*THEORIE DES NOMBRES* *Année 1988/1989*

*BESANCON* *Fascicule 2*

*ALGEBRE ELEMENTAIRE EN TEMPS POLYNOMIAL*

- *Calculabilité dans les structures algébriques dénombrables (65 p.)*
- *Sous-résultants, suite de Sturm, spécialisation (38 p.)*
- *Nombres algébriques et approximations (66 p.)*

*Henri LOMBARDI*

# CALCULABILITE DANS LES STRUCTURES ALGEBRIQUES DENOMBRABLES


Introduction ..... 2

## A) GENERALITES

a) Quelques classes de constructions intéressantes ..... 5
b) $\mathfrak{C}$-ensembles-discrets , $\mathfrak{C}$-fonctions , $\mathfrak{C}$-structures algébriques ..... 7
c) Entiers naturels ..... 11
d) Présentations des entiers relatifs et des nombres rationnels ..... 14

## B) STRUCTURES ALGEBRIQUES COMPLETEMENT $\mathfrak{P}$-CALCULABLES

a) Généralités sur les structures algébriques complètement $\mathfrak{P}$-calculables et sur les structures naturellement c-$\mathfrak{P}$-c ..... 17
b) Espaces vectoriels et modules libres ..... 19
c) Algèbres $\mathbf{M_n}(\mathbb{Z})$ , $\mathbf{M_n}(\mathbb{Q})$ , $\mathbb{Z}[X]$ , $\mathbb{Z}[X_1,X_2,...,X_n]$ , $\mathbb{Q}(X)$, $\mathbb{Q}(X_1,X_2,...,X_n)$ comme $\mathfrak{P}_0$-structures naturellement c-$\mathfrak{P}$-c ..... 20
d) Groupes et monoïdes complètement $\mathfrak{P}$-calculables ..... 24
e) Présentations "en magma" ou "par formules" ..... 25
f) Algèbre d'un monoïde A[M] , ..... 27
g) Pourquoi $\mathbb{Z}$ marche-t-il si bien ? ..... 29

## C) ALGEBRE LINEAIRE EN TEMPS POLYNOMIAL

Introduction ..... 33
a) Calcul matriciel sur un $\mathfrak{P}$-anneau ..... 34
b) Cas commutatif : déterminants, formules de Cramer et inversions de matrices ..... 40
c) Systèmes linéaires à coefficients dans un $\mathfrak{P}$-corps commutatif ..... 44
d) Evolution des coefficients dans la méthode du pivot (méthode de Bareiss) ..... 48

Notes ..... 57
Bibliographie ..... 62
Index ..... 63

# CALCULABILITE DANS LES STRUCTURES ALGEBRIQUES DENOMBRABLES


## Abstract

### Computability in Countable algebraic structures

We study the computability in discrete enumerable algebraic structures from the viewpoint of a given class of constructions $\mathfrak{C}$ . Our walk is to relativize for the class $\mathfrak{C}$ the methods of constructive mathematics. The most important class we study is $\mathfrak{P}$ : the class of polynomial time computable functions.

We introduce the notion of *completely $\mathfrak{C}$-computable algebraic structure* (the $\mathfrak{C}$-computability of evaluation of formulas). We prove that the most elementary algebraic structures are *completely $\mathfrak{P}$-computable in a natural sense*. For example the natural completely $\mathfrak{P}$-computable presentation of the ring of polynomials with integer coefficients is the usual one (dense presentation with integers in binary).

We study the $\mathfrak{P}$-computability of linear algebra. For commutative rings, we give strong links between the three notions:

- $\mathfrak{P}$-computability of determinants
- $\mathfrak{P}$-computability of the product of a list of matrices
- $\mathfrak{P}$-computability of addition, multiplication, exact division, and $\mathfrak{P}$-majoration of determinants (in the case of an integral domain).

(these links are often given for the arithmetic complexity only).



## Résumé

Cette étude est consacrée aux ensembles discrets énumérables lorsqu'on adopte le point de vue des constructions d'une classe donnée $\mathfrak{C}$ . La classe que nous avons essentiellement en vue est celle des constructions en temps polynomial. La démarche générale que nous suivons est de relativiser à une classe de construction donnée les méthodes de mathématiques constructives.

Dans le A , nous donnons les définitions de base, et quelques résultats élémentaires.

Dans le B , nous nous intéressons aux structures algébriques dénombrables effectives.

Les notions de "calcul algébrique", "calcul algébrique formel" et "calcul de classe $\mathfrak{C}$" interfèrent alors entre elles. Cela nous amène à la notion de "structure algébrique complètement $\mathfrak{C}$-calculable", qui s'avère être une bonne notion. Par définition, une présentation d'une structure algébrique est complètement $\mathfrak{C}$-calculable lorsque l'évaluation des formules est $\mathfrak{C}$-calculable. Nous établissons donc un lien entre la notion introduite et les présentations "par formule" ou "en magma" ( § e ).

Nous montrons que les structures algébriques les plus élémentaires sont complètement $\mathfrak{P}$-calculables, et en général "de manière naturelle". Cela implique qu'il y a une $\mathfrak{P}$-présentation naturellement attachée à une structure algébrique élémentaire. Par exemple la $\mathfrak{P}$-présentation naturelle de $(\mathbb{N}, 0, 1, +)$ est la présentation en unaire, tandis que la présentation naturelle de $(\mathbb{N}, 0, 1, +, \times)$ est la présentation en binaire.

De nombreuses structures algébriques "libres de type fini" sont complètement $\mathfrak{P}$-calculables et ont une structure de $\mathfrak{P}$-calculabilité naturelle. Par exemple les algèbres de polynômes (à un nombre fini d'indéterminées) sur $\mathbb{Z}$ , $\mathbb{Q}$ ou sur un anneau fini. Un quotient d'une de ces algèbres sera également complètement $\mathfrak{P}$-calculable lorsque l'idéal noyau est une partie $\mathfrak{P}$-détachable de l'algèbre des polynômes.

Il semble très improbable que la clôture algébrique de $\mathbb{Q}$ puisse être présentée de manière complètement $\mathfrak{P}$-calculable; nous donnons néanmoins un exemple (en B.g) d'extension algébrique infinie de $\mathbb{Q}$ présentée de manière que l'addition et le produit y soient complètement $\mathfrak{P}$-calculables.

Dans le C , qui peut être lu à peu près indépendamment du B , l'objectif est de montrer que l'algèbre linéaire "classique" est une algèbre en temps polynomial.

Nous construisons un bon stock d'anneaux commutatifs sur lesquels "le calcul des déterminants est en temps polynomial", à peu de chose près les mêmes que ceux qui ont été montrés complètement $\mathcal{P}$-calculables dans la partie B . Nous mettons en évidence le lien étroit entre la calculabilité des déterminants en temps polynomial d'une part et la calculabilité du produit d'une liste de matrices en temps polynomial d'autre part.

Enfin, nous étudions en détail la méthode du pivot améliorée à la Bareiss et sa calculabilité en temps polynomial.

# INTRODUCTION

Les mathématiques "ordinaires" ont un contenu constructif. Telle est du moins la thèse des mathématiques constructives (cf. [CA] et [CAL] pour une mise en pratique de cette thèse). Cette affirmation peut être interprétée de la manière suivante: tout théorème de mathématiques "ordinaires" , affirmant l'existence de certains objets "concrets" vérifiant certaines propriétés sous certaines hypothèses, doit pouvoir être *réalisé* sous forme d'un algorithme construisant l'objet en question à partir des données fournies dans les hypothèses. En général une preuve constructive d'un théorème fournit de manière immédiate un algorithme primitif récursif qui réalise le théorème en question.

Cette étude (et d'autres dans la même série) se situe dans le contexte général suivant: expliciter, dans les mathématiques ordinaires, les théorèmes qui peuvent être réalisés par des algorithmes de complexité "faible" (en temps polynomial par exemple). Il nous a semblé naturel de prendre une base de mathématiques constructives pour développer ce travail.

Dans cette étude, consacrée aux structures algébriques dénombrables, nous poursuivons les deux objectifs suivants:

- expliciter les interférences entre calcul algébrique et calculabilité au sens de la complexité
- expliciter dans quelle mesure l'algèbre linéaire ordinaire est en temps polynomial.

La démarche générale que nous suivons est de relativiser à une classe de construction donnée les méthodes de mathématiques constructives.

Nous utiliserons les mathématiques constructives de manière informelle[1] (c.-à-d. comme un mathématicien classique utilise la théorie des ensembles).

Disons en très bref que les mathématiques constructives dans le style Bishop n'énoncent que des théorèmes ayant une signification algorithmique. Elles fournissent donc, selon nous, une base naturelle pour tout travail mathématique centré sur la discussion d'algorithmes.

Nous indiquons maintenant deux ou trois définitions sensibles de mathématiques constructives, parce qu'elles nous guideront lorsque nous restreindrons les constructions à une classe préétablie.

---

[1] Les logiciens ont pour leur part beaucoup travaillé sur des systèmes formels qui peuvent rendre compte des mathématiques pratiquées dans un livre tel que [CA] (le premier livre de Bishop date de 1967, et la logique intuitionniste de Heyting date de 1930). On pourra par exemple consulter [FCM] à ce sujet.

Les notions de construction ou opération sont considérées comme des *notions premières, non définies,* au même titre que la notion d'entier naturel .

Constructivement, un ensemble $(X, \neq_X)$ est donné lorsque:

- on décrit ce qu'il faut faire pour construire un objet de l'ensemble $X$ .
- on décrit, concernant les objets de $X$ , une **relation de séparation** , notée $\neq_X$ , et vérifiant les propriétés suivantes (axiomes pour une relation de séparation):

  pour tous $x, y, z$ dans $X$

  i. $x \neq_X x$ est absurde

  ii. $x \neq_X y$ équivaut à $y \neq_X x$

  iii. $x \neq_X y$ implique $x \neq_X z$ ou $y \neq_X z$

**NB**: le "ou" , dans iii. , est un "ou" constructif, c.-à-d. doit pouvoir être constaté comme résultat d'une construction.

On définit alors une **relation d'égalité**, notée $x =_X y$ , par : " $x \neq_X y$ est absurde ". Cette égalité de $X$ est une relation d'équivalence[2].

Un ensemble $(X, \neq_X)$ est appelé **discret** , lorsque, pour tous $x$ et $x'$ dans $X$ , on a : $x \neq_X x'$ ou $x =_X x'$ . L'ensemble des nombres réels "n'est pas" discret dans la mesure où on ne sait pas décider "en général" si 2 réels donnés sont égaux ou séparés.

On appelle **fonction** de l'ensemble $(X, \neq_X)$ vers l'ensemble $(Y, \neq_Y)$ une opération de $X$ vers $Y$ qui vérifie la propriété d'extensionnalité suivante :

$$F(x) \neq_Y F(x') \Rightarrow x \neq_X x'.$$

Une fonction $f : X \longrightarrow Y$ est dite **surjective** si on connait une opération $r$ de $Y$ vers $X$ vérifiant : pour tout $y \in Y$ , $f(r(y)) =_Y y$ . Notez que $r$ n'est pas nécessairement une fonction.

Une **énumération** d'un ensemble $X$ est à très peu de choses près une application surjective de $\mathbb{N}$ , ensemble des entiers naturels, sur $X$ : elle est donnée précisément comme suit :

- une fonction $f : \mathbb{N} \longrightarrow X \cup \{u\}$ et une opération $r : X \longrightarrow \mathbb{N}$ qui vérifient : pour tout $x$ de $X$ , on a $f(r(x)) =_X x$ .

L'objet $u$ est extérieur à $X$, il a été rajouté pour le cas où on ne sait pas a priori si $X$ est vide ou non . Si $X$ est "habité", c.-à-d. si on connait un élément de $X$ , il revient au même de dire qu'il existe une application surjective de $\mathbb{N}$ sur $X$ . L'opération $r$ n'est pas nécessairement une fonction. Un ensemble qui possède une énumération est dit **énumérable**.

Si maintenant nous considérons une classe de constructions $\mathfrak{C}$ , et que nous estimons que seules les constructions de cette classe sont acceptées, nous obtenons la notion correspondante de $\mathfrak{C}$-ensemble, ou ensemble $\mathfrak{C}$-présenté.

Par exemple, si nous considérons les constructions faisables par une machine de Turing, nous aurons une notion d'ensemble "récursivement présenté".

Pour ce qui concerne une version relativisée à $\mathfrak{C}$ de la notion d'ensemble discret, nous demanderons que l'alternative $x \neq_X x'$ ou $x =_X x'$ puisse être tranchée au moyen d'une $\mathfrak{C}$-construction à partir des entrées $x$ et $x'$ . Si nous voulions relativiser à $\mathfrak{C}$ la notion de

---

[2] Cette étude est consacrée aux ensembles discrets énumérables ; dans ce cas (et dans celui des espaces métriques), il y a une relation de séparation au sens constructif. Dans [CA] , Bishop donne une définition de la notion d'ensemble basée sur une relation d'égalité plutôt que de séparation.

relation de séparation, sans hypothèse de discrétion, le problème serait plus délicat, et la réponse à apporter n'est peut-être pas unique[3].

[3] Signalons néanmoins que la notion d'espace métrique séparable complet se laisse relativiser à une classe $\mathfrak{C}$ de manière simple et directe, ce qui permet de traiter dans ce cadre une grande partie de l'analyse. Par exemple on a des définitions naturelles de $\mathfrak{C}$-nombre réel ou de $\mathfrak{C}$-fonction continue de $[0,1]$ vers $\mathbb{R}$.

# A) GENERALITES SUR LES $\mathfrak{C}$-ENSEMBLES-DISCRETS

Nous supposerons que la classe de constructions $\mathfrak{C}$ concerne des objets du type "mots sur un alphabet fini". Plus précisément, pour tous alphabets finis A et B , si A* désigne le langage engendré par A , nous supposons définies les constructions de classe $\mathfrak{C}$ de $A^{*k}$ vers B*.

## a) Quelques classes de constructions intéressantes

### Stabilité par composition

Comme nous avons en vue des classes de constructions qui fournissent des opérations de $A^{*k}$ vers B* , la question de la stabilité de ces opérations par composition se pose naturellement. Ce n'est pas le cas lorsqu'on étudie les algorithmes comme "sélecteurs de langage".

Or, des classes de complexité comme **DTIME**($n^2$) ne sont pas stables par composition. Nous introduisons donc pour remédier à cet inconvénient des classes de complexité où la taille de la sortie est mieux majorée que le temps de calcul, ou l'espace de calcul.

Nous notons **SPACERES**(f) la classe des algorithmes où la taille de la sortie (space résultat) est majorée par f(n) , où n est la taille de l'entrée.

Précisons ici quelques abréviations, certaines très classiques, que nous utiliserons:

**DTIME**(O(f)) pour $\cup_{c,a}$ **DTIME**(c+a.f)

| | | | |
|---|---|---|---|
| **LINTIME** | = **DT1** = **DTIME**(O(n)) | **DT0** | = $\cup_c$ **DTIME**(n+c) |
| $\mathcal{P}$ | = $\cup_b$ **DTIME**($O(n^b)$) | **DTNLG** | = $\cup_b$ **DTIME**($O(n.\lg^b(n))$) |
| **DSP1** | = **DSPACE**(O(n)) | **PSPACE** | = $\cup_b$ **DSPACE**($O(n^b)$) |
| **RES0** | = $\cup_c$ **SPACERES**(n+c) | **RES1** | = **SPACERES**(O(n)) |
| **RESP** | = $\cup_b$ **SPACERES**($O(n^b)$) | | |
| $\mathcal{P}_0$ | = **RES0** $\cap$ $\mathcal{P}$ | $\mathbf{DTNLG}_0$ | = **DTNLG** $\cap$ **RES0** |
| $\mathbf{DTIME}_0(O(n^k))$ | = **RES0** $\cap$ **DTIME**($O(n^k)$) etc... | | |
| $\mathcal{P}_1$ | = **RES1** $\cap$ $\mathcal{P}$ | $\mathbf{DTNLG}_1$ | = **DTNLG** $\cap$ **RES1** |
| $\mathbf{DTIME}_1(O(n^k))$ | = **RES1** $\cap$ **DTIME**($O(n^k)$) etc... | | |

Nous ferons souvent référence également à la classe $\mathbb{P}\mathbf{r}$ des fonctions primitives récursives, et à la classe **Rec** des fonctions récursives.

Ce sont toutes des classes stables par composition . Et on a l'inclusion évidente:

**PSPACE** $\subset$ **RESP**.

Lorsque $\mathbb{N}$ est présenté en binaire, une opération f de $\mathbb{N}$ vers $\mathbb{N}$ est **RES0** ssi f(n) = O(n) , et elle est **RES1** ssi il existe un k tel que $f(n) = O(n^k)$

## Mesures de la grandeur des entrées et sorties

Par ailleurs, on a parfois intérêt à considérer une mesure de l'entrée qui ne soit pas directement la taille de l'objet (pour un type de description choisi), tout en étant polynomialement relié à la taille.

Expliquons-nous sur un exemple : considérons les algèbres $\mathbf{M_n}(\mathbb{Z})$ . Pour une matrice $A = (a_{ij})$ , la taille $s(A)$ dans une présentation "naturelle" , sera : $s(A) = n^2 + \Sigma s(a_{ij})$

Cependant, si nous considérons $t(A) := n + s(\Sigma |a_{ij}|)$, on peut vérifier facilement que, pour 2 matrices A et B , on obtient l'inégalité : $t(AB) \leq t(A) + t(B)$ . Par ailleurs les "mesures" t et s sont polynomialement reliées. Mais avec la mesure t le produit des matrices est **RES0**, ce qui n'est pas le cas avec la mesure "naturelle" s .

Ainsi, un ensemble sera toujours présenté avec une mesure de la grandeur des objets qui le composent.

Si la mesure n'est pas précisée, c'est qu'il s'agit de la taille "naturelle" au sens de la longueur du mot utilisé pour représenter l'objet .

Notons que 2 objets de X , distincts en tant que mots de A* , mais égaux dans X , ont en général 2 mesures distinctes: par exemple un même nombre rationnel peut être représenté par 2 fractions distinctes, de tailles distinctes .

Lorsque la classe de construction $\mathfrak{C}$ considérée est une classe de complexité, il faudra la comprendre au sens de la mesure considérée lorsqu'est définie la présentation de l'ensemble étudié (comme nous venons de le faire en affirmant que le produit des matrices est **RES0** lorsqu'on utilise la mesure t ) .

## Hypothèses concernant la classe de constructions $\mathfrak{C}$

Nous devons expliciter quelques hypothèses générales concernant la classe $\mathfrak{C}$ des constructions considérées.

Ces hypothèses seront en quelque sorte nos "axiomes de la théorie des $\mathfrak{C}$-ensembles-discrets". Elles seront immédiatement vérifiées pour les classes que nous avons en vue. Elles permettent de faire fonctionner les constructions élémentaires concernant les $\mathfrak{C}$-ensembles-discrets. Comme nous envisageons dans nos applications essentiellement les classes $\mathcal{P}_0$ , $\mathcal{P}_1$ , $\mathcal{P}$ , **DTNLG**, **PSPACE**, $\mathbb{P}\mathbf{r}$ , **Rec** , on pourrait très bien se passer de ce paragraphe, qui manifeste un souci de généralité peut-être abusif.

Nous allons formuler nos hypothèses de manière assez lâche, renvoyant un exposé plus détaillé en note (n.1).

Nous abrègerons " construction de classe $\mathfrak{C}$ " en **$\mathfrak{C}$-construction** .

Nous désignerons par A et B des alphabets finis, A* et B* les langages qu'ils engendrent.

L'ensemble **Lst** (A*), des listes d'éléments de A* (ou encore : suites finies d'éléments de A*), peut être réalisé comme une partie d'un langage A°* (où A° est obtenu en rajoutant à A des symboles représentant [ , ] et ; ) . Si $X_1$ , $X_2$ ,...., $X_n$ sont des parties de A*, l'ensemble $X_1 \times X_2 \times ... \times X_n$ peut être réalisé comme une partie de **Lst**(A*) (listes convenables de n éléments).

Les $\mathfrak{C}$-constructions doivent permettre d'accomplir 2 tâches :

- définir les $\mathfrak{C}$-parties des ensembles A* , et
- définir les $\mathfrak{C}$-opérations entre $\mathfrak{C}$-parties X et Y d'ensembles A* et B*, lorsqu'on a défini pour X et Y une mesure de la grandeur de leurs objets.

La mesure de la grandeur d'un objet de X ( $\mathfrak{C}$-partie de A*) est toujours supposée vérifier les propriétés suivantes:

- c'est un entier naturel $> 0$ , et
- elle est polynomialement reliée à la taille naturelle (qui est la longueur du mot, sauf pour le mot vide v de taille 1)
- l'identité $I : x \to x$ de $(X, \| \ \|_{A^*})$ vers $(X, \| \ \|_X)$ est une $\mathfrak{C}$-opération

Voici maintenant la formulation de nos hypothèses:

- **constructions élémentaires appartenant à $\mathfrak{C}$:**
  toutes les constructions de la classe $\mathbf{DTNLG_0}$ sont dans $\mathfrak{C}$.
- **rapport entre $\mathfrak{C}$-parties et $\mathfrak{C}$-opérations:**
  une partie X de A* est une $\mathfrak{C}$-partie si et seulement si sa fonction caractéristique (opération de A* vers {**oui**,**non**} ) est une $\mathfrak{C}$-opération (ici A* est muni de la mesure naturelle).
- **propriétés de stabilité pour les $\mathfrak{C}$-opérations:**
  * stabilité pour la composition.
  * stabilité pour la définition par cas :
    f et g sont 2 $\mathfrak{C}$-opérations de X vers Y , sl est une $\mathfrak{C}$-opération de X vers {**oui**,**non**} , on définit l'opération $h : X \to Y$ , par :

    $h(x) := f(x)$ si $sl(x) =$ **oui** , et $h(x) := g(x)$ sinon
  * stabilité pour **Lst** :
    si $f : X \to Y$ est une $\mathfrak{C}$-opération, il en est de même pour l'opération
    g: **Lst** (X) $\to$ **Lst** (Y) , définie par
    $g([x_1,x_2,...,x_n]) := [f(x_1),f(x_2), ... ,f(x_n)]$ .

