mathcal{P}$-majorer les $|\partial R/\partial X_i|$ à partir de l'entrée (R,B) ❑

Remarque : on peut définir comme en dimension 1 un majorant-évident et un minorant-évident d'un polynôme sur un pavé. Mais si le pavé contient l'origine en son intérieur, il faudra le décomposer en 2^n sous-pavés. Il peut donc sembler préférable de choisir un majorant plus grossier (obtenu à partir des valeurs absolues des coefficients et des **sup**($|a_i|$, $|b_i|$)) .

Proposition A.b6 :

Il existe une $\mathcal{P}$-opération $Ev_n : \mathbb{Q}[X_1,X_2,...,X_n] \times \mathbb{R}_{alg}^{\ n} \to \mathbb{R}_{alg}$ telle que :

$Ev_n(R , (\xi_1,\xi_2,...,\xi_n)) = R(\xi_1,\xi_2,...,\xi_n)$ (égalité au sens de $\mathbb{R}_{alg}$)

preuve> on raisonne exactement comme en 2 variables (Th A.b3) ❑

NB: a priori, le degré d'un élément de $\mathbb{Q}[\xi_1,\xi_2,...,\xi_n]$ est inférieur ou égal à $d_1.d_2. ... .d_n$.

Remarque : Lorsque les polynômes P_i (où $\xi_i = (P_i, a_i, b_i)$) et R sont des polynômes unitaires, le calcul du signe de $R(\xi_1,\xi_2,...,\xi_n)$ peut être obtenu de manière plus rapide que par le calcul de $R(\xi_1,\xi_2,...,\xi_n)$ dans $\mathbb{R}_{alg}$. En effet, soit m_i un majorant des modules des racines de P_i (i = 1,...,n). Il est alors facile de calculer un majorant m_R pour les $| R(\zeta_1,\zeta_2,...,\zeta_n) |$ où les ζ_i sont des racines arbitraires des P_i . Le produit des $R(\zeta_1,\zeta_2,...,\zeta_n)$ non nuls est un entier algébrique, donc un entier, ce qui donne l'implication :

$$R(\xi_1,\xi_2,...,\xi_n) \neq 0 \Rightarrow | R(\xi_1,\xi_2,...,\xi_n) | \geq 1/m_R^{(d_1.d_2...d_n-1)} .$$

Par suite, il suffit de calculer une approximation rationnelle convenable de $R(\xi_1,\xi_2,...,\xi_n)$ pour connaître son signe. Or d'après la proposition A.a12 , le calcul d'approximations rationnelles est très rapide.

c) Situation des racines réelles d'un polynome de $\mathbb{Q}[X]$

Nous examinons dans ce paragraphe deux démonstrations du théorème :

Théorème A.c1:

Il existe une $\mathcal{P}$-opération $\mathbb{Q}[X] \rightarrow \mathbf{Lst}(\mathbb{R}_{alg})$ qui calcule la liste ordonnée des racines réelles d'un polynôme à coefficients rationnels.

Recherche de racine par dichotomie sur un intervalle réel algébrique

Proposition A.c2 : Soit S la $\mathcal{P}$-partie de $\mathbb{Q}[X] \times \mathbb{R}_{alg} \times \mathbb{R}_{alg}$ formée par les (P, x, y) vérifiant : $x < y$, $P(x).P(y) < 0$.

Il existe une $\mathcal{P}$-opération $S \rightarrow \mathbb{R}_{alg}$ qui calcule à partir de P, x, y une racine de P sur l'intervalle $[x, y]$.

preuve> On peut $\mathcal{P}$-calculer des dyadiques a et b tels que $x < a < b < y$, $P(a).P(x') > 0$ pour $x' \in [x, a]$, $P(b).P(y') > 0$ pour $y' \in [b, y]$: en effet :
si M est un majorant de $|P'|$ sur un intervalle contenant $[x, y]$ et si $|P(x)| > m > 0$ il suffira que $|x' - x| < m/M$ pour que $P(x')$ ait même signe strict que $P(x)$, or m est $\mathcal{P}$-calculable d'après les propositions A.b2 et A.a4, et a s'en déduit par les propositions A.a4 et A.a6. On est donc ramené au cas où x et y sont rationnels (proposition A.a8). ❑

Méthode élémentaire (tableau de variation)

On procède "par recurrence" sur le degré du polynôme P.
Soit r rationnel positif tel que l'intervalle $]-r, r[$ contienne toutes les racines réelles de P.
Si on connaît la liste ordonnée des racines $x_1, \ldots, x_k$, (éventuellement vide), du polynôme dérivé P' sur l'intervalle $[-r, r]$, on pose $x_0 = -r$, $x_{k+1} = r$, on connaît le signe de P' sur chacun des intervalles $]x_i, x_{i+1}[$, donc le tableau de variation de P sur l'intervalle $[-r, r]$. On calcule ensuite les réels algébriques $P(x_i)$ (cf proposition A.b2), ou au moins leurs signes. On garde les x_i qui sont racines de P ; et il faut enfin calculer les racines sur les intervalles $]x_i, x_{i+1}[$ où P change de signe strict (cf proposition A.c2).

Il est clair que le calcul (décrit ci-dessus) de la liste des racines de P sur l'intervalle $[-r, r]$ à partir de celle des racines de P', est un $\mathcal{P}$-calcul.

Il suffit donc de vérifier que l'on peut polynomialement majorer la taille des polynômes dérivés successifs et de leurs tableaux de racines à partir de la taille du polynôme de départ P. La majoration pour $P \rightarrow [P', P^{(2)}, \ldots, P^{(d)}]$ ($d = \deg(P)$) est immédiate. La majoration pour la taille des racines s'en déduit par la proposition A.a7.

Si on utilise le théorème A.a11 et la proposition A.a12, on peut conduire tout l'agorithme en utilisant uniquement des calculs de valeurs approchées des réels algébriques considérés, qui sont de la forme $P^{(j)}(\zeta)$ où ζ est une racine réelle de $P^{(j+1)}$.

Notons que l'algorithme calcule en fait tous les zéros des polynômes $P, P', P^{(2)}, \ldots, P^{(d)}$. On peut en déduire sans se fatiguer beaucoup plus un tableau complet de signes et de variations pour la liste $[P, P', P^{(2)}, \ldots, P^{(d)}]$.

Méthode à la Sturm

Rappelons une version du théorème de Sturm : la *suite de Sturm* du polynôme P (supposé sans facteur carré) est la liste $S = [P_0, P_1, ..., P_k]$ définie de manière récurrente par:

$P_0 := P$, $P_1 := P'$, $P_{i+1} := -\mathbf{Rst}(P_i, P_{i-1})$ (on arrête au reste de degré 0)

On note $\mathbf{W_{Stu}}(a)$ le nombre de changements de signes dans la suite des $P_i(a)$ (en sautant les 0). Le théorème de Sturm affirme :

> Si $a < b$ le nombre de zéros réels de P sur l'intervalle $]a, b]$ est égal à $\mathbf{W_{Stu}}(a) - \mathbf{W_{Stu}}(b)$

La difficulté pour appliquer le théorème de Sturm est qu'on n'est pas sûr de pouvoir calculer en temps polynomial la suite de Sturm de P . Cette difficulté peut être tournée en calculant une liste de polynômes S ' équivalente à la suite de Sturm au sens que $\mathbf{W_{Stu}}(a) - \mathbf{W_{Stu}}(b) = \mathbf{W_{S'}}(a) - \mathbf{W_{S'}}(b)$. (Cf par exemple **[Lom2]** ou **[GLRR]**, le plus simple est de prendre la suite de Sturm-Habicht en introduisant une convention particulière pour le compte du nombre de changements en un point a racine de l'un des polynômes non identiquement nuls de la suite)

La 1ère partie de l'algorithme consiste donc à calculer $Q = P/pgcd(P,P')$ puis une liste de poynômes S ' équivalente à la suite de Sturm de Q .

La 2ème partie de l'algorithme consiste à isoler les racines sur des intervalles rationnels ouverts disjoints.

Si r est un entier majorant des valeurs absolues des racines de P , on calcule $\mathbf{W_{S'}}(-r) - \mathbf{W_{S'}}(r) = k$. Si $k = 1$ ou 0 , la 2ème partie de l'algorithme est terminée.

Sinon, on pose $d_0 = -r$, $d_1 = r$, $m_1 = k$ pour démarrer une dichotomie. A la fin de chaque étape, on aura une liste croissante de dyadiques $d_0 , ... , d_j$ et des entiers $m_1 , ... , m_j$ tels que : m_i est le nombre de racines réelles de P sur l'intervalle $] d_{i-1} , d_i]$.

Une étape (dichotomie) consiste à diviser en 2 ceux des intervalles de la liste précédente où il y a plus d'une racine : si l'un des demi-intervalles obtenus ne porte aucune racine et s'il jouxte un intervalle qui ne porte pas de racine non plus , on fusionne ces 2 intervalles : cela assûre qu'il n'y a jamais plus de 2.k intervalles.

La 2ème partie de l'algorithme se termine lorsque tous les m_i obtenus sont égaux à 0 ou 1. Les racines sont alors isolées les unes des autres, sur des intervalles dyadiques semi-ouverts. Les racines d_i éventuelles sont évidentes.

On peut donc s'attaquer à la *3ème partie de l'algorithme* : diviser en 2 les intervalles ouverts de la liste portant une racine, déterminer à chaque fois le bon demi-intervalle par le calcul de $\mathbf{W_{S'}}(c)$ (ou par le calcul du signe de P(c) si P change de signe sur l'intervalle), ceci jusqu'à ce que P ' soit évidemment-de-signe-constant sur l'intervalle portant la racine.

Le nombre d'étapes, aussi bien dans la 2ème partie que dans la 3ème partie de l'algorithme, est convenablement majoré grâce au lemme 1. De plus la taille des d_i successifs est majorée par $\mathbf{lg}(r)$ + nombre d'étapes. Comme le calcul d'une suite équivalente à la suite de Sturm est un $\mathcal{P}$-calcul, l'ensemble de l'algorithme est en temps polynomial.

Notons qu'il est possible de conduire les calculs en évitant que P s'annule en une borne d'un des intervalles considérés : si k est l'exposant de 2 dans cd(P) , aucune racine de P ne peut avoir un dénominateur avec un exposant de 2 supérieur à k . Si donc une borne, obtenue en dichotomisant un intervalle, s'avère être un zéro de P , il suffit de la décaler de $1/2^{k'}$ avec $k' > k$, et k' assez grand pour que la borne reste à l'intérieur de l'intervalle dichotomisé.

Généralisation

Signalons le théorème suivant, qui renforce le théorème A.c1, mais qui ne sera démontré que dans la partie B (cf. la remarque qui suit la proposition B.b4).

Théorème A.c3:

Soient $\xi_1, \xi_2, ..., \xi_n$ des éléments de $\mathbb{R}_{alg}$ racines de polynômes $Q_1, ..., Q_n$ de $\mathbb{Z}[X]$ et de degrés $d_1, ..., d_n$.
Alors les racines réelles du polynôme $X^n + \xi_1 X^{n-1} + ... + \xi_n$ peuvent être calculées comme éléments de $\mathbb{R}_{alg}$ en temps uniformément polynomial par rapport à $n.d_1...d_n$ et à la taille de la liste $[Q_1, ..., Q_n]$.

d) Deux mots sur $\mathbb{C}_{alg}$

Nous désignerons par $\mathbb{C}_{alg}$ l'ensemble des nombres complexes algébriques présenté sous la forme $\mathbb{R}_{alg}[\sqrt{-1}]$, cad $\mathbb{R}_{alg}{}^2$. On obtient alors les résultats suivants:

Théorème A.d1 :

Il existe une $\mathcal{P}$-opération $Ev_n : \mathbb{Q}[\sqrt{-1}][Z_1,Z_2,...,Z_n] \times \mathbb{C}_{alg}{}^n \to \mathbb{C}_{alg}$ avec :
$Ev_n(R, (\xi_1,\xi_2,...,\xi_n)) = R(\xi_1,\xi_2,...,\xi_n)$ (égalité au sens de $\mathbb{C}_{alg}$)

Théorème A.d2:

Il existe une $\mathcal{P}$-opération $\mathbb{Q}[X] \to \mathbf{Lst}(\mathbb{C}_{alg})$ qui calcule une liste des racines complexes d'un polynôme à coefficients rationnels, avec leurs multiplicités.

preuve>
théorème A.d1 : soit $P \in \mathbb{Q}[\sqrt{-1}][Z_1,Z_2,...,Z_n]$, et faisons $Z_i = X_i + \sqrt{-1}\ Y_i$. On obtient $P(Z_1,Z_2,...,Z_n) = P_1(X_1,X_2,...,X_n,Y_1,Y_2,...,Y_n)$. L'application $P \to P_1$ est une $\mathcal{P}$-fonction (cf par exemple dans **[Lom1]** les propositions B.a1 , B.c4 et B.f13). Si maintenant on sépare la partie réelle et la partie imaginaire de P_1 , on obtient 2 polynômes à 2n variables, qu'il s'agit d'évaluer dans $\mathbb{R}_{alg}$.
théorème A.d2 : on peut supposer les coefficients dans $\mathbb{Z}[\sqrt{-1}]$, on écrit comme ci-dessus
$P(Z) = Q(X,Y) + \sqrt{-1}\ R(X,Y)$ avec Q et R dans $\mathbb{Z}[X,Y]$,
on élimine Y entre Q et R, on obtient un polynôme T(X) , (calcul du déterminant d'une matrice à coefficients dans $\mathbb{Z}[X]$)
on élimine X entre Q et R, on obtient un polynôme S(Y) ,
si α est une racine de T et β une racine de S , il reste à tester si $P(\alpha + \sqrt{-1}\ \beta) = 0$: on peut appliquer la proposition précédente,
enfin la multiplicité des racines se teste en évaluant les dérivées successives. ❑

Question ouverte : un problème interessant consiste à trouver une borne inférieure de complexité pour les automorphismes non triviaux (c.-à-d. distincts de Id et de la conjugaison) de $\mathbb{C}_{alg}$. On peut conjecturer que tout automorphisme non trivial est de complexité en temps au moins exponentiel.
Notons qu'il en existe en temps primitif récursif : considérer par exemple l'automorphisme de $\mathbb{Q}[\sqrt{2}]$ qui échange $\sqrt{2}$ et $-\sqrt{2}$ et le prolonger de proche en proche à $\mathbb{C}_{alg}$ tout entier

en rajoutant une à une les racines des polynômes de $\mathbb{Z}[X]$ (à la $i^{ème}$ étape on obtient un isomorphisme explicite d'un sous corps $\mathbb{Q}[\zeta_i]$ de $\mathbb{C}_{alg}$ vers un sous corps $\mathbb{Q}[\xi_i]$ où ζ_i et ξ_i sont conjugués).

Signalons l'analogue du théorème A.b1:

Théorème A.d3 : Soit A un $\mathcal{P}$-ensemble et f une fonction de A vers $\mathbb{C}_{alg}$.
Pour que f soit $\mathcal{P}$-calculable, il faut et suffit que les 2 conditions suivantes soient vérifiées :

- la fonction $f : A \to \mathbb{C}$ est une $\mathcal{P}$-fonction
- il existe une $\mathcal{P}$-opération $G : A \to \mathbb{Z}[X] - \{0\}$ telle que :
  si $G(z) = S$, alors $f(z)$ est racine de S.

preuve> on calcule à partir de S deux polynômes de $\mathbb{Z}[X] - \{0\}$ annulant respectivement la partie réelle et la partie imaginaire de f(z) et on conclut par le théorème A.b1 . ❑

Notons enfin le résultat suivant qui concerne l'approximation et l'isolation des racines complexes d'un polynôme à coefficients entiers.

Proposition A.d4 :
Il existe une $\mathcal{P}$-opération $\mathbb{Z}[X] \to \mathbf{Lst}(\mathbb{Q}^2 \times \mathbb{Q}^+)$ qui, à partir d'un polynôme P sans facteur carré, calcule une liste d'éléments (x_i, y_i, r_i) vérifiant :

- chaque disque de centre (x_i, y_i) et de rayon r_i contient exactement une racine de P
- le processus itératif de Newton démarrant avec (x_i, y_i) converge vers la racine en question

preuve> cela résulte facilement du théorème A.d2 , d'une minoration de l'écart entre 2 racines complexes d'un polynôme sans facteur carré de $\mathbb{Z}[X]$ et des majorations classiques pour l'itération à la Newton. ❑

Remarque : La présentation de $\mathbb{C}_{alg}$ via la partie réelle et la partie imaginaire n'est pas en fait une présentation très naturelle . Par exemple la proposition A.d4 peut être réalisée par un algorithme, beaucoup plus performant que celui proposé ici, qui ne calcule pas en tant que telles les parties réelles et imaginaires des racines de P (cf par exemple [**Sch**] ou [**Pan**]). En conséquence, les racines de P sont représentées de manière nettement plus agréable sous la forme $(P, [x_i, y_i, r_i])$. Nous généralisons ce genre de présentation dans le § B b) .

e) Une généralisation

Soit $\mathbf{Q}'$ un corps ordonné dénombrable dans une présentation telle que :

(i) la relation d'ordre est $\mathcal{P}$-décidable

(ii) les lois de corps sont décrites par des $\mathcal{P}$-opérations[1]

(iii) les déterminants sont $\mathcal{P}$-calculables[2]

(iv) il existe une $\mathcal{P}$-opération qui, à partir d'un $x \in \mathbf{Q}'$, calcule un entier $m(x) \in \mathbb{N}$ majorant x dans $\mathbf{Q}'$[3].

On démontre alors facilement que:

(v) l'homomorphisme injectif $\mathbb{Q} \to \mathbf{Q}'$ est une $\mathcal{P}$-fonction

(vi) l'homomorphisme injectif croissant $\mathbf{Q}' \to \mathbb{R}$ est une $\mathcal{P}$-fonction

On peut alors chercher à avoir une représentation raisonnable de l'ensemble $\mathbf{R}'_{alg}$ des réels algébriques sur $\mathbf{Q}'$. On relit les § a, b, c, d en essayant de remplacer partout $\mathbb{Q}$ et $\mathbb{Z}$ par $\mathbf{Q}'$. On présente un élément de $\mathbf{R}'_{alg}$ par un triplet (P , a , b) où P est un polynôme unitaire de $\mathbf{Q}'[X]$, $a,b \in \mathbf{Q}'$. On s'aperçoit que presque toutes les démonstrations restent valables (sauf quelques unes que nous signalons ensuite).

En particulier, on obtient :

Théorème A.e1 : Sous les hypothèses (i), (ii), (iii), (iv) ci-dessus :

a) On peut construire une $\mathcal{P}$-opération pour l'homorphisme d'évaluation de $\mathbf{Q}'[X_1,X_2,...,X_n] \times \mathbf{R}'^{n}_{alg}$ vers $\mathbf{R}'_{alg}$:

$$Ev_n(R , (\xi_1,\xi_2,...,\xi_n)) = R(\xi_1,\xi_2,...,\xi_n)$$

b) On peut construire une $\mathcal{P}$-opération $\mathbf{Q}'[X] \to \mathbf{Lst}(\mathbf{R}'_{alg})$ qui calcule la liste ordonnée des racines réelles d'un polynôme à coefficients dans $\mathbf{Q}'$.

On notera que dans le a) la dépendance polynomiale est pour n fixé. Ce qui relativise le résultat obtenu.

Les démonstrations qui ne sont plus valables sont les suivantes :

- les majorations explicites qui tiennent comptent de faits particuliers à $\mathbb{Z}$: le fait qu'un résultant non nul entier est en valeur absolue ≥ 1 (théorème A.a11),
- le fait que les produits dans $\mathbb{Z}$ sont en O(n log(n) loglog(n)) (prop A.a12) ,
- et le fait qu'un élément de $\mathbb{Z}$ est exactement connu à partir d'une approximation à 1/4 près, dans la preuve du corollaire de la proposition A.a6 .

Signalons un "non-résultat" analogue dans le cas récursif : si $\mathbf{Q}'$ est un corps récursivement présenté, il n'y a pas automatiquement un test récursif pour la question " $u \in \mathbf{Q}'$? " lorsque $u \in \mathbf{R}'_{alg}$: on peut par exemple considérer une extension algébrique de $\mathbb{Q}$ obtenue en rajoutant les $\sqrt{p_n}$ où la suite p_n est une suite récursivement énumérable mais non récursive de nombre premiers. Le test $\sqrt{m} \in \mathbf{Q}'$? , pour $m \in \mathbb{N}$ n'est donc pas récursif.

---

[1] (i) et (ii) signifient que, dans la présentation considérée, $\mathbf{Q}'$ est un $\mathcal{P}$-corps-ordonné.

[2] on dit encore que $\mathbf{Q}'$ est **det-c$\mathcal{P}$c**

[3] on dira alors que $\mathbf{Q}'$ est $\mathcal{P}$-archimédien

Un exemple: Si α est un réel transcendant, le corps $\mathbb{Q}(\alpha)$ est un $\mathfrak{P}$-corps **det-c$\mathfrak{P}$c**. Ce corps est $\mathfrak{P}$-archimédien si et seulement si le nombre α est $\mathfrak{P}$-transcendant au sens suivant :

il existe une $\mathfrak{P}$-opération $\text{Min} : \mathbb{Z}[X] - \{0\} \rightarrow \mathbb{D}$ telle que

$0 < \text{Min}(P) \leq | P(\alpha) |$

En outre, la relation d'ordre dans $\mathbb{Q}(\alpha)$ est alors $\mathfrak{P}$-décidable, et on peut appliquer le théorème A.e1 .

Notons enfin que le fait pour α d'être $\mathfrak{P}$-transcendant peut encore s'exprimer au moyen de la majoration polynomiale suivante:

il existe $c , k , h \in \mathbb{N}$ tels que pour tout polynôme non nul P de $\mathbb{Z}[X]$ on ait : $| P(\alpha) | \geq 1/2^{c\, d^h\, ls(P)^k}$

où d est le degré de P et $ls(P) = \mathbf{lg(sup(}|$ coeffs de P $|\mathbf{))}$

B) DISCUSSION A PROPOS DE DIFFERENTES PRESENTATIONS DES NOMBRES ALGEBRIQUES

a) Position du problème

Un exemple : systèmes d'affectations polynomiales en cascade

Supposons que nous présentions un entier relatif sous forme du résultat d'un système d'affectations polynomiales en cascade. De ce point de vue, une liste

$[x_1, P_1, P_2, \ldots , P_k]$ avec

$P_1 \in \mathbb{Z}[X_1]$, $P_2 \in \mathbb{Z}[X_1,X_2]$, ... , $P_k \in \mathbb{Z}[X_1,X_2, \ldots ,X_k]$,

(où les P_i sont écrits en présentation creuse ordinaire) sert à représenter l'élément $x = x_{k+1}$ de $\mathbb{Z}$ défini par les affectations en cascade:

$x_2 := P_1(x_1)$

$x_3 := P_2(x_1,x_2)$

.....

$x_{k+1} := P_k(x_1,x_2 \ldots ,x_k)$

Nous noterons $\mathbb{Z}_{pol}$ la présentation ainsi obtenue de l'ensemble des entiers relatifs. Plus généralement si A désigne un anneau donné dans une présentation concrètement spécifiée, nous désignerons par **Sap**(A) la nouvelle présentation obtenue de manière analogue à $\mathbb{Z}_{pol}$, par système d'affectations polynomiales en cascade.

La présentation $\mathbb{Z}_{pol}$ est par bien des égards supérieure à la présentation standard en binaire $\mathbb{Z}$. En particulier, on peut effectuer *en temps linéaire* (c.-à-d. dans la classe **LINTIME**) les affectations polynomiales en cascade dans $\mathbb{Z}_{pol}$. Il suffit en effet de "juxtaposer" les différentes affectations après avoir changé quelques noms de variables. Et à vrai dire, on a fait exactement ce qu'il fallait pour rendre les affectations polynomiales en cascade calculables en temps linéaire dans la présentation qu'on a construit[1]. On peut exprimer le résultat précédent comme suit : l'évaluation $\mathbf{Sap}(\mathbb{Z}_{pol}) \rightarrow \mathbb{Z}_{pol}$ est en temps linéaire. Par contre, si on considère l'élément de $\mathbb{Z}_{pol}$ défini par $[2, X_1^2, X_2^2, \ldots , X_k^2]$, il est égal à 2^{2^k} et cela implique que l'évaluation $\mathbb{Z}_{pol} \rightarrow \mathbb{Z}$ est en espace exponentiel (au moins). Cela montre que la présentation $\mathbb{Z}_{pol}$ est nettement préférable à la présentation $\mathbb{Z}$ en ce qui concerne les affectations polynomiales en cascade.