## b) $\mathfrak{C}$-ensembles-discrets , $\mathfrak{C}$-fonctions , $\mathfrak{C}$-structures algébriques

### Présentation d'un ensemble énumérable

D'un point de vue constructif, les objets d'un ensemble X sont en général représentés par des mots écrits sur un alphabet fini déterminé A . Seuls certains mots de A* représentent des objets de X .

S'il existe un test (une opération) P de A* vers {**oui**,**non**} indiquant si le mot m représente ou non un objet de X , l'ensemble est alors énumérable . Lorsqu'on a ainsi décrit les objets d'un ensemble énumérable X , on dit qu'on a défini une **présentation** de X.

Tout ensemble énumérable peut naturellement être "présenté".

## La catégorie des $\mathfrak{C}$-ensembles-discrets

**Définition A.b1 :**

Un **$\mathfrak{C}$-ensemble-discret** (ou ensemble-discret-$\mathfrak{C}$-présenté) est donné lorsque:

- on considère un alphabet fini A
- on considère une opération $P_X$ de classe $\mathfrak{C}$ de $A^*$ vers {**oui**,**non**} acceptant un langage $X \subset A^*$ : les mots de X seront les objets de l'ensemble.
- on a défini une opération $V_X$ de classe $\mathfrak{C}$ , de $X \times X$ vers {**oui**,**non**}, qui vérifie, pour tous x, y, z de X :

  $V_X(x,x) =$ **oui** , $V_X(x,y) = V_X(y,x)$ ,

  $V_X(x,y) = V_X(y,z) =$ **oui** $\Rightarrow$ $V_X(x,z) =$ **oui**

  (l'égalité de x et y comme éléments de X est définie par $V_X(x,y) =$ **oui**) .
- on a défini une mesure de la grandeur des mots de X , polynomialement reliée à la taille. (la mesure doit toujours être un entier $> 0$ )

**Notations :** La mesure de la grandeur de l'objet x du $\mathfrak{C}$-ensemble-discret X sera en général notée $\|x\|_X$ , ou plus simplement $\|x\|$ .
En toute rigueur, le $\mathfrak{C}$-ensemble-discret X devrait être noté $(A, P_X, V_X, \| \|_X)$.

**Remarque :** l'identité entre mots de $A^*$ peut être $\mathfrak{C}$-testée; c'est donc une relation d'égalité possible.

Rappelons qu'une fonction de X vers Y est par définition une opération extensionnelle (c.-à-d.: qui se comporte bien par rapport aux relations de séparation définies sur X et Y). Ceci nous amène à définir la catégorie des $\mathfrak{C}$-ensembles discrets comme suit.

**Définition A.b2 :** Etant donnés deux $\mathfrak{C}$-ensembles-discrets X et Y , on appellera **$\mathfrak{C}$-fonction** de X vers Y une opération de classe $\mathfrak{C}$ de X vers Y qui est une fonction de X vers Y .

On a donc défini la catégorie des $\mathfrak{C}$-ensembles-discrets.

On appellera **$\mathfrak{C}$-équivalence** un isomorphisme dans cette catégorie. Lorsque on a deux classes de constructions $\mathfrak{C}_1$ et $\mathfrak{C}_2$ avec $\mathfrak{C}_1 \subset \mathfrak{C}_2$ , il y a un foncteur d'oubli de la catégorie des $\mathfrak{C}_1$-ensembles-discrets vers celle des $\mathfrak{C}_2$-ensembles-discrets.

## Quelques $\mathfrak{C}$-équivalences (plus de détails en note n.2)

Pour un entier n (abstrait) nous noterons **lg**(n) sa longueur lorsqu'il est écrit en binaire. L'ensemble $\mathbb{N}$ des entiers naturels présentés en binaire est une **DT0**-partie de $\{0,1\}^*$ , qui est **DT0**-équivalente à $\{0,1\}^*$ , donc $\mathfrak{C}$-équivalente à $\{0,1\}^*$ ($\mathfrak{C}$ contient $\mathbf{DTNLG_0}$ ). De même, l'ensemble $\mathbb{N}$ présenté en base b est une **DT0**-partie de $A^*$ , qui est **DT0**-équivalente à $A^*$, où A est un alphabet à b lettres : {0,1,...,b-1} . Le changement de base de numérotation est une fonction de classe $\mathbf{DTNLG_1}$ . En prenant **lg**(n) pour mesure de l'entier n écrit en base b , les présentations de $\mathbb{N}$ en binaire et en base b sont donc $\mathfrak{C}$-équivalentes. De même, en modifiant convenablement la mesure des mots dans $A^*$ , les ensembles $A^*$ sont 2 à 2 $\mathfrak{C}$-équivalents .

Soit par ailleurs $X = (A, P_X, V_X, \| \|_X)$ un $\mathfrak{C}$-ensemble-discret. Notons X' le $\mathfrak{C}$-ensemble-discret $X' := (A, P_X, V_X, \| \|_{X'})$, où seule la mesure de la grandeur des objets a été modifiée. Soit la fonction $I : x \rightarrow x$ , définie de X vers X': I est une $\mathfrak{C}$-équivalence

si la classe $\mathfrak{C}$ contient $\mathcal{P}$, puisque les mesures sont polynomialement reliées entre elles. Il est donc bien clair que l'introduction d'une mesure de la taille des objets n'a d'intérêt pratique que pour les classes de constructions strictement plus petites que $\mathcal{P}$. Dans le cas contraire, la catégorie obtenue sans introduire de mesure de la taille des objets, équivalente à la catégorie des $\mathfrak{C}$-ensembles-discrets, est bien suffisante.

## Sous-$\mathfrak{C}$-ensembles-discrets, applications $\mathfrak{C}$-surjectives, $\mathfrak{C}$-quotients

Si X est le $\mathfrak{C}$-ensemble-discret $(A, P_X, V_X, \| \|_X)$, un **sous-$\mathfrak{C}$-e-d** Y de X est défini lorsqu'on a donné une $\mathfrak{C}$-fonction f de X vers {**oui**,**non**}. Cela définit la $\mathfrak{C}$-partie $Y := \{ x \in X ; f(x) = \textbf{oui} \}$ de A*.

On définit l'égalité et la mesure sur Y comme induites par celles de X. L'injection canonique $Y \to X$ est alors une $\mathfrak{C}$-fonction. Les sous-$\mathfrak{C}$-e-d de X sont stables par intersection, réunion et différence. Un sous-$\mathfrak{C}$-e-d de X est encore appelé une **partie $\mathfrak{C}$-détachable** de X, ou une **$\mathfrak{C}$-partie** de X.

Notez que toute $\mathfrak{C}$-partie Z de A* contenue dans X ne définit pas nécessairement une partie $\mathfrak{C}$-détachable de X, parce que l'égalité dans X peut être plus lâche que celle dans A*, et Z n'est pas forcément saturée pour la relation d'égalité dans X.

Une $\mathfrak{C}$-fonction f de X vers Y est dite **$\mathfrak{C}$-surjective** lorsqu'on connait une $\mathfrak{C}$-opération $r : Y \to X$ qui vérifie, pour tout y de Y : $f(r(y)) =_Y y$ .
Notez que r n'est pas nécessairement une fonction.
La composée de deux fonctions $\mathfrak{C}$-surjectives est une fonction $\mathfrak{C}$-surjective.

Un **$\mathfrak{C}$-quotient** de $X = (A, P_X, V_X, \| \|_X)$ est par définition un $\mathfrak{C}$-ensemble-discret de la forme $X' = (A, P_X, V_{X'}, \| \|_X)$, où l'on a, pour tous x, y de X :

$$x =_X y \quad \Rightarrow \quad x =_{X'} y \ .$$

La projection canonique de X sur X' est alors une $\mathfrak{C}$-fonction $\mathfrak{C}$-surjective. Notez que $V_{X'}$ est une $\mathfrak{C}$-fonction de $X \times X$ vers {**oui**,**non**}. (voir le § qui suit pour $X \times X$ comme $\mathfrak{C}$-ensemble-discret)

## Produit de 2 $\mathfrak{C}$-ensemble-discrets, $\mathfrak{C}$-structures algébriques

On a une notion naturelle de produit de 2 $\mathfrak{C}$-ensembles-discrets : on écrit les mots représentant les éléments x et y de X et Y l'un à la suite de l'autre, séparés par un symbole ne faisant pas partie des alphabets utilisés. Et on pose:

$$\| (x,y) \| = \| x \| + \| y \| \ .$$

Il s'agit d'ailleurs du produit dans la catégorie des $\mathfrak{C}$-ensembles-discrets pour des classes $\mathfrak{C}$ comme $\mathcal{P}$, $\mathcal{P}_1$, $\mathbf{DTIME}_1(O(n^k))$, $\mathbb{P}\mathbf{r}$, **Rec.** (n.3)

A partir de ces notions de sous-$\mathfrak{C}$-e-d et de produit de 2 $\mathfrak{C}$-ensembles-discrets, nous pouvons parler de $\mathfrak{C}$-lois de composition, de $\mathfrak{C}$-relations binaires etc... et donc de $\mathfrak{C}$-monoïdes, $\mathfrak{C}$-groupes, $\mathfrak{C}$-anneaux, $\mathfrak{C}$-ensembles-ordonnés et plus généralement de **$\mathfrak{C}$-structure algébrique**[1] d'un type donné .

Il faut noter qu'il s'agit de structures algébriques sur des ensembles discrets énumérables.

---

[1] Nous ne chercherons pas ici à donner la définition précise la plus générale possible de la notion de $\mathfrak{C}$-structure algébrique. Disons que cette structure ne doit impliquer qu'un nombre fini d'ensembles, avec un nombre fini de lois de compositions, de constantes, et de relations (unaire ou binaire ou ternaire ou ...).

Chaque fois que c'est possible , nous considèrerons que les axiomes de la structure algébrique sont présentés comme purement universels: par exemple pour les groupes, anneaux et corps . Ainsi, dans un $\mathfrak{C}$-groupe, non seulement la loi produit, mais aussi la loi unaire : $x \rightarrow x^{-1}$ , doivent être des $\mathfrak{C}$-fonctions. (n.4)

**Remarques :**

1 - Dans le cas de la classe **Rec**, notre notion de **Rec**-structure est exactement équivalente à la notion de structure algébrique récursivement présentée définie dans [**F-S**].(n.5)

2 - Tout $\mathfrak{C}$-ensemble-discret définit évidemment un ensemble au sens constructif. Lorsqu'un ensemble X est déjà défini constructivement, une **$\mathfrak{C}$-présentation** de cet ensemble est donnée par : - un $\mathfrak{C}$-ensemble-discret X' d'une part, - une bijection entre X et X', d'autre part. Ainsi un $\mathfrak{C}$-ensemble-discret peut-il être considéré comme un ensemble "abstrait" muni d'une structure de $\mathfrak{C}$-calculabilité additionnelle.
De la même manière, lorsqu'un ensemble est muni d'une structure algébrique précise, nous parlerons de **$\mathfrak{C}$-présentation de cette structure algébrique** pour une $\mathfrak{C}$-présentation de l'ensemble X qui fait des lois de composition des $\mathfrak{C}$-fonctions (et des relations unaires, binaires etc ... des $\mathfrak{C}$-relations).

## Structures algébriques naturellement primitives récursives

Notons $\mathbb{N}$ pour l'ensemble des entiers naturels présentés en binaire.

Si nous considérons la classe $\mathbb{P}\mathbf{r}$ des fonctions primitives récursives, nous avons immédiatement le résultat suivant (les axiomes de Peano sont là pour ça en quelque sorte...):

Si $\mathbb{N}'$ est une $\mathbb{P}\mathbf{r}$-présentation de la structure algébrique ( $\mathbb{N}$ , 0 , $n \rightarrow n + 1$ ) , alors la fonction "identité" de $\mathbb{N}$ vers $\mathbb{N}'$ est une $\mathbb{P}\mathbf{r}$-fonction[1].

De manière générale nous dirons qu'une **structure algébrique est naturellement primitive récursive** lorsque il existe une $\mathbb{P}\mathbf{r}$-présentation "naturelle" de cette structure au sens qu'elle est $\mathbb{P}\mathbf{r}$-initiale parmi toutes les $\mathbb{P}\mathbf{r}$-présentations de cette structure. (c.-à-d.: la bijection "identité" qui va de la $\mathbb{P}\mathbf{r}$-structure naturelle vers une autre $\mathbb{P}\mathbf{r}$-présentation est une $\mathbb{P}\mathbf{r}$-fonction). Il est clair que la $\mathbb{P}\mathbf{r}$-présentation "naturelle" est alors unique à $\mathbb{P}\mathbf{r}$-isomorphisme près.

Pour une autre classe de constructions $\mathfrak{C}$ , nous pourrions parler de **structure algébrique naturellement de type $\mathfrak{C}$** . En fait, il s'avère que ce n'est pas "la bonne" notion. La bonne notion est celle de structure algébrique "naturellement complètement $\mathfrak{C}$-calculable", que nous étudierons au B .

Si une structure algébrique est naturellement primitive récursive, tous les automorphismes de la $\mathfrak{C}$-structure naturelle sont primitifs récursifs.

Les structures algébriques "de type fini" qui peuvent être $\mathbb{P}\mathbf{r}$-présentées possèdent une $\mathbb{P}\mathbf{r}$-présentation naturelle (la seule qu'on considère en général). (n.6) . Mais il y a des groupes de présentation finie pour lesquels l'égalité n'est pas récursivement décidable (Théorème de Novikoff), donc qui ne peuvent pas être **Rec**-présentés.

---

[1] Divertissement mathématique : toute $\mathbb{P}\mathbf{r}$-présentation de la structure algébrique ($\mathbb{N}$ , 0 , $n \rightarrow n+1$ , $n \rightarrow n \dot{-} 1$ ) est-elle $\mathbb{P}\mathbf{r}$-équivalente à la présentation standard ?
Voici par ailleurs un exemple de structure algébrique constructivement définie qui "n'est pas" constructivement isomorphe à ( $\mathbb{N}$ , $n \rightarrow n + 1$ ): l'ensemble sous-jacent est $\mathbb{N}$ , le successeur de a est a + 1 sauf éventuellement dans les 2 cas suivants : si a est le 1er contre-exemple à la conjecture de Machin-Bidule, le successeur de a est 0, et le successeur de a – 1 est a + 1 . Dans cet exemple, la fonction successeur est bien définie, mais on ne sait pas déterminer un élément n'ayant pas de prédécesseur tant qu'on n'a pas résolu la conjecture de Machin-Bidule.

Notez que $(\mathbb{N}, \times)$ n'est pas naturellement primitive récursive puisqu'il existe des automorphismes non primitifs récursifs de cette structure.

**Problème ouvert :** construire un groupe de présentation finie pour lequel l'égalité est récursive mais pas primitive récursive. (si la réponse est positive cela donne un exemple de groupe discret **Rec**-présenté mais qui ne peut pas être $\mathbb{P}\mathbf{r}$-présenté)
**NB**: comme la plupart des problèmes ouverts signalés dans ce texte, celui-ci n'est pas "garanti ouvert" par l'auteur.

## Sous-structures et structures quotients

Etant donnée une $\mathfrak{C}$-structure algébrique X, si Y est une partie $\mathfrak{C}$-détachable qui est une sous-structure, on obtient de manière évidente une $\mathfrak{C}$-présentation de la structure algébrique Y, on dit que Y est une **$\mathfrak{C}$-sous-structure** de X.

On définit de même une notion de **$\mathfrak{C}$-structure-quotient** lorsqu'un $\mathfrak{C}$-quotient est une structure quotient.

$\mathfrak{C}$-sous-structures et $\mathfrak{C}$-structures-quotients vérifient les propriétés caractéristiques universelles habituelles.

Pour qu'un quotient d'un $\mathfrak{C}$-groupe soit un $\mathfrak{C}$-quotient il faut et suffit que le noyau de la projection soit un $\mathfrak{C}$-sous-groupe, c.-à-d. un sous-groupe $\mathfrak{C}$-détachable.

ooooooooooooooooooooooooooo

Désormais, sauf mention explicite du contraire, nous utiliserons "ensemble" pour "ensemble discret", et "$\mathfrak{C}$-ensemble" pour "$\mathfrak{C}$-ensemble-discret".

ooooooooooooooooooooooooooo

# c) Entiers naturels

## Présentation en unaire

Nous noterons $\mathbb{N}_1$ l'ensemble des entiers naturels présenté en unaire, par exemple sous forme $\{1\}^*$ ou sous forme **Lst** (c.-à-d. **Lst**(alphabet vide) ), ou sous toute autre forme $\mathfrak{C}$-équivalente.

La structure algébrique $(\mathbb{N}_1, 0, 1, +, \dot{-}, \mathbf{div}, \mathbf{mod}, >)$ est une $\mathfrak{P}_0$-structure; le produit est $\mathfrak{P}$ mais pas **RES1**. ($a \dot{-} b$ est égal à $a - b$ si $a > b$, et à 0 sinon)

## Présentation en binaire

Nous noterons $\mathbb{N}$ l'ensemble des entiers naturels présenté en binaire, ou de toute autre manière $\mathfrak{C}$-équivalente. Par exemple présenté en base b, (dès que $\mathfrak{C}$ contient $\mathbf{DTNLG}_0$), mais en prenant pour mesure, au lieu de la longueur $t_0$ du mot : $1 + \mathrm{Ent}(t_0 . (\log(2)/\log(b)))$.

Du point de vue de la théorie des langages, le $\mathfrak{C}$-ensemble $\mathbb{N}$ joue un rôle essentiel du fait qu'il est $\mathfrak{C}$-équivalent à $\{0,1\}^*$, ou encore à $B^*$ pour n'importe quel alphabet fini B ayant au moins 2 lettres (dès que $\mathfrak{C}$ contient $\mathbf{DTNLG}_0$).

La structure algébrique $(\mathbb{N}, 0, 1, +, \dot{-}, \times, \mathbf{div}, \mathbf{mod}, >)$ est une $\mathfrak{P}_0$-structure.

**Notation:** nous réservons la notation **lg**(n) pour "longueur de l'entier n s'il était écrit en binaire" (même si l'entier n considéré à ce moment-là n'est pas exprimé en binaire).

Le $\mathfrak{C}$-ensemble $\mathbb{N}_1$ est $\mathfrak{C}$-équivalent à la $\mathfrak{C}$-partie $\mathbf{N}_1$ de $\mathbb{N}$ formée des puissances de 2 . Mais il n'existe pas de $\mathfrak{P}$-fonction injective de $\mathbb{N}$ vers $\mathbb{N}_1$ . Les $\mathfrak{P}$-ensembles $\mathbb{N}$ et $\mathbb{N}_1$ ne sont pas $\mathfrak{P}$-équivalents. La fonction $n \to n$ de $\mathbb{N}_1$ vers $\mathbb{N}$ est une $\mathfrak{P}$-fonction, mais pas une $\mathfrak{P}$-équivalence.

## Autres présentations

Il existe bien d'autres présentations de l'ensemble des entiers naturels, 2 à 2 non $\mathfrak{P}$-équivalentes.

Par exemple on peut noter $\mathbb{N}_b$ le $\mathfrak{P}$-ensemble obtenu par une présentation "en bibase b" :
un entier n est présenté sous forme d'une liste de couples ( i , $a_i$ ), où i est un entier écrit en base b , et $a_i$ est un chiffre de cette base, les i arrivant en ordre croissant, et avec : $n = \Sigma\, a_i b^i$. La structure algébrique ( $\mathbb{N}_b$ , + , × , > ) est une $\mathfrak{P}_1$-structure, mais rien ne va plus avec la soustraction ou la division. (cf., dans $\mathbb{N}_b$ la soustraction $b^i - 1$ ). La fonction $n \to n$ de $\mathbb{N}$ vers $\mathbb{N}_b$ est une $\mathfrak{P}$-fonction, mais non pas une $\mathfrak{P}$-équivalence.

Nous allons voir maintenant comment la notion classique de dénombrabilité se scinde en plusieurs notions bien distinctes du point de vue constructif et du point de vue des $\mathfrak{C}$-ensembles.

## Enumérations[1]:

Rappelons qu'une énumération d'un ensemble X est donnée par une fonction $f : \mathbb{N} \to X \cup \{u\}$ et une opération $r : X \to \mathbb{N}$ qui vérifient : pour tout x de X , on a $r(f(x)) =_X x$ . (l'objet u est extérieur à X ).

Lorsque X est un $\mathfrak{C}$-ensemble , f une $\mathfrak{C}$-fonction et r une $\mathfrak{C}$-opération, nous disons que X est **$\mathfrak{C}$-énumérable**, et que **f est une $\mathfrak{C}$-énumération** de X .

Toute $\mathfrak{C}$-fonction surjective de $\mathbb{N}$ sur un $\mathfrak{C}$-ensemble X n'est pas forcément une $\mathfrak{C}$-énumération car elle peut n'être pas $\mathfrak{C}$-surjective. (cf. par exemple la fonction $n \to \mathbf{lg}(n)$ de $\mathbb{N}$ vers $\mathbb{N}$ : c'est une $\mathfrak{P}$-fonction surjective qui n'est pas $\mathfrak{P}$-surjective)

Par contre on a :

Tout $\mathfrak{P}$-ensemble X est $\mathfrak{P}$-énumérable.
Plus généralement , si $\mathfrak{C}$ est une classe de constructions contenant $\mathfrak{P}$ , tout $\mathfrak{C}$-ensemble X est $\mathfrak{C}$-énumérable.

En effet, remarquons tout d'abord que la mesure de la grandeur des objets de X n'intervient pas, puisque nous raisonnons à une $\mathfrak{C}$-équivalence près, et que $\mathfrak{C}$ contient $\mathfrak{P}$. D'autre part, si X est construit sur l'alphabet A , on pourra composer une $\mathfrak{P}$-équivalence $\mathbb{N} \to A^*$ avec la $\mathfrak{C}$-fonction de $A^*$ dans $X \cup \{u\}$ définie comme suit:

- si $x \in X$ , $x \to x$ , sinon $x \to u$ .

On vérifie que la composée est bien une $\mathfrak{C}$-énumération.

**Remarque** : Un mathématicien classique qui veut se faire une idée de ce que peut bien signifier un ensemble énumérable discret pour un constructiviste peut se tenir le discours suivant : admettons une notion a priori d'effectivité (ce qui est plus facile que d'admettre une notion a priori d'ensemble à la Cantor - Zermelo - Frankel) ; notons **Constr** la classe de toutes les fonctions effectivement calculables portant sur des langages $A^*$ ; alors la catégorie

[1] La terminologie énumération, dénombrement, numérotation choisie ici est "assez" arbitraire, et ne prétend naturellement pas être exhaustive.

des ensembles énumérés discrets pour un constructiviste est équivalente à celle des **Constr**-ensembles-discrets , au sens des définitions ci-dessus, qui peuvent être lues avec des lunettes "classiques".

## Dénombrements

Un **dénombrement** d'un ensemble X est par définition une énumération (f,r) de X telle que r soit une fonction.

Un ensemble discret qui possède une énumération (f,r) possède un dénombrement (f,r') : r'(x) est le plus petit entier n inférieur ou égal à r(x) tel que : $x =_X f(n)$ . Par ailleurs un ensemble dénombrable X est nécessairement discret. Autrement dit, "dénombrable" équivaut à "énumérable et discret".

La notion de dénombrement, relativisée à la classe de constructions $\mathfrak{C}$, donne les notions de **$\mathfrak{C}$-dénombrement**, et de $\mathfrak{C}$-ensemble **$\mathfrak{C}$-dénombrable**.

Par le même argument que ci-dessus, tout **PSPACE**-ensemble-discret est **PSPACE**-dénombrable. Et de même pour toute classe $\mathfrak{C}$ stable par récurrence bornée (une définition par récurrence bornée est une définition par récurrence primitive où on astreint la fonction définie à rester majorée par une fonction donnée préalablement). Par contre l'ensemble énumérable $\mathbb{P}\mathbf{r}(\mathbb{N},\mathbb{N})$ des $\mathbb{P}\mathbf{r}$-fonctions de $\mathbb{N}$ vers $\mathbb{N}$ n'est pas **Rec**-dénombrable. (l'égalité n'y est pas **Rec**-décidable)

Si (f,r) est un $\mathfrak{C}$-dénombrement du $\mathfrak{C}$-ensemble X , la $\mathfrak{C}$-opération $x \to f(r(x))$ "choisit" un élément particulier dans chaque classe d'équivalence de la relation $=_X$ . Autrement dit, les différents "représentants" d'un élément de $(X, =_X)$ possèdent une "forme réduite canonique", qui peut être $\mathfrak{C}$-calculée.

Un $\mathfrak{P}$-ensemble-discret n'est pas "a priori" $\mathfrak{P}$-dénombrable: cette question a manifestement à voir avec le fameux problème $\mathfrak{P} = \mathfrak{NP}$ ? . (cf. n.7)

## Numérotations:

Par définition, une **numérotation** d'un ensemble X est une énumération (f,r) qui vérifie:

i. si $f(n) = u$ , alors pour tout $m > n$ , $f(m) = u$

ii. si $f(p) \neq u$ et $f(p) =_X f(q)$ , alors $p = q$

Toute numérotation est un dénombrement.

Les ensembles finis sont numérotables. Les ensembles infinis dénombrables sont numérotables. L'ensemble des contre-exemples à la conjecture de Goldbach "n'est pas" numérotable.

La notion de numérotation, relativisée à la classe de constructions $\mathfrak{C}$, donne les notions de **$\mathfrak{C}$-numérotation**, et de $\mathfrak{C}$-ensemble **$\mathfrak{C}$-numérotable**.

L'ensemble $\mathbb{Q}$ des nombres rationnels est de manière naturelle un $\mathfrak{P}$-ensemble, qui est $\mathfrak{P}$-dénombrable, mais qui ne semble pas $\mathfrak{P}$-numérotable. Cela confirmerait l'impression intuitive que l'ensemble $\mathbb{Q}$ est un petit peu plus compliqué que $\mathbb{N}$ ou que l'ensemble des nombres décimaux.

Si $f : \mathbb{N} \to \mathbb{N}$ est une fonction récursive qui croît plus vite que toute fonction $\mathbb{P}\mathbf{r}$ , construite par récurrence double, son image peut être une partie $\mathbb{P}\mathbf{r}$-détachable de $\mathbb{N}$ , (et donc un $\mathbb{P}\mathbf{r}$-ensemble), mais elle n'est pas $\mathbb{P}\mathbf{r}$-numérotable. (n.8)

De la même manière, et plus simplement, $\mathbb{N}_1$ est un ensemble $\mathfrak{P}$-dénombrable qui n'est pas $\mathfrak{P}$-numérotable.

### $\mathcal{P}$-ensembles $\mathcal{P}$-réductibles

Nous introduisons enfin une notion qui est une version affaiblie de la $\mathcal{P}$-dénombrabilité. Elle nous sera utile dans certains théorèmes par la suite.

Un $\mathcal{P}$-ensemble X est dit **$\mathcal{P}$-réductible** si on a un polynôme Q et une $\mathcal{P}$-opération

r : $X \to X$ , qui vérifient :

- pour tout x de X , $r(x) =_X x$
- si $y =_X x$ , alors $\|r(x)\| < Q(\| y \|)$

On peut dire que l'opération r remplace le représentant x par un autre représentant r(x) , mais de taille raisonnable: c'est une sorte de forme réduite non canonique, mais utilisable pour les calculs de classe $\mathcal{P}$.

La notion de $\mathcal{P}$-réductibilité est une notion qui apparaît naturellement dans certaines preuves de $\mathcal{P}$-calculabilité. Néanmoins, il semble que tous les exemples utiles d'ensembles $\mathcal{P}$-réductibles soient également, de manière immédiate, des ensembles $\mathcal{P}$-dénombrables. La notion de $\mathcal{P}$-réductibilité n'est donc pas nécessaire pour les applications les plus courantes des théorèmes où elle intervient. Elle constitue sans doute un raffinement peu utile de la notion de $\mathcal{P}$-dénombrabilité.

## d) Présentations des entiers relatifs et des nombres rationnels

### Symétrisation d'un $\mathcal{C}$-monoïde commutatif régulier

La construction du symétrisé du monoïde commutatif régulier M , en munissant $M \times M$ de la relation d'égalité convenable, fonctionne sans problème du point de vue des $\mathcal{C}$-structures algébriques.

En termes savants: le foncteur d'oubli des $\mathcal{C}$-groupes abéliens vers les $\mathcal{C}$-monoïdes commutatifs réguliers possède un adjoint à gauche.

On notera que lorsqu'un $\mathcal{P}$-monoïde commutatif n'est pas régulier, l'égalité dans le groupe obtenu classiquement par symétrisation peut ne pas être décidable[1].

Soit M un $\mathcal{C}$-monoïde commutatif régulier, G un $\mathcal{C}$-groupe, $f : M \to G$ un homomorphisme qui fait de G le symétrisé de M . Pour que f fasse de G le $\mathcal{C}$-symétrisé de M , il faut et suffit que :

- f est une $\mathcal{C}$-fonction, et
- il existe 2 $\mathcal{C}$-opérations $g_1$ et $g_2$ de G vers M telles que, pour tout x dans G , on ait : $x =_G f(g_1(x)) - f(g_2(x))$.

---

[1] Soit $(u_p)$ une $\mathcal{P}$-suite d'entiers, d'image non récursive. Considérons le monoïde commutatif librement engendré par une suite $(a_n)$ et codé par la partie de **Lst**($\mathbb{N}$) formée par les listes croissantes d'entiers. Introduisons la relation d'équivalence stable engendrée par les relations $a_{3n+1}.a_{3p+2} = a_{3n}.a_{3p+2}$ si $n = u_p$ . On obtient un $\mathcal{P}$-monoïde commutatif. Mais dans le symétrisé, $a_{3n+1} = a_{3n}$ si et seulement si n est une valeur prise par la suite $(u_p)$ .

Les propriétés d'admettre un $\mathfrak{P}$-dénombrement ou une $\mathfrak{P}$-numérotation ne passent pas "a priori" d'un $\mathfrak{P}$-monoïde commutatif régulier à son symétrisé[1].

Nous dirons qu'un $\mathfrak{C}$-monoïde commutatif M noté multiplicativement est **$\mathfrak{C}$-divisible** lorsqu'il existe une $\mathfrak{C}$-opération D de $M \times M$ vers $M \cup \{u\}$ vérifiant :

si $D(a,b) = u$ , alors pour tout $x \in M$ : $a.x \neq b$ , et , si $D(a,b) \in M$ ,

alors : $a.D(a,b) = b$ .

Un $\mathfrak{C}$-monoïde commutatif régulier M est $\mathfrak{C}$-divisible si et seulement si M "est" une partie $\mathfrak{C}$-détachable de son symétrisé: plus précisément: si l'homomorphisme $f : M \to G$ est une $\mathfrak{C}$-équivalence entre M et une $\mathfrak{C}$-partie de G .

## La présentation standard $\mathbb{Z}$

La présentation des entiers relatifs sous forme d'un nombre en binaire avec un signe, sera considérée comme la présentation standard, et sera notée $\mathbb{Z}$ .

Elle fait de ( $\mathbb{Z}$ , + , − , × , **div** , **mod** , < ) une $\mathfrak{P}_0$-structure. De plus ce $\mathfrak{P}_0$-groupe est $\mathfrak{P}_0$-isomorphe au $\mathfrak{P}_0$-symétrisé de $\mathbb{N}$ [(2)].

De manière générale nous noterons $\mathbb{Z}$ toute présentation des entiers relatifs $\mathfrak{P}_1$-isomorphe à la présentation standard et faisant de ( $\mathbb{Z}$ , + , − , × , **div** , **mod** , < ) une $\mathfrak{P}_0$-structure.

C'est le cas par exemple pour la présentation en base 3 avec les chiffres 0 , 1 , -1 ou encore en base 2 avec les chiffres 0 , 1 , -1 et la relation d'égalité convenable (cette présentation peut être utile pour l'écriture de valeurs approchées successives de nombres réels).

## Autres présentations des entiers relatifs

Les autres présentations de l'ensemble des entiers naturels que nous avons décrites donnent par symétrisation des présentations des entiers relatifs non $\mathfrak{P}$-isomorphes à la présentation standard. Nous noterons $\mathbb{Z}_1$ le symétrisé de $\mathbb{N}_1$ : on obtient une présentation $\mathfrak{P}_1$-isomorphe en prenant un entier codé en unaire avec un signe[3].

## Corps des fractions d'un $\mathfrak{C}$-anneau intègre

La construction du corps des fractions d'un anneau intègre M , en munissant $M \times ( M - \{0\})$ de la relation d'égalité convenable, fonctionne sans problème du point de vue des $\mathfrak{C}$-structures algébriques pour les classe $\mathfrak{C}$ suivantes : $\mathfrak{P}$ , $\mathfrak{P}_1$ , **PSPACE** , **RES1** , $\mathbb{P}r$ , **Rec** .

On a par contre de petits ennuis avec l'addition pour la classe $\mathfrak{P}_0$: par exemple dans le corps des fractions de $\mathbb{Z}$ l'addition est seulement dans **SPARES**(2.n) (additionner 1/1 et $1/2^n$ pour s'en convaincre).

Un $\mathfrak{P}$-anneau intègre est dit $\mathfrak{P}$-divisible lorsque le monoïde multiplicatif M - {0} est $\mathfrak{P}$-divisible . L'anneau est alors identifiable à une $\mathfrak{P}$-partie de son corps de fractions.

---

[1] La première question ($\mathfrak{P}$-dénombrement) aura une réponse positive si $\mathfrak{P} = \mathfrak{N}\mathfrak{P}$ (cf. n.7) . La deuxième question pourrait faire l'objet d'un divertissement mathématique.

[2] Divertissement mathématique : Soit $\mathbb{Z}'$ une autre présentation des entiers relatifs et supposons que la structure :
( $\mathbb{Z}'$ , + , − , × , div , mod , < ) soit une $\mathfrak{P}_0$-structure et $\mathbb{Z}'$ un $\mathfrak{P}$-ensemble $\mathfrak{P}$-réductible, alors la fonction $z \to z$ de $\mathbb{Z}$ vers $\mathbb{Z}'$ est-elle nécessairement un $\mathfrak{P}$-isomorphisme ?

[3] Divertissement mathématique : notons $\mathbb{Z}_2$ le symétrisé de $\mathbb{N}_2$ . On voit facilement que $\mathbb{N}_2$ est $\mathfrak{P}$-numérotable. Est-ce que $\mathbb{Z}_2$ est $\mathfrak{P}$-numérotable ?

## $\mathbb{Q}$ comme $\mathcal{P}_0$-structure

Si on mesure la grandeur de la fraction a/b par: $\|a/b\| := \lg(|a|+b)$ , on constate immédiatement que:

En notant $\mathbb{Q}$ l'ensemble des rationnels présenté comme $\mathbb{Z} \times \mathbb{N}^+$ muni de la relation d'égalité convenable et de la mesure définie ci-dessus, on obtient une $\mathcal{P}_0$-présentation $\mathcal{P}_1$-équivalente à celle obtenue avec la mesure naturelle, et la structure ( $\mathbb{Q}$ , + , − , × , / , Ent , < , numérateur de la fraction réduite ) est une $\mathcal{P}_0$-structure[1].

On notera $\mathbb{D}$ l'ensemble des nombres dyadiques dans sa présentation naturelle (binaire avec virgule et signe) et avec une mesure qui fait de :

$$(\mathbb{D} , + , - , \times , \text{Ent} , < , (x, n) \to x/2^n : \mathbb{D} \times \mathbb{N}_1 \to \mathbb{D} )$$

une $\mathcal{P}_0$-structure. (par exemple la mesure héritée de celle de $\mathbb{Q}$ ).

---

[1] $\mathbb{Q}$ est $\mathcal{P}_0$-dénombrable puisque le calcul de la réduite d'une fraction est $\mathcal{P}_0$ .
$\mathbb{Q}$ est $\mathbb{P}\mathbf{r}$-numérotable et c'est un corps naturellement primitif récursif. Donner une numérotation de $\mathbb{Q}$ revient à donner une numérotation de $\mathbb{Z} \times \mathbb{N}^+$ pour laquelle:
- (a) on sait numéroter en ordre croissant les fractions réduites, et :
- (b) on sait pour chaque fraction réduite le numéro qui lui est attribué.

$\mathbb{Q}$ est **PSPACE**-numérotable. Problème : $\mathbb{Q}$ est-il $\mathcal{P}$-numérotable ?

# B) STRUCTURES ALGEBRIQUES COMPLETEMENT $\mathfrak{P}$-CALCULABLES

## a) Généralités sur les structures algébriques complètement $\mathfrak{P}$-calculables et sur les structures naturellement c-$\mathfrak{P}$-c

### Structures algébriques complètement $\mathfrak{C}$-calculables

Si X est une $\mathfrak{C}$-structure algébrique , avec *un nombre fini* de lois de composition, on peut définir un $\mathfrak{C}$-ensemble **Calc**(X) dont les éléments sont les écritures de calculs à effectuer dans cette $\mathfrak{C}$-structure. Par exemple, dans le corps $\mathbb{Q}$ :

$$( 1/2 +1/( 3/4 +1/( 5 /( 6 - 17/7 )))) \times ( 3/5 + 5/7 - 15 /(1+ 13/4 ))$$

**Définition B.a1 :**

On dira que la structure algébrique X est **complètement $\mathfrak{C}$-calculable** si l'opération naturelle : "faisons la calcul indiqué" qui transforme un élément de **Calc**(X) en un élément de $X \cup \{u\}$ (union disjointe; u vaut pour "non défini") est une $\mathfrak{C}$-opération[1].

Une $\mathfrak{P}$-structure n'est pas nécessairement complètement $\mathfrak{P}$-calculable, comme nous le verrons sur plusieurs exemples (les polynômes en présentation creuse, ou les réels algébriques en présentation naïve notamment). Cela tient à une possible explosion de la taille des objets lors des calculs successifs. On démontre par contre immédiatement.

**Proposition B.a1 :**

Pour qu'une $\mathfrak{P}$-structure algébrique X soit complètement $\mathfrak{P}$-calculable il faut et suffit que l'opération naturelle : "faisons la calcul indiqué" , de **Calc**(X) vers $X \cup \{u\}$ soit **RESP** (c.-à-d. polynomialement majorée en taille).

Toute $\mathfrak{P}_0$-structure est complètement $\mathfrak{P}_1$-calculable.

Toute $\mathbb{P}$r-structure est complètement $\mathbb{P}$r-calculable .

**Remarque** : Dans la plupart des exemples de $\mathfrak{P}_0$-structures que nous étudions, on a en fait une majoration de la mesure de la sortie par la mesure de l'entrée (sans avoir à rajouter une constante), ce qui implique que la structure est en fait complètement $\mathfrak{P}_0$-calculable.

**Exemple** : fractions continues dans $\mathbb{Q}$

Considérons **Lst**($\mathbb{Q}$) , $\mathfrak{P}_0$-ensemble des listes d'éléments de $\mathbb{Q}$. On a une application "fraction continue" de **Lst**($\mathbb{Q}$) vers $\mathbb{Q}$ [2] , donnée par :

---

[1] On aurait une définition analogue pour une $\mathfrak{C}$-structure algébrique impliquant plusieurs $\mathfrak{C}$-ensembles. Par exemple avec 3 $\mathfrak{C}$-ensembles X, Y, Z l'opération "faisons le calcul indiqué" a pour source l'ensemble **Calc**(X,Y,Z) analogue de **Calc**(X) , et pour but l'ensemble $X \cup Y \cup Z \cup \{u\}$ (union disjointe).

[2] En fait, cette application est définie sur des $\mathfrak{P}$-parties convenables de **Lst**($\mathbb{Q}$) ; par exemple: tous les $q_i$ sont $> 0$ à partir du $2^{ème}$.

$$(q_1,q_2,...,q_n) \rightarrow q_1+1/(q_2+1/(...+1/q_n)..)$$

En notant $[q_1;q_2,...,q_n]$ la fraction continue du 2ème membre, on obtient :

$$\| [q_1;q_2,...,q_n] \| \leq \| q_1 \| + \| q_2 \| + ... + \| q_n \|$$

puisque $\| r \| = \| 1/r \|$ et $\| a+b \| \leq \| a \| + \| b \|$ . Donc, l'application "fraction continue" est une $\mathcal{P}_0$-fonction .

On peut également montrer que l'application "fraction continue" réalise une $\mathcal{P}$-équivalence entre $\mathbb{Q}$ et le $\mathcal{P}$-ensemble constitué par les développements en fraction continue standards (tous les $q_i$ sont entiers, $>0$ à partir du 2ème, et le dernier $\neq 1$). (n.9)

## Caractérisation des structures algébriques complètement $\mathcal{P}$-calculables

**Théorème B.a1 :**

Pour qu'une structure algébrique soit complètement $\mathcal{P}$-calculable, il faut et suffit qu'elle soit $\mathcal{P}$-isomorphe à une $\mathcal{P}_0$-structure.

La condition est évidemment suffisante (cf. prop. B.a1 ). Supposons réciproquement que la structure (X,Y,Z) soit complètement $\mathcal{P}$-calculable, alors l'ensemble $X \cup Y \cup Z \cup \{u\}$ peut être $\mathcal{P}$-présenté par **Calc**(X,Y,Z) muni de la relation d'égalité convenable (les expressions une fois calculées sont égales), et les lois de composition peuvent être (**DT1** $\cap$ **RES0**)-présentées en les écrivant entièrement sous forme de calculs à effectuer: bref les calculs ne sont effectués qu'au moment de tester l'égalité.

**Remarques :**

(1) En fait seule la partie "directe" est réellement utilisée; on transforme la $\mathcal{P}$-structure en une $\mathcal{P}_0$-structure en modifiant seulement la mesure de la grandeur des objets. La partie réciproque a un intérêt plutôt théorique.