Mais il y a un prix à payer ....

La raison poour laquelle on préfère en général travailler avec $\mathbb{Z}$ est qu'on ne sait pas bien faire dans $\mathbb{Z}_{pol}$ un certain nombre d'opérations et de tests considérés comme indispensables :

Questions ouvertes :

(1) Peut-on tester en temps polynomial l'égalité 2 éléments de $\mathbb{Z}_{pol}$?

---

[1] De la même manière que les présentations en magma sont exactement ce qu'il faut pour rendre la partie "lois de composition" d'une structure algébrique calculable en temps linéaire. (cf [**Lom1**] § B.e)

(2) Peut-on calculer en temps polynomial le signe d'un élément de $\mathbb{Z}_{pol}$?
(3) L'opération de division euclidienne est-elle une $\mathcal{P}$-opération dans $\mathbb{Z}_{pol}$?
La réponse à la troisième question semble presque sûrement négative : lorsqu'on divise 2 nombres dont l'ordre de grandeur est $2^{2^{n+1}}$ et 2^{2^n} le reste de la division est a priori du même ordre de grandeur, les 2 premiers nombres peuvent être choisis de manière à être présentés par une liste de taille environ c.n (où c est constant) dans $\mathbb{Z}_{pol}$, mais, pour un polynôme Q fixé, les listes de taille $\leq Q(c.n)$ ne représentent pas plus de $N^{Q(c.n)}$ nombres distincts[1], et il n'y a donc "aucune chance" pour que le reste de la division puisse être écrit *en espace polynomial* dans $\mathbb{Z}_{pol}$.

Un système d'affectations polynomiales en cascade peut être vu sous forme d'un programme à exécuter, dans lequel seul un jeu fini d'instructions est autorisé, sans aucune boucle. C'est ce que l'on appelle encore un *straight-line program* dans la littérature : les présentations par straight-line program ont surtout été étudiées pour des anneaux de polynômes à coefficients dans $\mathbb{Z}$ ou dans $\mathbb{Q}$, en général les seules instructions autorisées sont les instructions d'affectation : $Z \leftarrow c$ (c un élément de l'anneau donné dans une présentation "ordinaire"), $Z \leftarrow X$, $Z \leftarrow X + Y$, $Z \leftarrow X \times Y$, $Z \leftarrow X$ divisé par Y (le programme avorte si $Y = 0$ ou si X n'est pas divisible par Y). Un exemple typique d'un tel straight-line program est le programme permettant de calculer un déterminant par la méthode de Bareiss, sans recherche de pivot non nul. Si on exclut les divisions, on obtient une présentation $\mathcal{P}$-équivalente à la présentation par systèmes d'affectations polynomiales. Les résultats obtenus dans l'étude de la présentation par straight-line programs sont essentiellement probabilistes : par exemple on peut tester rapidement avec une très faible probabilité d'erreur si 2 entiers de $\mathbb{Z}_{pol}$ sont égaux en les calculant modulo quelques nombre premiers. On pourra par exemple consulter l'article de Kaltofen: [**Kal**] .

Si on augmente le nombre d'instructions autorisées, on assouplit la présentation des objets considérés, au détriment de la facilité à exécuter certains tests ou opérations. Si on autorise les boucles **Répéter** i **fois** (où i est la valeur prise par une variable du programme) on obtiendra une présentation $\mathbb{Z}_{prim}$ pour les entiers (nous mettons prim en indice pour indiquer que l'algorithme qui calcule l'entier est de type primitif récursif).
Divertissement mathématique: Le test d'égalité dans $\mathbb{Z}_{prim}$ est-il primitif récursif ?

La structure algebrique de $\mathbb{R}_{alg}$

Notons $\mathbf{R}_{alg}$ (avec un **R** gras) l'ensemble des nombres réels algébriques *abstrait* (c.-à-d. abstraction faite de tout présentation particulière de cet ensemble).

En tant qu'ensemble, c'est un ensemble *dénombrable* c.-à-d. *énumérable (1)* et *discret (2)* (nous reprenons la terminologie utilisée dans [**Lom1**]) .

Du point de vue de sa structure algébrique, nous retenons tout d'abord que c'est un *corps réel clos archimédien* , ce qui signifie :

(3) une strucure de corps
(4) une relation d'ordre total compatible avec la structure de corps
(5) la majoration de tout élément par un entier
(6) l'existence d'un zéro pour un polynôme P sur un intervalle où il change de signe.

En outre, nous avons :

(7) une injection naturelle $\mathbf{R}_{alg} \rightarrow \mathbb{R}$

---

[1] N est le nombre de symboles dans l'alphabet utilisé

(8) pour tout élément x de $\mathbf{R}_{alg}$ *l'existence d'un polynôme non nul de* $\mathbb{Z}[X]$ *qui annule x .*

Cela suffit pour une description abstraite de $\mathbf{R}_{alg}$, c.-à-d. à isomorphisme unique près. D'un point de vue constructif, la traduction des éléments de structure numérotés de (1) à (8) ci-dessus doit être entièrement faite en termes d'opérations, tests, fonctions, au sens constructif de ces termes. Dressons un tableau pour expliciter ceci :

élément de la structure	opération correspondante
(1) ensemble énumérable	construction d'objets concrets représentant les nombres réels algébriques abstraits : tout processus analogue à la construction des entiers naturels. On notera désormais R_a le préensemble ainsi construit.
(2) discret	on donne un test d'égalité dans R_a : c'est désormais un ensemble énumérable discret
(3) structure de corps	on donne les constantes 0 et 1 ainsi que les fonctions correspondant aux 4 opérations de la structure de corps (la fonction $x \to 1/x$ est définie pour $x \neq 0$)
(4) relation d'ordre	on donne un test pour $x < y$? en termes constructifs, on dit que la relation d'ordre est discrète
(5) archimédien	on donne une opération Maj : $R_a \to \mathbb{N}$ telle que $x \leq Maj(x).1$ pour tout x
(6) réel clos	on donne une opération Rac : $R_a[X] \times R_a \times R_a \to R_a$ telle que : Si $a < b$ et $P(a).P(b) < 0$, alors Rac(P , a , b) = u avec : u est une racine de P sur] a , b [(a priori Rac n'est pas une fonction)
(7) injection canonique $R_a \to \mathbb{R}$	on donne une opération $F : R_a \times \mathbb{N}_1 \to \mathbb{D}$ telle que $F(x,n) = r/2^n$ (avec $r \in \mathbb{Z}$) est une approximation de x avec la précision 2^{-n} .
(8) tous les éléments sont algébriques sur $\mathbb{Q}$	on donne une opération Pol : $R_a \to \mathbb{Z}[X] - \{0\}$ telle que : $Pol(x)(x) = 0$ pour tout x (a priori Pol n'est pas une fonction)

Considérons maintenant la présentation naïve $\mathbb{R}_{alg}$ définie en A.a . C'est une $\mathcal{P}$-présentation de la structure dans la mesure où les éléments de structure (2) (3) (4) (5) (7) (8) sont réalisables comme des $\mathcal{P}$-opérations. Mais il y a manifestement deux points faibles. D'une part, lorsqu'on fait des opérations arithmétiques en chaîne, par exemple lorsqu'on évalue un polynôme avec un nombre d'indéterminées non fixé a priori, il y a une croissance exponentielle inévitable de la taille des réels calculés, à cause de l'explosion de leur degré. Autrement dit, $\mathbb{R}_{alg}$ n'est pas une présentation complètement-$\mathcal{P}$-calculable de la structure de corps. D'autre part, pour ce qui concerne la recherche des racines d'un polynôme à coefficients dans $\mathbb{R}_{alg}$ on a le même problème d'explosion exponentielle du degré donc de la taille.

En fait, nous allons voir qu'on ne peut avoir les opérations (2) → (8) simultanément en temps polynomial.

Présentations $\mathcal{P}$-équivalentes à la présentation naïve

Proposition B.a1 :

Soit R_a un corps ordonné, donné dans une présentation telle que :

(3') les lois de corps sont $\mathcal{P}$-calculables

(4') le signe d'un élément est $\mathcal{P}$-calculable

(6') il existe une $\mathcal{P}$-opération Rac : $\mathbb{Z}[X] \times \mathbb{Q} \times \mathbb{Q} \to R_a$ telle que :

$(a < b\ ,\ P(a).P(b) < 0) \Rightarrow$ Rac(P , a , b) = u est une racine de P sur] a , b [

(8') Il existe une opération Pol : $R_a \to \mathbb{Z}[X] - \{0\}$ telle que :

Pol(x)(x) = 0 pour tout x .

Alors R_a et $\mathbb{R}_{alg}$ sont deux présentations $\mathcal{P}$-isomorphes de $\mathbf{R}_{alg}$

preuve>

On commence par remarquer que R_a est le corps des nombres réels algébriques, d'après (6') et (8').

La propriété (6') implique évidememnt que l'isomorphisme unique de $\mathbb{R}_{alg}$ vers R_a est une $\mathcal{P}$-fonction. En particulier, l'injection $\mathbb{Q} \to R_a$ est une $\mathcal{P}$-fonction.

Pour montrer que l'isomorphisme de R_a vers $\mathbb{R}_{alg}$ est une $\mathcal{P}$-fonction, on applique le théorème A.b1 . Vu (8'), il nous suffit de démontrer que l'injection $R_a \to \mathbb{R}$ est une $\mathcal{P}$-fonction. Nous voulons donc $\mathcal{P}$-calculer, pour $(x,n) \in R_a \times \mathbb{N}_1$ un rationnel b approchant x avec la précision 2^{-n} . On teste $x > 0$? (si non on travaillera avec -x). On majore x par un entier m (obtenu à partir des coefficients de Pol(x)). On procède ensuite par dichotomie à partir de 0 et m : le nombre de tests et la taille des rationnels intervenant dans les tests " $x > r$? " sont convenablement majorés. Comme l'injection $\mathbb{Q} \to R_a$ est une $\mathcal{P}$-fonction, le temps necessaire pour ces tests est lui même convenablement majoré. ❑

Ceci ne signifie pas pour autant que certaines présentations $\mathcal{P}$-isomorphes à $\mathbb{R}_{alg}$ ne soient pas préférables à d'autres.

Par exemple, nous pouvons accepter de représenter un nombre algébrique sous forme $R(\xi_1,\xi_2,...,\xi_n) / d$, où n est a priori majoré par un n_0 fixe, où les ξ_i sont des éléments de $\mathbb{R}_{alg}$ définis comme racines de polynômes unitaires, où R est à coefficients dans $\mathbb{Z}$, et où d est un entier . D'après la proposition A.b6 on obtient une présentation $\mathcal{P}$-isomorphe à $\mathbb{R}_{alg}$, mais les calculs y sont plus souples (voir notamment la remarque qui suit la proposition A.b6).

Nous étudions dans le § b) une présentation très souple, par systèmes d'équations emboitées, non $\mathcal{P}$-isomorphe à $\mathbb{R}_{alg}$, mais pour laquelle les majorations de taille et de temps de calcul sont essentiellement les mêmes que dans $\mathbb{R}_{alg}$.

Autres présentations

L'espoir de tout réaliser en temps polynomial étant exclu, la tentative raisonnable serait de laisser tomber (8) (la $\mathcal{P}$-calculabilité d'un polynôme annulant x) en ne conservant qu'une caratérisation indirecte de l'algébricité de x .

Même dans ce cas, il semble cependant improbable qu'une autre présentation du corps des réels algébriques puisse rendre à la fois le test de comparaison $\mathcal{P}$-décidable et l'addition et le produit c-$\mathcal{P}$-c . Une réponse définitivement négative serait obtenue si on démontrait un résultat analogue à celui énoncé ci-dessous:

Question ouverte :

? l'opération qui, à partir d'une liste d'entiers $[x_1, \dots, x_n]$, (les x_i dans $\mathbb{Z}$ écrits en binaire) calcule le signe de la somme $\sum x_i^{1/3}$ *n'est pas* calculable en temps polynomial.

Nous discutons maintenant la difficulté de réaliser simultanément les 3 exigences suivantes :

- un test de comparaison en temps polynomial
- calcul d'approximations rationnelles en temps polynomial
- définition des réels algébriques par systèmes d'équations emboitéees

Nous commençons par remarquer que la possibilité donnée a priori de définir un réel algébrique par un système d'équations en cascade n'offre, prise isolément, aucune difficulté particulière (cf $\mathbb{Z}_{pol}$ dans un contexte analogue). Par ailleurs, elle implique que la structure d'anneau ainsi présentée est complètement-$\mathcal{P}$-calculable, et même que les affectations polynomiales en cascade sont calculables en temps linéaire dans la présentation retenue (car une affectation polynomiale $z := P(\xi_1,\xi_2,\dots,\xi_n)$ n'est jamais que le calcul de la racine d'une équation de degré 1 en z). Or il semble déjà bien problématique de donner un test de comparaison en temps polynomial dans $\mathbb{Z}_{pol}$.

Rappelons par ailleurs que si on a une présentation de $\mathbf{R}_{alg}$ où le test de comparaison est en temps polynomial, alors on aura le calcul d'approximations rationnelles en temps polynomial pour tout nombre de valeur absolue "pas trop grande" (par dichotomie).

Nous donnons 2 exemples pour mettre en évidence la difficulté a priori de l'entreprise (si elle n'est pas déjà vouée à l'échec par une réponse positive à la question ouverte supra marquée d'un gros point d'interrogation). On considère pour cela le polynôme $P(X,Y) := Y^3 - X^2 - 1$. Pour tout x réel, l'équation $P(x,y) = 0$ admet une racine unique en y . Pour tout $y \geq 2$, l'équation $P(x,y) = 0$ admet une racine $x > 2$ unique.

Considérons d'abord le système d'équations :

$$P(x_1,x_2) = 0 , P(x_2,x_3) = 0 , \dots , P(x_{n-1},x_n) = 0 , \; x_1 = 2$$

On peut alors calculer en temps polynomial à partir de l'entrée $(n,m) \in \mathbb{N}_1^2$ un rationnel $a_{n,m}$ tel que $| x_n - a_{n,m} | \leq 1/2^m$. Par contre, il n'est pas du tout évident qu'on ait un test en temps polynomial pour " $x_n - b > 0$? ", à partir de l'entrée $(n,b) \in \mathbb{N}_1 \times \mathbb{Q}$. En effet, le degré du réel x_n est a priori 3^n , et son irrationalité très forte ne permet pas *a priori* de le situer rapidement par rapport à un rationnel par un simple calcul de valeurs approchées.

Si maintenant, on considère les mêmes équations "prises à l'envers" :

$$P(x_n,x_{n-1}) = 0 , \dots , P(x_3,x_2) = 0, P(x_2,x_1) = 0 , \; x_1 = 4 , \; x_2 > 0 , \dots , x_n > 0$$

système qui donne un caractérisation semi-algébrique de x_n , il est impossible de calculer en temps polynomial à partir de l'entrée $n \in \mathbb{N}_1$ un rationnel qui approche x_n à 1 près, tout simplement parce que $x_n > 2 + 2^{(3/2)^n}$. On peut en revanche espérer avoir un test de comparaison à un nombre rationnel en temps polynomial , car cette comparaison n'est délicate que lorsque le rationnel à comparer est très grand, de sorte que la taille de l'entrée est elle-même très grande.

b) Systèmes d'équations en cascade, avant la levée de l'ambiguité

Position du problème, notations

Signalons pour commencer qu'il est nettement plus agréable, plutôt que travailler dans $\mathbb{R}_{alg}$, de travailler avec les *entiers algébriques réels* en considérant la partie $\mathbb{R}_{e,alg}$ formée des triplets (P , a , b) où P est un polynôme unitaire et où a et b sont de la forme $(c - 1)/2^n$ et $(c + 1)/2^n$ avec $c \in \mathbb{Z}$. Par ailleurs tout calcul dans $\mathbb{R}_{alg}$ se ramène facilement à un calcul dans $\mathbb{R}_{e,alg}$. On peut enfin noter $\mathbb{C}_{e,alg}$ la présentation des entiers algébriques complexes via leurs parties réelles et imaginaires (présentées dans $\mathbb{R}_{e,alg}$).

Nous étudions dans ce paragraphe une présentation des entiers algébriques réels ou complexes, que nous notons $\mathbb{C}_{sae,N}$ et qui est directement inspirée du système D5 ([**DD**]) . Ce dernier utilise des systèmes d'équations emboitées: de tels systèmes d'équations peuvent avoir plusieurs solutions et il y a donc ambiguité quant au nombre algébrique décrit . Le problème le plus immédiat qui se pose avec D5 est celui de la levée des ambiguités "en cours de calcul" . Cette levée des ambiguités, avec en sortie tous les cas possibles, peut manifestement prendre un temps exponentiel, si par exemple on demande d'additionner 2n nombres racines de l'équation $X^2 = 2$, et qu'on pose le problème de savoir si la somme obtenue est nulle . Par ailleurs, le maintien des ambiguités *aussi longtemps que possible* peut très bien être vu aussi comme le principal avantage de D5 . L'ambition de D5 est d'être utilisable pour tous calculs usuels sur les nombres algébriques, à la demande, un peu comme on utilise des entiers de longueur arbitraire dans n'importe quel système de calcul formel.
Nous étudions ici ce qui se passe lorsqu'on lève a priori l'ambiguité en donnant une approximation rationnelle (dans $\mathbb{Q}[\sqrt{-1}]$) convenable de la solution. En ce qui concerne les entiers algébriques réels, ils sont simplement obtenus lorsqu'on impose à l'approximation rationnelle d'être réelle. Il va de soi que le système pourrait être adapté pour des calculs avec des entiers algébriques p-adiques.

Le résultat auquel on arrive est celui-ci :

> *tout* calcul *raisonnable* dans $\mathbb{C}_{sae,N}$ peut être mené en temps uniformément polynomial par rapport à, d'une part la taille de l'entrée, d'autre part les "degrés a priori" (voir définition un peu plus loin) des entiers algébriques entrés.

Et ces calculs raisonnables comprennent le calcul de valeurs approchées, le test de comparaison, la recherche des racines d'une équation et la résolution de certains systèmes d'équations linéaires (ceux dont les coefficients restent dans un sous-corps convenablement contrôlé).

On pourra objecter que, finalement, on n'obtient rien de fondamentalement meilleur qu'avec $\mathbb{R}_{e,alg}$. La réponse est que D5 possède beaucoup plus de souplesse, ce qui permet dans bien des cas d'avoir un calcul en temps polynomial par rapport à la seule taille des entrées, qui peut être beaucoup plus petite que la taille des entrées analogues dans $\mathbb{R}_{e,alg}$. D'autre part, la meilleure méthode pour démontrer les majorations correspondantes dans $\mathbb{R}_{alg}$ est sans doute via la présentation D5 .

Systèmes d'équations algébriques emboitées

Un *système d'équations algébriques emboitées* (ou encore "en cascade") est donné par une liste de polynômes $\mathbf{P} := [P_1,P_2,...,P_k]$ avec

$$P_1 \in \mathbb{Z}[X_1] , P_2 \in \mathbb{Z}[X_1,X_2] , ... , P_k \in \mathbb{Z}[X_1,X_2,...,X_k]$$

chaque P_j étant unitaire de degré d_j en tant que polynôme en X_j

Le système est dit *normalisé* si les conditions suivantes sur les degrés sont réalisées

$$d_j \geq 2 \text{ pour tout } j \text{ et } d_{X_h}(P_j) < d_h \text{ pour tout } h < j$$

Dans un système normalisé, on évite les affectations polynomiales en cascade pour se concentrer sur l'aspect "solution d'équations algébriques".

Une *solution réelle (resp. complexe) du système défini par la liste* $\mathbf{P}$ est un k-uple $\xi = [\xi_1,\xi_2,...,\xi_k]$ de nombres réels (resp. complexes) vérifiant

$$P_1(\xi_1) = 0 , P_2(\xi_1,\xi_2) = 0 , ... , P_1(\xi_1,\xi_2,...,\xi_k) = 0 .$$

On est alors amené naturellement à travailler dans l'anneau $\mathbb{Z}[\xi_1,\xi_2,...,\xi_k]$. Nous noterons $\mathbb{A}_\xi$ cet anneau.

Le problème de la levée de l'ambiguité

Un système normalisé d'équations algébriques emboitées étant donné, se pose le problème de la levée de l'ambiguité, c.-à-d. comment coder une solution particulière du sytème.
Dans le cas des solutions réelles, on peut envisager pour cela plusieurs méthodes:
– codage de la racine ξ_i de $P_i(\xi_1,...,\xi_{i-1},X_i)$ via les signes que prennent les dérivées successives de P_i (par rapport à la variable X_i), en utilisant le lemme de Thom (cf [**CoR**])
– codage de la racine ξ_i de $P_i(\xi_1,...,\xi_{i-1},X_i)$ par son numéro d'ordre (le nombre de racines réelles est connu par le théorème de Sturm)
– on situe la racine sur un intervalle rationnel où le polynôme admet une seule racine réelle (de nouveau utilisation du théorème de Sturm)
– méthode naïve: on situe la racine sur un intervalle rationnel où le polynôme change de signe et où la dérivée reste de signe constant de manière évidente
– méthode analytique (ou "purement numérique"): on donne une approximation rationnelle[1] $(x_1,x_2,...,x_k)$ de $(\xi_1,\xi_2,...,\xi_k)$ avec l'assurance que le processus de Newton, appliqué pour la valeur initiale $(x_1,x_2,...,x_k)$ convergera vers $(\xi_1,\xi_2,...,\xi_k)$.

A priori, dans le cas réel, il semble que la meilleure solution doive être recherchée à l'une des 2 extrémités, selon que l'on se situe dans un cadre de géométrie algébrique réelle ou de géométrie analytique réelle.

Nous étudierons ici les résultats de complexité quand on adopte le dernier point de vue, et nous nous situerons d'emblée dans le cas complexe. Signalons quelques avantages qui sautent immédiatement au regard:
– comme la solution $(\xi_1,\xi_2,...,\xi_k)$ est traitée globalement, on n'aura pas de récurrence sur k à assumer, et la taille des calculs sera plus aisée à maîtriser
– tous les calculs " dans $\mathbb{C}$ " sont a priori très aisés (grande efficacité de la méthode de Newton)
– la méthode est facilement généralisable au cas réel ou p-adique; et dans ce dernier cas, Hensel (c.-à-d. Newton p-adique) est encore plus facile à contrôler.
Signalons également deux désavantages (liés entre eux d'ailleurs)

---

[1] Comme dèjà signalé, dans le cas complexe, nous parlons d'approximation rationnelle pour une approximation dans $\mathbb{Q}[\sqrt{-1}]^k$

– seules les racines *simples* d'un système d'équations donné (c.-à-d.: chaque ξ_i est racine simple du polynôme correspondant) sont *immédiatement* codables, c.-à-d. sans changer de système d'équations.(en fait, voir l'extension du codage donnée dans la définition B.c10)
– certaines racines d'un système "peu encombrant" peuvent avoir un code "relativement encombrant" (en particulier les racines "presque doubles")
Il semble clair que les désavantages sont exactement symétriques des avantages. Sans doute à l'autre extrémité, avec la méthode à la Thom, la situation serait elle renversée.