(2) Nous avons raisonné avec une structure algébrique donnée par un nombre fini de lois de compositions. Il est clair que PrB.a1 et ThB.a1 restent vrais si la structure algébrique comporte en outre un nombre fini de constantes et des relations, et si les lois de composition sont définies sur des parties définies par ces relations.

**Notation abrégée** : nous noterons désormais c-$\mathcal{P}$-c comme abréviation pour "complètement $\mathcal{P}$-calculable".

## Propriétés de stabilité élémentaires

Les structures algébriques c-$\mathcal{P}$-c sont stables pour les constructions suivantes : produit de 2 structures ; $\mathcal{P}$-quotient d'une structure ; sous-structure $\mathcal{P}$-détachable .

Si K est un $\mathcal{P}$-anneau intègre où l'addition et le produit sont c-$\mathcal{P}$-c , il en est de même pour son corps de fractions.

**NB**: Pour un $\mathcal{P}$-anneau K , la condition "être c-$\mathcal{P}$-c en tant qu'anneau" est a priori plus forte que la condition "addition et produit, chacune prise isolément, sont c-$\mathcal{P}$-c dans K"[1].

## Structures algébriques naturellement c-$\mathcal{P}$-c

Considérons un $\mathcal{P}$-anneau A dans lequel l'addition et la multiplication sont c-$\mathcal{P}$-c . Alors l'unique homomorphisme $f : \mathbb{Z} \rightarrow A$ est une $\mathcal{P}$-fonction: en effet f(10011001) , par

---

[1] Dans le premier cas, on évalue une formule comportant des additions, soustractions ou produits. Dans le deuxième cas, on évalue le produit ou la somme d'une liste.

exemple, est égal à :

$$1_A + 2_A \times 2_A \times 2_A + 2_A \times 2_A \times 2_A \times 2_A + 2_A \times 2_A \times 2_A \times 2_A \times 2_A \times 2_A \times 2_A .$$

Or il est clair que l'écriture ci-avant a une taille polynomialement reliée à la taille de 10011001.

En langage savant, $\mathbb{Z}$ est objet initial dans la catégorie des $\mathfrak{P}$-anneaux complètement $\mathfrak{P}$-calculables. (en fait l'addition et la multiplication de A ont seulement besoin d'être, chacune de leur coté, c-$\mathfrak{P}$-c ) .

En particulier, si nous considérons la structure algébrique abstraite "anneau des entiers relatifs", nous voyons que la $\mathfrak{P}$-présentation standard est $\mathfrak{P}$-initiale parmi les $\mathfrak{P}$-présentations qui en font un anneau c-$\mathfrak{P}$-c . En ce sens la $\mathfrak{P}$-présentation standard est naturellement c-$\mathfrak{P}$-c .

Cette terminologie est à rapprocher de celle que nous avons introduite en A.b lorsque nous avons parlé des structures naturellement primitives récursives.

De manière générale, nous posons les définitions suivantes:

**Définition B.a2** : Une structure algébrique énumérable discrète "abstraite" (c.-à-d. abstraction faite de toute présentation de cette structure) sera dite **naturellement c-$\mathfrak{P}$-c** lorsqu'il existe un objet initial dans la catégorie suivante: les objets sont les $\mathfrak{P}$-présentations de cette structure qui la rendent c-$\mathfrak{P}$-c , les flèches sont les applications "identité" qui sont des $\mathfrak{P}$-fonctions.
Cet objet initial est alors défini de manière unique à $\mathfrak{P}$-équivalence près, et nous l'appellerons **la présentation c-$\mathfrak{P}$-c naturelle** de la structure abstraite.

**Exemples** : La présentation en unaire de (entiers naturels, + ), les présentations en binaire de (entiers naturels, + , × ) et de (entiers relatifs, + , − , × ) sont les présentations c-$\mathfrak{P}$-c naturelles de ces structures algébriques. (cf. raisonnement ci-dessus pour les présentations en binaire et PrB.d2 pour la présentation en unaire)

## b) Espaces vectoriels et modules libres

### Généralités

Lorsque K est un anneau énumérable discret présenté, et X un ensemble énumérable discret présenté , nous réservons la notation $K^{(X)}$ pour la présentation suivante du K-module librement engendré par X : c'est la partie de **Lst**(K × X) , obtenue en ne gardant que les listes où tous les "coefficients" $k_i \in K$ sont non nuls, et où les $x_i \in X$ sont sans répétition (et on précise un couple (0,c) pour représenter l'élément nul).

Pour $X = \mathbb{N}_1$, (dans le cas de $\mathfrak{P}$-ensembles) , on obtient une $\mathfrak{P}$-présentation $\mathfrak{P}$-équivalente à celle obtenue sous la forme **Lst**(K) . Et dans le cas où X est fini et a n éléments, on obtient une $\mathfrak{P}$-présentation $\mathfrak{P}$-équivalente à $K^n$ .

Les 2 propositions suivantes sont de démonstration immédiate.

**Proposition B.b1** : Soit K un $\mathfrak{P}$-anneau où l'addition est c-$\mathfrak{P}$-c , et X un $\mathfrak{P}$-ensemble-discret . Alors $K^{(X)}$ est le $\mathfrak{P}$-K-module librement engendré par X dans la catégorie des $\mathfrak{P}$-K-modules où l'addition est c-$\mathfrak{P}$-c .

**Proposition B.b2** : Soit K un $\mathcal{P}$-anneau intègre où l'addition et la multiplication sont c-$\mathcal{P}$-c , et L son corps de fractions . Alors $L^{(X)}$ est $\mathcal{P}$-équivalent à la présentation en forme "réduite au même dénominateur" , c.-à-d. à $K^{(X)} \times (K - \{0\})$ muni de la relation d'égalité convenable .

**Remarque** : Lorsque la dimension est finie et fixée, les hypothèses "c-$\mathcal{P}$-c" dans les propositions ci-dessus sont inutiles. De plus, toute application linéaire de $K^n$ vers un $\mathcal{P}$-K-module est une $\mathcal{P}$-fonction , et, lorsque K est un $\mathcal{P}$-corps, tout sous-espace libre de $K^n$ est un $\mathcal{P}$-sous-espace $\mathcal{P}$-libre (par définition **un $\mathcal{P}$-espace est $\mathcal{P}$-libre s'il est $\mathcal{P}$-isomorphe à un espace** $K^{(X)}$ ) .

Modules libres sur $\mathbb{Z}$ et $\mathbb{Q}$ comme $\mathcal{P}_0$-structures.

Les $\mathbb{Z}$-modules $\mathbb{Z}^n$ et $\mathbb{Z}^{(\mathbb{N})}$ (présenté sous la forme **Lst**($\mathbb{Z}$) ) sont des $\mathcal{P}_0$-$\mathbb{Z}$-modules lorsqu'on les munit des mesures : $\|( a_1,a_2,...,a_n) \| = n + \mathbf{lg}( \Sigma \mid a_i \mid )$ (pour le cas **Lst**, n est la longueur de la liste)

Les $\mathbb{Q}$-espaces vectoriels $\mathbb{Q}^n$ (présenté sous la forme $\mathbb{Z}^n \times \mathbb{N}^+$ ) et $\mathbb{Q}^{(\mathbb{N})}$ (présenté sous la forme **Lst**($\mathbb{Z}$)$\times \mathbb{N}^+$ ) sont des $\mathcal{P}_0$-$\mathbb{Q}$-espaces vectoriels lorsqu'on les munit des mesures :

$\| ( ( a_1,a_2,...,a_n), d ) \| = n + \mathbf{lg}( d + \Sigma \mid a_i \mid )$

## c) Algèbres $\mathbf{M_n}(\mathbb{Z})$ , $\mathbf{M_n}(\mathbb{Q})$ , $\mathbb{Z}[X]$ , $\mathbb{Z}[X_1,X_2,...,X_n]$ , $\mathbb{Q}(X)$, $\mathbb{Q}(X_1,X_2,...,X_n)$ comme $\mathcal{P}_0$-structures naturellement c-$\mathcal{P}$-c

Algèbres $\mathbf{M_n}(\mathbb{Z})$ et **Flin**($\mathbb{Z}$) comme $\mathcal{P}_0$-structures

Les notions de $\mathbb{Z}$-algèbre unitaire et d'anneau sont identiques parce que $\mathbb{Z}$ est objet initial dans la catégorie des anneaux.

De même donc pour les notions de $\mathcal{P}$-$\mathbb{Z}$-algèbre unitaire c-$\mathcal{P}$-c et de $\mathcal{P}$-anneau c-$\mathcal{P}$-c.

Nous pouvons faire de l'anneau $\mathbf{M_n}(\mathbb{Z})$ une $\mathcal{P}_0$-structure si nous prenons pour mesure de la matrice $A = (a_{ij})$ : $\|A\| = n + \mathbf{lg}(\Sigma \mid a_{ij} \mid)$. Nous réservons la notation $\mathbf{M_n}(\mathbb{Z})$ pour cette présentation de cet anneau (ou pour une présentation $\mathcal{P}$-équivalente).

C'est de plus une structure d'anneau naturellement c-$\mathcal{P}$-c : en effet, toute présentation de cet anneau qui en fait un anneau c-$\mathcal{P}$-c en fait aussi un $\mathbb{Z}$-module c-$\mathcal{P}$-c , et la présentation $\mathbf{M_n}(\mathbb{Z})$ est librement engendrée par sa base canonique dans la catégorie des $\mathbb{Z}$-modules c-$\mathcal{P}$-c.

Avec la même mesure, nous pouvons considérer l'algèbre **Flin**($\mathbb{Z}$) , réunion emboîtée des $\mathbf{M_n}(\mathbb{Z})$ (cf. § C.a pour les détails), et nous obtiendrons une $\mathcal{P}_0$-structure.

Le corps $\mathbb{Q}$ comme objet initial

Considérons un $\mathcal{P}$-corps K , de caractéristique nulle, et dans lequel l'addition et la multiplication sont c-$\mathcal{P}$-c . Alors l'unique homomorphisme $f : \mathbb{Q} \to K$ est une $\mathcal{P}$-fonction: cela résulte immédiatement de la proposition analogue pour $\mathbb{Z}$. En particulier:

le corps des rationnels est naturellement c-$\mathcal{P}$-c en tant que corps.

## Algèbres $\mathbf{M_n}(\mathbb{Q})$ et $\mathbf{Flin}(\mathbb{Q})$ comme $\mathcal{P}_0$-structures

Nous pouvons faire de l'anneau $\mathbf{M_n}(\mathbb{Q})$ une $\mathcal{P}_0$-structure si nous le présentons sous forme "réduite au même dénominateur" et prenons pour mesure de la matrice A représentée par le couple $((a_{ij}), d)$ de $\mathbf{M_n}(\mathbb{Z}) \times \mathbb{N}^+$: $\|A\| = n + \mathbf{lg}(d + \Sigma\,|\,a_{ij}\,|)$. Nous réservons la notation $\mathbf{M_n}(\mathbb{Q})$ pour cette présentation de cet anneau (ou pour toute présentation $\mathcal{P}$-équivalente, par exemple sous forme d'une liste de $n^2$ éléments de $\mathbb{Q}$ avec pour mesure la taille effective du mot utilisé pour représenter la matrice)[1].

Avec la même mesure, nous pouvons considérer l'algèbre $\mathbf{Flin}(\mathbb{Q})$, réunion emboîtée des $\mathbf{M_n}(\mathbb{Q})$, et nous obtiendrons une $\mathcal{P}_0$-structure.

## $\mathbb{Z}[X]$ comme $\mathcal{P}_0$-structure et comme objet libre à un générateur

Nous notons $\mathbb{Z}[X]$ l'anneau des polynômes à coefficients dans $\mathbb{Z}$ présenté sous la forme $\mathbf{Lst}(\mathbb{Z})$ (en arrêtant la liste au coefficient dominant $a_d$), et avec la mesure suivante:

$$\| \Sigma a_i X^i \| = d + \mathbf{lg}(\Sigma\ |a_i|).$$

Il est clair que cette mesure est polynomialement reliée à la taille naturelle $\Sigma \|a_i\|$. Un calcul immédiat montre de plus que $(\mathbb{Z}[X], +, \times)$ est alors une $\mathcal{P}_0$-structure, donc c-$\mathcal{P}$-c.

Remarquons maintenant que la taille naturelle du polynôme $P = \Sigma a_i X^i$ (écrit sous forme de la liste de ses coefficients) est elle-même polynomialement reliée à la taille de l'écriture:

$$a_0 + a_1 \times X + a_2 \times X \times X + \ldots + a_d \times X \times \ldots \times X \quad \text{dans } \mathbf{Calc}(\mathbb{Z}[X]).$$

Si maintenant A est un $\mathcal{P}$-anneau où l'addition et la multiplication sont c-$\mathcal{P}$-c, on a une $\mathcal{P}$-fonction de $\mathbb{Z}[X] \times A$ vers $\mathbf{Calc}(A)$ :

$(P,b) \rightarrow \hat{a}_0 + \hat{a}_1 \times b + \hat{a}_2 \times b \times b + \ldots + \hat{a}_d \times b \times \ldots \times b$, où $\hat{a}$, pour a dans $\mathbb{Z}$, est la "valeur" de a dans A (cf. § précédent). Il ne reste plus qu'à "faire le calcul indiqué" pour obtenir le résultat suivant:

**Proposition B.c1** : Si A est un $\mathcal{P}$-anneau où l'addition et la multiplication sont c-$\mathcal{P}$-c, l'homomorphisme d'évaluation de $\mathbb{Z}[X] \times A$ vers A : $(P,b) \rightarrow P(b)$ est une $\mathcal{P}$-fonction.
De plus, si A est un $\mathcal{P}$-anneau où l'addition est c-$\mathcal{P}$-c, pour b fixé, l'homomorphisme d'évaluation en b, de $\mathbb{Z}[X]$ vers A : $P \rightarrow P(b)$ est une $\mathcal{P}$-fonction si et seulement si la fonction $n \rightarrow b^n$, de $\mathbb{N}_1$ vers B, est une $\mathcal{P}$-fonction.

En particulier, $\mathbb{Z}[X]$ est une structure naturellement c-$\mathcal{P}$-c, et dans la catégorie des $\mathcal{P}$-anneaux où l'addition et la multiplication sont c-$\mathcal{P}$-c, $\mathbb{Z}[X]$ est l'objet librement engendré par X.

**Remarque** : Il existe une autre présentation utile de l'anneau des polynômes à coefficients dans $\mathbb{Z}$, non $\mathcal{P}$-équivalente à la précédente. Nous notons $\mathbb{Z}[X]_c$ cet anneau lorsque le polynôme est représenté par la liste des couples $(i, a_i)$, en ne mentionnant que les coefficients non nuls, et i étant écrit en binaire. C'est une présentation adéquate pour calculer sur des polynômes "creux", c.-à-d. avec peu de coefficients non nuls. $\mathbb{Z}[X]_c$ est un $\mathcal{P}$-anneau, et

---

[1] Divertissement mathématique: $\mathbb{Q}$, $\mathbf{M_n}(\mathbb{Q})$ sont-il naturellement c-$\mathcal{P}$-c en tant qu'anneau ?

l'homomorphisme : $P \to P$ de $\mathbb{Z}[X]$ vers $\mathbb{Z}[X]_c$ est une $\mathcal{P}$-fonction bijective, mais non une $\mathcal{P}$-équivalence.

En fait $\mathbb{Z}[X]_c$ n'est pas c-$\mathcal{P}$-c, comme le montre l'exemple du produit:

$( 1 + X ).( 1 + X^2).( 1 + X^4 )....( 1 + X^{2^n} )$, qui, dans $\mathbb{Z}[X]_c$, prend "beaucoup plus" de place après avoir été effectué qu'avant.

## $\mathbb{Z}[X,Y]$ comme $\mathcal{P}_0$-structure et comme objet libre à deux générateurs

Nous notons $\mathbb{Z}[X,Y]$ l'anneau des polynômes à 2 indéterminées X et Y et à coefficients dans $\mathbb{Z}$, présenté sous la forme **Lst(Lst($\mathbb{Z}$))** (c.-à-d. comme une liste d'éléments de $\mathbb{Z}[X]$, lorsqu'on voit $\mathbb{Z}[X,Y]$ sous la forme $\mathbb{Z}[X]\,[Y]$ ), et avec la mesure suivante:

$$\Sigma a_{ij}.X^iY^j = \text{degré total} + \mathbf{lg}( \Sigma\ |a_{ij}| ) .$$

Il est clair que cette mesure est polynomialement reliée à la taille naturelle $\Sigma \parallel a_{ij} \parallel$. Un calcul immédiat montre de plus que $( \mathbb{Z}[X,Y] , + , \times )$ est alors une $\mathcal{P}_0$-structure, donc c-$\mathcal{P}$-c.

Nous pouvons poursuivre exactement comme au paragraphe précédent, et nous obtiendrions la proposition analogue de la même manière. En particulier :

- $\mathbb{Z}[X,Y]$ est une structure d'anneau naturellement c-$\mathcal{P}$-c (et la présentation décrite est la présentation c-$\mathcal{P}$-c naturelle)
- dans la catégorie des $\mathcal{P}$-anneaux commutatifs où l'addition et la multiplication sont c-$\mathcal{P}$-c, $\mathbb{Z}[X,Y]$ est l'objet librement engendré par X et Y.
- si A est un $\mathcal{P}$-anneau où l'addition est c-$\mathcal{P}$-c, pour b et c fixés qui commutent entre eux, l'homomorphisme d'évaluation en b et c, de $\mathbb{Z}[X,Y]$ vers A : $P \to P(b,c)$ est une $\mathcal{P}$-fonction si et seulement si la fonction

  $(n,p) \to b^n.c^p$, de $\mathbb{N}_1 \times \mathbb{N}_1$ vers B, est une $\mathcal{P}$-fonction.

On obtient les mêmes résultats pour $\mathbb{Z}[X_1,X_2,...,X_n]$. Cette méthode ne peut cependant pas se généraliser à une infinité d'indéterminées : la mesure qui fait de $\mathbb{Z}[X_1,X_2,...,X_n]$ une $\mathcal{P}_0$-structure est toujours "la même" :

degré total + lg(Σ |coefficients |);

et elle est toujours polynomialement reliée à la taille concrète (longueur du mot utilisé), *mais* elle l'est de moins en moins bien au fur et à mesure que le nombre de variables augmente. (Voir cependant dans le §f : Deux remarques sur l'anneau $\mathbb{Z}[(X_i)_{i\in\mathbb{N}}]$).

## $\mathbb{Q}(X)$ comme $\mathcal{P}_0$-structure naturellement c-$\mathcal{P}$-c

**Proposition B.c2** : Lorsqu'on présente $\mathbb{Q}(X)$ comme corps des fractions de $\mathbb{Z}[X]$, on obtient une structure c-$\mathcal{P}$-c, et c'est la présentation c-$\mathcal{P}$-c naturelle sur le corps abstrait $\mathbb{Q}(X)$.

Soient par ailleurs un $\mathcal{P}$-corps K, de caractéristique nulle, et dans lequel l'addition et la multiplication sont c-$\mathcal{P}$-c, et b un élément de K transcendant sur son sous corps premier: alors l'homomorphisme d'évaluation $F \to F(b)$ de $\mathbb{Q}(X)$ vers K est une $\mathcal{P}$-fonction.

*preuve>* Dans le corps des fractions de $\mathbb{Z}[X]$ , la mesure de la fraction rationnelle $F = P/Q$ est :

$$\| P \| + \| Q \| = \deg(P) + \deg(Q) + \mathbf{lg}( \Sigma \mid a_i \mid) + \mathbf{lg}( \Sigma \mid b_j \mid) ,$$

elle est polynomialement reliée à la taille concrète : $\Sigma \| a_i \| + \Sigma \| b_j \|$ , mais aussi à la mesure

$$\sup(\deg(P),\deg(Q)) + \mathbf{lg}(\Sigma \mid a_i \mid + \Sigma \mid b_j \mid) .$$

Or cette dernière fait de $\mathbb{Q}(X)$ une $\mathcal{P}_0$-structure. Ainsi, $\mathbb{Q}(X)$ , présenté comme corps des fractions de $\mathbb{Z}[X]$ , est c-$\mathcal{P}$-c.
Si maintenant K et b sont comme dans l'alinéa 2 , on a l'homomorphisme d'évaluation $P \rightarrow P(b)$ , avec comme source $\mathbb{Z}[X]$ , qui est un $\mathcal{P}$-homomorphisme d'après la proposition B.c1 , d'où on déduit l'analogue lorsqu'on prolonge à $\mathbb{Q}(X)$.
Enfin, en prenant comme cas particulier : pour K une $\mathcal{P}$-présentation de $\mathbb{Q}(X)$ qui en fait une structure c-$\mathcal{P}$-c , et pour b l'indéterminée X, on voit que l'on avait bien défini la présentation c-$\mathcal{P}$-c naturelle du corps abstrait $\mathbb{Q}(X)$. ❑

On généraliserait sans peine ces résultats à $\mathbb{Q}[X_1,X_2,...,X_n]$ et à $\mathbb{Q}(X_1,X_2,...,X_n)$. Une mesure utilisable pour faire de $\mathbb{Q}(X_1,X_2,...,X_n)$ une $\mathcal{P}_0$-structure est la suivante :

$$\| F \| = \| P/Q \| = \sup(\deg(P),\deg(Q)) + \mathbf{lg}( \Sigma \mid \text{coefficients} \mid )$$

(P et Q étant dans $\mathbb{Z}[X_1,X_2,...,X_n]$ et deg représentant le degré total)

## D'autres $\mathcal{P}_0$-structures d'anneaux

**Proposition B.c3** : Toute $\mathbb{Z}$-algèbre libre K de dimension finie k (comme $\mathbb{Z}$-module) est un anneau naturellement c-$\mathcal{P}$-c . On peut obtenir la structure c-$\mathcal{P}$-c naturelle de K en présentant K comme un sous-module libre de $\mathbf{M}_k(\mathbb{Z})$ , et alors K est une $\mathcal{P}_0$-structure.
De plus K [X] et K [$X_1,X_2,...,X_n$] peuvent être présentées de manière à obtenir des $\mathcal{P}_0$-structures , et ce sont les structures c-$\mathcal{P}$-c naturelles de ces anneaux.