Nous noterons $\mathbb{C}_{sae,N}$ l'ensemble des entiers algébriques complexes dans la présentation via des systèmes d'équations algébriques emboitées, la levée de l'ambiguité étant faite à la Newton (nous précisons plus loin exactement cette présentation et en particulier comment on assure la convergence). Nous dirons que *le couple* $(\mathbf{P},[x_1,...,x_k])$ *constitue une présentation de la liste* $\xi = [\xi_1,\xi_2,...,\xi_k]$ dans $\mathbb{C}_{sae,N}$. Enfin, si le polynôme R de $\mathbb{Z}[X_1,X_2,...,X_k]$ a son degré en chaque X_i inférieur à d_i nous dirons que *le triplet* $(\mathbf{P},[x_1,...,x_k],R)$ *constitue une présentation de l'entier algébrique* $R(\xi_1,\xi_2,...,\xi_k)$ dans $\mathbb{C}_{sae,N}$.

L'entier $\mathbf{d} := d_1\, d_2 \ldots d_k$ est par définition le *degré a priori* des entiers algébriques ξ_k et $R(\xi_1,\xi_2,...,\xi_k)$ dans cette présentation. Dans un contexte où plusieurs systèmes d'équations emboitées interviennent, on notera $dg(\mathbf{P})$ pour $d_1\, d_2 \ldots d_k$.

Enfin, nous noterons $\mathbb{R}_{sae,N}$ l'ensemble des entiers algébriques réels, donnés dans la présentation analogue à $\mathbb{C}_{sae,N}$ (l'approximation rationnelle étant dans $\mathbb{Q}^k$).

L'anneau $\mathbb{A}_{\mathbf{P}}$

Un système *normalisé* d'équations algébriques emboitées étant donné par la liste $\mathbf{P}$, l'anneau $\mathbb{A}_{\mathbf{P}}$ est par définition le quotient $\mathbb{Z}[X_1,X_2,...,X_k] / <\mathbf{P}>$, où $<\mathbf{P}>$ est l'idéal engendré par $P_1(X_1)$, $P_2(X_1,X_2)$, ... , $P_k(X_1,X_2,...,X_k)$. C'est un $\mathbb{Z}$-module libre de dimension $\mathbf{d}$ dont une base canonique est donnée par les monômes unitaires de $\mathbb{Z}[X_1,X_2,...,X_k]$ de degré $< d_j$ en chaque variable X_j . Cet anneau (variable au cours des calculs puisqu'on doit pouvoir introduire de nouveaux nombres algébriques à volonté) est le cadre de travail naturel dans D5 . C'est à la fois parce que cet anneau est "variable" et parce que les calculs raisonnables sont (relativement) bien maîtrisés dans cet anneau que la présentation $\mathbb{C}_{sae,N}$ est (relativement) efficace.
Un élément α *de l'anneau* $\mathbb{A}_{\mathbf{P}}$ *est toujours considéré comme présenté via ses coordonnées sur la base canonique, a priori en présentation creuse.* Si $\mathbf{lg}(\alpha)$ est sa taille en présentation creuse, sa taille en présentation dense est majorée par $\mathbf{d}\ \mathbf{lg}(\alpha)$. Il est cependant "exceptionnel" que la taille en présentation creuse reste significativement plus petite que la taille en présentation dense après que quelques calculs (des produits, notamment) aient été effectués dans $\mathbb{A}_{\mathbf{P}}$.
Nous notons $\mathbf{lg}(\mathbf{P})$ la taille de la liste $\mathbf{P}$, les entiers étant écrits en binaire et la présentation des polynômes pouvant être creuse. La taille de la liste en présentation dense est alors majorée par $\mathbf{k}\ \mathbf{d}\ \mathbf{lg}(\mathbf{P})$.

Si $\xi = [\xi_1,\xi_2,...,\xi_k]$ est une solution réelle (ou complexe, ou p-adique) du système défini par la liste $\mathbf{P}$, l'anneau $\mathbb{A}_\xi = \mathbb{Z}[\xi_1,\xi_2,...,\xi_k]$ est évidemment un quotient de $\mathbb{A}_{\mathbf{P}}$. Alors que dans $\mathbb{A}_{\mathbf{P}}$ nous avons une écriture unique pour chaque élément, il n'en est pas de même pour $\mathbb{A}_\xi$, et cela pose quelques problèmes pour majorer la taille des calculs dans $\mathbb{A}_\xi$.

Un des buts essentiels de ce § est de montrer le résultat suivant :

la recherche des solutions (comme éléments de $\mathbb{C}_{e,alg}^{k}$ *ou comme éléments de* $\mathbb{C}_{sae,N}^{k}$*) d'un système normalisé d'équations algébriques emboitées* **P** *peut être réalisée en temps uniformément polynomial par rapport à* **d** *et* **lg(P)** .

En tant que résultat général, on ne peut évidemment espérer mieux, vu que le degré a priori de l'entier algébrique dans la présentation $\mathbb{C}_{sae,N}$ est bien souvent son vrai degré, et vu le nombre de solutions possibles a priori. Si la taille d'*une* solution dans $\mathbb{C}_{e,alg}^{k}$ est en règle générale contrôlée polynomialement par **lg(P)** et **d** , il semble relativement fréquent que la taille dans $\mathbb{C}_{sae,N}^{k}$ soit, elle, contrôlée seulement par **lg(P)** , ce qui montrerait la supériorité des présentations à la D5 .

Majorations polynomiales uniformes pour les calculs dans $\mathbb{A}_{\mathbf{P}}$

Les techniques de majoration que nous utilisons ici sont celles données dans [**Lom1**] à propos des $\mathcal{P}_0$-anneaux. Nous sommes cependant obligés de redémontrer certains résultats dans la mesure où nous souhaitons des majorations uniformes (avec **P** variable, donc $\mathbb{A}_{\mathbf{P}}$ variable). Nous rappelons que nous travaillons avec un système **P** normalisé.

Notations pour différentes grandeurs reliées à la taille d'une matrice

Lorsque **M** est une matrice, ou un polynôme, ou une liste de matrices etc... à coefficients dans $\mathbb{Z}$, donné dans une présentation précisée (creuse ou dense) nous noterons :

$| M |_1 := \mathbf{lg}(\Sigma |$ coeffs de M $|)$ $\qquad | M |_2 := \mathbf{lg}\left(\sqrt{\Sigma |\text{ coeffs de M }|^2}\right)$

$| M |_\infty := \mathbf{lg}(\sup |$ coeffs de M $|)$

dim(M) := le nombre de coefficients dans la présentation dense "naturelle"

Par exemple, pour la liste de polynômes **P** considérée ici :

$\dim(\mathbf{P}) = d_1 + d_1 d_2 + \ldots + d_1 d_2 \ldots d_k$

On a : $| M |_\infty \leq | M |_2 \leq | M |_1 \leq \mathbf{lg}(\dim(M)) + | M |_\infty$ et la taille en présentation dense est majorée par $\dim(M) | M |_\infty$.

Le résultat vraiment utile est le suivant : si **M** et **N** sont 2 polynômes, ou 2 matrices (de dimensions convenables), alors $| MN |_1 \leq | M |_1 + | N |_1$.

Majoration pour l'addition et le produit dans $\mathbb{A}_{\mathbf{P}}$

Nous noterons α , β , γ ... des éléments de $\mathbb{A}_{\mathbf{P}}$. Nous noterons $\lambda_1,\lambda_2,\ldots,\lambda_k$ les variables $X_1,X_2,\ldots,X_k$ vues comme éléments de $\mathbb{A}_{\mathbf{P}}$.

L'addition dans $\mathbb{A}_{\mathbf{P}}$ est simplement l'addition coefficient par coefficient, ce qui donne la majoration :

$$\boxed{| \alpha + \beta |_1 \leq \mathbf{sup}(| \alpha |_1, | \beta |_1) + 1} \qquad (1)$$

et donc également :

$$\boxed{\left| \sum_{i=1}^{n} \alpha_i \right|_1 \leq \mathbf{sup}_{i=1,\ldots,n} (| \alpha_i |_1) + \mathbf{lg}(n)} \qquad (2)$$

Le produit dans $\mathbb{A}_{\mathbf{P}}$ est à peine plus compliqué. Notons P_α et P_β les polynômes (de degrés $\leq d_i - 1$ en X_i (i=1,...,k)) qui correspondent à α et β . Le produit $\alpha.\beta$ dans $\mathbb{A}_{\mathbf{P}}$ s'obtient en réduisant modulo l'idéal $< \mathbf{P} >$ (engendré par la liste **P**) le polynôme $P_\alpha.P_\beta$.

On a déjà $| P_\alpha.P_\beta |_1 \le | P_\alpha |_1 + | P_\beta |_1 = | \alpha |_1 + | \beta |_1$. Le polynôme $P_\alpha.P_\beta$ est de degré $\le 2d_i - 2$ en X_i (i=1,...,k) .
Notons $\mathcal{P}_\mathbf{d}$ le $\mathbb{Z}$-module libre des polynômes de degré $\le 2d_i - 2$ en X_i (i=1,...,k) , c'est un module de dimension :

$$\boxed{\mathbf{d'} = \prod_{i=1}^{k}(2d_i - 1) \le \mathbf{d}^2} \qquad (3)$$

La réduction modulo $< \mathbf{P} >$ pour un polynôme Q de $\mathcal{P}_\mathbf{d}$ revient à réécrire Q sur la base $\mathfrak{B}_\mathbf{P}$ définie ci-après, et à garder les coordonnées utiles. La base $\mathfrak{B}_\mathbf{P}$ est formée des monômes de degré $\le d_i - 1$ en X_i (i=1,...,k) , puis de produits $P_i.M_{i,h}$ où les $M_{i,h}$ sont tous les monômes de degrés majorés selon le tableau suivant, où on a noté $r_i = 2d_i - 2$:

i	degré en X_1	degré en X_2	degré en X_3	degré en X_4	.....	degré en X_k
1	$\le r_1 - d_1$	$\le r_2$	$\le r_3$	$\le r_4$	.....	$\le r_k$
2	$< d_1$	$\le r_2 - d_2$	$\le r_3$	$\le r_4$	.....	$\le r_k$
3	$< d_1$	$< d_2$	$\le r_3 - d_3$	$\le r_4$	.....	$\le r_k$
.	...	...	...	...	.....	...
.	...	...	...	...	.....	...
k	$< d_1$	$< d_2$	$< d_3$	$< d_4$	.....	$\le r_k - d_k$

Si $M = c\, X_1^{s_1} X_2^{s_2} \dots X_k^{s_k}$ est un monôme, nous dirons que le k-uple $\mathbf{s} := [s_0,s_1,...,s_k]$ est l'exposant du monôme, et nous noterons $c\, \mathbf{X^s}$ ce monôme. Nous ordonnons les exposants selon l'ordre lexicographique suivant :

$[s_0,s_1,...,s_k]$ précède $[t_0,t_1,...,t_k]$ ssi $\exists\, i$ $s_i < t_i$, $s_{i+1} = t_{i+1}$, ... , $s_k = t_k$
Alors le monôme dominant de P_i est $X_i^{d_i}$ et le monôme dominant de $P_i\, \mathbf{X^s}$ est $X_i^{d_i}\, \mathbf{X^s}$. Par ailleurs, pour tout exposant s pour un monôme $\mathbf{X^s}$ de $\mathcal{P}_\mathbf{d}$, ou bien $\mathbf{X^s}$ est dans la base canonique de $\mathcal{A}_\mathbf{P}$, ou bien il existe un i unique tel que : $d_1 < s_1$, ... , $d_{i-1} < s_{i-1}$, $d_i \ge s_i$; de sorte que $\mathbf{X^s}$ est le monôme dominant d'un unique polynôme de la base $\mathfrak{B}_\mathbf{P}$. En d'autres termes, la base $\mathfrak{B}_\mathbf{P}$ est triangulaire par rapport à la base canonique de $\mathcal{P}_\mathbf{d}$, formée des monômes $\mathbf{X^s}$ rangés selon l'ordre lexicographique défini ci-dessus. Réduire le polynôme $P_\alpha.P_\beta$ modulo $< \mathbf{P} >$, revient à multiplier la matrice $\Gamma_\mathbf{P}$, inverse de la matrice de la base $\mathfrak{B}_\mathbf{P}$, par le vecteur colonne correspondant au polynôme $P_\alpha.P_\beta$. Cette matrice inverse a pour coefficients des cofacteurs de la matrice de la base $\mathfrak{B}_\mathbf{P}$. D'après l'inégalité de Hadamard pour majorer les déterminants, on a donc :

$$| \Gamma_\mathbf{P} |_\infty \le (\mathbf{d'} - \mathbf{d})\, \mathbf{sup}(\, | P_i |_2 \,) \le (\mathbf{d'} - \mathbf{d})\, | \mathbf{P} |_2 \le (\mathbf{d'} - \mathbf{d})\, | \mathbf{P} |_1$$

d'où : $$| \Gamma_\mathbf{P} |_1 \le 2\, \mathbf{lg}(\mathbf{d'}) + (\mathbf{d'} - \mathbf{d})\, \mathbf{sup}(\, | P_i |_2 \,)$$

nous noterons $$\boxed{m_\mathbf{P} = 2\, \mathbf{lg}(\mathbf{d'}) + (\mathbf{d'} - \mathbf{d})\, \mathbf{sup}(\, | P_i |_2 \,)} \qquad (4)$$

d'où enfin $$\boxed{| \alpha \times \beta |_1 \le m_\mathbf{P} + | \alpha |_1 + | \beta |_1} \qquad (5)$$

et $$\boxed{\left| \prod_{i=1}^{n} \alpha_i \right|_1 \le \sum_{i=1}^{n} | \alpha_i |_1 + (n-1)\, m_\mathbf{P}} \qquad (6)$$

NB: Le calcul du prouit $\alpha \times \beta$ dans $\mathbf{A_P}$ est donc en temps uniformément polynomial par rapport à **d** et la taille des entrées, c.-à-d. encore par rapport à $|\alpha|_1$, $|\beta|_1$, **d** et **lg(P)** . Par exemple en résolvant le système triangulaire par substitutions successives.

Remarques :

1) Si on pose $|\alpha|_P := m_P + |\alpha|_1$, alors on a les 2 majorations :
$|\alpha+\beta|_P \le \sup(|\alpha|_P, |\beta|_P) + 1$ et $|\alpha \times \beta|_P \le |\alpha|_P + |\beta|_P$
Pour **P** fixé, les calculs de majorations dans $\mathbf{A_P}$ sont donc entièrement analogues à ceux dans $\mathbb{Z}$.

2) Si on a "beaucoup" de calculs à faire dans $\mathbf{A_P}$ avec **P** fixé, on peut construire une fois pour toutes la table de multiplication de $\mathbf{A_P}$, c.-à-d. évaluer une fois pour toutes les expressions $\lambda_1^{n_1}...\lambda_h^{n_h}$ où $n_1 \le 2d_1 - 1$, $n_2 \le 2d_2 - 1$, ... , $n_{h-1} \le 2d_{h-1} - 1$, $d_h < n_h \le 2d_h - 1$ $(1 \le h \le k)$.

On déduit des majorations précédentes les 2 propositions qui suivent :

Théorème B.b1 :

a) Soit $[\alpha_i]_{i=1,...,m}$ une liste d'éléments de $\mathbf{A_P}$ et $Expr(Y_1,...,Y_m)$ une expression algébrique écrite explicitement avec des + , × , des entiers écrits en binaire, et les variables Y_i , alors l'évaluation de $Expr(\alpha_1,...,\alpha_m)$ dans $\mathbf{A_P}$ est en temps uniformément polynomial par rapport à **d** et la taille des entrées, c.-à-d. encore par rapport à $\Sigma|\alpha_i|_1$, **lg(Expr)** , **d** et **lg(P)**

b) En particulier si $R \in \mathbb{Z}[Y_1,...,Y_m]$, $d_i = d_{X_i}(R)$, $d_R = d_1 + ... + d_m$, n_R = nombre de coefficients non nuls de R, on a :

$$\boxed{|R(\alpha_1,...,\alpha_k)|_1 \le (m_P + \sup(|\alpha_j|_1))\, d_R + |R(Y_1,...,Y_m)|_\infty + \lg(n_R)} \quad (7)$$

preuve> *a)* Le nombre d'opérations élémentaires de $\mathbf{A_P}$ est majoré par **lg(Expr)** . Le résultat final et chaque résultat intermédiaire sont convenablement majorés en appliquant (1) et (5) , chaque opération élémentaire de $\mathbf{A_P}$ est donc en temps convenablement majoré.

b) Si $c X_1^{r_1} ... X_m^{r_m}$ est un monôme de R , la majoration (6) donne :

$$|c\,\alpha_1^{r_1} ... \alpha_m^{r_m}|_1 \le |c|_\infty + (r_1+...+r_m-1)\, m_P + r_1|\alpha_1|_1 + ... + r_m|\alpha_m|_1$$
$$\le |R|_\infty + (d_R - 1)\, m_P + d_R \sup(|\alpha_j|_1)$$

et on conclut par l'inégalité (2) ❑

Remarque: La méthode qui consisterait à évaluer l'expression dans $\mathbb{Z}[X_1,X_2,...,X_k]$ et à la réduire ensuite modulo <P> ne permet pas d'obtenir une majoration en temps uniformément polynomial, à cause du trop grand nombre de monômes qui apparaissent dans l'expression avant sa réduction modulo <P> . Ceci complique nettement la tâche pour certains calculs à venir (les calculs de déterminants notamment) parce que l'anneau $\mathbf{A_P}$, contrairement à $\mathbb{Z}[X_1,X_2,...,X_k]$, n'est pas intègre.

Proposition B.b2 :

Soit $R \in \mathbb{Z}[X_1,X_2,...,X_k]$ de degré r_i en X_i , et soit

$$r = \sup_{j \in \{1,2,...k\}} (\mathrm{Ent}(r_j/(d_j - 1))) \qquad r = \prod_{j=1}^{k} r_j$$

a) On note $\lambda_1, \lambda_2, ..., \lambda_k$ les variables $X_1, X_2, ..., X_k$ vues comme éléments de $\mathbf{A_P}$. Alors, l'évaluation de $R(\lambda_1,...,\lambda_k)$ dans $\mathbf{A_P}$ est en temps

uniformément polynomial par rapport à $\mathbf{d}$, $\mathbf{r}$ et la taille des entrées, c.-à-d. encore par rapport à $|\ R\ |_1$, $\mathbf{r}$, $\mathbf{d}$ et $\mathbf{lg(P)}$

b) On a la majoration

$$\boxed{|\ R(\lambda_1,...,\lambda_k)\ |_1 \leq\ r\,(1 + m_P) + |\ R(X_1,X_2,...,X_k)\ |_\infty + \mathbf{lg}(r_1...r_k)} \qquad (8)$$

preuve> Montrons la majoration (8).
On considère un monôme $c.X_1^{s1}.X_2^{s2}....X_k^{sk}$ de R , on l'écrit sous forme d'un produit de facteurs qui sont des monômes d'exposants inférieurs ou égal à $(d_1\text{-}1,...,d_k\text{-}1)$, le premier de ces facteurs a sa $|\ |_1$ majorée par $|\ c\ |_1 \leq |\ R\ |_\infty$ les autres par 1 . Le nombre des facteurs est $r + 1$. On conclut par (6) que :

$$|\ c.X_1^{s1}.X_2^{s2}....X_k^{sk}\ |_1 \leq r\ m_P + r + |\ R(X_1,X_2,...,X_k)\ |_\infty$$

Enfin, il y a au plus $r_1...r_k$ monômes à ajouter.
La majoration du temps de calcul est claire. Elle peut être sensiblement améliorée si le polynôme R est creux, puisqu'il y a peu d'addition de monômes qui interviennent. ❑

Majoration de la taille dans $\mathbb{C}_{alg}^{\ k}$ des solutions d'un système normalisé d'équations algébriques emboitées

Proposition B.b3 :

Le calcul du polynôme minimum dans $\mathbb{Z}[X]$ d'un élément α de $\mathbf{A_P}$ est en temps uniformément polynomial par rapport à $|\ \alpha\ |_1$, $\mathbf{d}$ et $\mathbf{lg(P)}$.

preuve> On calcule la liste $[\alpha^i]_{i=0,...,d-1}$ dans $\mathbf{A_P}$ (proposition précédente), il reste à établir la première relation de dépendance $\mathbb{Q}$-linéaire entre ces vecteurs, par exemple en triangulant à la Bareiss dans $\mathbb{Z}$ la matrice $\mathbf{d} \times \mathbf{d}$ dont les colonnes sont les α^i écrits sur la base canonique de $\mathbf{A_P}$ convenablement ordonnée. ❑

Remarque : Un élément α de $\mathbf{A_P}$ est inversible (dans $\mathbf{A_P} \otimes \mathbb{Q}$) si et seulement si il est non diviseur de 0 , si et seulement si son polynôme minimum T vérifie $T(0) \neq 0$. *Il y a donc un test d'inversibilité dans* $\mathbf{A_P} \otimes \mathbb{Q}$ *en temps uniformément polynomial par rapport à* $|\ \alpha\ |_1$, $\mathbf{d}$ *et* $\mathbf{lg(P)}$.
De plus, lorsque α est inversible, $T(0)\ \alpha^{-1}$ s'exprime comme polynôme en α de degré $< d(T)$ (avec les coefficients de T en ordre inverse) et peut donc lui aussi être calculé en temps uniformément polynomial.
Une autre méthode consiste à regarder l'équation en β ($\beta \in \mathbf{A_P} \otimes \mathbb{Q}$) : $\alpha\,\beta = 1$, comme un système de $\mathbf{d}$ équations à $\mathbf{d}$ inconnues dans $\mathbb{Q}$ (les coefficients de β sur la base canonique de $\mathbf{A_P}$). Ce système d'équations est facile à écrire une fois qu'on a construit la table de multiplication de $\mathbf{A_P}$.

Proposition B.b4 :

La taille de toute solution dans $\mathbb{R}_{alg}^{\ k}$ ou $\mathbb{C}_{alg}^{\ k}$ d'un système normalisé d'équations algébriques emboitées défini par la liste $\mathbf{P}$ est uniformément majorable par un polynôme en $\mathbf{d}$ et $\mathbf{lg(P)}$.

preuve> Soit ξ_1 , ..., ξ_k une solution du système. On applique la proposition précédente à ξ_1 , ..., ξ_k vus comme éléments de $\mathbf{A_P}$. Dans le cas réel on termine en appliquant la proposition A.a4 . La proposition analogue s'applique de manière immédiate dans le cas complexe. ❑

Remarque :
On peut obtenir immédiatement le résultat suivant (qui sera amélioré par la suite)

Le calcul de toutes les solutions dans $\mathbb{R}_{alg}^{\,k}$ d'un système normalisé d'équations algébriques emboitées défini par la liste $\mathbf{P}$ peut être effectué en temps uniformément polynomial à partir de $\mathbf{d}^k$ et $\mathbf{lg(P)}$.

preuve > notons λ_i la valeur de X_i dans $\mathbb{A}_{\mathbf{P}}$. Pour $i = 1,...,k$, on peut calculer en temps uniformément polynomial les polynômes minimaux T_i des λ_i dans $\mathbb{A}_{\mathbf{P}}$, puis, en appliquant le théorème A.c1, toutes les racines de ces polynômes T_i. Il s'agit de tester ensuite chaque k-uple $\xi_1, ..., \xi_k$ pour savoir s'il est une solution du système emboité $\mathbf{P}$.

Pour cela considérons le système $T := [T_1(X_1),T_2(X_2),...,T_k(X_k)]$, pour lequel $dg(\mathbf{T}) = d_1^{\,k}.d_2^{\,k-1}...d_k$. Le polynôme $P_j(X_1,X_2,...,X_j)$ définit un élément π_j de $\mathbb{A}_{\mathbf{T}}$ dont on peut calculer le polynôme minimum S_j. Le réel algébrique $P_j(\xi_1,\xi_2,...,\xi_j)$ est une racine de S_j, et on sait rapidement en calculer une bonne approximation rationnelle.