*preuve>* Si $e_1,e_2,...,e_k$ est une base de K comme $\mathbb{Z}$-module et si b est un élément de K , nous considérons la matrice B de l'application linéaire "multiplication par b " exprimée sur la base $e_1,e_2,...,e_k$ , nous notons $\sigma(b)$ la somme $\Sigma$ |coeffs de B| , et nous considérons la mesure $\| b \| = \mathbf{lg}(\sigma(b))$ . L'homomorphisme $b \rightarrow B$ est un isomorphisme de K sur une sous algèbre K' de $\mathbf{M}_k(\mathbb{Z})$ . Il faut voir que K' est une partie $\mathcal{P}$-détachable de $\mathbf{M}_k(\mathbb{Z})$ . Cela se déduit facilement du fait que tout sous espace vectoriel libre de $\mathbf{M}_k(\mathbb{Q})$ est $\mathcal{P}$-détachable et $\mathcal{P}$-libre . Ainsi K peut être présenté de manière à être une $\mathcal{P}_0$-structure .

Voyons pourquoi K , ainsi présenté , est naturellement c-$\mathcal{P}$-c en tant qu'anneau : soit K° une présentation de l'anneau en question qui en fasse un anneau c-$\mathcal{P}$-c . Tout d'abord, en tant qu'anneau c-$\mathcal{P}$-c , K° est une $\mathcal{P}$-$\mathbb{Z}$-algèbre donc un $\mathcal{P}$-$\mathbb{Z}$-module ; par ailleurs la présentation choisie est $\mathcal{P}$-équivalente à celle de K comme $\mathbb{Z}$-module $\mathcal{P}$-libre .
Ainsi l'application "identité" de K vers K° est-elle nécessairement une $\mathcal{P}$-fonction.

Décrivons maintenant une $\mathcal{P}_0$-présentation de l'anneau $K[X_1,X_2,...,X_n]$ : nous présentons K comme ci dessus et $K[X_1,X_2,...,X_n]$ comme **Lst**(**Lst**( ... (**Lst**(K)) ...)) (cf. $\mathbb{Z}[X,Y]$ ) , et nous prenons pour mesure du polynôme P le nombre $\| P \|$ = degré total + $\mathbf{lg}( \Sigma \sigma(b_i) )$ , où les $b_i$ sont tous les coefficients , et avec $\sigma(b)$ défini ci-dessus.
Nous laissons au lecteur le soin de vérifier que c'est une $\mathcal{P}_0$-structure , et au lecteur courageux celui de vérifier que c'est la structure d'anneau c-$\mathcal{P}$-c naturelle. ❑

**Remarques :**

1) à partir de ces $\mathcal{P}_0$-anneaux on peut en construire d'autres en utilisant les sous-anneaux $\mathcal{P}$-détachables et les $\mathcal{P}$-quotients . La théorie des bases standards conduit d'ailleurs au résultat suivant: tout idéal donné comme de type fini dans $\mathbb{Z}[X_1,X_2,...,X_n]$ est $\mathcal{P}$-détachable.

2) la preuve doit être légèrement modifiée lorsque K ne possède pas de neutre pour la multiplication, car il se peut qu'on ait un b non nul pour lequel la multiplication par b soit néanmoins égale à la fonction partout nulle: on se ramène au "bon cas" en "rajoutant" un élément 1 à K (c.-à-d. en considérant K comme une sous algèbre de l'anneau $\mathbb{Z} \times K$)

**Proposition B.c4** : Soit K une $\mathbb{Q}$-algèbre qui est en outre un $\mathbb{Q}$-espace vectoriel libre de dimension finie : alors K peut être présentée de manière à être une $\mathcal{P}_0$-structure d'anneau.
De plus $K[X]$ , $K[X_1,X_2,...,X_n]$ et (si K est intègre) $K(X_1,X_2,...,X_n)$ , peuvent être présentées de manière à obtenir des $\mathcal{P}_0$-structures .

Même principe de présentation (et de démonstration du résultat) que pour la proposition B.c3 et pour l'algèbre $\mathbb{Q}(X_1,X_2,...,X_n)$ : notamment, dans le cas où K est un corps, on utilisera la présentation d'un élément de K sous forme d'une fraction de 2 "vecteurs" à coordonnées dans $\mathbb{Z}$ sur la base considérée pour obtenir un $\mathcal{P}_0$-corps.

## d) Groupes et monoïdes complètement $\mathcal{P}$-calculables

### $\mathcal{P}_1$-monoïdes et $\mathcal{P}_1$-groupes

L'associativité permet de donner une condition suffisante affaiblie pour le fait d'être c-$\mathcal{P}$-c.

**Proposition B.d1** : Tout $\mathcal{P}_1$-monoïde est une structure algébrique complètement $\mathcal{P}$-calculable. Même chose pour un $\mathcal{P}_1$-groupe.

*preuve*> Soit tout d'abord $(M , \times)$ un $\mathcal{P}_1$-monoïde, et soit $(u_1,u_2,...,u_n)$ dans **Lst**(M) une liste dont on veut calculer le produit. Soit $c \geq 1$ et $d \geq 0$ tels que

$$\| u \times v \| \leq c(\| u \| + \| v \|) + d.$$

Quitte à remplacer, le temps de la démonstration, la mesure $\| u \|$ par la mesure $\| u \|_1 = \| u \| + d$ , on peut supposer que $d = 0$ . Ensuite, supposons tout d'abord $n = 2^p$ : on voit par récurrence sur p , en regroupant les facteurs 2 par 2 que :

$$\| u_1 \times u_2 \times ... \times u_n \| \leq c^p . \Sigma \| u_i \| .$$

Si n est quelconque, on est facilement ramené au calcul précédent (par exemple en multipliant par le neutre un nombre convenable de fois) et on a la majoration analogue :

$$\| u_1 \times u_2 \times ... \times u_n \| \leq c^{\lg(n)} . \Sigma \| u_i \| .$$ Or

$c^{\lg(n)} \leq c^{1+\log_2(n)} \leq c.n^{\log_2(c)}$ . Cela montre que le produit, en tant qu'opération de **Lst**(M) vers M est **RESP** : il n'y a pas explosion de la taille lors du calcul.

Soit maintenant $(G, \times , x \rightarrow x^{-1})$ un $\mathcal{P}_1$-groupe . On remarque que dans **Calc**(G) toute écriture peut être remplacée par une écriture de valeur égale (une fois le calcul effectué) et où les exposants -1 n'interviennent qu'au niveau le plus bas (c.-à-d. accolés à des éléments de

G et non à des écritures comportant des produits).
On est donc ramené au cas des monoïdes. ❑

Groupes et monoïdes libres (dans la catégorie c-$\mathfrak{P}$-c )

On a immédiatement une caractérisation de $\mathbb{N}_1$ comme objet libre à un générateur dans la catégorie des $\mathfrak{P}$-monoïdes c-$\mathfrak{P}$-c et, plus généralement, de **Lst**(X) comme librement engendré par X :

**Proposition B.d2** : Si ( M , × ) est un monoïde complètement $\mathfrak{P}$-calculable, alors la fonction $M \times \mathbb{N}_1 \to M : (b, n) \to b^n$ est une $\mathfrak{P}$-fonction. Pour b fixé, on obtient un $\mathfrak{P}$-homomorphisme : $\mathbb{N}_1$ est objet librement engendré par 1 dans la catégorie des monoïdes c-$\mathfrak{P}$-c . Et $\mathbb{N}_1$ est la présentation c-$\mathfrak{P}$-c naturelle des entiers naturels en tant que monoïde additif.

**Proposition B.d3** : Soit X un $\mathfrak{P}$-ensemble-discret. Alors le monoïde **Lst**(X) est l'objet librement engendré par X dans la catégorie des monoïdes c-$\mathfrak{P}$-c (avec pour flèches les $\mathfrak{P}$-homomorphismes). Si X est fini, on obtient ainsi la présentation c-$\mathfrak{P}$-c naturelle du monoïde.
De même, **Lst**(X $\cup$ X') (où X' est une copie de X, disjointe de X ) est le groupe librement engendré par X dans la catégorie des groupes c-$\mathfrak{P}$-c ; et pour X fini on obtient une structure naturellement c-$\mathfrak{P}$-c . En particulier $\mathbb{Z}_1$ est la présentation c-$\mathfrak{P}$-c naturelle du groupe des entiers relatifs.

On laisse le soin au lecteur d'énoncer les résultats analogues pour les monoïdes commutatifs et les groupes abéliens.

## e) Présentations "en magma" ou "par formules"

Structures libres du point de vue complètement-$\mathfrak{C}$-calculable

Les résultats obtenus pour les anneaux $\mathbb{Z}[X_1,X_2,...,X_n]$ et pour les groupes libres sont en fait des cas particuliers d'un résultat général dans le cadre de l'"algèbre universelle" , c.-à-d. la théorie des structures algébriques qui ne font intervenir que des constantes et des lois de composition partout définies obéissant à des axiomes purement universels.

On obtient, en maths classiques, une structure librement engendrée par L en considérant le "magma" des "écritures de calculs à effectuer dans la structure qui n'impliquent que des constantes et des éléments de L " (ouf!) et en prenant pour relation d'égalité la relation d'équivalence la plus fine qui rende vraie les axiomes de la structure. Du point de vue constructif, on ne sait pas a priori si la relation d'égalité ainsi définie est la négation d'une relation de séparation, ou non. Enfin, même si la relation d'égalité est la négation d'une relation de séparation, on ne sait pas a priori si elle est décidable (c.-à-d.: si l'ensemble obtenu est discret).

Si L est un $\mathfrak{C}$-ensemble-discret, on obtiendra de la même manière une $\mathfrak{C}$-structure complètement $\mathfrak{C}$-calculable si et seulement si la relation d'équivalence obtenue (sur le $\mathfrak{C}$-ensemble formé par "les écritures de calculs etc...") est $\mathfrak{C}$-calculable. C'est de plus la structure librement engendrée par L dans la catégorie des structures algébriques "complètement

$\mathfrak{C}$-calculables" du type voulu (groupe, anneau,...) . Lorsque L est fini on obtient une structure "naturellement complètement $\mathfrak{C}$-calculable" . Et l'on est évidemment intéressé à ce que la classe $\mathfrak{C}$ soit "la plus petite possible" , c.-à-d. que les calculs y soient le plus simples possibles.
Lorsque L est fini on n'a pas ipso facto, pour la structure librement engendrée par L un résultat affirmant qu'elle soit $\mathfrak{C}$-numérotable, ou même $\mathfrak{C}$-dénombrable.

Notons que notre traitement des structures c-$\mathfrak{P}$-c libres a utilisé une voie plus directe bien que moins générale: au lieu de quotienter le magma des "écritures de calculs à effectuer ...." par la bonne relation d'équivalence, et de prouver que cette relation d'équivalence était $\mathfrak{P}$-calculable , et de chercher enfin un système de représentants canoniques, nous avons utilisé des présentations déjà connues où chaque objet de la structure est représenté par un (ou des) élément(s) d'un $\mathfrak{P}$-langage $Y \subset A^*$ , ensuite nous avons montré que la présentation en question était une $\mathfrak{P}$-présentation qui rendait la structure c-$\mathfrak{P}$-c , et enfin qu'elle était libre du point de vue de la catégorie c-$\mathfrak{P}$-c.

**Remarque** : Nous comprenons pourquoi la notion de structure "naturellement primitive récursive" est si efficace, en comparaison d'autres classes de constructions (je ne pense pas qu'il existe en dehors des ensembles finis de structures "naturellement de type $\mathfrak{P}$"), c'est parce que pour la classe $\mathbb{P}\mathbf{r}$, la $\mathbb{P}\mathbf{r}$-calculabilité implique la $\mathbb{P}\mathbf{r}$-complète-calculabilité (proposition B.a1), et que la bonne notion est bel et bien la complète-$\mathfrak{C}$-calculabilité (pour les structures algébriques).
Quant à la classe $\mathfrak{P}_{00} := \mathfrak{P} \cap \mathbf{SPACERES}(n)$ , elle vérifie la même propriété que la classe $\mathbb{P}\mathbf{r}$ (remarque après la prop. B.a1) mais les lois de composition dans "le magma" en question sont seulement a priori $\mathfrak{P}_0$, ce qui empêche un traitement général de la question. Notons cependant que $( \mathbb{N} , + , \times )$ est "presque" naturellement de type $\mathfrak{P}_{00}$ : si $\mathbf{N}'$ est une $\mathfrak{P}_{00}$-présentation de cette structure où la mesure de 1 est 1 , alors l'application identité de $\mathbb{N}$ vers $\mathbf{N}'$ est une $\mathfrak{P}_1$-fonction : en effet, $2 = 1 + 1$ est de mesure au plus 2 dans $\mathbf{N}'$ , ensuite on calcule dans $\mathbf{N}'$ un nombre en binaire par l'algorithme de Horner et on voit qu'un nombre de longueur n en binaire sera au maximum de mesure 3.n dans $\mathbf{N}'$.

## Présentations "en magma" ou "par formules"

Les structures libres construites précédemment sont des cas particuliers des présentations "en magma", définies ci-après.

Etant donné un ensemble dénombrable X avec une structure algébrique abstraite donnée par un nombre fini de lois de composition et de constantes, et par des relations, si X' est une partie $\mathfrak{C}$-présentée de X , qui engendre X , on peut présenter X en utilisant la partie de $\mathbf{Calc}(X)$ ne faisant intervenir que des constantes et des éléments de X' . Pour cette présentation, les lois de composition sont dans $\cup_c \mathbf{DTIME}(n+c)$ . Mais il reste le problème de l'égalité dans X et des autres relations faisant partie de la structure, qui peuvent ne pas être $\mathfrak{C}$-décidables, ni même **Rec**-décidables

Nous appellerons cette présentation la présentation en magma (ou encore "par formules") sur le système générateur X' .

Si on a un ensemble X donné par une $\mathfrak{C}$-présentation, et si on le munit d'une structure algébrique au moyen de lois de composition, on appellera présentation par formules de X la présentation par formules sur le système générateur X lui-même. La structure algébrique est complètement $\mathfrak{C}$-calculable si et seulement si le $\mathfrak{C}$-ensemble X est $\mathfrak{C}$-équivalent à sa présentation par formules.

**Exemples** :

(1) La présentation par formules de $(\mathbb{Z}[T] ; +, -, \times)$ sur le système générateur $(0, 1, T)$ est $\mathcal{P}$-équivalente à la présentation naturelle.

(2) Si nous considérons sur le $\mathcal{P}$-ensemble précédent $\mathbb{Z}[T]$ la structure algébrique $(+, -, \times, P \to P^2)$ avec la présentation par formules correspondante, nous obtenons une "généralisation" de la présentation creuse (on autorise en effet des exposants en binaire non seulement pour $T$, mais pour tout polynôme déjà écrit). Nous noterons $\mathbb{Z}[T]_m$ cette présentation. Elle n'est sans doute pas bien adaptée aux calculs formels généraux, notamment pour ce qui concerne la relation d'égalité et la division euclidienne. Elle est par contre très bien adaptée à l'évaluation dans un anneau fini $K$, ou, plus généralement, dans un $\mathcal{P}$-anneau $K$ où les lois $+, -, \times$ et l'élévation au carré seraient $\mathcal{P}_0$.

Ceci confirme l'appréciation selon laquelle la présentation d'une structure algébrique doit être choisie en fonction des calculs qu'on désire effectuer.

## f) Algèbre d'un monoïde A[M], algèbres de polynômes

### Algèbre d'un monoïde A[M]

**Proposition B.f1** : Soit $A$ un $\mathcal{P}$-anneau commutatif où l'addition est c-$\mathcal{P}$-c et $M$ un $\mathcal{P}$-monoïde. On présente l'anneau $A[M]$ comme la $\mathcal{P}$-partie de $\mathbf{Lst}(A \times M)$ formée par les listes d'éléments $(a_i, m_i)$ avec $a_i \neq 0$ (sauf pour représenter 0) et sans répétition sur les $m_i$. On obtient ainsi un $\mathcal{P}$-anneau où l'addition est c-$\mathcal{P}$-c. De plus, si $B$ est un $\mathcal{P}$-anneau où l'addition est c-$\mathcal{P}$-c, si $f : A \to B$ est un $\mathcal{P}$-homomorphisme d'anneau, $g : M \to B$ un $\mathcal{P}$-homomorphisme de $M$ dans $(B, \times)$ et si tout $f(a)$ commute avec tout $g(m)$, alors l'homomorphisme canonique ("d'évaluation") de $A[M]$ vers $B$ qui factorise $f$ et $g$ est un $\mathcal{P}$-homomorphisme.

*Démonstration immédiate.* Le caractère c-$\mathcal{P}$-c de l'addition est indispensable pour montrer que le produit dans $A[M]$ est bien une $\mathcal{P}$-fonction (lorsqu'on "regroupe" les coefficients d'un même $m$) et pour montrer que l'homomorphisme "d'évaluation" est une $\mathcal{P}$-fonction.

La présentation proposée pour $A[M]$ est donc "naturelle", à $\mathcal{P}$-isomorphisme près. En langage des catégories (et dans le cas des $A$-algèbres unitaires) : dans la catégorie des "$\mathcal{P}$-$A$-algèbres unitaires où l'addition est c-$\mathcal{P}$-c", le foncteur d'oubli vers les $\mathcal{P}$-monoïdes (obtenu en ne conservant que la structure multiplicative) possède un adjoint à gauche.

**Terminologie**: Nous utilisons $A$-algèbre pour $A$-algèbre associative, non forcément unitaire. L'anneau $A$, lui, est supposé commutatif.

**Remarque** : En fait la commutativité de $A$ n'est pas essentielle: lorsque $A$ n'est pas commutatif, on peut construire de la même manière un anneau $A[M]$ où les $m \in M$ commutent avec les $a \in A$. La proposition B.f1 resterait inchangée.

## Algèbres A[X]

Nous supposons toujours que A est un $\mathfrak{P}$-anneau commutatif où l'addition est c-$\mathfrak{P}$-c .

Il existe au moins 2 présentations de A[X] intéressantes: la première est sous la forme **Lst**(A) (on écrit tous les coefficients jusqu'à celui de degré maxi) et la seconde est la présentation creuse, pour le cas où on estime que la plupart des coefficients sont nuls, sous la forme **Lst**( $A \times \mathbb{N}$ ) , et on écrit uniquement les coefficients non nuls, en signalant l'exposant en binaire.

Nous réservons la notation A[X] pour la première présentation: cette présentation donne une $\mathfrak{P}$-A-algèbre $\mathfrak{P}$-isomorphe à $A[\mathbb{N}_1]$ . La 2ème , que nous noterons $A[X]_c$ donne une $\mathfrak{P}$-A-algèbre $\mathfrak{P}$-isomorphe à $A[\mathbb{N}]$ . (calculs immédiats)

**Proposition B.f2** : Soient A un $\mathfrak{P}$-anneau commutatif où l'addition est c-$\mathfrak{P}$-c et B une $\mathfrak{P}$-A-algèbre unitaire où l'addition est c-$\mathfrak{P}$-c .

1) si $b \in B$ est tel que la fonction $\mathbb{N}_1 \to B : n \to b^n$ soit une $\mathfrak{P}$-fonction, alors l'homomorphisme "d'évaluation en b" : $P \to P(b)$ de A[X] vers B est une $\mathfrak{P}$-fonction.
2) si l'application $\mathbb{N}_1 \times B \to B : (n,b) \to b^n$ est une $\mathfrak{P}$-fonction, alors la fonction d'évaluation : $(P,b) \to P(b)$ de $A[X] \times B$ vers B est une $\mathfrak{P}$-fonction. (ce sera le cas si la multiplication dans B est c-$\mathfrak{P}$-c ).

*preuve*> Le 1) est un cas particulier de la proposition B.f1 . Le 2) revient à introduire b comme paramètre: démonstration immédiate. ❑

**Remarques** :

1) il ne semble pas que la propriété pour le produit d'être c-$\mathfrak{P}$-c passe de A à A[X]
2) en remplaçant $\mathbb{N}_1$ par $\mathbb{N}$ et A[X] par $A[X]_c$ on obtient une proposition analogue ... mais il est bien rare que la fonction $\mathbb{N} \to B : n \to b^n$ soit une $\mathfrak{P}$-fonction. (il y a le cas des corps finis)
3) la commutativité de A n'est pas indispensable: cf. la remarque après la prop. B.f1

## Algèbres $A[X_1,X_2,...,X_n]$

Les algèbres A[X][Y] , A[Y][X] sont $\mathfrak{P}$-isomorphes à l'algèbre du monoïde $\mathbb{N}_1 \times \mathbb{N}_1$. On notera A[X,Y] pour toute $\mathfrak{P}$-présentation $\mathfrak{P}$-isomorphe à l'une de ces 3 .