Or une racine non nulle de S_j est minorée par $\dfrac{|\,c_h\,|}{|\,c_h\,| + \mathbf{sup}(|\,c_i\,|)}$ où c_h est le coefficient non nul de degré minimum de S_j. ❑

La même preuve donne le résultat suivant au niveau de $\mathbb{R}_{alg}$:

Soient $\xi_1, \xi_2, ..., \xi_n$ *des éléments de* $\mathbb{R}_{alg}$ *racines de polynômes* $Q_1, ..., Q_n$ *de* $\mathbb{Z}[X]$ *et de degrés* $d_1, ..., d_n$. *Alors les racines réelles du polynôme* $X^n + \xi_1 X^{n-1} + ... + \xi_n$ *peuvent être calculées comme éléments de* $\mathbb{R}_{alg}$ *en temps uniformément polynomial par rapport à* $d_1...d_n$ *et à* $\Sigma\,|\,Q_i\,|_1$

Produit d'une liste de matrices à coefficients dans $\mathbb{A}_P$

Proposition B.b5 :

a) Soient Δ et Δ' deux matrices de dimensions $n \times p$ et $p \times q$. On a la majoration

$$\boxed{|\,\Delta \times \Delta'\,|_1 \leq m_{\mathbf{P}} + |\,\Delta\,|_1 + |\,\Delta'\,|_1 + \mathbf{lg}(n) + \mathbf{lg}(p) + \mathbf{lg}(q)} \qquad (9)$$

b) Soit $\Gamma = [\Gamma_i]_{i=1,...,m}$ une liste de matrices à coefficients dans $\mathbb{A}_{\mathbf{P}}$, de dimensions adéquates pour qu'on puisse calculer le produit $\prod \Gamma_i$. Alors le calcul dans $\mathbb{A}_{\mathbf{P}}$ du produit $\prod \Gamma_i$ peut être effectué en temps uniformément polynomial par rapport à $\mathbf{d}$, $\dim(\Gamma)$ et la taille des entrées, c.-à-d. encore par rapport à $|\,\Gamma\,|_1$, $\dim(\Gamma)$, $\mathbf{d}$ et $\mathbf{lg(P)}$.

preuve> *Le produit de 2 matrices tout d'abord* (Δ et Δ' de dimensions $n \times p$ et $p \times q$) : le nombre d'opérations élémentaires dans $\mathbb{A}_{\mathbf{P}}$ est polynomialement majoré à partir n, p, q. De plus les inégalités (2) et (6) montrent que la taille des résultats intermédiaires est convenablement contrôlée et donnent pour chaque coefficient $\gamma_{i,j}$ du produit $\Delta \times \Delta'$ le majoration : $|\,\gamma_{i,j}\,|_1 \leq m_{\mathbf{P}} + |\,\Delta\,|_1 + |\,\Delta'\,|_1 + \mathbf{lg}(p)$ d'où on déduit immédiatement (9)

Pour le produit de m matrices : l'inégalité (9) montre que la taille des matrices intermédiares est bien contrôlée. ❑

Calculs de déterminants dans $\mathbb{A}_P$

Théorème B.b6 :

Soit Γ une matrice carrée à coefficients dans $\mathbb{A}_{\mathbf{P}}$, de dimension $m \times m$. Alors le calcul dans $\mathbb{A}_{\mathbf{P}}$ du déterminant de Γ est en temps uniformément polynomial par rapport à $\mathbf{d}$, m et la taille des entrées, c.-à-d. encore par rapport à $|\,\Gamma\,|_1$, m, $\mathbf{d}$ et $\mathbf{lg(P)}$.

preuve> On pourrait songer à utiliser la méthode de Bareiss, mais il y a un risque qu'à une certaine étape tous les coefficients "candidats pivots" soient diviseurs de 0 sans que le déterminant soit nul. Par contre, la méthode de Leverrier, vue la proposition B.b5 , fonctionne correctement en temps uniformément majoré. Même démonstration que pour le Théorème B.b1 dans **[Lom1]**. La méthode de Fadeev peut également être utilisée[1]. ❑

On notera qu'il est également possible d'utiliser la méthode de Samuelson (cf. **[Sam]** et **[Ber]**) pour calculer les déterminants puisque tous les résultats intermédiaires sont convenablement majorés en taille. L'avantage est que cette méthode peut se généraliser en caractéristique p , en particulier si on veut travailler dans la clôture algébrique d'un corps fini **F** ou dans la clôture algébrique du corps des fractions rationnelles correspondant **F**(X).

c) Systèmes d'équations en cascade, après une levée de l'ambiguité à la Newton

Le cadre de travail

Si $(\mathbf{P},[x_1,...,x_k])$ est une présentation dans $\mathbb{C}_{sae,N}$ de la liste $[\xi_1,\xi_2,...,\xi_k]$ nous cherchons à travailler dans l'anneau $\mathbf{A}_\xi$ quotient de $\mathbf{A}_\mathbf{P}$. Si $\alpha \in \mathbf{A}_\mathbf{P}$ nous noterons α_ξ l'élément correspondant de $\mathbf{A}_\xi$ et nous dirons que *le triplet* $(\mathbf{P},[x_1,...,x_k],\alpha)$ *est une présentation de l'entier algébrique* α_ξ *dans* $\mathbb{C}_{sae,N}$.
Nous disons que $\mathbf{d} = d_1...d_k$ est *le degré a priori* de l'entier algébrique α_ξ .
Nous notons $\mathbf{lg}(\alpha_\xi)$ la taille de $(\mathbf{P},[x_1,...,x_k],\alpha)$ (**P** et α peuvent être donnés en présentation creuse).

Les remarques qui suivent les propositions B.b3 et B.b4 montrent (à très peu près) qu'on pourrait systématiquement "désemboiter" les systèmes d'équations algébriques emboitées et garder des bornes de complexité "en temps uniformément polynomial par rapport à **d** et $\mathbf{lg}(\alpha_\xi)$". Le but est cependant justement de désemboiter le moins possible en espérant que la complexité effective soit plus faible (ce qui serait à peu près exclu si on désemboitait systématiquement), c'est en tout cas là la philosophie de D5 .

Précisions concernant les conditions de convergence du processus de Newton

Une étude particulièrement détaillée des conditions de convergence du processus de Newton est donnée dans **[Ost]** notamment chap 38 → 42 .
Nous nous en tiendrons à des conditions plus classiques quoique moins fines données dans **[DM]** chap XIII § 3, 4, 5, 6 .
On considère un système réel de n équations (algébriques ou transcendantes) à n inconnues $f_i(z_1,...,z_n) = 0$ (i = 1,..., n) , où les f_i sont 2 fois continument dérivables. Nous notons encore **f** l'application de U (ouvert de $\mathbb{R}^n$) vers $\mathbb{R}^n$ définie par les f_i .
Le processus de Newton démarre en un n-uple $\mathbf{x} = (x_1,...,x_n)$ tel que la matrice jacobienne

---

[1] La méthode de Fadeev est une version améliorée de la méthode de Leverrier. Comme l'anneau $\mathbf{A}_\mathbf{P}$ est traité à travers une représentation sans ambiguité, la condition de $\mathbf{P}$-réductibilité exigée dans **[Lom1]** pour une majoration correcte de la taille des objets manipulés pendant l'exécution de l'algorithme de Fadeev est automatiquement vérifiée.

de $\mathbf{f}$ soit inversible en $\mathbf{x}$. On suppose que l'ouvert U contient la boule fermée $\overline{B_\rho}(\mathbf{x})$ de centre $\mathbf{x}$ et de rayon ρ . On note Γ_0 la matrice inverse de la matrice jacobienne de $\mathbf{f}$ en $\mathbf{x}$. On choisit pour norme dans $\mathbb{R}^n$, $\| \mathbf{z} \| := \mathbf{sup}(| z_i |)$. On utilise pour les matrices la norme correspondante, plus précisément :
Si a_{ij} sont les coefficients de Γ_0 , on note $\| \Gamma_0 \| := \sup_i (\sum_j | a_{ij} |)$.

On suppose que les majorations suivantes sont vérifiées :

$$\| \Gamma_0 \| \leq A_0$$

$$\| \Gamma_0 \mathbf{f}(\mathbf{x}) \| \leq B_0 \leq \rho/2$$

$$\sum_{k=1,\ldots,n} \left| \frac{\partial^2 f_i(z)}{\partial z_j \, \partial z_k} \right| \leq C \quad \text{pour } z \in \overline{B_\rho}(\mathbf{x}),\ i,j \in \{1,\ldots n\}$$

$$2\, n\, A_0\, B_0\, C = \mu_0 < 1$$

Alors on est assuré que le processus itératif converge vers un point ξ de la boule $\overline{B_\rho}(\mathbf{x})$ qui est l'unique solution de $\mathbf{f}(\mathbf{x}) = 0$ dans cette boule. En fait, si $\mathbf{x}^{(p)}$ est le p-ème itéré, on a $\| \xi - \mathbf{x}^{(p)} \| \leq (1/2^{p-1}\ \mu_0^{2^p-1}\ B_0$. En outre si $2B_0/\mu_0 \leq \rho$, alors tout point $\mathbf{x}'$ de la boule de centre $\mathbf{x}$ et de rayon $(1 - \mu_0)\, B_0 / 2\, \mu_0$ peut être choisi comme début du processus itératif.

Pour le cas qui nous intéresse, les f_i sont les polynômes de la liste $\mathbf{P}$. La matrice jacobienne de $\mathbf{f}$ est triangulaire, et son déterminant, calculé au point $\mathbf{x}$ est égal à :

$$\frac{\partial P_1(x_1)}{\partial X_1} \; \frac{\partial P_2(x_1,x_2)}{\partial X_2} \; \ldots \; \frac{\partial P_k(x_1,\ldots,x_k)}{\partial X_k}$$

Dans les majorations désirées, la plus difficile à contrôler est a priori celle de $\| \Gamma_0 \|$ or les coefficients de Γ_0 sont égaux à des cofacteurs de la matrice jacobienne divisés par le déterminant. Une *minoration* contrôlée des dérivées partielles ci-dessus est donc la clef du problème.

Dans le cas d'un système complexe de n équations à n inconnues dans $\mathbb{C}$ on obtient des résultats tout à fait semblables: le système peut d'ailleurs être traité en le considérant comme un système de 2n équations réelles à 2n inconnues réelles.

Majorations polynomiales pour les calculs dans $\mathbb{C}_{sae,N}$

Conséquences des majorations dans $\mathbb{A}_P$

Tous les calculs dans $\mathbb{A}_P$ peuvent être considérés comme des calculs dans $\mathbb{A}_\xi$ et les majorations obtenues dans $\mathbb{A}_P$ sont ipso facto des majorations pour les calculs dans $\mathbb{A}_\xi$. Bien que nous n'ayions pas encore les moyens de montrer la calculabilité en temps convenable des solutions d'un système normalisé d'équations algébriques emboitées, nous avons la possibilité de majorer convenablement la taille des solutions, comme conséquence de la proposition B.b4 :

Proposition B.c1 :

Il existe une majoration polynomiale uniforme en fonction de $\mathbf{d}$ et $\mathbf{lg(P)}$ pour la taille de $[x_1,x_2,\ldots,x_k]$ où $(\mathbf{P},[x_1,\ldots,x_k])$ est une présentation dans $\mathbb{C}_{sae,N}$

de $[\xi_1,\xi_2,...,\xi_k]$ solution complexe *simple* du système normalisé d'équations algébriques emboitées défini par la liste **P** .

preuve> Dans toute cette preuve, nous dirons "convenable" pour "polynomiale uniforme en fonction de **d** et **lg(P)**". Nous donnons la preuve pour le cas d'une solution réelle. L'adaptation au cas complexe ne présente pas de difficulté.
D'après la proposition B.b4 la taille de $[\xi_1,\xi_2,...,\xi_k]$ vus comme éléments de $\mathbb{R}_{alg}$ est correctement contrôlée. Si ξ_i est représenté par (T_i,a_i,b_i) dans $\mathbb{R}_{alg}$, on peut se situer a priori sur le pavé produit des $[\ a_i\ ,\ b_i\ ]$. On calcule une majoration convenable des dérivées partielles secondes sur ce pavé. Cela fournit en particulier une majoration du taux de variation des dérivées partielles premières intervenant dans le déterminant de la jacobienne. Par ailleurs posons $\alpha_j := \partial P_j(\xi_1,...,\xi_j) \,/\, \partial X_j$. La taille de α_j dans $\mathbb{A}_{\mathbf{P}}$ est simplement celle du polynôme $\partial P_j(X_1,...,X_j) \,/\, \partial X_j$. On a donc une majoration convenable des coefficients du polynôme minimum de α_j dans $\mathbb{A}_{\mathbf{P}}$, ce qui fournit une minoration convenable de $|\ \alpha_j\ |$ (qui est par hypothèse non nul). En couplant ce renseignement avec la majoration du taux de variation de la dérivée partielle, on aura alors une minoration convenable du déterminant de la jacobienne sur un nouveau pavé "pas trop minuscule" autour de ξ et donc une majoration convenable de $\|\ \Gamma_0\ \|$ sur ce pavé, d'où on déduit une majoration convenable de l'écart entre **x** et $[\xi_1,\xi_2,...,\xi_k]$ pour que le processus de Newton converge, ce qui majore convenablement la taille de **x** . ❑

Théorème B.c2 :

Soit $(\mathbf{P},[x_1,...,x_k],\alpha)$ une présentation de l'entier algébrique α_ξ dans $\mathbb{C}_{sae,N}$.

a) Le calcul d'une approximation de α_ξ avec la précision $1/2^n$ est en temps uniformément polynomial par rapport à **d** et la taille des entrées, c.-à-d. encore par rapport à n , $|\ \alpha\ |_1$, **d** et **lg(P)**[1] .

b) Le calcul de α_ξ dans $\mathbb{C}_{alg}$ est en temps uniformément polynomial par rapport à $|\ \alpha\ |_1$, **d** et **lg(P)** .

preuve> *pour le a)* on utilise la méthode de Newton pour calculer ξ avec une approximation arbitraire, en ne conservant à chaque étape que la partie significative du développement en base 2 du rationnel obtenu, ce qui permet de contrôler la taille des calculs intermédiaires[2].

pour le b) cela résulte du a) et du fait qu'on sait calculer en temps convenable un polynôme non nul de $\mathbb{Z}[X]$ annulant α_ξ : on termine en appliquant la proposition A.d3 . ❑

Signe d'un élément de $\mathbb{A}_\xi$ et calcul de son inverse

Proposition B.c3 :

Soit $(\mathbf{P},[x_1,...,x_k],\alpha)$ une présentation de l'entier réel algébrique α_ξ dans $\mathbb{C}_{sae,N}$.

Alors le test d'égalité à 0 pour α_ξ , le calcul du signe de α_ξ dans le cas réel et, lorsque $\alpha_\xi \neq 0$, le calcul de l'inverse de α_ξ (dans $\mathbb{A}_\xi \otimes \mathbb{Q}$) est en temps uniformément polynomial par rapport à **d** et la taille des entrées, c.-à-d. encore par rapport à $|\ \alpha\ |_1$, **d** et **lg(P)** .

---

[1] En langage plus imagé, on pourrait dire que l'évaluation $\mathbb{A}_\xi \to \mathbb{C}$ est en temps uniformément polynomial par rapport à **d** et la taille des entrées (cf la définition A.a5 dans un contexte voisin.

[2] Ceci mériterait un développement détaillé à soi tout seul.

preuve> *pour le signe ou le test d'égalité à 0* on majore les coefficients du polynôme minimum de α . Si $\alpha_\xi \neq 0$, on a donc une majoration des coefficients d'un polynôme de $\mathbb{Z}[X]$ annulant $1/\alpha_\xi$, ce qui nous donne la précision avec laquelle il faut calculer α_ξ pour être assuré de son signe. En pratique, on a intérêt à mener en parallèle le calcul de plus en plus approché de α_ξ d'une part, et celui de la précision souhaitée d'autre part, le premier calcul pouvant aboutir à un résultat effectif bien avant la limite de précision imposé.
Une autre méthode pour déterminer un degré de précision suffisant (et cependant pas trop grand) pour connaître le signe de α_ξ est la suivante :

- on majore les modules des conjugués des ξ_i (ce qui peut être fait une fois pour toutes
- on en déduit une majoration m des modules des conjugués de α_ξ
- comme α_ξ est un entier algébrique de degré $\leq \mathbf{d}$ on a :

$$\alpha_\xi \neq 0 \Rightarrow |\alpha_\xi| > 1/m^{d-1}$$

pour l'inverse il semble difficile de se passer en général du calcul du polynôme minimum P_α de α (sauf si on sait par un argument quelconque que α est inversible dans $\mathcal{A}_\mathbf{P} \otimes \mathbb{Q}$). A partir de P_α on obtient (en le divisant par une puissance de X) un polynôme $Q \in \mathbb{Z}[X]$ tel que $Q(0) \neq 0$ et $Q(\alpha_\xi) = 0$, ce qui permet alors de calculer l'inverse de α_ξ sous forme β_ξ / n , où $\beta \in \mathcal{A}_\mathbf{P}$. ❑

Calculs de déterminants dans $\mathcal{A}_\xi$

Pour le calcul du déterminant d'une matrice à coefficients dans $\mathcal{A}_\xi$ on peut donc hésiter entre, d'une part, la méthode de Leverrier (ou celle de Fadeev) dans $\mathcal{A}_\mathbf{P}$ et, d'autre part, la méthode de Bareiss dans $\mathcal{A}_\xi$.
Cependant, il faut noter que la méthode de Bareiss est a priori peu sûre : lorsque l'homomorphisme d'évaluation $\mathcal{A}_\mathbf{P} \rightarrow \mathcal{A}_\xi$ n'est pas injectif (ce qui doit être considéré comme le cas général), un même élément de $\mathcal{A}_\xi$ peut être représenté par des éléments de $\mathcal{A}_\mathbf{P}$ de taille arbitrairement grande. Quand la méthode de Bareiss exige une division exacte dans $\mathcal{A}_\xi$ avec un dénominateur non inversible dans $\mathcal{A}_\mathbf{P} \otimes \mathbb{Q}$, il faudrait donc préciser quel algorithme de division exacte dans $\mathcal{A}_\xi$ on utilise, et démontrer qu'aucune explosion de la taille des objets manipulés n'en résulte.
En outre, le calcul du déterminant directement dans $\mathcal{A}_\mathbf{P}$ présente l'avantage de pouvoir être spécialisé pour toute solution du système d'équations emboitées considéré.

Relation de Bezout complète entre deux polynômes de $\mathcal{A}_\xi[X]$

Proposition B.c4 :

Soit $(\mathbf{P},[x_1,\ldots,x_k])$ une présentation dans $\mathbb{C}_{sae,N}$ de la solution $\xi_1,\xi_2,\ldots,\xi_k$ du système d'équations algébriques emboitées $\mathbf{P}$. Soient Q et R 2 polynômes de $\mathcal{A}_\xi[X]$. On désire calculer G , U , V , Q_1 , R_1 dans $(\mathcal{A}_\xi \otimes \mathbb{Q})[X]$ qui vérifient :

$$U\,Q + V\,R = G\,,\quad Q_1\,G = Q\,,\quad R_1\,G = R$$

Ce calcul est en temps uniformément polynomial par rapport à $\mathbf{d}$ et la taille des entrées, c.-à-d. plus précisément par rapport à $|Q|_1$, $|R|_1$, deg(Q) , deg(R) , $\mathbf{d}$ et $\mathbf{lg(P)}$.

preuve> C'est de l'algèbre linéaire dans le corps $\mathcal{A}_\xi \otimes \mathbb{Q}$ ❑

Remarque : En fait, si on prend pour G le dernier polynôme sous-résultant non nul de Q et R , les polynômes U et V sont à coefficients dans $\mathcal{A}_\xi$. Cependant, on n'arrivera pas en

général à avoir tous les polynômes simultanément à coefficients dans $\mathcal{A}_\xi$.
En outre, lorsque Q est unitaire, et qu'on prend pour Q_1 *le* polynôme unitaire (il n'y en a qu'un possible, puisqu'il est défini à une constante multiplicative près dans $\mathcal{A}_\xi \otimes \mathbb{Q}$), on sait que les coefficients de Q_1 sont dans la clôture intégrale de $\mathcal{A}_\xi$, mais justement en général $\mathcal{A}_\xi$ n'est pas intégralement clos.

Calculs dans la clôture intégrale de $\mathcal{A}_\xi$

Il serait donc intéressant de donner une bonne majoration explicite des dénominateurs possibles pour un élément de $\mathcal{A}_P \otimes \mathbb{Q}$ dont l'image dans $\mathcal{A}_\xi \otimes \mathbb{Q}$ est dans la clôture intégrale de $\mathcal{A}_\xi$. Notons $\mathfrak{B}_\xi$ cette clôture intégrale.
Ceci donnerait une version améliorée de $\mathbb{C}_{sae,N}$ où on représenterait tous les éléments de $\mathfrak{B}_\xi$ plutôt que les seuls éléments de $\mathcal{A}_\xi$, tout en gardant le même genre de majorations pour les temps calculs :
En effet, si n_P peut servir de dénominateur commun à tous les éléments de $\mathcal{A}_P \otimes \mathbb{Q}$ dont l'image est dans $\mathfrak{B}_\xi$, alors si $\alpha/n_P = \beta/n$ irréductible, avec $\alpha_\xi/n_P \in \mathfrak{B}_\xi$, on peut noter

$| \alpha_\xi/n_P |_{\mathfrak{B}_\xi} = | \beta_\xi/n |_{\mathfrak{B}_\xi} = | \alpha |_1$ et on obtient :

$| \beta_\xi/n |_{\mathfrak{B}_\xi} \leq | \beta |_1 + | n_P |_1$

$| \beta_\xi/n + \gamma_\xi/m |_{\mathfrak{B}_\xi} \leq \sup(| \beta_\xi/n |_{\mathfrak{B}_\xi} , | \gamma_\xi/m |_{\mathfrak{B}_\xi}) + 1$

$| \beta_\xi/n \,.\, \gamma_\xi/m |_{\mathfrak{B}_\xi} \leq | \beta_\xi/n |_{\mathfrak{B}_\xi} + | \gamma_\xi/m |_{\mathfrak{B}_\xi} + m_P$ (cf § b pour m_P)

Recherche des solutions simples d'un système d'équations algébriques emboitées

Nous traitons tout d'abord le cas des racines simples, qui est naturel dans notre cadre de travail.

Proposition B.c5 :

Soit $(P,[x_1,...,x_k])$ une présentation dans $\mathbb{C}_{sae,N}$ de la solution simple $[\xi_1,\xi_2,...,\xi_k]$ du système d'équations algébriques emboitées P .

Soit Q un polynôme unitaire de $\mathcal{A}_\xi[X]$.

On désire calculer les racines simples de Q sous la forme suivante:

si λ est une de ces racines, alors $[\xi_1, \xi_2, ..., \xi_k, \lambda]$ est représenté dans $\mathbb{C}_{sae,N}$ par $(R,[y_1,...,y_k,y_{k+1}])$ où R est la liste P prolongée par Q

Ce calcul peut être réalisé en temps uniformément polynomial par rapport à d et la taille des entrées, c.-à-d. plus précisément par rapport à $| Q |_1$, deg(Q) , d et lg(P) .

preuve> On utilise l'algorithme de Schönage ou celui de Victor Pan (cf. [**Sch**] ou [**Pan**]) pour calculer des approximations arbitraires des racines de Q . Cet algorithme ne nécessite que la connaissance des valeurs approchées des coefficients de Q . Ces évaluations sont en temps convenablement contrôlé grâce au théorème B.c2 a) . Par ailleurs, dans la mesure où on ne s'intéresse qu'aux racines simples, la précision requise est convenablement contrôlée grâce à la proposition B.c1 . Plus précisément, soit T le polynôme minimum de $\lambda_{k+1} \in \mathcal{A}_R$. Le polynôme T est calculable en temps convenable. On en déduit une minoration convenable, soit ε , pour l'écart entre 2 racines distinctes de T , et donc aussi entre 2 racines distinctes de $Q(\xi_1,\xi_2,...,\xi_k,X)$. Si l'algorithme de Victor Pan situe 2 ou plusieurs racines de ce polynôme à une distance inférieure à ε on est assuré qu'il s'agit d'une racine multiple. ❑

Remarques :

1) En cas de racine multiple, cette racine multiple peut alors être explicitée sous la forme

suivante: c'est l'unique racine du polynôme située dans un certain disque de centre $a + \sqrt{-1}\ b$ et de rayon r, où a, b, r sont des rationnels calculables en temps convenable.

2) Dans le cas réel, on peut utiliser plusieurs autres méthodes pour déterminer les racines simples de $\mathbf{P}$:

a) la méthode des tableaux de signes approchés (cf § C.b) pour trouver des intervalles contenant chacun exactement une racine de Q en utilisant uniquement des évaluations approchées de Q et de ses dérivées, et assez petits pour que Newton fonctionne à partir d'un bord de l'intervalle.

b) une méthode à la Sturm améliorée genre Sturm-Habicht (cf **[GLRR]** ou **[Lom2]**) : la taille des polynômes sous-résultants est bien contrôlée puisque les coefficients de ces polynômes sont des déterminants. Noter l'intérêt qu'il y a à calculer ces coefficients dans $\mathbf{A}_{\mathbf{P}}$, puisque le même calcul servira pour toutes les solutions réelles de $\mathbf{P}$: ce n'est qu'au moment de l'évaluation des signes que le calcul se particularise.

c) ou la méthode élémentaire (vrais tableaux de signes, cf § A.c). Là encore, on aura des calculs de résultants lors des tests de signes, mais on n'a plus l'avantage signalé en b) d'un "précalcul" commun à toutes les solutions réelles de $\mathbf{P}$.

Les solutions b) et c) sont semble-t-il beaucoup plus coûteuses que la solution a), dans la mesure où ces 2 méthodes utilisent systématiquement des calculs de déterminants et des évaluations exactes de signes, alors que la méthode a) se contente de calculs d'évaluations approchées du polynôme et de ses dérivées.

Théorème B.c6 :

Soit $\mathbf{P}$ un système normalisé d'équations algébriques emboitées.

On désire calculer les solutions simples du système sous la forme présentée dans $\mathbb{C}_{sae,N}$:

plus précisément toute solution $\xi_1, \xi_2, ..., \xi_k$ doit être explicitée sous forme $(P,[x_1,...,x_k])$.

Ce calcul est en temps uniformément polynomial par rapport à $\mathbf{d}$ et $\mathbf{lg(P)}$.

preuve> On applique la proposition B.c5 de manière itérative. Pour avoir une majoration du temps convenable, il suffit de montrer que la taille de tous les objets utilisés comme "entrées" lors des différentes applications de B.c5 est convenablement majorée. Il suffit pour cela de s'assurer que la taille de toutes les approximations rationnelles $[y_1,...,y_j]$ $(j \leq k)$ obtenues au cours du calcul est correctement maîtrisée, ce qui est donné par B.c1 . ❑

En combinant le résultat précédent et le théorème B.c2 b) on obtient :

Corollaire B.c7:

Soit $\mathbf{P}$ un système normalisé d'équations algébriques emboitées. Les solutions simples du système peuvent être calculées comme éléments de $\mathbb{C}_{alg}{}^{k}$ en temps uniformément polynomial par rapport à $\mathbf{d}$ et $\mathbf{lg(P)}$.

La recherche des solutions non simples

Nous discutons maintenant la question des racines multiples. Nous donnons tout d'abord l'analogue, beaucoup moins joli, de la proposition B.c5 .

Proposition B.c8 :

Soit $(\mathbf{P},[x_1,...,x_k])$ une présentation dans $\mathbb{C}_{sae,N}$ de la solution $\xi_1,\xi_2,...,\xi_k$ du système d'équations algébriques emboitées $\mathbf{P}$.

Soit Q un polynôme de $\mathbb{A}_\xi[X]$.

On désire calculer les racines de Q sous la forme suivante:

si λ est une de ces racines, alors on détermine un entier c, un diviseur Q_2 de $Q(X/c)$, unitaire et à coefficients dans $\mathbb{A}_\xi$, et des dyadiques $y_1, ..., y_k, y_{k+1}$ tels que:

$[\xi_1, \xi_2, ..., \xi_k, c\lambda]$ est représenté dans $\mathbb{C}_{sae,N}$ par $(\mathbf{R},[y_1,...,y_k,y_{k+1}])$ où $\mathbf{R}$ est la liste $\mathbf{P}$ prolongée par Q_2

Ce calcul est en temps uniformément polynomial par rapport à $\mathbf{d}$ et la taille des entrées, c.-à-d. encore par rapport à $|Q|_1$, $\mathbf{d}$ et $\mathbf{lg(P)}$.

preuve> On applique la proposition B.c4 pour calculer un polynôme Q_1 unitaire et à coefficients dans $\mathbb{A}_\xi \otimes \mathbb{Q}$ qui est la partie sans facteur carré de Q . On prend pour c le ppcm des dénominateurs. On fait le changment de variable $Y = c X$ et on obtient un polynôme unitaire Q_2 à coefficients dans $\mathbb{A}_\xi$ qu'on traite par la proposition B.c5 .❑

Nous pourrions maintenant appliquer B.c8 de manière itérative, *à condition de* fournir un argument de majoration pour la taille des listes de polynômes emboités qui apparaissent dans le processus itératif.

Il nous semble de toute manière préférable de continuer à travailler dans $\mathbb{A}_\mathbf{P}$ quitte à utiliser une autre caractérisation numérique pour les solutions non simples du système. Voici une proposition dans ce sens.

Une racine multiple de $P_{i+1}(\xi_1,\xi_2,...,\xi_i,X)$ peut être vue comme une racine simple de l'une des dérivées de P_{i+1} par rapport à X. Nous établissons donc tout d'abord la proposition analogue à la proposition B.c5 .

Proposition B.c9 :

Soit $(\mathbf{P},[x_1,...,x_k])$ une présentation dans $\mathbb{C}_{sae,N}$ de la solution simple $[\xi_1,\xi_2,...,\xi_k]$ du système d'équations algébriques emboîtées $\mathbf{P}$.

Soit Q un polynôme unitaire de $\mathbb{A}_\xi[X]$.

On désire calculer les racines multiples de Q sous la forme suivante:

si λ est une de ces racines, alors on détermine son ordre de multiplicité $i+1$ et $[\xi_1, \xi_2, ..., \xi_k, \lambda]$ est représenté dans $\mathbb{C}_{sae,N}$ par $(\mathbf{R},[y_1,...,y_k,y_{k+1}])$ où $\mathbf{R}$ est la liste $\mathbf{P}$ prolongée par $Q^{(i)}$

Ce calcul peut être réalisé en temps uniformément polynomial par rapport à $\mathbf{d}$ et la taille des entrées, c.-à-d. plus précisément par rapport à $|Q|_1$, $\deg(Q)$, $\mathbf{d}$ et $\mathbf{lg(P)}$.

preuve> on raisonne comme à la proposition B.c5, la conclusion est qu'on connaît la multiplicité de chacune des racines de Q , ce qui nous ramène au cas d'un racine simple de $Q^{(i)}$. ❑

Ceci justifie que nous étendions la présentation $\mathbb{C}_{sae,N}$, par exemple de la manière suivante :

Définition B.c10 :

Nous dirons qu'une liste $[\xi_1,\xi_2,...,\xi_k]$ est une solution d'ordre $s = [s_1,s_2,...,s_k]$ du sytème emboîté P si chaque ξ_i est racine d'ordre s_i de l'équation correspondante. Nous noterons $P^{(s)}$ l'application de $\mathbb{C}^k$ vers $\mathbb{C}^k$ définie par les $P_i^{(s_i)}$ (dérivée par rapport à X_i).

Nous dirons que $(P, [(x_1, s_1),...,(x_k,s_k)])$ *est une présentation de la solution* $[\xi_1,\xi_2,...,\xi_k]$ *de* P *dans* $\mathbb{C}_{sae,N}$ *(étendue)* si la méthode de Newton appliquée à $P^{(s)}$ et initialisée à $[x_1,...,x_k]$ converge vers $[\xi_1,\xi_2,...,\xi_k]$.

En outre si α est un élément de $\mathbb{A}_P$ nous dirons que $(P, [(x_1, s_1),...,(x_k,s_k)], \alpha)$ *est une présentation de l'entier algébrique* α_ξ *dans* $\mathbb{C}_{sae,N}$ *(étendue)* .

Avec cette extension de la présentation $\mathbb{C}_{sae,N}$, il n'est pas difficile de vérifier que tous les résultats du § d jusqu'à la proposition B.c9 restent valables. D'où finalement, avec les mêmes arguments que pour la preuve du théorème B.c6 :

Théorème B.c11 :

Soit P un système normalisé d'équations algébriques emboîtées.

On désire calculer toutes les solutions (simples ou multiples) du système dans la présentation $\mathbb{C}_{sae,N}$ étendue (définition d.8) :

plus précisément toute solution $\xi_1, \xi_2, ..., \xi_k$ doit être explicitée sous forme $(P, [(x_1, s_1),...,(x_k,s_k)])$.

Ce calcul est en temps uniformément polynomial par rapport à d et $lg(P)$.

Corollaire B.c12 :

Soit P un système normalisé d'équations algébriques emboîtées. Les solutions du système peuvent être calculées comme éléments de $\mathbb{C}_{alg}^k$ en temps uniformément polynomial par rapport à d et $lg(P)$.

C) METHODES APPROXIMATIVES

Introduction

L'étude faite en B c) a montré l'efficacité assez bonne des méthodes approximatives pour calculer avec des nombres algébriques réels ou complexes.

Nous examinons dans ce chapitre deux théorèmes "en temps polynomial" qui relèvent par leur nature même de méthodes approximatives. Ces méthodes sont indispensables chaque fois qu'on a à résoudre un problème dont les variables sont dans $\mathbb{R}$, $\mathbb{C}$, ou un espace de fonctions.

Le théorème fondamental de l'algèbre est de ceux-là.

Quand on passe à la recherche des racines réelles d'un polynôme à coefficients réels, une méthode classique comme la méthode de Sturm devient impraticable dans un contexte constructif pour la simple raison qu'il n'y a pas de test d'égalité à 0 pour un nombre réel "en général". L'affirmation classique selon laquelle on peut situer les racines réelles d'un polynôme à coefficients réels devient *fausse* d'un point de vue constructif. Il y a néanmoins un substitut constructif à cette affirmation: la possibilité de dresser un tableau de signes "approché" pour un tel polynôme (cf § b) pour plus de précision).

Dans les deux cas envisagés ci-dessus, il serait possible de donner une solution algorithmique en traitant le cas "discret" (variables dans $\mathbb{Q}$) par des méthodes discrètes, et en concluant par un théorème de perturbation effectif.

Il est cependant naturel, et, en pratique, plus efficace, de chercher un algorithme utilisant directement des méthodes approximatives.

Pour ce qui concerne le théorème fondamental de l'algèbre, l'algorithme de Victor Pan résout la question au mieux. Nous nous sommes contentés, pour l'essentiel, dans le § a) d'une discussion générale à propos de la solution en temps polynomial du théorème fondamental de l'algèbre. Les résultats dans le cas discret obtenus dans le chapitre A au moyen de la présentation naïve, donneraient une réponse positive au problème posé, sans recours à l'algorithme de Victor Pan, moyennant cependant un bon théorème de perturbation des racines d'un polynôme à coefficients dans $\mathbb{C}$ (par exemple le théorème d'Ostrowski).

Pour ce qui concerne la possibilité de dresser un tableau de signes approché pour un polynôme à coefficients réels, nous avons utilisé au § b) la méthode "la plus naïve qui soit", celle des tableaux de signes et de variations approchés des dérivées successives du polynôme en commençant par le plus bas degré. Nous obtenons malgré tout un temps de calcul assez honorable. Le défaut de l'algorithme est qu'il calcule "beaucoup" de tableaux de signes, plus que ce qui est strictement nécessaire. L'avantage est qu'il est par nature approximatif, et qu'il s'applique donc à toute fonction qui se laisse bien approcher (pour la norme uniforme) par des fonctions polynômes.

Ceci nous amène à faire une brève étude, au § c) , de la classe des fonctions "approchables en temps polynomial par des polynômes, pour la norme uniforme, sur un intervalle compact". Cette classe est en fait celle des fonctions Gevrey $\mathcal{P}$-calculables. Tous les calculs élémentaires dans cette classe de fonction s'avèrent être en temps polynomial, pour des raisons tout à fait

immédiates. Nous obtenons ainsi une amélioration des théorèmes de Ko-Friedman ([**KF1**], [**KF2**]) et Müller ([**Mü2**]) concernant les fonctions analytiques et $\mathcal{P}$-calculables, et une simplification de leurs preuves. Les théorèmes s'appliquent en fait aux fonctions Gevrey $\mathcal{P}$-calulables.
Nous concluons par quelques perspectives de travail dans le cadre ainsi tracé : la géométrie algébrique réelle exacte dans la clôture réelle de $\mathbb{Q}$ pourrait, selon nous, être avantageusement remplacée par une géométrie algébrique réelle approximative dans tous les problèmes appliqués. En fin de compte on se situerait alors inévitablement dans un cadre de géométrie "analytique approximative" ou "Gevrey approximative".

a) Le théorème fondamental de l'algèbre est en temps polynomial

Position du problème, une première solution

L'affirmation qui fait le titre du § a demande à être précisée.

Enoncé sous forme constructive, le théorème fondamental de l'algèbre dit qu'il existe une opération qui permet de décomposer un polynôme non nul de $\mathbb{C}[X]$ en un produit de facteurs linéaires (c.-à-d. de la forme aX+b). En 1882, Kronecker a donné une preuve constructive que les nombres algébriques complexes forment un corps algébriquement clos (cf [**Kro**]). En 1924 , Brouwer et de Loor donnent une preuve constructive du théorème fondamental de l'algèbre dans le cas d'un polynôme unitaire, d'où on peut déduire aisément le théorème fondamental de l'algèbre général (on trouve un x tel que $P(x) \neq 0$ puis on fait une homographie de la droite projective complexe qui envoie x à l'infini, ce qui nous ramène au cas unitaire). On trouve un exposé analogue particulièrement clair dans [**BB**] . Par ailleurs, une méthode "purement" algébrique est donnée en exercice dans [**MRR**] .

Insistons sur 2 points :
- l'opération qui réalise le théorème fondamental de l'algèbre traite en entrée un élément de $\mathbb{C}^{n+1} - \{0,...,0\}$ et donne en sortie n éléments de $\mathbb{C}^2 - \{0,0\}$: mais les seules opérations effectives connues traitant une entrée de ce genre la traitent toujours via ses approximations rationnelles. (en outre le fait que le n-uple est distinct de {0,...,0} doit être fourni explictement en entrée sous forme d'une approximation rationnelle qui le montre). Donc, il est indispensable qu'il y ait un "théorème de perturbation" affirmant qu'une faible variation des coefficients induit une faible variation des facteurs linéaires. Ce théorème de perturbation est nécessairement fourni en filigrane dans toute preuve constructive du théorème fondamental de l'algèbre.
- cependant, il est impossible de réaliser le théorème fondamental de l'algèbre au moyen d'une opération *extensionnelle* de $\mathbb{C}^{n+1} - \{0,...,0\}$ vers $(\mathbb{C}^2 - \{0,0\})^n$, car il n'y a pas moyen d'obtenir les racines chacune séparément comme fonction continue des coefficients du polynôme. C'est uniquement avec une sortie sous forme d'une liste "non ordonnée" de n éléments de $\mathbf{P}_1(\mathbb{C})$ (droite projective complexe ou sphère de Riemann) que l'on a une formulation vraiment agréable.

Pour ce qui concerne une version "en temps polynomial" du théorème fondamental de l'algèbre, le premier énoncé remonte à [**KF1**] , sous forme d'une conjecture. Peu de temps après, Schönage ([**Sch**]), signale que la conjecture est mal formulée mais très certainement vraie si bien reformulée. Une preuve est donnée dans [**Hoo**] . L'article de Victor Pan ([**Pan**]) fournit une preuve non dite "de fait" dans la mesure où les approximations des zéros

de f sont calculées en temps polynomial en utilisant uniquement des approximations convenablement maitrisées de valeurs de f .

Quant au fond, la preuve du théorème fondamental de l'algèbre en temps polynomial peut se ramener à ceci : *primo*, une preuve que la recherche des racines dans $\mathbf{P}_1(\mathbb{C})$ d'un polynôme à coefficients dans $\mathbb{Q}[\sqrt{-1}]$ est en temps polynomial ; *secundo* un théorème de perturbation effectif où le $\forall\ \varepsilon\ \exists\ \delta$ ... soit réalisable en temps polynomial, ce qui signifie qu'il existe un polynôme R à coefficients entiers positifs tel que : $\forall\ \varepsilon = 2^{-k}\ \exists\ \delta = 2^{-R(k)}$ ...
Il suffit donc de rechercher dans la littérature un bon théorème de perturbation effectif, ce qu'on trouve dans [**Ost**][1] .

Voyons maintenant l'énoncé de Ko et Friedman :

> Il existe une machine de Turing à oracle , M , qui , lorsqu'on lui donne en entrée 2 entiers d et n en unaire (d étant le degré du polynôme et n le degré de précision souhaité sur les racines) fournit en temps polynomial une liste de d éléments de $\mathbb{D}[\sqrt{-1}]$ (sous forme: 2n éléments de $\mathbb{D}$) qui approchent à 2^{-n} près les racines du polynôme. Les coefficients du polynôme sont fournis par l'oracle de la manière suivante : la machine "pose la question" *tel coefficient avec telle précision* ? (sous forme de 2 entiers en unaire: numéro du coefficient et degré de précision souhaité) et l'oracle répond en donnant 2 dyadiques (la partie réelle et la partie imaginaire) avec le bon nombre de digits après la virgule, le temps compté pour la réponse étant simplement le nombre de digits affichés.

Le problème est "mal formulé" dans la mesure où la machine doit disposer d'une entrée supplémentaire : un entier (en unaire) m tel que 2^m majore la valeur absolue des coefficients et tel que 2^{-m} minore la valeur absolue du coefficient dominant. On est alors immédiatement ramené au cas d'un polynôme unitaire avec une majoration des coefficients donnée en entrée. Le temps de calcul doit être polynomial par rapport à $m + d + n$. Modulo le fait que la recherche des racines complexes d'un polynôme de $\mathbb{Q}[\sqrt{-1}][X]$ est en temps polynomial (cf par exemple le théorème A.d2 couplé avec la proposition A.a6) il nous reste à vérifier que le théorème de perturbation d'Ostrowski est "polynomial".

Théorème de perturbation d'Ostrowski ([**Ost**] p 221)

Soient $f(z) := z^d + a_1 z^{d-1} + \ldots + a_d$, $g(z) := z^d + b_1 z^{d-1} + \ldots + b_d$

$$\gamma := 2 \sup_{j \in \{1,2,\ldots d\}} (|a_j|^{1/j}, |b_j|^{1/j})$$

$$\varepsilon := \left(\sum_{j=1}^{d} |a_j - b_j| \gamma^{d-j} \right)^{1/d}$$

Alors les zéros α_i de f et les zéros β_i de g peuvent être ordonnés de manière que pour tout i $|\alpha_i - \beta_i| \leq 2 d \varepsilon$

---

[1] On notera cependant qu'une preuve constructive du théorème fondamental de l'algèbre ne saurait être fournie par "une preuve non constructive dans le cas général + une preuve constructive pour le cas "discret" (coefficients dans $\mathbb{Q}[\sqrt{-1}]$) + un théorème de perturbation"

Si les $|a_i|$ et $|b_i|$ sont majorés par 2^m ($m \geq 0$) et si $|a_j - b_j| \leq 2^{-h}$ on obtient $2\,d\,\varepsilon \leq 2^r$ avec $r = -(h/d)+m+2+2\log_2(d)$ et pour avoir $r \leq -n$ il suffit de prendre $h := d\,(n + m + 2 + 2.\log_2(d))$. cqfd

Un deuxième énoncé

Nous pouvons considérer le théorème fondamental de l'algèbre comme fournissant une fonction uniformément continue de $\mathbf{P}_n(\mathbb{C})$ (espace projectif complexe de dimension n) vers $\mathbf{Sym}_n(\mathbf{P}_1(\mathbb{C}))$, où nous notons $\mathbf{Sym}_n(A)$, pour un espace métrique A , l'espace séparé complété de A^n muni de la métrique :

$d([x_1,...,x_n],[y_1,...,y_n]) = \mathbf{inf}\,(\Sigma_i\,|\,x_i - y_{\sigma(i)}\,|)$ où σ parcourt les permutations.

Pour avoir une version " en temps polynomial" du théorème fondamental de l'algèbre, il nous faut donner un sens naturel à "fonction uniformément continue $\mathcal{P}$-calculable d'une espace métrique compact vers un autre".

Tout ceci se situe naturellement dans le cadre des fonctions $\mathcal{P}$-calculables de [a , b] vers $\mathbb{R}$. Supposons que la fonction f soit M-lipschitzienne et qu'on veuille calculer un $z/2^n$ ($z \in \mathbb{Z}$) approchant f(x) avec la précision $1/2^n$. Il suffit pour cela de calculer un rationnel r approchant f(x) avec la précision $1/2^{n+1}$, et pour cela de calculer un rationnel r' approchant f(x') avec la précision $1/2^{n+2}$, où $|x' - x| \leq 1/(M.2^{n+2})$. Avec $M = 2^h$, cela donne la possibilité de choisir x' à une distance $< 1/2^{n+h+2}$ de x; donc, si x est donné comme rationnel, de choisir x' sous la forme $x''/2^{n+h+1}$ ($x'' \in \mathbb{Z}$) . Si x est donné comme dyadique, cela revient à garder n + h + 1 chiffres après la virgule. La machine qui calcule avec des dyadiques de longueur arbitraire a normalement son pointeur "en tête de x" (avant la virgule) au moment de lire x , et elle se contentera donc de lire la partie utile de x . De sorte qu'on peut poser comme raisonnable la définition suivante qui ne fait pas intervenir le module de Lipschitz en tant que tel et s'étend donc sans changement à une fonction continue arbitraire de [a , b] vers $\mathbb{R}$ [1] :

Définition C.a1 :

a) Une fonction de [a , b] vers $\mathbb{R}$ est dite $\mathcal{P}$-calculable s'il existe un polynôme P et un programme Prog tels que : pour tout dyadique $x \in [a , b]$ et tout entier en unaire m , le programme calcule à partir de l'entrée (x , m) une approximation $F(x,m) = z/2^m$ ($z \in \mathbb{Z}$) de f(x) avec la précision $1/2^m$ en temps inférieur à P(m)
NB : cette définition sous-entend que le dyadique en entrée x n'est pas necessairement lu en entier, seuls les premiers digits réellement utiles du développement binaire sont lus, selon les besoins du programme. On obtient ainsi un substitut à la notion d'oracle

b) Une suite de fonctions f_n de [a , b] vers $\mathbb{R}$ est dite $\mathcal{P}$-calculable s'il existe un polynôme P et un programme Prog tels que : pour tout dyadique $x \in [a , b]$ et tous entier en unaire n , m , le programme calcule à partir de l'entrée (x , n , m) une approximation $F(x,n,m) = z/2^m$ ($z \in \mathbb{Z}$) de $f_n(x)$ avec la précision $1/2^m$ en temps inférieur à P(n,m) .

---

[1] La définition proposée ici est équivalente à la notion usuelle de fonction $\mathcal{P}$-calculable donnée dans la littérature (cf par exemple [K-F])

Insistons bien sur le fait que la majoration du temps de calcul ne dépend que de la précision désirée sur le résultat, et pas de x : l'entrée est (x , n) mais le temps de calcul est majoré par P(n) . Ce qui est bien agréable.

Par ailleurs, il est clair qu'une fonction $\mathcal{P}$-calculable de [a , b] vers $\mathbb{R}$ est uniformément continue avec un module de continuité particulier:

$\forall\, \varepsilon = 1/2^m \quad \exists\, \delta = 1/2^{P(m)}$ ....

où P est un polynôme à coefficients entiers positifs

La définition A.a1 s'étend immédiatement aux fonctions d'un pavé de $\mathbb{R}^n$ vers $\mathbb{R}^m$, et donc presque immédiatement au cas qui nous intéresse. Il nous faudra fournir pour $\mathbf{P}_n(\mathbb{C})$ un nombre fini de cartes sous forme de pavés de $\mathbb{R}^{2n}$ avec des changements de cartes qui soient des fonctions $\mathcal{P}$-calculables. Toute manière naturelle de réaliser cet atlas conduit à la même notion de fonction $\mathcal{P}$-calculable sur $\mathbf{P}_n(\mathbb{C})$.