On obtient immédiatement une proposition analogue à la précédente. Et on peut généraliser pour les anneaux de polynômes à un nombre fini de variables: en particulier:

**Proposition B.f3** : Soient A un $\mathfrak{P}$-anneau commutatif où l'addition est c-$\mathfrak{P}$-c et B une $\mathfrak{P}$-A-algèbre unitaire commutative où l'addition et le produit sont c-$\mathfrak{P}$-c . Alors l'évaluation :

$A[X_1,X_2,...,X_n] \times B^n \to B : \quad (P, (b_1,b_2,...,b_n)) \to P(b_1,b_2,...,b_n)$ est une $\mathfrak{P}$-fonction.

## Deux remarques sur l'anneau $\mathbb{Z}[(X_i)_{i \in \mathbb{N}}]$

Il s'agit de l'anneau des polynômes à coefficients dans $\mathbb{Z}$ avec une infinité d'indéterminées. C'est donc la $\mathbb{Z}$-algèbre $A = \mathbb{Z}[M]$ du monoïde $M = \mathbb{N}_1^{(\mathbb{N})}$, et nous pouvons considérer la présentation correspondante. On n'obtient pas une structure c-$\mathfrak{P}$-c: cf. le produit $(X_1+X_2) \times (X_3+X_4) \times ... \times (X_{2k-1}+X_{2k})$ qui occupe "beaucoup" d'espace une fois développé.

*La première remarque* est que, pour un $\mathcal{P}$-anneau K , il revient au même d'affirmer qu'addition et produit sont c-$\mathcal{P}$-c ou que l'évaluation des polynômes est une $\mathfrak{C}$-opération, c.-à-d. : $A \times K^{(\mathbb{N})} \to K$ : $(P, (k_i)_{i \in \mathbb{N}}) \to P((k_i)_{i \in \mathbb{N}})$ est une $\mathcal{P}$-opération.
On notera a contrario que le fait que K soit c-$\mathcal{P}$-c en tant qu'anneau signifie que l'évaluation de *toute formule* (et pas seulement les polynômes) est une $\mathcal{P}$-opération.

*Deuxième remarque:* si nous présentons l'anneau $\mathbb{Z}[(X_i)_{i \in \mathbb{N}}]$ au moyen de la présentation par formules sur le système générateur $\mathbb{Z} \cup \{ X_i ; i \in \mathbb{N} \}$ (lui même codé comme réunion disjointe de $\mathbb{Z}$ et $\mathbb{N}$ ), il y a moyen de deviner en temps polynomial si 2 expressions (indiquant des calculs à effectuer dans $\mathbb{Z}[(X_i)_{i \in \mathbb{N}}]$ représentent des éléments distincts de $\mathbb{Z}[(X_i)_{i \in \mathbb{N}}]$ , en effet :

a) le degré du polynôme en chaque $X_i$ ainsi que le nombre de $X_i$ est majoré par la taille de la formule

b) deux polynômes à k variables de degrés $\leq d$ en chaque variable sont distincts si et seulement si ils sont évalués distincts en un k-uple $\in \{0,1,...,d\}^k$

c) l'évaluation d'une formule dans $\mathbb{Z}$ est un $\mathcal{P}$-calcul.

Il suffit donc de deviner un "point" où les deux expressions prennent des valeurs différentes. Ainsi, si $\mathcal{P} = \mathcal{N}\mathcal{P}$ , l'anneau $\mathbb{Z}[(X_i)_{i \in \mathbb{N}}]$ possèderait une $\mathcal{P}$-présentation qui le rendrait c-$\mathcal{P}$-c [1].

## g) Pourquoi $\mathbb{Z}$ marche-t-il si bien ?

### Introduction

Nous avons établi le caractère c-$\mathcal{P}$-c de $\mathbb{Z}$-modules, d'anneaux et de corps construits à partir de $\mathbb{Z}$ en utilisant systématiquement l'argument suivant : lorsqu'on remplace la mesure "naturelle" par la mesure $n + \mathbf{lg}(\Sigma |\text{coefficients}|)$, où n est le nombre de coefficients dans $\mathbb{Z}$ nécessaires à la description de l'objet considéré, les lois de composition considérées deviennent $\mathcal{P}_0$. Tout cela marche bien "parce que" majorer la grandeur en valeur absolue d'un entier permet de majorer la place occupée par l'entier en écriture binaire. Le raisonnement par exemple serait complètement en défaut si on prenait des coefficients rationnels au lieu de prendre des coefficients entiers.

L'idée pour généraliser les résultats obtenus à partir de $\mathbb{Z}$ doit donc être cherchée un peu plus loin: nos raisonnements sont "trop" simples pour pouvoir être généralisés parce qu'ils utilisent des propriétés trop fortes de $\mathbb{Z}$.

Le rêve serait de démontrer un théorème du genre: si K est un anneau où l'addition et le produit sont c-$\mathcal{P}$-c, alors il en est de même pour **Flin**(K) et pour K[X] (on sait déjà que, pour K intègre, il en est de même pour le corps des fractions de K ). Cela semble improbable.

Nous pouvons cependant démontrer un théorème analogue pour une certaine classe de $\mathcal{P}$-anneaux qui possèdent des propriétés de majorations assez fortes pour la mesure de la somme de 2 éléments.

---

[1] Divertissement mathématique: si l'égalité est $\mathcal{P}$-testable (dans l'anneau $\mathbb{Z}[(X_i)_{i \in \mathbb{N}}]$ muni de la présentation par formules ci-dessus) alors $\mathcal{P} = \mathcal{N}\mathcal{P}$ .

## Majoration pour l'addition itérée

**Lemme 1** : Si (K,+) est un monoïde $\mathfrak{P}$-présenté où est vérifiée l'inégalité:

$$\| a + b \| \leq \sup( \| a \| , \| b \| ) + C$$

alors pour $(u_1, u_2, \ldots , u_n)$ dans **Lst**(K) , on a l'inégalité:

$$\| u_1 + u_2 + \ldots + u_n \| \leq \sup( \| u_i \| ) + C.\mathbf{lg}(n-1)$$

*preuve>*

– si $n = 2^k$ , alors $\mathbf{lg}(n-1) = k-1$, démonstration par récurrence sur k immédiat

– si $n = 2^k + m$ , avec $0 < m < 2^k$, on peut supposer (récurrence) l'inégalité vraie pour m , et elle passe à n , puisque $\mathbf{lg}(n-1) = k$ et $\mathbf{lg}(m) < k$ . ❑

## Définition des $\mathfrak{P}_0$-anneaux

**Définition B.g1** : Un $\mathfrak{P}$-anneau K sera appelé un **$\mathfrak{P}_0$-anneau** lorsque:

1) K est $\mathfrak{P}$-dénombrable
2) l'addition vérifie la majoration suivante:
$$\|a + b \| \leq \sup( \| a \| , \| b \| ) + C$$
3) la multiplication est c-$\mathfrak{P}$-c

## Majoration pour les déterminants

**Proposition B.g1** : Soit K un $\mathfrak{P}_0$-anneau commutatif. On représente une matrice carrée de $\mathbf{M}_n$ (K) par la liste de ses $n^2$ coefficients, (on peut prendre pour mesure de la matrice $(u_{ij})$ le nombre $s = \max(n,\sup(\|u_{ij}\|))$ . La fonction déterminant est **RESP** (cette fonction est définie sur l'ensemble des listes à $n^2$ éléments de K , pour n variable, ensemble qui représente la réunion disjointe des $\mathbf{M}_n$ (K) )

**Remarque** : on établira en C que, sous certaines hypothèses supplémentaires, la fonction déterminant est alors une $\mathfrak{P}$-fonction. (cf. prop. C.b2, Th C.b1 et C.d1)

*preuve>* Puisque le produit est c-$\mathfrak{P}$-c, il existe un polynôme Q tel que Q(s) majore la mesure de tout produit de n facteurs pris parmi les $u_{ij}$. Donc on obtient par le lemme 1 les majorations:

$$\| \det((u_{ij})) \| \leq \mathbf{lg}(n!).Q(s) \leq n.\mathbf{lg}(n).Q(s) \leq (X^2Q)(s)$$ ❑

## Propriétés de stabilité des $\mathfrak{P}_0$-anneaux

**Proposition B.g2** :

Si ( K , +, × , 0 , 1 ) est un $\mathfrak{P}_0$-anneau, alors il en est de même pour:

- tout sous anneau qui est une partie $\mathfrak{P}$-détachable de K
- tout quotient $\mathfrak{P}$-dénombrable de K par un idéal $\mathfrak{P}$-détachable
- $\mathbf{M}_n$ (K) et **Flin**(K) (réunion emboîtée des $\mathbf{M}_n$ (K) )
- $K[X_1,X_2,\ldots,X_n]$
- K [M] , si M est un monoïde $\mathfrak{P}$-dénombrable c-$\mathfrak{P}$-c qui possède "peu d'objets de petite taille" (c.-à-d. : il existe un polynôme fixé Q tel que les objets sous forme réduite canonique et de taille inférieure à p sont en nombre au plus Q(p) )

(Dans les trois derniers cas la mesure de l'objet x est précisée dans la démonstration)

**Remarques** :

1) le caractère $\mathcal{P}_0$, ne se conserve par contre pas par passage au corps des fractions

2) en fait **Flin**(K) ne possède pas d'élément neutre pour la multiplication, ce n'est donc pas un anneau, ... mais il est $\mathcal{P}_0$,

*preuve*> Les 2 premières stabilités sont évidentes.
Voyons $\mathbf{M_n}$ (K) : reprenons la mesure décrite dans la prop. B.g1 . Pour l'addition de 2 matrices, la majoration est immédiate.
Il reste à voir que le produit :

$((U_1, U_2, ..., U_p)) \rightarrow U_1 \times U_2 \times ... \times U_p$ , d

e $\mathbf{Lst}(\mathbf{M_n}(K))$ vers $\mathbf{M_n}(K)$ est **RESP** (cf. prop. B.a1). Or chaque coefficient du produit $U_1 \times U_2 \times ... \times U_p$ est égal à une somme de $n^{p-1}$ produits de p coefficients des matrices $U_k$, et $\mathbf{lg}(n^{p-1}) \simeq (p-1).\mathbf{lg}(n)$. Le calcul de majoration est donc analogue à celui fait pour le déterminant.
Comme nous avons intégré n au calcul de majoration, il vaut également pour le produit dans **Flin**(K), réunion emboîtée des $\mathbf{M_n}(K)$.
Raisonnements et calculs analogues pour K[X] , avec la mesure

$$\|Q\| = \max(\deg(Q), \sup(\|\text{coefficients}\|)).$$

Puis récurrence pour $K[X_1,X_2,...,X_n]$ .
Le dernier cas est une généralisation du précédent, puisque $K[X_1,X_2,...,X_n]$ peut être considéré comme la K-algèbre du monoïde additif $M = \mathbb{N}_1^n$ qui vérifie les hypothèses convenables.
Nous présentons K[M] comme expliqué en f), avec pour mesure:

$$\|\Sigma\, a_i\, m_i\| = \sup(\|a_i\|, \|m_i\|, n).$$

La majoration pour l'addition est immédiate. Il nous faut de plus une majoration polynomiale pour la mesure du produit itéré: soit

$$\prod_{i=1}^{k} x_i = \prod_{i=1}^{k} \sum_{j=1}^{n_i} a_{i,j} \cdot m_{i,j} = \sum_{j=1}^{n} b_j \cdot m_j$$

avec $b_j = \sum a_{1,j_1} . a_{2,j_2} .... a_{k,j_k}$

somme étendue aux $(j_1,j_2,...,j_k)$ vérifiant $m_{1,j_1}.m_{2,j_2}....m_{k,j_k} = m_j$

Le coefficient $b_j$ de $m_j$ est une somme possédant moins de $n_1.n_2...n_k$ termes, ce qui donne une majoration convenable de sa taille:

$$\|b_j\| \leq P(\max(k, \sup(\|a_{i,j'}\|))) + C.k.\sup(\mathbf{lg}(n_{j'})) \leq T(\max(k,\sup(\|x_i\|)))$$

où P et T sont des polynômes fixés.
Enfin, il nous faut une majoration de n , c.-à-d. : qu'il n'y ait pas trop de termes $b_j.m_j$ , nous utilisons pour cela le fait que M possède "peu d'objets de petite taille" : en effet on a

$\|m_i\| \leq R(\max(k, \sup(\|m_{i,j}\|)))$ , où R est un polynôme fixé, ceci parce que M est un monoïde c-$\mathcal{P}$-c, et donc: $n < Q(R(\max(k,\sup(\|m_{i,j}\|))))$ , et Q(R) est un polynôme fixé. ❑

**Remarque** : le calcul de majoration dans $\mathbf{M_n}(K)$ ci-dessus suggère qu'un exemple d'anneau K où + et × seraient c-$\mathcal{P}$-c sans que la même propriété soit vérifiée pour $\mathbf{M_2}(K)$, pourrait être cherché du coté d'un anneau où l'on aurait "assez souvent" $\|x + y\| \geq \|x\| + \|y\|$.

## Quelques exemples de $\mathcal{P}_0$-anneaux

Nous pouvons compter parmi les $\mathcal{P}_0$-anneaux les anneaux suivants:

– $\mathbb{Z}$ , $\mathbf{M_n}(\mathbb{Z})$ , $\mathbf{Flin}(\mathbb{Z})$ , $\mathbb{Z}[X_1,X_2,...,X_k]$ pour leur présentation naturelle, et avec les mesures définies en c) (indicateur polynomialement relié à: nombre de coefficients + $\lg(\Sigma|\text{coefficients}|)$)

– tout anneau fini

– tout anneau qui est un $\mathbb{Z}$-module libre de dimension finie: il est en effet isomorphe à un sous anneau $\mathcal{P}$-détachable de $\mathbf{M_n}(\mathbb{Z})$ (cf. prop. B.c3), on le munit de la présentation correspondant à cette sous-structure: en particulier les anneaux d'entiers dans les corps de nombres.

– tout anneau construit à partir de l'un des précédents par l'une des constructions autorisées par la proposition B.g2

Nous donnons maintenant un exemple de $\mathcal{P}_0$-anneau qui est une $\mathbb{Q}$-algèbre mais qui n'est pas de type fini en tant que $\mathbb{Q}$-algèbre.

Ceci en application de la dernière propriété de stabilité énoncée en B.g2 : si nous considérons le monoïde additif de $\mathbb{Q}_1$ (corps des fractions de $\mathbb{Z}_1$), nous voyons qu'il répond à l'hypothèse "peu d'objets de petite taille"; cependant l'addition n'est pas c-$\mathcal{P}$-c ; on peut néanmoins considérer des sous groupes additifs de $\mathbb{Q}_1$ qui sont c-$\mathcal{P}$-c : par exemple les sous-groupes $T_r$ obtenus en imposant que le dénominateur de la fraction soit une puissance d'un entier r fixé. L'algèbre $K[T_r]$ obtenue est l'algèbre des polynômes à coefficients dans K et avec des exposants $k/r^n$ (r fixé) pour X. Lorsque $K = \mathbb{Q}$ , certains quotients de cette algèbre sont des extensions infinies de $\mathbb{Q}$ où l'addition et le produit sont c-$\mathcal{P}$-c : par exemple, en remplaçant X par 2 , c.-à-d. en faisant le quotient par l'idéal engendré par $X - 2$ . En quotientant par un idéal convenable contenant $X - 1$ , on obtient une extension infinie engendrée par des racines de l'unité.

## Structure de $\mathcal{P}$-calculabilité naturelle dans les anneaux commutatifs de présentation finie

D'après la théorie des bases standards, tout idéal de type fini de $\mathbb{Z}[X_1,X_2,...,X_n]$ est une $\mathcal{P}$-partie de cet anneau (pour sa $\mathcal{P}$-présentation naturelle).

Comme par ailleurs l'anneau $\mathbb{Z}[X_1,X_2,...,X_n]$ est l'objet librement engendré par les $X_i$ dans la catégorie des $\mathcal{P}$-anneaux où addition et produit sont c-$\mathcal{P}$-c, on obtient le résultat suivant:

Tout anneau commutatif A de présentation finie est naturellement c-$\mathcal{P}$-c; on peut prendre comme $\mathcal{P}$-présentation naturellement c-$\mathcal{P}$-c , sa $\mathcal{P}$-présentation comme quotient de $\mathbb{Z}[X_1,X_2,...,X_n]$ .
Plus généralement: si A est un anneau commutatif engendré par n éléments $a_1,a_2,...,a_n$ qui est $\mathcal{P}$-présenté de manière que l'addition et le produit soient c-$\mathcal{P}$-c, alors la $\mathcal{P}$-présentation de A comme quotient de $\mathbb{Z}[X_1,X_2,...,X_n]$ (via l'homomorphisme d'évaluation $X_i \to a_i$ ) est naturellement c-$\mathcal{P}$-c.
De plus cette présentation en fait à la fois un $\mathcal{P}_0$-anneau où les déterminants sont $\mathcal{P}$-calculables[1], et, s'il est $\mathcal{P}$-dénombrable, un $\mathcal{P}_0$-anneau.

---

[1] En effet les déterminants sont $\mathcal{P}$-calculables dans $\mathbb{Z}[X_1,X_2,...,X_n]$ comme nous le verrons dans le § C.

# C) ALGEBRE LINEAIRE EN TEMPS POLYNOMIAL

## Introduction

L'algèbre linéaire est essentiellement l'étude des systèmes d'équations linéaires.

Lorsqu'on se situe dans un corps où l'égalité est décidable, le problème est entièrement résolu, du point de vue mathématique, soit par la méthode du pivot de Gauss, soit, dans le cas commutatif, par la méthode des déterminants et les formules de Cramer.

Que se passe-t-il du point de vue calculatoire ?

La méthode du pivot utilise peu de calculs : le nombre d'opérations arithmétiques élémentaires (addition, multiplication, division, test d'égalité à 0) pour traiter une matrice $n \times n$ est majoré par un polynôme en $n$ .

Les formules de Cramer donnent quant à elles, dans un $\mathcal{P}_0$-corps, une solution **DTIME**((n+2)!) du problème, si l'on s'en tient aux formules de définition explicites des déterminants. Par exemple elle est totalement impraticable pour une matrice $20 \times 20$ . Par contre elle donne des renseignements d'ordre théorique décisifs, à savoir que le fonctionnement des systèmes d'équations linéaires est gouverné par des polynômes en les coefficients du système (les déterminants justement).

Par ailleurs, les formules de Cramer montrent que, lorsque la fonction déterminant est $\mathcal{P}$-calculable, les systèmes d'équations linéaires sont $\mathcal{P}$-résolubles, même si la méthode est incontestablement plus lourde que celle du pivot.

Il reste à déterminer des conditions suffisantes faciles à vérifier pour la $\mathcal{P}$-calculabilité des déterminants. On trouve grosso-modo que la $\mathcal{P}$-calculabilité des déterminants équivaut au fait que le produit des matrices est complètement $\mathcal{P}$-calculable (cf. prop. C.b1 et Th C.b1 ). Or cette propriété est vérifiée pour les anneaux les plus couramment étudiés (cf. § B f ), et elle est de plus très stable.

Vu l'intérêt "pratique" de la méthode du pivot, il est intéressant d'étudier sous quelles conditions elle est praticable en temps polynomial, c.-à-d. sans explosion de la taille des coefficients. C'est ce que nous faisons au § d), où nous exposons la méthode de Bareiss.

La méthode du pivot, quoiqu'utilisant peu d'opérations arithmétiques élémentaires, n'est pas garantie a priori contre une explosion de la taille des coefficients. Si par exemple on se situe dans le $\mathcal{P}$-corps $\mathbb{Q}((X_i)_{i \in \mathbb{N}})$ , et qu'on veuille traiter par la méthode du pivot une matrice $n \times n$ dont les coefficients sont des $X_i$ distincts, on fera un calcul purement formel, et on retombera sur les formules de Cramer, avec un calcul plus long que celui donné par les formules de définition des déterminants[1]. Phénomène plus grave encore: si on applique la méthode du pivot à une matrice à coefficients dans $\mathbb{Q}$ sans simplifier les fractions obtenues au fur et à mesure des calculs, la taille des coefficients explosera.

Or, si on examine l'évolution des coefficients d'une matrice triangulée par la méthode du pivot, on s'aperçoit que tous les coefficients successifs peuvent s'exprimer comme quotients de

[1] Le corps $\mathbb{Q}((X_i)_{i \in \mathbb{N}})$ cité en exemple ne semble pas pouvoir être rendu complètement $\mathcal{P}$-calculable par une présentation adéquate: cf. fin du § B f

déterminants extraits de la matrice de départ. La morale est qu'on est a priori garanti contre une explosion de la taille des coefficients dans la méthode du pivot si on sait majorer la taille des déterminants.

Autrement dit: dans un $\mathcal{P}$-corps $\mathcal{P}$-dénombrable le seul obstacle à la $\mathcal{P}$-résolubilité des systèmes d'équations linéaires (et au calcul en temps polynomial des déterminants) par la méthode du pivot provient éventuellement de l'impossibilité de majorer polynomialement la taille des déterminants : l'algèbre linéaire sur un $\mathcal{P}$-corps $\mathcal{P}$-dénombrable est en temps polynomial si et seulement si la fonction dét est **RESP** (polynomialement majorée en taille du résultat). (Th C.d1)

Dans le § a) nous étudions la propriété pour un anneau d'être **Mat-c$\mathcal{P}$c**, c.-à-d. que le produit matriciel y est complètement $\mathcal{P}$-calculable. Elle équivaut à l'inversibilité en temps polynomial des matrices triangulaires avec des 1 sur la diagonale, et possède une forte stabilité.