On obtient alors la version suivante du théorème fondamental de l'algèbre :

Le théorème fondamental de l'algèbre est réalisable par une fonction $\mathcal{P}$-calculable de $\mathbf{P}_n(\mathbb{C})$ vers $\mathbf{Sym}_n(\mathbf{P}_1(\mathbb{C}))$

La preuve est la suivante (en raccourci): étant donné un polynôme non nul de degré majoré par n , il est certainement non nul dans un des cercles centrés en un $\alpha \in \{0,1,...n\}$ et de rayon r = 1/3 . En envoyant cet α à l'infini par l'homographie $z \to 1/(z-\alpha)$ le polynôme est transformé en un polynôme de degré sûrement n qui admet toutes ses racines dans un cercle de rayon 3 et on peut appliquer la version du théorème fondamental de l'algèbre en temps polynomial donnée dans le premier paragraphe.

b) Méthode des tableaux de signes approchés

Position du problème, définitions

La méthode élémentaire repose sur des évaluations de signes des polynômes $P^{(k-1)}$, $P^{(k-2)}$,... , $P^{(2)}$, P' , P en des nombres algébriques de taille raisonnable.

En fait il faut évaluer le signe de $P^{(i)}(x)$ en un x qui est racine de $P^{(i+1)}$. Quand il s'agit du signe + ou du signe − , le calcul est facile : il suffit de calculer un rationnel approchant suffisamment $P^{(i)}(x)$. Or, il est a priori étonnant qu'un zéro multiple de P' , ou de P'' ,... ou de $P^{(k-2)}$ puisse constituer un obstacle au "calcul facile" d'un zéro simple de P. Bref, on doit pouvoir entièrement se passer de calculs de résultants et PGCD si le polynôme P est sans facteur carré, et n'utiliser que des évaluations approchées de P et de ses dérivées successives en des dyadiques de taille raisonnable. D'où ce qui suit.

Définition C.b1 : Soit $f : [\,a\,,\,b\,] \to \mathbb{R}$ une fonction continue, et ε un rationnel positif. On appellera ε - tableau de signes de f sur $[\,a\,,\,b\,]$, un découpage de l'intervalle en un nombre fini d'intervalles $[\,a_0\,,\,a_1\,]$, $[\,a_1\,,\,a_2\,]$, ... , $[\,a_{k-1}\,,\,a_k\,]$, marqués par + , 0 , ou − ; les conditions suivantes étant vérifiées :

* $f(x) > 0$ sur tout intervalle marqué +
* $f(x) < 0$ sur tout intervalle marqué −
* $-\varepsilon < f(x) < \varepsilon$ sur tout intervalle marqué 0
* 2 signes consécutifs sont toujours distincts
* $a_0 = a < a_1 < \ldots < a_{k-1} < a_k = b$
* $a_1 \ldots a_{k-1}$ sont des dyadiques

On remarquera qu'une fois sur 2 exactement le signe marqué est 0 .

Si f est continument dérivable, un ε-tableau de signes pour f' nous donne beaucoup de renseignements sur f : sur les intervalles marqués 0 , la fonction f varie très peu; et sur les autres, elle est strictement monotone.

Si f est un polynôme dont le discriminant est non nul et possède une minoration de taille raisonnable, alors, pour un ε de taille raisonnable, la connaissance d'un ε-tableau de signes pour f suffit à situer les racines de f puisque sur un intervalle marqué 0 la fonction f est necessairement monotone (on se reporte à la preuve du lemme 1 du A a : si $U f + V f' = \mathbf{Res}(f,f')$, et si N majore $|\,U\,|$ et $|\,V\,|$ sur $[\,a\,,\,b\,]$ on a en tout x de l'intervalle : $|\,f(x)\,| > |\,\mathbf{Res}(f,f')\,| \,/\, 2N$ ou $|\,f'(x)\,| > |\,\mathbf{Res}(f,f')\,| \,/\, 2N$) .

Nous allons maintenant établir la proposition désirée, qui nous dit que, si on sait calculer les ε-tableaux de signes pour f' , alors on sait calculer ceux de f. Voyons d'abord la preuve, ce qui nous permettra d'y voir plus clair pour énoncer la proposition.

Un algorithme

On suppose qu'on a $|\,b - a\,| \le 2^k$, $\|\,f'\,\| < M \le 2^{n_1}$ ($a\,,\,b \in \mathbb{Q}$, $k\,,\,n_1 \in \mathbb{Z}$) , $\varepsilon = 1/2^n$, et que a et b sont des dyadiques. On veut calculer un ε-tableau de signes pour f.

* *préliminaires* : on pose $\varepsilon' = 1/2^{n+k+3}$, $\varepsilon'' = 1/2^{n+n_1+2}$. On calcule un ε'-tableau de signes pour f' sur $[\,a\,,\,b\,]$. Soient $a_0 = a$, a_1 , ... , a_{m-1} , $a_m = b$ les bornes des

intervalles. Pour $i = 1,2,...,m - 1$ on calcule une approximation de $f(a_i)$ à $\varepsilon/8$ près, sous forme $f_i = g_i/2^{n+3}$ ($g_i \in \mathbb{Z}$) , de sorte que $f(a_i) \in [(g_i - 1)/2^{n+3} , (g_i+1)/2^{n+3}]$

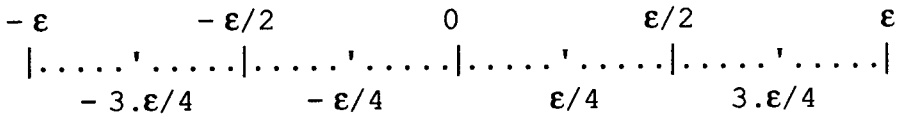
* *plan du travail* :

– 1ère étape : on marque pour f les intervalles marqués 0 pour f' , "avec un peu de large"

– 2ème étape : on remplace les bornes des intervalles par des dyadiques de taille raisonnable

– 3ème étape : on dichotomise les intervalles restants, puisque f est monotone sur ces intervalles

* *1ère étape* : la variation de f sur une intervalle marqué 0 pour f' est inférieure à $| b - a |.\varepsilon' \leq \varepsilon/8$. Si l'intervalle $[a_i , a_{i+1}]$ est marqué 0 pour f' , on obtient donc : $x \in [a_i , a_{i+1}] \Rightarrow f(x) \in [(g_i - 2)/2^{n+3} , (g_i+2)/2^{n+3}]$. Cet intervalle de longueur $\varepsilon/2$, centré en $g_i.\varepsilon/8$, est immédiatement situé sur un des intervalles $[- \infty, - \varepsilon/4]$, $[- 3.\varepsilon/4 , 3.\varepsilon/4]$, ou $[\varepsilon/4 , +\infty]$:



Ceci permet de marquer l'intervalle $[a_i , a_{i+1}]$ pour f , avec un peu de large.

* *2ème étape* : *élargissement des intervalles déjà marqués pour f* , par décalage des bornes, de manière à obtenir pour nouvelles bornes des dyadiques ayant au plus $n + n_1 + 2$ chiffres après la virgule. S'il s'agit de a ou b , on ne décale pas la borne. Pour une borne a_i ($i = 1,2,...,m - 1$) située en début d'intervalle déjà marqué (resp. en fin), on remplace a_i par le plus grand (resp. le plus petit) $m.\varepsilon"$ ($m \in \mathbb{Z}$) qui lui est inférieur (resp. supérieur). Comme a_i est décalé de moins que $\varepsilon"$, la variation sur $f(a_i)$ est inférieure à $M.\varepsilon"$ donc aussi à $\varepsilon/4$. De sorte que sur les intervalles agrandis, on a maintenant $f(x) > 0$ sur ceux qui étaient marqués + , $f(x) < 0$ sur ceux qui étaient marqués - , et $- \varepsilon < f(x) < \varepsilon$ sur ceux qui étaient marqués 0 . Quant aux intervalles non marqués, ils ont rétréci, et donc f reste strictement monotone sur ces intervalles.

* *petite explication concernant les dichotomies "sur réseau"* : à l'étape suivante, on va procéder à des dichotomies commençant avec des dyadiques qui sont tous dans $\mathbb{Z}.\varepsilon"$. Nous ne voulons pas quitter le réseau $\mathbb{Z}.\varepsilon"$, parce que nous savons a priori que ce n'est pas nécessaire, et que nous majorons ainsi la taille des dyadiques manipulés. Aussi, lorsque nous ferons dans une dichotomie : $c \leftarrow (a' + b')/2$, il faudra toujours sous-entendre : si $(a' + b')/2 = (r + 1/2).\varepsilon"$ ($r \in \mathbb{Z}$) , alors on fait $c \leftarrow r.\varepsilon"$. Si on démarre la dichotomie avec un intervalle de longueur $s.\varepsilon"$ avec $2^j < s \leq 2^{j+1}$ on aboutit en j ou j + 1 étapes à un intervalle de longueur $\varepsilon"$.

* *petite préparation pour la 3ème étape*. Si $a = a_0$ n'est pas sur un intervalle déjà marqué pour f , on va faire un changement de notation et un décalage : a_{-1} notera maintenant a , et a_0 va être décalé vers la droite (le moins possible) de manière à se retrouver sur $\mathbb{Z}.\varepsilon"$. La variation de f sur l'intervalle $[a_{-1} , a_0]$ étant inférieure à $\varepsilon/4$, il suffit de calculer $f(a_0)$ avec une précision meilleure que $\varepsilon/4$ pour situer les f(x) (x sur l'intervalle) sur un intervalle de longueur $< 3.\varepsilon/4$ et donc pour pouvoir marquer l'intervalle $[a_{-1} , a_0]$ pour f . On procède de manière symétrique avec b s'il n'est pas sur un intervalle déjà marqué pour f . Désormais on

peut affirmer, concernant un intervalle non encore marqué pour f :

- les bornes de l'intervalle sont sur $\mathbb{Z}.\varepsilon''$
- l'intervalle est entouré de 2 intervalles marqués pour f
- f est strictement monotone sur l'intervalle.

* *3ème étape proprement dite : traitement des intervalles restants.*

Supposons par exemple f strictement croissante sur un intervalle $[\, a_i , a_{i+1} \,]$ non encore marqué pour f. Sur les intervalles $[\, a_{i-1} , a_i \,]$ et $[\, a_i , a_{i+1} \,]$ les marques pour f peuvent être a priori dans l'un des 6 cas suivants, dont 3 donnent lieu à des "contractions" immédiates :

1 a_{i-1} ...0... a_i .....a_{i+1} ...0.....a_{i+2} contraction $\Rightarrow$ a_{i-1} ....0 ..... a_{i+2}
2 a_{i-1} ...+... a_i .....a_{i+1} ...+ ....a_{i+2} contraction $\Rightarrow$ a_{i-1} ....+..... a_{i+2}
3 a_{i-1} ...+... a_i .....a_{i+1} ...0.....a_{i+2} contraction $\Rightarrow$ a_{i-1}..0 ou +.. a_{i+2}
4 a_{i-1} ...0... a_i .....a_{i+1} ...+.....a_{i+2}
5 a_{i-1} ... −....a_i.....a_{i+1} ...0.....a_{i+2}
6 a_{i-1} ... −....a_i.....a_{i+1} ...+.....a_{i+2}

Il nous faut voir que dans les 3 cas restants, on peut obtenir en un temps raisonnable un déplacement des bornes ou une contraction. Dans les cas 4 et 5, a_i et a_{i+1} doivent fusionner en un c_i intermédiaire ("contraction"), et dans le cas 6, a_i et a_{i+1} doivent être rapprochés de manière qu'on puisse marquer 0 l'intervalle qui les sépare, tout en gardant les marques − et + sur les intervalles jouxtant.

Voyons d'abord le cas 4 : on va procéder à une dichotomie (sur le réseau $\mathbb{Z}.\varepsilon''$) à partir de a_i et a_{i+1} pour déterminer un $c \in [\, a_i , a_{i+1} \,] \cap \mathbb{Z}.\varepsilon''$ tel que $f(c) \in [\, \varepsilon/4 , \varepsilon \,]$. Les évaluations f_c (valeur approchée de f(c)) sont toujours faites en tant que valeur approchée à $\varepsilon/4$ près, prise sur le réseau $\mathbb{Z}.\varepsilon/4$. On peut conclure à coup sûr que $f(c) \in [\, \varepsilon/4 , \varepsilon \,]$ lorsque $f_c = \varepsilon/2$ ou $3.\varepsilon/4$. Lorsqu'on ne peut pas conclure à coup sûr, on a $f(c) > 3.\varepsilon/4$ ou $f(c) < \varepsilon/2$. Or la variation de f sur un intervalle du réseau $\mathbb{Z}.\varepsilon''$ est strictement inférieure à $\varepsilon/4$. Donc la dichotomie aboutit sûrement en au plus $n + k + n_1 + 3$ étapes.

Le cas 5 se traite comme le cas 4.

Le cas 6 est à peine plus compliqué : on détermine, par 2 dichotomies séparées un c et un $d \in [\, a_i , a_{i+1} \,] \cap \mathbb{Z}.\varepsilon''$ tels que $f(c) \in [\, -\varepsilon , -\varepsilon/4 \,]$ et $f(d) \in [\, \varepsilon/4 , \varepsilon \,]$. On peut conserver certains résultats intéressants la 2ème dichotomie pendant la 1ère, et on peut utiliser c comme la borne inférieure de départ pour la 2ème dichotomie.

* *majoration du temps de calcul*

Hypothèses: $h , n_0 , k , n_1 \in \mathbb{Z}$, $-2^h \le a < b \le 2^h$, $b - a \le 2^k$, la fonction f est continument dérivable sur $[\, a , b \,]$ avec $\| f' \|_\infty < 2^{n_1}$, $\| f \|_\infty \le 2^{n_0}$. On suppose qu'on sait calculer un $(1/2^n)$-tableau de signes pour f' en au plus T(n) étapes élémentaires, et que le nombre de bornes d'intervalles dans le tableau est majoré par S(n). On suppose enfin que l'on sait évaluer f(x) (pour $x \in \mathbb{D}$) avec la précision $1/2^n$ sous la forme $z/2^n$ (où $z \in \mathbb{Z}$) en au plus P(n) étapes élémentaires.

On pose $n' := n + k + 3$, $\varepsilon' := 1/2^{n'}$, $n'' := \mathbf{sup}(1, n + n_1 + 2) + h$, $\varepsilon'' := 1/2^{n+n_1+2}$.

La longueur d'un dyadique $x \in \mathbb{Z}\ \varepsilon'' \cap [\, a , b \,]$ est donc majorée par n".

Dans la 1ère étape du calcul on utilise (en étapes élémentaires) au plus :

T(n') pour le tableau de f'

S(n') P(n+3) pour les évaluations $f(a_i)$ avec la précision $\varepsilon/8$

S(n') c_1 (n+3+n_0) pour les "annexes" (marquer pour f les intervalles marqués 0 pour f' , les GOTO etc...)

La 2ème étape (décalage des bornes pour être sur le réseau $\mathbb{Z}\ \varepsilon''$) et le préliminaire de la 3ème étape (traitement éventuel des 2 bornes a et b) demandent au plus :

S(n') c_2 n" étapes élémentaires

La $3^{ème}$ étape demande au plus $S(n')$ dichotomies. Chaque dichotomie procède au plus en $n+k+n_1+2$ étapes. Chaque étape de dichotomie demande le calcul d'une demi-somme sur des dyadiques de longueur $\leq n''$, une évaluation de f à $1/2^{n+2}$ près, ce qui fait un nombre d'étapes élémentaires au plus égal à $n''+P(n+2)+c_3$. En tout la $3^{ème}$ étape demande donc au plus :

$S(n') ((n+k+n_1+2) (n''+P(n+2)+c_3) + c_4) + c_5$ étapes élémentaires

Nous notons $n_2 := \mathbf{sup}(n+k+n_1+2 , n+n_1+2+h , h+1 , n+3+n_0)$ et nous obtenons une majoration globale par $T(n+k+3) + S(n_2)\, O(n_2\, P(n+3))$. Si on considère n comme seule variable, alors $n_2 \leq n+c$

Remarque : L'algorithme et le calcul de majoration sont assez grossiers. Par exemple on demande de calculer un $((1/2^{n'})$-tableau de signes pour f', mais la précision n' n'est réellement utile que dans le cas où tout l'intervalle est marqué 0 pour f' , mais alors l'agorithme se termine à la première étape. Par ailleurs les évaluations f_c avec la précision $\varepsilon/4$ (dans les dichotomies) n'ont pas besoin en général d'être "complètes"; ce dont nous avons réellement besoin, pour une dichotomie dans le cas 4 par exemple, est d'obtenir un renseignement du type: $f_c \leq \varepsilon/4$, ou $f_c \geq \varepsilon$, ou $f_c = \varepsilon/2$ ou $f_c = 3\varepsilon/4$.

Quelques conséquences

Nous commençons par énoncer le résultat précédent

Théorème C.b2 :

Soit une fonction f continument dérivable sur $[\, a , b \,]$ $(a , b \in \mathbb{D})$.

On suppose $\| f' \|_\infty < 2^{n_1}$, $\| f \|_\infty \leq 2^{n_0}$, $-2^h \leq a < b \leq 2^h$, $b - a \leq 2^k$, avec h , n_0 , k , $n_1 \in \mathbb{Z}$.

On suppose qu'on sait calculer un $(1/2^n)$-tableau de signes pour f' en au plus $T(n)$ étapes élémentaires, et que le nombre de bornes d'intervalles dans le tableau est majoré par $S(n)$. On suppose enfin que l'on sait évaluer $f(x)$ (pour $x \in \mathbb{D}$) avec la précision $1/2^n$ sous la forme $z/2^n$ (où $z \in \mathbb{Z}$) en au plus $P(n)$ étapes élémentaires $(P(n) \geq n)$.

On note $n_2 := \mathbf{sup}(n+k+n_1+2 , n+n_1+2+h , h+1 , n+3+n_0)$.

Alors on peut calculer un $(1/2^n)$-tableau de signes pour f en au plus

$T(n+k+3) + S(n_2)\, O(n_2\, P(n+3))$ étapes élémentaires

On en déduit:

Théorème C.b3 :

Soit f une fonction r fois continument dérivable sur $[\, a , b \,]$ $(a , b \in \mathbb{D})$.

On suppose $\| f \|_\infty , \| f' \|_\infty , \ldots , \| f^{(r)} \|_\infty < 2^{n_1}$, $-2^h \leq a < b \leq 2^h$, $b - a \leq 2^k$, avec h , $n_1 \in \mathbb{N}$, $k \in \mathbb{Z}$.

On suppose que l'on sait évaluer $f(x) , f'(x) , \ldots , f^{(r)}(x)$ (pour $x \in \mathbb{D}$) avec la précision $1/2^n$ sous la forme $z/2^n$ (où $z \in \mathbb{Z}$) en au plus $P(n)$ étapes élémentaires $(P(n) \geq n$, croissante).

On suppose enfin que $f^{(r)}(x)$ est de signe constant sur $[\, a , b \,]$.

Alors on peut calculer un $(1/2^n)$-tableau de signes pour f en au plus

$O(r^2 (n+rc) P(n+rc))$ étapes élémentaires

où $c = \sup(k+n_1+2 , n_1+2+h , 3+n_1 , k+3) + c_0$.

preuve> Notons $T(n,s)$ une majoration du nombre d'étapes élémentaires pour calculer un $(1/2^n)$-tableau de signes pour $f^{(s)}$. En utilisant le théorème C.b2 , on a :

$$\begin{aligned} T(n,r) &\leq c_0 \\ T(n,r-1) &\leq c_0 + 2\,O((n+c)\,P(n+c)) \\ T(n,r-2) &\leq T(n+c,r-1) + 4\,O((n+c)\,P(n+c)) \\ &\leq c_0 + 2\,O((n+2c)\,P(n+2c)) + 4\,O((n+c)\,P(n+c)) \\ &\dots \\ T(n,r) &\leq c_0 + 2\,O((n+rc)\,P(n+rc)) + 4\,O((n+(r-1)c)\,P(n+(r-1)c)) + \dots \\ &\quad + 2r\,O((n+c)\,P(n+c)) \\ &\leq c_0 + 2\,(1+2+\dots+r)\,O((n+rc)\,P(n+rc)) \end{aligned}$$

(il s'agit du "même" O à chaque fois) ❑

Théorème C.b4 :

Soit f une fonction continue sur $[a , b]$ $(a , b \in \mathbb{D})$.

On suppose que f est limite uniforme d'une suite (P_n) de polynômes de la manière suivante :

a) $\| P_n - f \|_\infty \leq 1/2^n$

b) la suite P_n est Φ-calculable (de $\mathbb{N}_1$ vers $\mathbb{Q}[X]$)

Alors on peut calculer un $(1/2^n)$-tableau de signes pour f en temps polynomial (l'entrée est n en unaire)

preuve> Soit $\varepsilon := 1/2^n$, pour calculer un ε-tableau de signes pour f , on calcule un $(\varepsilon/2)$-tableau de signes pour chacun des deux polynômes $P_{n+1} + \varepsilon/2$ et $P_{n+1} - \varepsilon/2$ ❑

Remarques :

1) Nous étudions dans le §c) les fonctions qui vérifient les hypothèses du théorème C.b4 . Nous verrons en particulier que ce sont des fonctions indéfiniment dérivables.

2) Le théorème C.b4 est encore valable si on remplace la suite (P_n) de polynômes par une suite de fractions rationnelles $F_n = P_n/Q_n$ où les Q_n sont des polynômes minorés par 1 sur l'intervalle $[a , b]$. En effet, on obtient alors un ε-tableau de signes pour la fraction rationnelle en calculant un ε-tableau de signes pour le numérateur. Et la classe des fonctions continues qui se laissent "bien" approcher par des fractions rationnelles est nettement plus importante que celles qui se laissent "bien" approcher par des polynômes.

3) Notons $O(n^{h^+})$ pour $O(n^h \log^i(n))$ avec i non précisé. L'application du théorème C.b2 au cas des polynômes à coefficients entiers donne le résultat suivant :

Soit un polynôme f de degré r et à coefficients entiers de tailles majorées par m .

Alors on peut calculer un $(1/2^n)$-tableau de signes pour f en au plus

$O(r^{5^+} + r^{4^+} s^{1^+} + r^{3^+} s^{2^+})$ *étapes élémentaires où* $s = m + n$

On commence par remarquer que l'on peut se limiter à l'intervalle $[-1 , 1]$ puisqu'un $(1/2^n)$-tableau de signes pour f à l'extérieur de l'intervalle peut être obtenu à partir d'un $(1/2^n)$-tableau de signes pour g , polynôme aux inverses (les mêmes coefficients dans l'ordre inverse) sur l'intervalle. Par ailleurs, on peut calculer 8 tableaux de signes approchés sur des intervalles de longueur 1/8 et les recoller bout à bout. On est donc ramené au cas $k = -3$ dans le théorème C.b2 .

Notons alors $T(n,r,m)$ une majoration du temps de calcul pour le tableau de signe. Soit d'autre part $\varphi(n,r,m)$ le temps de calcul pour l'évaluation en un dyadique de $[-1 , 1]$ avec

la précision $1/2^n$ d'un polynôme de degré r et à coefficients entiers de tailles majorées par m. Une majoration de $|| f ||_\infty$ est donnée par $2^{m+\log_2(r+1)}$. Les coefficients de $f^{(q)}$ sont majorés par $2^{m+q\log_2(r)}$ et $|| f^{(q)} ||_\infty \leq 2^{m+(q+1).\log_2(r)}$. Ainsi, en appliquant le théorème C.b2 de manière récurrente, le coefficient n_2 ne dépassera jamais $n+m+(r+1)\log(r) = s+O(r^{1+})$ et le temps de calcul indiqué $P(n+3)$ ne dépassera jamais $\varphi(n,r,m+r\log(r))$. Pour majorer $T(n,r,m)$ il y aura donc r termes à additionner majorés par $r\, O((s+r^{1+}) \varphi(n,r,m+r\log(r)))$. En prenant $\varphi(n,r,m) = r.(n+m)^{1+}$, on trouve la majoration indiquée.

Appliquons le résultat précédent avec n assez grand pour que :

$| f | < 1/2^n$ sur un intervalle $\Rightarrow f'$ de signe constant sur l'intervalle.

La preuve du lemme 1 § A.a fournit $\varepsilon = | \mathrm{Res}(f,f') |/2N$ où N majore $| U |$ et $| V |$ avec $U f + V f' = \mathrm{Res}(f,f')$. Avec la majoration des coefficients de U et V donnée dans la preuve du lemme on obtient: $n = O(m\, r^{1+})$. Ce qui donne finalement

Soit un polynôme f de degré r sans facteur carré et à coefficients entiers de tailles majorées par m. Alors on peut isoler les racines réelles de f en temps majoré par $O(r^{5+} m^{2+})$.

Utilisant des méthodes assez sophistiquées, V. Pan ([**Pan**]) donne la majoration suivante (la meilleure actuellement connue) pour le temps du calcul permettant d'obtenir avec la précision $1/2^n$ les zéros complexes d'un polynôme à coefficients entiers majorés par 2^m :

$$O(r^3 s \log^2(r.s) \mathrm{loglog}(r.s)) \quad \text{où } s = m+n ,$$

cad avec notre notation en $O(r^{3+} (m+n)^{1+})$

Cela permet d'isoler les zéros réels d'un polynôme sans facteur carré à coefficients entiers : il suffit de calculer les zéros complexes avec une précision de $1/2^n$ où $n = (2r+1).(m+1+\log(r)) + 1$ (cf [**Sch**]), ce qui donne un calcul en $O(r^{4+} m^{1+})$. Il y a même une meilleure borne dans [**Sch**] (pour la première majoration, donc pour la seconde) mais Pan conteste la validité de la démonstration. On voit en tout cas que la méthode tout à fait élémentaire des tableaux de signes approchés fournit une majoration pas trop mauvaise.

c) Fonctions approchables en temps polynomial par des fonctions polynômes

Rappels de quelques résultats de la théorie de l'approximation uniforme par des polynômes

(Voir par exemple [**Riv**] et [**Che**])

Notations :

$\mathbf{C}[\,a\,,\,b\,]$ est l'espace des fonctions réelles continues sur le segment $[\,a\,,\,b\,]$.
$\mathbf{C}$ est l'espace $\mathbf{C}[-1\,,\,1\,]$, la norme uniforme sur cet intervalle est notée $\|\,f\,\|_\infty$ et la distance correspondante d_∞.
$\mathbf{C}^{(k)}$ est l'espace des fonctions k fois continument dérivables sur $[-1\,,\,1\,]$.
$\mathbf{C}^{(\infty)}$ est l'espace des fonctions indéfiniment dérivables sur $[-1\,,\,1\,]$.
$\mathcal{P}_n$ est l'espace des polynômes de degré $\leq n$.
T_n est le polynôme de Chebyshev de degré n :

$$T_n(\varphi(z)) = \varphi(z^n) \quad \text{avec} \quad \varphi(z) = \frac{1}{2}\left(z + \frac{1}{z}\right) \ ,$$

on peut également les définir par $T_n(\cos(x)) = \cos(nx)$ ou par

$$F(u,x) = \frac{1 - u.x}{1+u^2 - 2u.x} = \sum_{n\,=\,0}^{\infty} T_n(x)\ u^n$$

$E_n(f) = d_\infty(f, \mathcal{P}_n)$ pour $f \in \mathbf{C}$.
On considère sur $\mathbf{C}$ le produit scalaire

$$\langle g,h\rangle \ := \int_{-1}^{1} \frac{g(x).h(x)}{\sqrt{1 - x^2}}\ dx$$

On notera $\|\,f\,\|_2$ la norme au sens de ce produit scalaire.
Les polynômes T_i $(i = 0, 2, ..., n)$ forment une base orthogonale de $\mathcal{P}_n$ pour ce produit scalaire, avec $\langle T_0,T_0\rangle = \pi$ et $\langle T_i,T_i\rangle = \pi/2$ pour $i > 0$.

$$A_k = A_k(f) := \frac{2}{\pi} \int_{-1}^{1} f(x).T_k(x)\ \frac{dx}{\sqrt{1 - x^2}}$$

Les A_k sont appelés les *coefficients de Chebyshev* de f .

La fonction $s_n(f) := A_0/2 + \sum_{i=1}^{n} A_i\ T_i = \sum_{i=0}^{n}{}' A_i\ T_i$ est la projection orthogonale de f sur $\mathcal{P}_n$ au sens du produit scalaire considéré.

La série correspondante est appelée *la série de Chebyshev*[1] de f .
$S_n(f) = \|\,f - s_n(f)\,\|_\infty$, on a immédiatement $|\,A_{n+1}\,| \leq S_n(f) + S_{n+1}(f)$

---

[1] Elle converge au sens de $\mathbf{L}^2$ pour le produit scalaire considéré. La série de Chebyshev est aux fonctions continues sur $[-1\,,\,1\,]$ ce que la série de Fourier est aux fonctions continues périodiques, ce qui se comprend bien en considérant le "changement de variable" $z \to 1/2(\,z + 1/z\,)$ qui transforme le cercle unité du plan complexe en le segment $[-1\,,\,1\,]$ et la fonction $z \to z^n$ en le polynôme T_n.

Les zéros de T_n sont les $\xi_i^{[n]} = \cos\left(\frac{2i-1}{n} \frac{\pi}{2} \right)$ $i = 1, 2, ..., n$

et on a $T_n(x) = 2^{n-1} \prod_{i=1}^{n} \left(x - \xi_i^{[n]} \right)$

Les extrema de T_n sur $[-1, 1]$ sont les

$$\eta_i^{[n]} = \cos\left(\frac{i}{n} \frac{\pi}{2} \right) \qquad i = 0, 2, ..., n$$

Des valeurs approchées de $s_n(f)$ peuvent être calculées au moyen de formules d'interpolation: on pose

$$\alpha_k^{[m]} = \frac{2}{m} \sum_{i=1}^{m} f(\xi_i^{[m]})\ T_k(\xi_i^{[m]})$$

$u_n^{[m]} = \sum_{k=0}^{n}{}' \alpha_k^{[m]} T_k(x)$: $u_n^{[n]}$ est le polynôme qui interpole f aux zéros de T_{n-1}

Quelques résultats

Evaluation d'un polynôme $p(x) = \sum_{k=0}^{n}{}' A_k T_k$:

Les formules récurrentes $T_{m+1}(x) = 2 x T_m(x) - T_{m-1}(x)$ conduisent à un algorithme à la Horner:

$$B_{n+1} = B_{n+2} = 0\ ,\ B_k = 2 x B_{k+1} - B_{k+2} + A_k\ ,\ p(x) = \frac{B_0 - B_2}{2}$$

Théorème de Markov: Si $g \in \mathcal{P}_n$ et $\| g \|_\infty \le 1$, alors $\| g' \|_\infty \le n^2$ (1)

et $\| g^{(k)} \|_\infty \le T_n^{(k)}(1) = \frac{n^2 (n^2-1) ... (n^2-(k-1)^2)}{1.3.5....(2k-1)}$ pour $n \ge 2$ (2)

Comparaison de $E_n(f)$ et $S_n(f)$:

$$E_n(f) \le S_n(f) \le \left(4 + \frac{4}{\pi^2} \log(n)\right) E_n(f) \qquad (3)$$

Comparaison de $E_n(f)$ et $A_{n+1}(f)$:

Pour $n \ge 1$ on a $\int_{-1}^{1} | T_n(x) | \frac{dx}{\sqrt{1-x^2}} = 2$,

d'où on déduit $(\pi/4) A_{n+1}(f) \le E_n(f)$ (4)

Théorèmes de Jackson : Soit $f \in \mathbf{C}$

(i) $E_n(f) \le \pi\lambda / (2n+2)$ si $| f(x) - f(y) | \le \lambda | x - y |$ (5)

(ii) $E_n(f) \le (\pi / 2)^k \| f^{(k)} \|_\infty / [(n+1)(n)(n-1) ... (n-k+2)]$ (6)

si $f \in \mathbf{C}^{(k)}$ et $n \ge k$

Convergence de la série de Chebyshev d'une fonction

La série de Chebyshev d'une fonction $f \in \mathbf{C}^{(k)}$ converge uniformément vers f si $k \ge 1$, et elle est absolument convergente (pour la norme $\| f \|_\infty$) si $k \ge 2$. En outre on a alors

$$S_n(f) = \| s_n(f) - f \|_\infty \le \sum_{j=n+1}^{\infty} | A_j | \qquad (7)$$

$$\| s_n(f) - u_n^{[n]} \|_\infty \le \sum_{j=n}^{\infty} | A_j | + \sum_{i=1}^{\infty} | A_{(2i+1)n} | \qquad (8)$$

Approximation uniforme des fonctions $\in \mathbf{C}^{(\infty)}$ par des polynômes

Les propriétés suivantes sont équivalentes

(i) $\forall k \ \exists M>0 \ \forall n \quad E_n(f) \le M / n^k$

(ii) $\forall k \ \exists M>0 \ \forall n \quad S_n(f) \le M / n^k$

(iii) $\forall k \ \exists M>0 \ \forall n \quad | A_n(f) | \le M / n^k$

(iv) $\forall k \ \exists M>0 \ \forall n \quad \| u_n^{[n]} - f \|_\infty \le M / n^k$

(v) La fonction $f \in \mathbf{C}^{(\infty)}$

preuve> (i) et (ii) sont équivalents d'après (3) .

(iv) $\Rightarrow$ (i) trivialement.

(ii) $\Rightarrow$ (iii) parce que $| A_n(f) | \le S_n(f) + S_{n-1}(f)$

(iii) $\Rightarrow$ (iv) d'après (7) et (8).

(iii) $\Rightarrow$ (v) : la série $\Sigma' A_i T_i^{(h)}$ est absolument convergente d'après (2) et les majorations (iii) ; donc on peut dériver h fois terme à terme la série de Chebishev

(v) $\Rightarrow$ (i) d'après (6) ❑

Analyticité et approximation uniforme par des polynômes

Les propriétés suivantes sont équivalentes

(i) $\exists M>0, r<1 \ \forall n \quad E_n(f) \le M r^n$

(ii) $\exists M>0, r<1 \ \forall n \quad S_n(f) \le M r^n$

(iii) $\exists M>0, r<1 \ \forall n \quad | A_n(f) | \le M r^n$

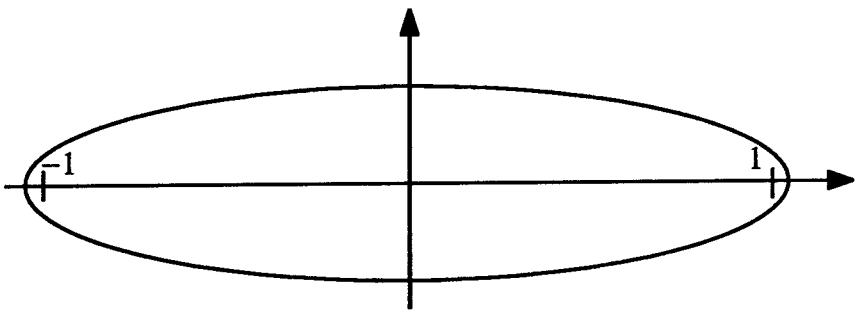
(iv) $\exists M>0, r<1 \ \forall n \quad \| u_n^{[n]} - f \|_\infty \le M r^n$

(v) $\exists r<1$ telle que f est analytique dans le plan complexe à l'intérieur de l'ellipse $\mathcal{E}_\rho$ de foyers 1 , -1 et dont le demi-somme des diamètres principaux est égale à $\rho = 1/r$

Et la limite inférieure des valeurs de r possibles est la même dans les 5 cas[1]

Ces propriétés sont équivalentes à l'analyticité de f sur l'intervalle, et aussi à:

(vi) $\exists M>0, R>0 \ \forall n \quad \| f^{(n)} \|_\infty \le M R^n n!$



Remarques :

1) L'espace des fonctions analytiques sur un intervalle compact possède donc une bonne description constructive, en termes de série de Chebyshev par exemple. Il apparaît comme une réunion dénombrable emboitée d'espaces métriques complets (ceux obtenus en utilisant la

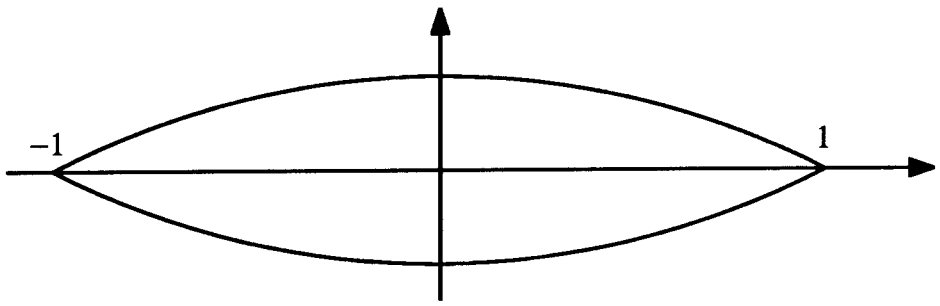
---

[1] Les équivalences (i) ... (iv) se montrent comme pour la proposition précédente. Pour l'équivalence avec (v) voir par exemple [**Riv**] . La condition (vi) représente à très peu près l'analyticité dans l'ouvert U_R formé des points dont la distance à l'intervalle est inférieure à 1/R .

définition (iii) et en fixant M et r rationnels par exemple). L'espace des fonctions $\mathbf{C}^\infty$ est beaucoup plus difficile à décrire constructivement, essentiellement parce qu'il n'existe pas de manière agréable d'engendrer les suites de rationnels à décroissance rapide[1].

2) La condition (i) peut être également lue comme suit : la fonction f peut être approchée à $1/2^n$ (pour la norme uniforme) par un polynôme de degré $\leq$ c.n , où c est une constante fixée, c.-à-d. encore : il existe un entier h tel que $E_{hn}(f) \leq 1/2^n$. Même remarque pour les conditions (ii) , (iii) et (iv) . Cela implique que la fonction f peut être approchée à $1/2^n$ par un polynôme à coefficients dyadiques dont la taille (en présentation dense sur la base des X^n ou sur la base des T_n) est en $O(n^2)$. La taille de la somme des valeurs absolues des coefficients est, elle, en O(n) .

Bakhvalov (cf [**Bak**] chap IV § 8 Th p 233) donne une condition suffisante du même genre pour qu'une fonction f soit analytique dans une lentille d'extrémités -1 et 1 du plan complexe (et non plus dans un voisinage du segment) : il suffit que la somme des valeurs absolues des coefficients d'un polynôme approchant f à $1/2^n$ soit majorée par $M\,2^{qn}$ (où M et q sont des constantes fixées). C.-à-d. encore: la taille de la somme des valeurs absolues des coefficients d'un polynôme approchant f à $1/2^n$ est en O(n) .



L'espace $\mathbf{C}[\,a\,,\,b\,]$ présenté via les polynômes à coefficients rationnels

L'espace $\mathbf{C}[\,a\,,\,b\,]$ muni de la norme uniforme peut être vu comme le complété de $\mathbb{Q}[X]$ muni de la norme correspondante. Nous obtenons alors une analogie entre $(\,\mathbb{R}\,,\,\mathbb{Q}\,,\,|\ |\,)$ et $(\,\mathbf{C}[\,a\,,\,b\,]\,,\,\mathbb{Q}[X]\,,\,\|\ \|_{\infty,[a,b]}\,)$.

Nous nous limiterons au cas $[\,a\,,\,b\,] = [\,-1\,,\,1\,]$. Le cas général s'y ramène immédiatement par changement de variable.

En ce qui concerne la présentation de $\mathbb{Q}[X]$ dans ce contexte, on peut raisonnablement hésiter entre la présentation dense ordinaire (sur la base des X^n) et la présentation sous forme de sommes de Chebyshev, c.-à-d. la présentation dense sur la base des T_n . (les coefficients dans $\mathbb{Q}$ étant toujours pris en binaire). Fort heureusement, le changement de base correspondant est en temps polynomial (dans les deux sens), et pour les problèmes de complexité en temps polynomial, les résultats obtenus sont les mêmes pour les deux présentations.

Remarquons également que la norme $\|\,f\,\|_\infty$ est $\mathfrak{P}$-calculable pour un polynôme f (avec la présentation dense) : plus précisément, la fonction $f \rightarrow \|\,f\,\|_\infty$, $\mathbb{Q}[X] \rightarrow \mathbb{R}$ est $\mathfrak{P}$-calculable au sens de la définition A.a3 . Le maximum de $|\,f\,|$ est en effet atteint en une

---

[1] Cela tient au $\forall\ k\ \exists\ M$ dans la définition de la décroissance rapide. Cette alternance de quantificateurs prend une forme explicite lorsqu'on donne M en fonction de k explicitement. Mais, en vertu de l'argument diagonal de Cantor, il n'y a pas de manière effective d'engendrer les fonctions effectives de $\mathbb{N}$ vers $\mathbb{N}$.

borne de l'intervalle ou en un zéro de f' , et il suffit de calculer le **sup** des $|f(x_i)|$ correspondants. (on applique donc le théorème A.c1 et la proposition A.a6 [1]).

Dans la définition C.a1 , c'est l'aspect "fonction" qui est pris en compte essentiellement, pour dire qu'une fonction est $\mathcal{P}$-calculable. Dans l'espace **C** , nous avons une autre notion intéressante (cf par exemple le théorème C.b4), qui est celle de fonction se laissant facilement approcher par des fonctions polynômes à coefficients rationnels. De la même manière qu'on définit *un réel de complexité* $\mathcal{P}$ (ou encore un $\mathcal{P}$-point de $\mathbb{R}$) comme un réel qui se laisse approcher à $1/2^n$ près par un rationnel en temps polynomial (en fonction de n pris en unaire), on peut définir, dans **C** un $\mathcal{P}$-point de $\mathbf{C_W}$. Notons que la définition des "points rationnels de **C** " comme étant les polynômes à coefficients rationnels possède une bonne part d'arbitraire[2], c'est pour cela notamment que nous spécifions le **W** en indice dans $\mathbf{C_W}$.

Définition C.c1 :

On notera $\mathbf{C_W}$ l'espace **C** lorsqu'il est présenté "à la Weierstrass" c.-à-d. via sa partie dénombrable dense constituée par les fonctions polynômes à coefficients rationnels, elles-mêmes vues comme éléments de $\mathbb{Q}[X]$ en présentation dense.

a) Une fonction $f \in \mathbf{C}$ est appelée un $\mathcal{P}$-point de $\mathbf{C_W}$ s'il existe une suite $m \to P_m$, $\mathcal{P}$-calculable (de $\mathbb{N}_1$ vers $\mathbb{Q}[X]$), vérifiant:
pour tout m $\quad \| P_m - f \|_\infty \leq 1/2^m$

b) Une suite f_m dans **C** est appelée une $\mathcal{P}$-suite s'il existe une suite $(n , m) \to P_{n,m}$, $\mathcal{P}$-calculable (de $\mathbb{N}_1 \times \mathbb{N}_1$ vers $\mathbb{Q}[X]$), vérifiant:
pour tous n , m $\quad \| P_{n,m} - f_n \|_\infty \leq 1/2^m$

Remarque:
Une définition immédiatement équivalente à la défintion a) est obtenue en demandant que f s'écrive comme somme d'une série $\Sigma\, Q_m$, où (Q_m) est une suite $\mathcal{P}$-calculable dans $\mathbb{Q}[X]$ vérifiant : $\| Q_m \|_\infty \leq 1/2^m$. Ceci donne une manière agréable de présenter les $\mathcal{P}$-points de $\mathbf{C_W}$.

Si f est un $\mathcal{P}$-point de $\mathbf{C_W}$ donné par une suite $m \to P_m$, $\mathcal{P}$-calculable, alors le degré de P_m est majoré par un polynôme en m , donc il existe un entier k et une constante B telles que le degré de P_m soit majoré par $(Bm)^k$. Soit alors n arbitraire, et considérons le plus grand entier m tel que $(Bm)^k \leq n$, c.-à-d. $m := \mathrm{Ent}(\sqrt[k]{n} / B)$. On a donc $m+1 \geq \sqrt[k]{n} / B$. En posant $r := 1/2^{1/B}$ et $\gamma := 1/k$ on obtient : $E_n(f) \leq 1/2^m \leq 2\, r^{n^\gamma}$, avec $r \in]\, 0, 1\, [$, $\gamma > 0$. En particulier, la suite $E_n(f)$ est à décroissance rapide et f est $\mathbf{C}^\infty$.

Ceci nous amène à étudier les fonctions f pour lesquelles ce genre de majoration est obtenu .

---

[1] En fait $\| f \|_\infty$ est un réel algébrique qui peut être calculé comme élément de $\mathbb{R}_{alg}$ en temps polynomial à partir de f

[2] On pourrait considérer les fonctions rationnelles à coefficients rationnels sans pole sur l'intervalle, ou encore les fonctions qui sont "polynômes par morceaux" pour des subdivisons finies et rationnelles de l'intervalle.

Fonctions dans la classe de Gevrey, ou polynomialement $\mathbf{C}^\infty$

Proposition C.c2 :

Soit $f \in \mathbf{C}$. Les propriétés suivantes sont équivalentes.

i)	$\exists M>0, r<1, \gamma>0$	$\forall n$	$E_n(f) \leq M r^{n^\gamma}$
ii)	$\exists M>0, r<1, \gamma>0$	$\forall n$	$S_n(f) \leq M r^{n^\gamma}$
iii)	$\exists M>0, r<1, \gamma>0$	$\forall n$	$\mid A_n(f) \mid \leq M r^{n^\gamma}$
iv)	$\exists M>0, r<1, \gamma>0$	$\forall n$	$\| u_n^{[n]} - f \|_\infty \leq M r^{n^\gamma}$
j)	$\exists c, \beta>0$	$\forall m \geq c n^\beta$	$E_m(f) \leq 1/2^n$
jj)	$\exists c, \beta>0$	$\forall m \geq c n^\beta$	$S_m(f) \leq 1/2^n$
jjj)	$\exists c, \beta>0$	$\forall m \geq c n^\beta$	$\mid A_m(f) \mid \leq 1/2^n$
jw)	$\exists c, \beta>0$	$\forall m \geq c n^\beta$	$\| u_m^{[m]} - f \|_\infty \leq 1/2^n$

Lorsque ces conditions sont vérifiées, nous dirons que :

f est polynomialement $\mathbf{C}^\infty$

preuve> i) $\Leftrightarrow$ ii) à partir de (3). i) $\Rightarrow$ iii) à partir de (4) . iv) $\Rightarrow$ i) est triviale. Les 4 équivalences du type i) $\Leftrightarrow$ j) résultent du même genre de calcul que celui qui a été fait avant la proposition.
L'implication jjj) $\Rightarrow$ jw) résulte d'un calcul de majoration simple, analogue à celui donné dans la preuve du théorème C.c5 d). ❑

Remarques :

1) L'espace des fonctions polynomialement $\mathbf{C}^\infty$ possède donc une présentation constructive agréable.

2) Pour $\gamma = 1$ on obtient les fonctions analytiques. Pour $\gamma > 1$, on obtient des fonctions entières.

3) Pour $\gamma \leq 1$, la limite supérieure des γ possibles est la même dans i), ii), iii) et iv) , la limite inférieure des β possibles est la même dans j), jj), jjj) et jw) , avec $\gamma = 1/\beta$.

4) Dans j), jj), jw) on peut supprimer le quantificateur $\forall$ m si on prend c et β entiers et $m = c\, n^\beta$.

En fait la classe des fonctions polynomialement $\mathbf{C}^\infty$ s'avère être une classe déjà étudiée, notamment dans la littérature concernant les solutions de certaines équations aux dérivées partielles : la classe de Gevrey.