Dans le § b) nous étudions la propriété pour un anneau commutatif d'être **Det-c$\mathcal{P}$c**, c.-à-d. avoir les déterminants $\mathcal{P}$-calculables , et nous montrons l'équivalence avec la propriété d'être **Mat-c$\mathcal{P}$c** dans le cas des $\mathcal{P}$-$\mathbb{Q}$-algèbres. Il est d'ailleurs probable que l'équivalence puisse être démontrée sans restriction aucune.

Dans le § c) consacré aux $\mathcal{P}$-corps commutatifs, nous montrons l'équivalence entre **Det-c$\mathcal{P}$c**, la $\mathcal{P}$-calculabilité des relations de dépendance linéaire entre vecteurs et la $\mathcal{P}$-résolubilité des systèmes d'équations linéaires. Nous montrons alors la $\mathcal{P}$-calculabilité de la géométrie des sous-espaces de dimension finie. Nous établissons enfin quelques nouvelles propriétés de stabilité pour les anneaux **Det-c$\mathcal{P}$c**.

Dans le § d) nous exposons la méthode de Bareiss, qui peut être vue soit comme une méthode de calculs de déterminants, soit comme une méthode du pivot "améliorée". Pour que cette méthode fonctionne en temps polynomial, il suffit que l'on soit dans un $\mathcal{P}$-anneau intègre $\mathcal{P}$-dénombrable où les divisions exactes peuvent être effectuées en temps polynomial.

## a) Calcul matriciel sur un $\mathcal{P}$-anneau

### Matrices sur un $\mathcal{P}$-anneau K

Nous nous intéressons à la complexité de calculs impliquant des matrices $A = (a_{ij})$ de $K^{hxc}$, les nombres de lignes et de colonnes h et c n'étant pas fixés a priori.

Il est naturel de prendre pour mesure de la grandeur de A le nombre $\| A \| = \Sigma \| a_{ij} \|$. Nous noterons $n_A$ le sup de h et c , et $s_A$ le sup des $\| a_{ij} \|$. La mesure A est polynomialement reliée à $n_A + s_A$. Si nécessaire, nous précisons $h_A$ et $c_A$ au lieu de h et c .

Il est pratique d'utiliser un ensemble **Mat**(K) , réunion disjointe des $K^{hxc}$. Nous pouvons par exemple réaliser cet ensemble sous la forme d'une $\mathcal{P}$-partie de **Lst**(**Lst**(K)) : une matrice est vue comme la liste de ses vecteurs colonnes $V_1 , V_2 , \ldots , V_c$ qui sont eux mêmes des éléments de **Lst**(K) : $V_j = (a_{1j} , a_{2j} , \ldots , a_{lj})$.

L'ensemble **Mat**(K) est muni de 2 lois de composition + et × non partout définies, mais définies sur des $\mathcal{P}$-parties convenables de **Mat**(K) × **Mat**(K).

Un élément de **Mat**(K) peut servir à représenter une application linéaire de $K^c$ vers $K^h$ ou un système de h vecteurs de $K^c$. En outre en modifiant la relation d'égalité, cette même matrice peut servir à représenter des objets de différents ensembles:

– une application linéaire de $K^{(\mathbb{N}_1)}$ vers $K^{(\mathbb{N}_1)}$, espace vectoriel[1] $\mathcal{P}$-présenté par **Lst**(K) muni de la relation d'égalité qui identifie une liste de c éléments à toute liste plus longue obtenue en rajoutant des 0 à la fin: nous noterons **Flin**(K) l'ensemble obtenu à partir de **Mat**(K) en changeant de la même manière la relation d'égalité . Lorsque K est commutatif, **Flin**(K) est une K-algèbre[2] sans élément neutre pour la multiplication. Il est clair que **Flin**(K) est un $\mathcal{P}$-ensemble.

– un sous-espace vectoriel de $K^h$ : celui engendré par les c vecteurs colonnes de la matrice. On ne peut dire a priori si l'ensemble obtenu , que nous noterons **Fsv**(K) , muni de la présentation ainsi décrite, est un $\mathcal{P}$-ensemble.

– un sous-espace vectoriel de $K^{(\mathbb{N}_1)}$ : nous noterons **Sv**(K) l'ensemble de ces sous-espaces, muni de cette présentation. C'est l'ensemble de tous les sous-espaces vectoriels finiment engendrés de $K^{(\mathbb{N}_1)}$.

Nous noterons **Lin**(K) l'ensemble des applications linéaires, de $K^{(\mathbb{N}_1)}$ vers $K^{(\mathbb{N}_1)}$, de la forme $a.I + M$ , où $a \in K$ et $M \in$ **Flin**(K), présenté sous la forme $K \times$ **Flin**(K) . Ceci revient à rajouter l'élément neutre manquant à **Flin**(K) . Lorsque K est commutatif c'est une K-algèbre unitaire.

**Proposition C.a1 :**

Les 4 propriétés suivantes du $\mathcal{P}$-anneau K sont équivalentes:

i. l'addition dans K est complètement $\mathcal{P}$-calculable
ii. le produit dans **Mat**(K) est une $\mathcal{P}$-fonction
iii. le produit dans **Flin**(K) est une $\mathcal{P}$-fonction
iv. **Lin**(K) est un $\mathcal{P}$-anneau
v. K[X] est un $\mathcal{P}$-anneau

Si ces propriétés sont vérifiées l'addition est c-$\mathcal{P}$-c dans **Mat**(K) , **Lin**(K) , **Flin**(K) et K[X].

*preuve>*

Il est clair que ii. , iii. et iv. sont équivalentes;

L'implication ii. $\Rightarrow$ i. se voit en multipliant un vecteur ligne ( 1 , 1 ,..., 1 ) par un vecteur colonne $(a_1 , a_2 , ... , a_c )$ ; l'implication i $\Rightarrow$ ii. s'obtient par un calcul immédiat.

L'implication i. $\Rightarrow$ v. a été démontrée en prop. B.f1; l'implication v. $\Rightarrow$ i. se démontre en considérant le coefficient de degré c-1 dans le polynôme produit du polynôme de coefficients ( 1 , 1 ,..., 1 ) par celui de coefficients $(a_1 , a_2 , ... , a_c )$ . ❑

## Inversion des matrices triangulaires et calculs de produits itérés de matrices

Nous noterons **Trimat**(K) l'ensemble des matrices carrées triangulaires supérieures de K , avec uniquement des 1 sur la diagonale . Une matrice $n \times n$ de **Trimat**(K) peut s'écrire $A = I - U$ avec $U^n = 0$ et admet une inverse $A^{-1} = I + U + U^2 + ... + U^{n-1}$.

---

[1] bien que K ne soit pas nécessairement un corps, nous employons le mot espace vectoriel que le lecteur voudra bien lire "module libre" (ou même "module à gauche libre" dans le cas non commutatif)

[2] rappel : nous utilisons "K-algèbre" pour K-algèbre associative, non forcément unitaire, et avec K commutatif seulement.

L'ensemble **Trimat**(K) ainsi présenté est manifestement un $\mathcal{P}$-ensemble.

**Théorème C.a1 :**

Les 4 propriétés suivantes du $\mathcal{P}$-anneau K sont équivalentes :

i. l'inversion dans **Trimat**(K) est une $\mathcal{P}$-fonction

ii. le produit dans **Mat**(K) est complètement $\mathcal{P}$-calculable

iii. le produit dans **Flin**(K) est complètement $\mathcal{P}$-calculable

iv. le produit dans **Lin**(K) est complètement $\mathcal{P}$-calculable

Lorsque c'est vérifié, l'addition et le produit dans K , **Mat**(K), **Flin**(K), **Lin**(K) et K[X] sont complètement $\mathcal{P}$-calculables.

*preuve>*

ii. $\Rightarrow$ i. : l'addition dans K est complètement $\mathcal{P}$-calculable d'après la proposition C.a1, elle l'est aussi dans **Mat**(K) , et dans ce cas, vu que le produit est également complètement $\mathcal{P}$-calculable, la formule ci-dessus donnant $A^{-1}$ est le programme d'un $\mathcal{P}$-calcul.

i. $\Rightarrow$ ii. :

– montrons d'abord que le produit dans K est complètement $\mathcal{P}$-calculable. On considère le système d'équations $x_1 = a_1 . x_2$ , $x_2 = a_2 . x_3$ , .... , $x_j = a_j . x_{j+1}$ , $x_{j+1} = 1$ ; il se résout en inversant la matrice triangulaire supérieure avec les $-a_i$ au dessus de la diagonale. Or cette matrice est $\mathcal{P}$-reliée à la liste $(a_1 , a_2 , ... , a_j)$.

– montrons ensuite que le produit dans **Mat**(K) est complètement $\mathcal{P}$-calculable: nous relisons la démonstration que nous venons de faire en interprétant les $a_i$ comme les matrices carrées dont le produit est à calculer et les $x_i$ comme des matrices carrées inconnues: on obtient la solution par inversion d'une matrice triangulaire constituée des blocs I sur la diagonale, $-a_i$ au dessus de la diagonale et 0 ailleurs.

On peut présenter ce même argument comme suit: si L est l'anneau des matrices carrées $n \times n$ sur K l'inversion dans **Trimat**(L) se déduit immédiatement de l'inversion dans **Trimat**(K).

On notera que le passage de la liste des matrices à multiplier à la matrice triangulaire constituée de blocs est bien de type $\mathcal{P}$.

– montrons maintenant que l'addition est complètement $\mathcal{P}$-calculable dans K: il suffit de multiplier les matrices triangulaires $2 \times 2$ ayant pour coefficient au dessus de la diagonale les $a_i$ qu'on veut additionner.

ii. et iii. sont clairement équivalents,

iv. implique iii. (clair)

iii $\Rightarrow$ iv. : supposons qu'on ait un produit d'éléments $(a_i.I + M_i)$ à calculer dans **Lin**(K) : soit n le sup des $n_{M_i}$, on effectue le produit des matrices $(a_i.I_n + M_i)$ dans l'anneau des matrices carrées $n \times n$ , on obtient une matrice P ; effectuons par ailleurs le produit des $a_i$ dans K , on obtient un élément a , on écrit P sous forme $a. I_n + M$ , et le produit à calculer dans **Lin**(K) n'est autre que $a.I + M$

Enfin si ii. est vérifiée on obtient que la multiplication dans K[X] est complètement $\mathcal{P}$-calculable comme suit:

si $P_1 , P_2 , ... , P_j$ sont les polynômes à multiplier, de degrés $d_1 , d_2 , ... , d_j$ , et si $n = 1 + d_1 + d_2 + ... + d_j$ , on considère, sur l'espace de dimension n formé par les polynômes de degré $n - 1$ , et pour la base canonique, les matrices des applications linéaires : multiplication par $P_i$ tronquée éventuellement au degré $n - 1$. Le produit des matrices correspond au produit des polynômes, et le polynôme produit se lit sur la 1ère colonne de la matrice produit. Il suffit donc de vérifier que le passage de la liste $[P_1 , P_2 , ... , P_j]$ à la liste des matrices $n \times n$ correspondantes est bien de type $\mathcal{P}$. ❑.

**Remarques :**

– divertissement mathématique : si le produit dans K[X] est c-$\mathcal{P}$-c, alors il en est de même dans **Mat**(K) ?

– lorsque K est un $\mathcal{P}$-corps vérifiant les propriétés équivalentes de la proposition, l'inversion des matrices est également une $\mathcal{P}$-fonction sur l'ensemble des matrices triangulaires avec des coefficients tous non nuls sur la diagonale.

Le théorème C.a1 nous conduit à poser la définition suivante:

**Définition C.a1** : Soit K un $\mathcal{P}$-anneau.

On dira que **le calcul des matrices dans K est complètement en temps polynomial,** ou encore que K est **Mat-c$\mathcal{P}$c** lorsqu'il vérifie les propriétés équivalentes énoncées au théorème C.a1.

On notera qu'en calcul purement formel le produit des matrices n'est pas c-$\mathcal{P}$-c (on peut le voir en multipliant n matrices $2 \times 2$ ) et que donc un anneau peut a priori être c-$\mathcal{P}$-c sans être **Mat-c$\mathcal{P}$c** . Problème ouvert : fournir un exemple concret où cette situation se produirait. (cf. à ce sujet le § B.g )

## Stabilité de la classe des anneaux Mat-c$\mathcal{P}$c

La classe des anneaux **Mat-c$\mathcal{P}$c** est très stable, comme en témoigne la proposition suivante.

**Proposition C.a2** : Soit K un anneau **Mat-c$\mathcal{P}$c** , alors sont également **Mat-c$\mathcal{P}$c** les anneaux suivants :

i. tout sous anneau de K qui est une partie $\mathcal{P}$-détachable de K
ii. tout $\mathcal{P}$-quotient de K
iii. l'anneau $\mathbf{M_n}$(K) des matrices carrées $n \times n$ à coefficients dans K
iv. **Lin**(K)
v. $K[X_1,X_2,...,X_n]$

et, lorsque K est commutatif

vi. le localisé en S de K , si S est une partie multiplicative $\mathcal{P}$-détachable de K ne contenant pas de diviseur de 0 .
vii. le corps des fractions de K , si K est intègre
viii. toute K-algèbre unitaire de dimension finie, si K est intègre et $\mathcal{P}$-divisible.

**Remarques :**

1) les anneaux que l'on peut construire à partir de $\mathbb{Z}$ et des anneaux finis par enchaînement de constructions correspondant aux stabilités ci-dessus énoncées forment un stock assez important.

2) pour que le quotient d'un $\mathcal{P}$-anneau K par un idéal bilataire soit un $\mathcal{P}$-quotient , il faut et il suffit que l'idéal soit $\mathcal{P}$-détachable.

*preuve>*

i. et ii. sont immédiats;

iii. se voit en "juxtaposant" les coefficients d'une matrice $h \times c$ à coefficients dans $\mathbf{M_n}$(K) de manière à en faire une matrice $h.n \times c.n$ à coefficients dans K , on est ramené à

effectuer des produits dans **Mat**(K) puis à réinterpréter le résultat en le décomposant en blocs $n \times n$.

vi. Tout d'abord l'addition dans $K_S$ est c-$\mathcal{P}$-c : pour additionner une liste de fractions, on les réduit au même dénominateur (or le produit dans S est c-$\mathcal{P}$-c ) puis on additionne les numérateurs. Cet argument revient à dire qu'en calcul purement formel, l'addition des fractions est c-$\mathcal{P}$-c.
Que le produit soit c-$\mathcal{P}$-c dans **Mat**($K_S$) résulte du fait qu'on obtient une $\mathcal{P}$-présentation $\mathcal{P}$-équivalente à **Mat**($K_S$) en prenant **Mat**(K) $\times$ S , où le couple (A,d) représente la matrice (1/d).A à coefficients dans $K_S$ . On termine en remarquant que le produit est c-$\mathcal{P}$-c dans **Mat**(K) et dans S.

vii. est un cas particulier de vi.

viii. se déduit de iii. et i. lorsque K est un $\mathcal{P}$-corps commutatif parce qu'une K-algèbre de dimension finie n est canoniquement isomorphe à une sous-algèbre de dimension n de $\mathbf{M_n}$(K): or toutes les applications linéaires entre espaces de dimensions finies sont des $\mathcal{P}$-fonctions et dans un espace vectoriel $K^m$ , tout sous espace de dimension n est $\mathcal{P}$-détachable parce qu'il est noyau d'une application linéaire.
Lorsque K est intègre et $\mathcal{P}$-divisible, l'argument s'applique au corps des fractions L de K. Une fois le calcul fait dans L , on sait revenir dans K puisque K est supposé $\mathcal{P}$-divisible.
**Remarque** : peut on trouver un argument plus général qui s'applique à toute K-algèbre unitaire de dimension finie ?

iv. On raisonne essentiellement comme pour le théorème C.a1 : pour effectuer le produit d'une liste de matrices à coefficients dans **Lin**(K), on considère le sup n des $n_{Mi}$ où les $a_i$. I + $M_i$ sont tous les coefficients de toutes les matrices de la liste et on effectue le produit de la liste des matrices correspondantes dont les coefficients "$a_i$. $I_n$ + $M_i$" sont dans $\mathbf{M_n}$(K) (cf. iii.) ; on écrit le résultat sous forme d'une matrice dont les coefficients sont de la forme a.$I_n$ + M , ( le a étant à chaque fois obtenu en effectuant le produit des matrices correspondantes et dont les coefficients sont les $a_i$. $I_n$ ) ; le résultat dans **Lin**(K) est la matrice "traduite" de celle obtenue en remplaçant chaque coefficient a. $I_n$ + M (qui est un élément de $\mathbf{M_n}$(K) ) par le coefficient a. I + M (qui est dans **Lin**(K) ). La chose importante à vérifier est que le passage de "la liste des matrices à coefficients dans **Lin**(K)" à "la liste des matrices à coefficients dans $\mathbf{M_n}$(K) " est bien de classe $\mathcal{P}$. Cela revient à rajouter tout plein de 0 et de $a_i$ ... mais on se convaincra sans difficulté que ça reste polynomialement majoré en taille.

v. On démontre le résultat pour K[X] et on raisonne comme au théorème C.a1 . Si $[A_1,A_2,...,A_n]$ est une liste de matrices de **Mat**( K[X] ) à multiplier, notons $p_{ijk}$ le coefficient en position i,j de la matrice $A_k$ et soit $d_k$ le degré maxi des $p_{ijk}$ pour k fixé. Alors dans le produit $A_1.A_2.....A_n$ , le degré maxi d'un coefficient est $d = d_1+d_2+...+d_n$ . Soit d' = d + 1 , on considère dans **Mat**($\mathbf{M_{d'}}$(K)) les matrices $B_1$, $B_2$, ..., $B_n$ dont les coefficients $b_{ijk}$ sont les matrices des applications linéaires " produit par $p_{ijk}$ tronqué éventuellement au degré d' " . On effectue le produit des $B_k$ puis on retraduit les coefficients du résultat dans K[X] . On vérifie que le passage de la liste $[A_1,A_2,...,A_n]$ à la liste $[B_1,B_2,...,B_n]$ est bien de classe $\mathcal{P}$ , c.-à-d. ici essentiellement que c'est **RESP**. ❑

## Exemples d'anneaux **Mat-c$\mathcal{P}$c**

Les $\mathcal{P}_{0'}$-anneaux sont **Mat-c$\mathcal{P}$c** d'après le proposition B.g2. Il y a donc $\mathbb{Z}$, les anneaux d'entiers dans les corps de nombre, les anneaux finis, puis les anneaux de polynômes sur ces anneaux, et tous $\mathcal{P}$-quotients de ces derniers.

Aux $\mathcal{P}_{0'}$-anneaux, nous pouvons ensuite rajouter tous ceux qu'on obtient en utilisant les

propriétés de stabilité énoncées en C.a2: notamment les localisés et corps de fractions qui ne peuvent être obtenus par B.g2.

## Inversion de matrices

**Définition C.a2** : On dira qu'un $\mathcal{P}$-anneau (commutatif ou non) K est **Inv-c$\mathcal{P}$c** lorsque les matrices carrées inversibles forment une $\mathcal{P}$-partie de **Mat**(K) , et que l'inversion des matrices carrées inversibles est une $\mathcal{P}$-fonction. On dira également : "**les matrices inversibles de Mat(K) sont $\mathcal{P}$-inversibles**" .

**Remarque** : On aura en particulier: les éléments inversibles de K forment une $\mathcal{P}$-partie de K , et le calcul de l'inverse d'un inversible est une $\mathcal{P}$-fonction.
De plus les matrices de **Trimat**(K) sont inversibles, et K est donc **Mat-c$\mathcal{P}$c**.

**Proposition C.a3** : Soit K un anneau **Inv-c$\mathcal{P}$c** , alors sont également **Inv-c$\mathcal{P}$c** les anneaux suivants :

i. tout sous anneau de K qui est une partie $\mathcal{P}$-détachable de K
ii. l'anneau $\mathbf{M_n}(K)$ des matrices carrées $n \times n$ à coefficients dans K
iii. **Lin**(K)
iv. $K[X_1,X_2,\ldots,X_n]$ : lorsque K est commutatif et intègre
v. toute K-algèbre unitaire de dimension finie K' (c.-à-d.: K' est un K-module libre de dimension finie): lorsque K est commutatif, intègre et $\mathcal{P}$-divisible

*preuve>*

le i. est immédiat.
ii. et iii. se vérifient en inversant une matrice composée de blocs $n \times n$.
v. se déduit de ii. et i. comme à la proposition C.a2

Nous montrons iv. tout d'abord pour K[X] . Si B est la matrice inverse d'une matrice A à coefficients dans K[X] , nous pouvons majorer a priori le degré des coefficients de B par un entier d polynomialement relié à la taille de A , puisque B est (à un scalaire multiplicatif près) la matrice transposée des cofacteurs de A . Nous procédons alors comme pour la proposition C.a2: nous "remplaçons" chaque coefficient polynôme Q par la matrice de l'application linéaire "produit par Q éventuellement tronqué au degré d". Nous avons donc remplacé la matrice A à coefficients dans K[X] par une matrice formée de blocs $d' \times d'$ ( $d' = d+1$ ) à coefficients dans K . Il nous reste à tester si la matrice obtenue est inversible dans **Mat**(K), et, si c'est le cas, à examiner les blocs $d' \times d'$ extraits de la matrice inverse pour vérifier s'ils sont de la forme voulue.
Pour $K[X_1,X_2,\ldots,X_n]$ , nous pouvons procéder par récurrence sur n. ❑

## Un résultat sur les anneaux commutatifs intègres $\mathcal{P}$-divisibles

**Proposition C.a4** : Supposons K intègre $\mathcal{P}$-divisible et **Mat-c$\mathcal{P}$c** , soit L son corps de fractions (K est identifiable à une partie $\mathcal{P}$-détachable de L). Alors:

- La division euclidienne dans L[X] est une $\mathcal{P}$-fonction de $L[X] \times L[X]$ vers $(L[X] \times L[X]) \cup \{u\}$ ( u pour le cas où on divise par 0 ) .
- Les anneaux K[X] et L[X] sont $\mathcal{P}$-divisibles .
- Les anneaux $K[X_1,X_2,\ldots,X_n]$ et $L[X_1,X_2,\ldots,X_n]$ sont $\mathcal{P}$-divisibles .