Définition C.c3 :[1]-

Une fonction f , $\mathbf{C}^\infty$ sur l'intervalle $[\,-1\,,\,1\,]$ est dite dans la classe de Gevrey d'ordre $\alpha > 0$ si ses dérivées vérifient une majoration :

$$\| f^{(n)} \|_\infty \leq M\, R^n\, n^{\alpha n}$$

La classe de Gevrey est obtenue lorsqu'on ne précise pas l'ordre α .

---

1 Cf par exemple Hormander: The Analysis of Linear Partial Differential Operators I p 281 (Springer 1983). Une fonction est Gevrey d'ordre 1 si et seulement si elle est analytique.

Proposition C.c4 :

Une fonction f , $\mathbf{C}^\infty$ sur l'intervalle [− 1 , 1] est dans la classe de Gevrey si et seulement si elle est polynomialement $\mathbf{C}^\infty$.

preuve> Supposons tout d'abord que f soit Gevrey d'ordre α . Le problème de majoration n'est délicat que pour $\alpha \geq 1$, ce qu'on supposera maintenant. En appliquant le théorème de Jackson, on obtient une majoration $E_n(f) \leq \pi^k \, || f^{(k)} ||_\infty / n^k$ dès que $n \geq 2k$, ce qui donne avec la majoration de Gevrey $E_n(f) \leq A \, (C \, k^\alpha / n)^k$. On peut supposer $C^{1/\alpha} \geq 2$ et on prend pour k un entier proche de $(n/2C)^{1/\alpha}$ ($\leq n/2$), d'où à très peu près $E_n(f) \leq A \, (1/2)^{(n/2C)^{1/\alpha}} = A \, r^{n^\gamma}$, avec $\gamma = 1/\alpha$.
Supposons maintenant que f soit polynomialement $\mathbf{C}^\infty$. Le problème de majoration n'est délicat que pour $\beta \geq 1$ (proposition C.c2), ce qu'on supposera maintenant. On écrit $f^{(k)} = \Sigma' A_m \, T_m^{(k)}$. D'où $|| f^{(k)} ||_\infty \leq \Sigma' | A_m | \, m^{2k}$, d'après le théorème de Markov. On utilise maintenant la majoration jjj) de la proposition C.c2. On prend c et β entiers pour simplifier (ce n'est pas une restriction). Dans la somme ci-dessus, on regroupe les termes pour m compris entre $c \, n^\beta$ et $c(n+1)^\beta$; dans le paquet obtenu, on majore chaque terme par $(1/2^n) \, m^k$, et on majore le nombre de termes par $c(n+1)^\beta$, d'où :

$$|| f^{(k)} ||_\infty \leq \Sigma_n \left(c(n+1)^\beta / 2^n \right) \left(c(n+1)^\beta \right)^{2k} \leq 2 \, c^{2k+1} \, \Sigma_n \, (n+1)^{\beta(2k+1)} / 2^n$$
$$\leq 4 \, c^{2k+1} \, \Sigma_n \, n^h / 2^n \quad \text{où } h = \beta(2k+1)$$

On majore cette série par la série obtenue en dérivant h fois la série $\Sigma_n \, x^n$ (puis en faisant $x = 1/2$) et on obtient que f est Gevrey d'ordre 2β . ❑

Remarque: Si on se base sur le cas des fonctions analytiques ($\alpha = \beta = \gamma = 1$), on peut espérer, dans le dernier cas, obtenir que f soit Gevrey d'ordre β au moyen d'un calcul de majoration plus sophistiqué.

Théorème C.c5 :

Soient f un $\mathcal{P}$-point de $\mathbf{C}_W$, a et b des $\mathcal{P}$-réels . Alors :

a) f est une fonction $\mathcal{P}$-calculable

b) f est polynomialement $\mathbf{C}^\infty$ c.-à-d.
$\exists \, M > 0$, $r < 1$, $k > 0$ tels que : $\forall \, n \quad E_n(f) \leq M \, r^{n^\gamma}$

c) la suite A_n est $\mathcal{P}$-calculable (l'entrée est $n \in \mathbb{N}_1$)

d) la suite $f^{(n)}$ est une $\mathcal{P}$-suite de $\mathbf{C}_W$ [1]

e) les nombres $|| f ||_\infty$, $|| f ||_2$, $|| f ||_1$ $\sup_{x \in [a,b]} (f(x))$ et $\int_a^b f(x) \, dx$ sont des $\mathcal{P}$-réels,
les suites de réels $|| f^{(n)} ||_\infty$, $|| f^{(n)} ||_2$, $|| f^{(n)} ||_1$, $\sup_{x \in [a,b]} (f^{(n)}(x))$
et $\int_a^b f^{(n)}(x) \, dx$ sont $\mathcal{P}$-calculables

preuve> f est un $\mathcal{P}$-point de $\mathbf{C}_W$ donné par une suite $m \to P_m$, $\mathcal{P}$-calculable
a) pour calculer f(x) avec la précision $1/2^n$ on calcule $P_{n+1}(x)$ avec la précision $1/2^{n+1}$.
b) déjà vu avant la proposition C.c2 .

---

[1] C'est une suite $\mathcal{P}$-calculable en tant que suite dans $\mathbf{C}_W$ donc a fortiori en tant que suite de fonctions (cf la définition C.a1)

c) la suite double $A_k(P_n)$ est $\mathcal{P}$-calculable (entrées k et n en unaire). Comme s_k est une projection orthogonale, on a $\| s_k(P_n) - s_k(f) \|_2 \leq \| P_n - f \|_2$.
Donc $| A_k(P_n) - A_k(f) | \leq \| P_n - f \|_2 \leq \pi \| P_n - f \|_\infty \leq 1/2^{n-2}$.
d) la suite double $P_n^{(k)}$ est $\mathcal{P}$-calculable (entrées en unaire). Il existe un entier h et une constante a telles que le degré de P_m soit majoré par $(2^a m)^h$. Donc, d'après le théorème de Markov (1) on a la majoration :
$\| P_n^{(k)} - P_{n-1}^{(k)} \|_\infty \leq (2^a n)^{2hk} \| P_n - P_{n-1} \|_\infty \leq (2^a n)^{2hk} / 2^{n-2} = 1/2^{n-(k.(2h(a+\log_2(n)))+2)}$.
On détermine alors aisément une constante n_0 telle que, pour $n \geq 2\ n_0\ k$, on ait $n \geq 2\ (k.(2h\ (a+\log_2(n)))+2)$ et donc $\| P_n^{(k)} - P_{n-1}^{(k)} \|_\infty \leq 1/2^{n/2}$, de sorte qu'avec $m(n) := 2\ \mathbf{sup}(n_0\ k\ ,\ n)$, on a, pour $q \geq m(n)$, $\| P_q^{(k)} - P_{q+1}^{(k)} \|_\infty \leq 1/2^n$, et donc, puisque $m(n+1) = m(n)$ ou $m(n)+2$, $\| P_{m(n)}^{(k)} - P_{m(n+1)}^{(k)} \|_\infty \leq 1/2^{n-1}$, d'où enfin $\| P_{m(n)}^{(k)} - f^{(k)} \|_\infty \leq 1/2^{n-2}$. On termine en notant que la suite double $(n,k) \rightarrow P_{m(n+2)}^{(k)}$ est $\mathcal{P}$-calculable.
e) pour ce qui concerne les $\mathcal{P}$-réels $\| f \|_\infty$, $\| f \|_2$, $\| f \|_1$ etc... on en calcule une approximation convenable sous la forme $\| P_n \|_\infty$, $\| P_n \|_2$, $\| P_n \|_1$ (le calcul dans le cas d'un polynôme est en temps polynomial) , même principe pour le cas des $\mathcal{P}$-suites en utilisant le d) . ❑

Théorème C.c6 :

Soit $f \in \mathbf{C}$. Les propriétés suivantes sont équivalentes.

i) f est une fonction $\mathcal{P}$-calculable et polynomialement $\mathbf{C}^\infty$

i') f est une fonction $\mathcal{P}$-calculable et
$\exists\ M > 0\ ,\ r < 1,\ \gamma > 0$ tels que : $\forall\ n\quad E_n(f) \leq M\, r^{n^\gamma}$

i") f est une fonction $\mathcal{P}$-calculable et Gevrey

ii) la suite $A_n(f)$ est $\mathcal{P}$-calculable et f est Gevrey

ii') la suite $A_n(f)$ est $\mathcal{P}$-calculable et
$\exists\ M > 0\ ,\ r < 1,\ \gamma > 0$ tels que : $\forall\ n\quad | A_n(f) | \leq M\, r^{n^\gamma}$

iii) f est un $\mathcal{P}$-point de $\mathbf{C_W}$.

preuve> Gevrey équivant à polynomialement $\mathbf{C}^\infty$ (proposition C.c4)
(iii) $\Rightarrow$ (i) et (ii) d'après le théorème précédent
(i) $\Leftrightarrow$ (i') et (ii) $\Leftrightarrow$ (ii') d'après la proposition C.c2
(ii') $\Rightarrow$ (iii) : Un polynôme (en présentation dense sur la base des T_n) approchant f avec la précision $1/2^{n+1}$, est obtenu avec la somme partielle extraite de la série de Chebyshev de f en s'arrêtant à l'indice $(Bn)^h$ (où B et h se calculent à partir de M et γ) . Il reste à remplacer chaque coefficient de Chebyshev par un rationnel l'approchant à $1/[(Bn)^h\ 2^{n+1}] = 1/2^{n+1+h.\log_2(Bn)}$.
(i) $\Rightarrow$ (iii) Un polynôme approchant f avec la précision $1/2^{n+1}$, est obtenu avec $u_m^{[m]}$, (où $m = (Cn)^k$, C et k se calculant à partir de M et γ , en tenant compte de (7) et (8)) . La formule définissant $u_m^{[m]}$ fournit ses coefficients sur la base des T_n et on peut calculer (en temps polynomial) une approximation à $1/2^{n+1+k.\log_2(Cn)}$ près de ces coefficients en profitant du fait que la suite double $\xi_i^{[n]}$ est une $\mathcal{P}$-suite de réels et que la fonction f est $\mathcal{P}$-calculable . ❑

Morale: Tout calcul usuel concernant les fonctions Gevrey $\mathcal{P}$-calculables est en temps polynomial

Fonctions $\mathcal{P}$-analytiques sur un intervalle compact

Corollaire C.c7 :

Soit $f \in \mathbf{C}$. Les propriétés suivantes sont équivalentes.

i) f est une fonction analytique et $\mathcal{P}$-calculable

ii) la suite $A_n(f)$ est $\mathcal{P}$-calculable et vérifie une majoration

$$| A_n(f) | \leq M r^n \quad (M > 0 , r < 1)$$

iii) f est une fonction analytique et est un $\mathcal{P}$-point de $\mathbf{C_W}$

Lorsque ces propriétés sont vérifiées, on dira que

f *est* $\mathcal{P}$*-analytique sur l'intervalle* $[-1 , 1]$

preuve> immédiat d'après le théorème C.c6 et la caractérisation des fonctions analytiques ❑

Remarque : De manière générale, les preuves fournies sont constructives, ce qui signifie que si les hypothèses sont vérifiées de manière explicite, on sait construire un algorithme qui réalise la conclusion. Par exemple dans le théorème ci-dessus l'implication (i) $\Rightarrow$ (iii) est réalisable par un algorithme lorsque: *primo* l'analyticité de la fonction f est connue de manière explicite (connaissance d'un couple de rationnels (M,r) vérifiant une des caractérisations des fonctions analytiques), et *secundo* la $\mathcal{P}$-calculabilité de f est donnée explicitement par un programme calculant " f(x) avec la précision $1/2^n$ " *et* une majoration polynomiale explicite du temps de calcul du programme .

Morale: Tout calcul usuel concernant les fonctions analytiques $\mathcal{P}$-calculables sur un intervalle compact est en temps polynomial

Remarque : les théorèmes C.c5, C.c6 et le corollaire C.c7 améliorent les résultats de **[KF1]**, **[KF2]** et **[Mü2]** sur les fonctions analytiques $\mathcal{P}$-calculables.

d) Extensions possibles

Signalons pour terminer quelques résultats qu'on peut espérer obtenir sans trop de difficulté dans la même direction de travail.

Explicitation des zéros réels d'une fonction analytique comme racines d'un polynôme

Considérons le théorème classique suivant :

> Si f est une fonction analytique non nulle sur l'intervalle $[-1, 1]$, il existe un polynôme $g(x)$ et une fonction analytique $h(x)$ tels que :
> $h(x) \geq 1$, et $f(x) = g(x)\,h(x)$ sur tout l'intervalle

En mathématiques classiques on peut prendre pour g un polynôme $c \prod (x - x_i)^{n_i}$ où les x_i sont les zéros de f (avec multiplicité n_i) sur l'intervalle $[-1, 1]$.
Constructivement, on ne peut espérer un polynôme g aussi précis, parce que g dépendrait de f de manière à la fois extensionnelle et discontinue (au voisinage d'une fonction possédant un zéro multiple ou un zéro en une extrémité de l'intervalle).
Néanmoins, sous la forme indiquée, le théorème est sûrement démontrable constructivement, et la démonstration fournit alors un algorithme dont les entrées et les sorties sont de la forme suivante :

> *entrées:*
> une fonction analytique f sur l'intervalle, donnée explicitement :
> par un "module d'analyticité" $(M,r) \in (\mathbb{N}, \mathbb{Q} \cap [0, 1[)$
> et par une suite de polynômes $P_n \in \mathbb{Q}[X]$, vérifiant :
> $\deg(P_n) \leq n$, $\| P_n - P_m \|_\infty \leq M (r^n + r^m)$ et $\| f - P_n \|_\infty \leq M\, r^n$
> un point $x_0 \in \mathbb{Q} \cap [-1, 1]$ et un rationnel $c > 0$ tels que $f(x_0) > c$
> *sorties:*
> le polynôme g à coefficients réels
> la fonction analytique h (donnée sous la même forme que f)

On peut enfin espérer une version "en temps polynomial" du théorème, ce qui signifie que l'algorithme ci-dessus doit travailler en temps polynômial, en un sens raisonnable:

* les entrées discrètes sont M, r, c, x_0 et m degré de précision souhaité sur la sortie
* les polynômes P_n sont fournis par un oracle, à la demande[1]
* les sorties sont :
  le degré d de g, M', r' (module d'analyticité de h), calculés indépendamment de m,
  les coefficients de g calculés avec la précision $1/2^m$, et
  un polynôme Q_m (de degré n) vérifiant $\| f - Q_m \|_\infty \leq M'\, r'^n \leq 1/2^m$

Le temps d'exécution de l'algorithme doit être polynomial par rapport à la taille des entrées discrètes

---

[1] La question posée est "polynôme numéro n ? ", n étant écrit en unaire. La taille des coefficients, donnés sous forme dyadique par exemple, peut être linéairement majorée à partir de n log(n)

Théorème de Sturm-Sylvester approximatif

Soient P et Q deux polynômes réels unitaires tels que **Res**(P,P') $\neq 0$ et **Res**(P,Q) $\neq 0$. Soit [a , b] un intervalle aux extrémités duquel P ne s'annule pas. Alors, en calculant un ε-tableau de signes pour P et un autre pour Q (où ε est donné à partir d'une minoration des 2 résultants) on peut calculer le nombre n_+ de zéros de P avec $Q > 0$ sur l'intervalle et le nombre n_- de zéros de P avec $Q < 0$ sur l'intervalle.
En mathématiques classiques les nombres $n_+ + n_-$ et $n_+ - n_-$ sont obtenus par apllication du théorème de Sturm à P et du théorème de Sylvester à P et Q (cf [**GLRR**]).
Voyons le cas du théorème de Sylvester (le théorème de Sturm est un cas particulier). Il faut calculer la suite des restes de l'algorithme d'Euclide démarrant avec P'Q et P , en modifiant convenablement les signes (suite des restes signés).
Constructivement, la suite des restes signés est en général impossible à calculer, parce qu'il n'y a pas de test d'égalité à 0 pour un réel, et qu'on est amené à hésiter sur le degré exact des polynômes dans la suite. En d'autres termes, la suite de Sturm-Sylvester est instable par rapport aux données.
La version "Habicht" du théorème de Sylvester utilise une version formelle de la suite de Sturm-Sylvester. Les restes signés sont remplacés par les polynômes sous-résultants, avec des modifications de signes convenables. En conséquence, la suite obtenue (suite de Sturm-Habicht) est toujours calculable, parce que les coefficients des polynômes sous-résultants sont des déterminants extraits de la matrice de Sylvester de P et **Rst**(P'Q,P).
On ne peut cependant pas appliquer "en général" le théorème de Sylvester "tel que" parce qu'il faut connaître quels sont les polynômes sous-résultants identiquement nuls, et évaluer le signe des autres aux bornes de l'intervalle considéré.
Il serait donc agréable d'avoir une version "approximative" et en temps polynomial du théorème de Sylvester. Cela signifierait par exemple qu'on est capable de déterminer en temps polynomial un ε tel que, chaque fois qu'on se pose le problème dévaluer exactement le signe d'un nombre c calculé dans le cours de l'algorithme de Sturm-Sylvester-Habicht, on peut faire comme si c était nul chaque fois qu'il est inférieur à ε en valeur absolue.

Régionnement approximatif du plan réel par une courbe algébrique

Considérons un polynôme f(x,y) à coefficients réels, la courbe algébrique réelle qu'il définit et le régionnement du plan réel qui en résulte. Il est a priori impossible de calculer la topologie de ce régionnement sans recourir à des tests de signe sûrs.
On doit pouvoir cependant obtenir un régionnement approximatif, à ε près, en temps polynomial, sous la forme d'un algorithme qui fournisse les informations suivantes. Tout d'abord l'algorithme indique le nombre de régions trouvées et, dans chaque région, un point représentatif. A partir d'une donnée (a,b) telle que $| f(a,b) | > \varepsilon$, l'algorithme calcule le numéro de la région où (a,b) est situé ainsi qu'un chemin menant de (a,b) au point représentatif de la région, chemin entièrement situé dans la partie $\{(x,y); | f(x,y) | > \varepsilon/2 \}$ du plan réel. Enfin s'il existe un chemin joignant 2 points dans la partie $\{(x,y); | f(x,y) | > \varepsilon \}$ les 2 points doivent être situés dans la même région par l'algorithme.
De manière imagée, si f(x,y) est la profondeur du fond marin, on cherche à savoir si un chemin sûr pour un bateau mène d'un point à un autre ($| f(x,y) | > \varepsilon/2$), et à déterminer ce chemin, en n'omettant que des chemins "pas tout à fait assez sûrs".

Il est presque certain qu'une méthode approximative sera nettement plus performante qu'un algorithme basé sur une méthode sûre dans le cas discret, agrémentée d'un théorème de perturbation.
La clé d'une telle méthode approximative pourrait être cherchée du coté d'une détermination des racines approximatives d'un polynôme qui serait donnée au moyen de fonctions $\mathcal{P}$-analytiques par morceaux, ou quelque chose du même genre ($\mathcal{P}$-points d'un espace de fonctions présenté par une partie dénombrable dense "agréable").
On pourrait chercher de la même manière à obtenir une réalisation du théorème fondamental de l'algèbre où les racines seraient données en fonction des coefficients au moyen de fonctions de $\mathbf{P}_n(\mathbb{C})$ vers $\mathbf{Sym}_n(\mathbf{P}_1(\mathbb{C}))$ $\mathcal{P}$-analytiques par morceaux, ou quelque chose du même genre.

Bibliographie, références

[Bak] Bakhvalov : Methodes Numériques. Editions MIR. Moscou. (1973, traduction française, 1976) .

[Bar] Bareiss E. H. : Sylvester's Identity and Multistep Integer-Preserving Gaussian Elimination . Math. Comp. 22 565-578 (1968) .

[Ber] Berkovitz S. J. : On computing the determinant in small parallel time using a small number of processors . Information Processing Letters 18 n°3 147-150 (1984) .

[BB] E. Bishop, D. Bridges : Constructive Analysis (Springer-Verlag; 1985) .

[BL] L. E. J. Brouwer, B. de Loor : Intuitionisticher Beweis des Fundamentalsatzes der Algebra. Proc. Acad. Amsterdam 27, 186-188 (1924) .

[Che] E. W. Cheney : Introduction to Approximation Theory. Mc Graw Hill Book Company. 1966 .

[CL] Collins G. E., Loos R. : Real Zeros of Polynomials p 83-94 dans Computer Algebra, Symbolic and Algebraic Computation édité par Buchberger, Collins, Loos . Springer Verlag 1982 .

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

[DD] Dominique Duval, Claire Dicrescenzo : Le système D5 de calcul formel avec des nombres algébriques. in Thèse (de D. Duval) présentée à l'Université Scientifique, Technologique et Médicale de Grenoble. (1987) .

[DM] Demidovitch, Maron : Eléments de calcul numérique. Editions MIR (1973) .

[GLRR] Gonzalez L., Lombardi H., Recio T., Roy M.F.: Spécialisation de la suite de Sturm et sous-résultants. 1988 . A paraître au RAIRO Informatique théorique. Version détaillée dans ce même numéro de CALSYF.

[Hoo] Hoover J. H. : Feasibly constructive analysis (1987) .

[Kal] Erich Kaltofen : GCD divisors of polynomials given by straight-line programs. JACM , v 35 n°1 , Jan 88 , 231-264 .

[KF1] Ker-I. KO, Harvey Friedman : Computational complexity of real functions Theoretical Computer Science 20, (1982), 323-352 .

[KF2] Ker-I. KO, Harvey Friedman : Computing power series in polynomial time. Adv. Appl. Math. 9, 40-50 (1988) .

[KLL] Kannan R., Lenstra A. K., Lovasz L. : Polynomial Factorisation and Nonrandomness of Bits of Algebraic and Some Transcendental Numbers. Mathematics of Computation, vol 50, n°181, Jan 88, 235-250 .

[Kro] L. Kronecker : Grundzüge einer arithmetischen Theorie des algebraischen Grossen (section 4), Journal für due reine und angewandte Mathematik 92, 1-122 (1822) .

[LLL] Lenstra A. K., Lenstra H. W. Jr. , Lovasz L. : Factoring polynomials with rational coefficients . Math Ann. v 261, 1982, 513-534 .

[Lom1] Lombardi Henri. : Calculabilité dans les structures algébriques dénombrables. Prépublication. Besançon. Juil 88 . 1ère partie de cette thèse.

[Lom2] Lombardi Henri : Sous-résultants, suite de Sturm, spécialisation Prépublication. Besançon. Dec 88 . 2ème partie de cette thèse.

[MRR] R. Mines, F. Richman, W. Ruitenburg : A Course in Constructive Algebra (Springer-Verlag; Universitext; 1988) .

[Mü1] N. Th. Müller : Subpolynomial complexity classes of real functions and real numbers Proc 13th ICALP LNCS 226 (1986) 284-293 .

[Mü2] N. Th. Müller : Uniform computational complexity classes of Taylor series. Lecture Notes in Computer Science n°267 (1987) .

[Ost] A. M. Ostrowski : Solution of Equations in Euclidean and Banach Spaces: 3ème édition de: Solution of equations and systems of equations (Academic Press; 1973) .

[Pan] Pan Victor : Algebraic complexity of computing polynomial zeros. Comput. Math. Applic. vol 14, n°4, 1987, 285-304 .

[Riv] Th. J. Rivlin : The Chebyshev Polynomials. A Wiley Interscience Publication. Wiley & Sons. New York 1974 .

[Sam] Samuelson P. A. : A method for determining explicitly the coefficients of the characteristic equation . Ann. Math. Stat. 13 (1942) 424-429.

[Sch] Schönage A. : The Fundamental Theorem of Algebra in Terms of Computational Complexity . Preliminary Report. Math. Inst. der Univ. Tübingen 1982 .

[Val] Brigitte Vallée. Un problème central en Géométrie algorithmique des Nombres: La réduction des réseaux (autour de l'algorithme LLL). Publication de l'Université de Caen, UFR des Sciendes. Juin 87 .

Henri LOMBARDI
Université de Franche-Comté
UFR des Sciences et Techniques
Laboratoire de Mathématiques
25030 BESANCON CEDEX