*preuve>* Soient 2 polynômes A et B , de degrés d et d + d' , avec d' ≥ 0 . Diviser A par B revient à exprimer linéairement A sur la base $1 , X , ... , X^{d-1} , B , B.X , ... , B.X^{d'}$ (dans l'espace des polynômes de degré ≤ d + d' ) . Or cette base est triangulaire par rapport à la base canonique (sur laquelle est exprimé A ), avec des coefficients tous non nuls sur la diagonale ( 1 ou le coefficient dominant de B ). La matrice correspondante est donc $\mathcal{P}$-inversible .
Le reste suit immédiatement. (la récurrence fonctionne grâce à la proposition précédente C.a2 v. ) ❑

## b) Cas commutatif : déterminants, formules de Cramer et inversions de matrices

### $\mathcal{P}$-calculabilité des déterminants

**Définition C.b1** : On dira qu'un $\mathcal{P}$-anneau commutatif K est **Det-c$\mathcal{P}$c** lorsque la fonction "déterminant" , (définie sur la $\mathcal{P}$-partie convenable de **Mat**(K) ), est une $\mathcal{P}$-fonction. On dira également : "**les déterminants sont $\mathcal{P}$-calculables dans K**".

**Définition C.b2** : On dira qu'un $\mathcal{P}$-anneau commutatif K est **Inv-1-c$\mathcal{P}$c** lorsque les matrices de déterminant égal à 1 forment une $\mathcal{P}$-partie de **Mat**(K) , et que l'inversion des matrices carrées de déterminant égal à 1 est une $\mathcal{P}$-fonction. On dira également : "**les matrices de Mat(K) de déterminant 1 sont $\mathcal{P}$-inversibles**" .

**Proposition C.b1** : Soit K un anneau commutatif **Det-c$\mathcal{P}$c** , alors :

i K est également **Inv-1-c$\mathcal{P}$c** et **Mat-c$\mathcal{P}$c**

De plus sont alors également **Det-c$\mathcal{P}$c** :

ii. tout sous anneau de K qui est une partie $\mathcal{P}$-détachable de K

iii. tout $\mathcal{P}$-quotient de K

iv. le localisé en S de K , si S est une partie multiplicative $\mathcal{P}$-détachable de K ne contenant pas de diviseur de 0 .

*preuve>*

le i. se déduit des formules de Cramer et du théorème C.a1 ;

le ii. et le iii. sont immédiats

le iv. résulte du fait que la "réduction au même dénominateur" de tous les coefficients d'une matrice de **Mat**( $K_S$ ) est une $\mathcal{P}$-opération. ❑

### Le cas des $\mathcal{P}$-$\mathbb{Q}$-algèbres : la méthode de Leverrier

**Théorème C.b1** : Soit K une $\mathcal{P}$-$\mathbb{Q}$-algèbre unitaire et commutative.
Alors les propriétés suivantes sont équivalentes:

i. le produit dans **Mat**(K) est c-$\mathcal{P}$-c (c.-à-d. K est **Mat-c$\mathcal{P}$c**)

ii. les déterminants sont $\mathcal{P}$-calculables dans K

iii. l'inversion des matrices triangulaires supérieures avec des 1 sur

la diagonale est un $\mathcal{P}$-calcul

iv. les matrices de **Mat**(K) de déterminant 1 sont $\mathcal{P}$-inversibles

Lorsque les inversibles de K forment une $\mathcal{P}$-partie, sur laquelle le calcul de l'inverse est un $\mathcal{P}$-calcul, ces propriétés sont en outre équivalentes à la suivante:

v. les matrices inversibles de **Mat**(K) sont $\mathcal{P}$-inversibles

*preuve>*

i. et iii. sont équivalents d'après le théorème C.a1.

ii. implique iv. et v. par les formules de Cramer

iv. et v. impliquent séparément iii.

Il reste à voir que i. implique ii., c.-à-d. en gros à calculer un déterminant en n'utilisant que des produits de matrices: la **méthode de Leverrier**, dont le seul inconvénient est d'utiliser des divisions par des entiers, convient pour les $\mathbb{Q}$-algèbres. Soit A une matrice carrée $n \times n$ à coefficient dans K (supposé **Mat-c$\mathcal{P}$c**), soit $P(X) = X^n - ( a_1.X^{n-1} + a_2.X^{n-2} + .... + a_n)$ son polynôme caractéristique, et soit $S_i$ la "somme des puissances i-èmes des valeurs propres" (cela "a un sens" même si K n'est pas intègre: $S_i$ est une fonction polynôme (à coefficients entiers) des $a_j$ donnée par les formules de Newton).

Alors $S_i$ est la trace de $A^i$ (il s'agit d'une identité algébrique, et il suffit donc qu'elle soit vraie dans le cas d'un corps algébriquement clos). Le calcul $A \longrightarrow [S_1,S_2,... ,S_n]$ est donc un $\mathcal{P}$-calcul. Ensuite il reste à calculer les $a_j$ par les formules de Newton, qui s'écrivent matriciellement:

$$\begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ S_1 & 2 & 0 & \cdots & 0 \\ S_2 & S_1 & 3 & \cdots & 0 \\ S_3 & S_2 & S_1 & \cdots & 0 \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \end{vmatrix} \begin{vmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ \cdot \\ \cdot \end{vmatrix} = \begin{vmatrix} S_1 \\ S_2 \\ S_3 \\ S_4 \\ \cdot \\ \cdot \end{vmatrix}$$

etc

Il s'agit donc de résoudre un système linéaire triangulaire dont les coefficients sont $\mathcal{P}$-donnés en fonction de $[S_1,S_2,... ,S_n]$.

En multipliant la i-ème ligne par 1/i, on est ramené à inverser une matrice de **Trimat**(K), puis à multiplier l'inverse obtenue par le vecteur colonne du second membre.

La multiplication par 1/i est un $\mathcal{P}$-calcul parce que K est une $\mathcal{P}$-$\mathbb{Q}$-algèbre. L'inversion dans **Trimat**(K) et le produit dans **Mat**(K) sont des $\mathcal{P}$-calculs parce que K est **Mat-c$\mathcal{P}$c**. ❑

**Problème ouvert** : Dans tout anneau commutatif **Mat-c$\mathcal{P}$c** les déterminants sont $\mathcal{P}$-calculables ?

**Remarques :**

1) tout $\mathcal{P}$-corps **Mat-c$\mathcal{P}$c** de caractéristique nulle est une $\mathcal{P}$-$\mathbb{Q}$-algèbre puisque $\mathbb{Q}$ est objet initial dans la catégorie (cf. B.c))

2) en combinant les théorèmes C.b1, C.b2 et les propriétés de stabilité pour le caractère **Det-c$\mathcal{P}$c** (cf. prop. C.b1, C.c5 et C.c6), on obtiendra un bon stock d'anneaux **Det-c$\mathcal{P}$c**.

3) la méthode de Leverrier est souvent présentée sous une forme améliorée dite **méthode de Faddev** :

$$\begin{array}{lll} A_1 = A & a_1 = tr(A_1) & B_1 = A_1 - a_1 I \\ A_2 = A B_1 & a_2 = tr(A_2) / 2 & B_2 = A_2 - a_2 I \\ A_3 = A B_2 & a_3 = tr(A_3) / 3 & B_3 = A_3 - a_3 I \\ \ldots\ldots\ldots & \ldots\ldots\ldots & \ldots\ldots\ldots \\ A_n = A B_{n-1} & a_n = tr(A_n) / n & B_n = A_n - a_n I = 0 \end{array}$$

Cette méthode peut être utilisée sous les mêmes hypothèses qui rendent la méthode de Leverrier en temps polynomial, en rajoutant toutefois une hypothèse de $\mathcal{P}$-réductiblité. En effet, les $a_i$ sont convenablement majorés en taille et on a l'égalité :

$$\begin{aligned} A_i &= A ( A ( \ldots A ( A - a_1 I ) - a_2 I ) \ldots ) - a_{i-1} I ) \\ &= A^i - ( a_1 A^{i-1} + a_2 A^{i-2} + \ldots + a_{i-1} A ) \end{aligned}$$

Donc si on suit l'algorithme de Fadeev en réduisant chaque $A_i$ et chaque $a_i$ intermédiaires obtenus, la taille du calcul est convenablement maîtrisée.

4) la méthode de Fadeev s'applique également avec un anneau commutatif **Mat-c$\mathcal{P}$c** $\mathcal{P}$-réductible K où la division par un entier, quand elle est possible, est unique[1], et réalisée par une opération en temps polynomial.

## La méthode de Samuelson

La méthode de Samuelson (cf. **[Sam]** ou **[Ber]**) est une méthode qui permet de calculer le polynôme caractéristique d'une matrice de manière récurrente.

Pour une matrice A à n lignes et n colonnes, on écrit :

$$A = \begin{array}{|c|c|} \hline a_{1,1} & R \\ \hline S & M \\ \hline \end{array}$$

Si $p(\lambda) = \sum_{i=0}^{n} p_{n-i}.\lambda^i = \det ( A - \lambda )$ $( p_0 = (-1)^n )$ et

$$q(\lambda) = \sum_{i=0}^{n-1} q_{n-1-i}.\lambda^i = \det ( M - \lambda.I ) \quad ( q_0 = (-1)^{n-1} )$$

La matrice adjointe de $( A - \lambda.I)$ est donnée par

$$adj( A - \lambda.I ) = - \sum_{k=2}^{n} ( M^{k-2}.q_0 + M^{k-3}.q_1 + \ldots + I.q_{k-2} ) \lambda^{n-k}$$

Et $p(\lambda)$ est donné par :

$$p(\lambda) = (a_{1,1} - \lambda) q(\lambda) + R . adj( A - \lambda.I ) . S$$

On en déduit immédiatement la :

---

[1] c.-à-d: K est sans torsion en tant que $\mathbb{Z}$-module

**Proposition C.b2** : Soit K un anneau commutatif **Mat-c$\mathcal{P}$c** , $\mathcal{P}$-réductible où les coefficients du polynôme caractéristique sont polynomialement majorés en taille , alors K est **Det-c$\mathcal{P}$c**.

Si les déterminants sont $\mathcal{P}$-majorés, ils sont $\mathcal{P}$-calculables ?

Le résultat proposé ci-dessus en interrogation fait l'objet du:

**Problème ouvert** : Si K est un $\mathcal{P}$-anneau commutatif $\mathcal{P}$-réductible où les déterminants sont polynomialement majorés en taille, alors ils sont $\mathcal{P}$-calculables ?

Nous démontrerons le résultat dans trois cas particuliers: lorsque K est intègre et $\mathcal{P}$-divisible (cf. Th. C.d1, via la méthode de Bareiss), lorsque K est un anneau où la division par un entier, quand elle est possible, est unique, et réalisée par une opération en temps polynomial (Th. C.b3, via la méthode de Fadeev), et lorsque K est un anneau où les coefficients du polynôme caractéristique sont polynomialement majorés en taille (Th C.b2, via la méthode de Samuelson).

**Proposition C.b3** : Soit K un $\mathcal{P}$-anneau commutatif $\mathcal{P}$-réductible où la fonction déterminant est **RESP** , alors K est **Mat-c$\mathcal{P}$c**.

*preuve*> Montrons tout d'abord que l'addition dans K est complètement $\mathcal{P}$-calculable:
Il suffit de montrer que l'addition itérée est **RESP**, or le déterminant de la matrice ci-dessous est égal à la somme a + b + c ... + k :

$$\begin{vmatrix} a & 1 & 1 & \cdots & 1 \\ -b & 1 & 0 & \cdots & 0 \\ -c & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ -k & 0 & 0 & \cdots & 1 \end{vmatrix}$$

Donc, l'addition itérée est **RESP** si on fait suivre chaque addition (dans une addition itérée) par une "réduction" du résultat (l'anneau est supposé $\mathcal{P}$-réductible).
On démontre de même que le produit dans K est c-$\mathcal{P}$-c en considérant le déterminant d'une matrice diagonale.
Voyons maintenant le produit matriciel. Tout d'abord le produit de 2 matrices est $\mathcal{P}$-calculable parce que l'addition est c-$\mathcal{P}$-c (prop. C.a1). Par ailleurs l'inversion dans **Trimat**(K) est **RESP** puisque la fonction déterminant est **RESP** . Le même raisonnement qui dans la preuve du Th C.a1, montrait que: " l'inversion dans **Trimat**(K) est $\mathcal{P}$ " implique " le produit dans **Mat**(K) est c-$\mathcal{P}$-c " ; ce même raisonnement montre maintenant que le produit dans **Mat**(K) est **RESP**. ❑

En combinant les propositions C.b2 et C.b3 on obtient immédiatement le :

**Théorème C.b2** : Soit K un $\mathcal{P}$-anneau commutatif $\mathcal{P}$-réductible où les coefficients du polynôme caractéristique sont polynomialement majorés en taille, alors K est **Det-c$\mathcal{P}$c** (par la méthode de Samuelson).

**Théorème C.b3 :**

Soit K un $\mathfrak{P}$-anneau commutatif $\mathfrak{P}$-réductible où la division par un entier, quand elle est possible, est unique, et réalisée par une opération en temps polynomial . Supposons en outre que les déterminants sont polynomialement majorés en taille. Alors les déterminants sont $\mathfrak{P}$-calculables dans K (par la méthode de Fadeev).

*preuve>* On applique la proposition C.b3 et la remarque 4) qui suit le théorème C.b1. ❑

## c) Systèmes linéaires à coefficients dans un $\mathfrak{P}$-corps commutatif

Dans ce § , K désignera un $\mathfrak{P}$-anneau commutatif intègre et $\mathfrak{P}$-divisible, et L son corps de fractions. Un cas particulier est celui où K est un $\mathfrak{P}$-corps, et L = K .

Dépendance linéaire, inversion de matrices, déterminants

Dans la définition qui suit, nous notons **Col**(K) la $\mathfrak{P}$-partie de **Mat**(K) formée des matrices à une seule colonne.

**Définition C.c1 :** On dira qu'un $\mathfrak{P}$-anneau commutatif intègre et $\mathfrak{P}$-divisible K est **Dep-c$\mathfrak{P}$c** lorsque les relations de dépendance linéaires entre vecteurs peuvent être $\mathfrak{P}$-calculées au sens suivant : il existe une $\mathfrak{P}$-opération D de **Mat**(K) vers **Col**(K) ∪ {u} vérifiant:

- si D(M) = u , les vecteurs colonnes de la matrice M sont linéairement indépendants
- si D(M) = V est un vecteur colonne, alors $V \neq 0$ et M.V = 0 : c.-à-d. que les coefficients de V définissent une relation de dépendance linéaire entre les vecteurs colonnes de la matrice M.

On dira encore "**la dépendance linéaire est $\mathfrak{P}$-calculable dans K**".

La morale de la proposition qui suit est essentiellement que tout calcul systématique d'inverses de matrices contient (de manière éventuellement cachée) un calcul de déterminants, et que ce fait s'étend aux calculs en temps polynomial.

**Proposition C.c1 :** Soient K et L comme précisés au début du § , alors les propriétés suivantes sont équivalentes :

i. K est **Det-c$\mathfrak{P}$c**
ii. K est **Dep-c$\mathfrak{P}$c**
j. L est **Det-c$\mathfrak{P}$c**
jj. L est **Dep-c$\mathfrak{P}$c**
jjj. L est **Inv-c$\mathfrak{P}$c**

*preuve>* i. et j. sont clairement équivalents (cf. prop. C.b1)
ii. et jj. de même (par les procédés de réduction au même dénominateur)
j. ⇒ jjj. par les formules de Cramer

jjj. $\Rightarrow$ L est **Mat-c$\mathcal{P}$c** (cf. Th. C.a1)

jjj. $\Rightarrow$ jj. : dans la matrice M à h lignes et c colonnes , on cherche tout d'abord un coefficient non nul ( h.c tests ) , puis une matrice $2 \times 2$ extraite de M , contenant ce coefficient non nul , et inversible ( (h-1).(c-1) tests) , puis, si on a trouvé, une matrice $3 \times 3$ extraite de M , contenant la matrice précédemment trouvée, et inversible ( (h-2).(c-2) tests ) . On continue jusqu'à constater l'indépendance linéaire des vecteurs colonnes , ou dans le cas contraire l'indépendance de $h' < h$ vecteurs colonnes et le fait que tout autre vecteur colonne dépend linéairement de ces h' qu'on a déterminés : dans ce cas, pour obtenir une relation de dépendance linéaire on utilise la dernière matrice inversible trouvée P (une matrice $h' \times h'$ ) , si W est un des vecteurs liés aux h' vecteurs en question , on extrait W' de W (en ne gardant que les coefficients des lignes intervenant dans P), et on calcule $W'.P^{-1}$ . (les détails laissés au lecteur)

jj. $\Rightarrow$ jjj. immédiat , donc en particulier jj. $\Rightarrow$ L est **Mat-c$\mathcal{P}$c**.

jj. $\Rightarrow$ j. : on considère la matrice carrée dont on veut trouver le déterminant comme la matrice d'une application linéaire f de $K^n$ vers $K^n$ , on va $\mathcal{P}$-construire une base (en utilisant le fait que L est **Mat-c$\mathcal{P}$c**) $f_1 , f_2 , \ldots , f_n$ , par rapport à laquelle l'expression de f sera particulièrement simple, et le déterminant immédiatement calculable. On commence par $f_1 = e_1$ , de la base canonique, on continue avec $f( f_1 ) = f_2$ , $f( f_2 ) = f_3$ , .... tant que ces vecteurs sont indépendants. Supposons qu'on ait déjà construit jusqu'à $f_j$ et que $f( f_j )$ soit dépendant des précédents . Alors on prendra pour $f_{j+1}$ le premier vecteur de la base canonique indépendant de $f_1 \ldots f_j$ , et on recommence à partir de là le même processus.

Les vecteurs de la base obtenue, sont tous (en tant que vecteurs colonnes) extraits des matrices A, $A^2,\ldots,A^n$. Il est donc clair que la matrice de f sur cette base est donnée par un $\mathcal{P}$-calcul (vu l'hypothèse jj.), et son déterminant est égal au produit de quelques uns de ces éléments. (cf. exemple ci-dessous, où on a mis des "blocs" en évidence en les séparant par des blancs)

A =

| | | |
|---|---|---|
| . . . . a | . . . x | . x |
| 1 . . . x | . . . x | . x |
| . 1 . . x | . . . x | . x |
| . . 1 . x | . . . x | . x |
| . . . 1 x | . . . x | . x |
| . . . . . | . . . b | . x |
| . . . . . | 1 . . x | . x |
| . . . . . | . 1 . x | . x |
| . . . . . | . . 1 x | . x |
| . . . . . | . . . . | . c |
| . . . . . | . . . . | 1 x |

les points représentent des 0

les croix représentent des coefficients non précisés

det(A) = a.(-b).(-c)

## Systèmes linéaires à coefficients dans un $\mathcal{P}$-corps commutatif

Un système de c équations linéaires à h inconnues dans L est donné par :

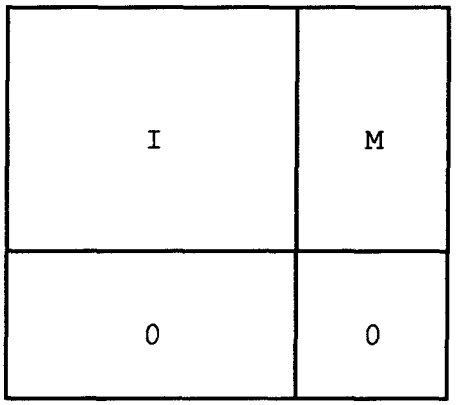
- une matrice $A = ( a_{ij} )$ de $L^{hxc} \subset$ **Mat**(L)
- un vecteur colonne $V = (b_i)$ de $L^h \subset$ **Col**(L)

Le système d'équations ainsi représenté est, pour i de 1 à h : $\Sigma_j \, a_{ij}. x_j = b_i$

Le système d'équations est équivalent à l'équation matricielle $A.X = V$ .

Les systèmes d'équations linéaires sur L forment un $\mathcal{P}$-ensemble, que nous noterons **Syslin**(L).

Un système linéaire sur L est dit **trivial** si la matrice A est de la forme:




De manière générale, lorsque L est un corps discret, on sait que tout système linéaire sur L est équivalent à un système trivial, à une permutation des inconnues près; résultat qu'on peut obtenir par la méthode du pivot de Gauss.

Un **système linéaire résolu** sur L est donné par :

- un système linéaire trivial sur L : ( T , W )
- l'indication de la permutation des colonnes à effectuer: par exemple sous la forme d'une matrice de permutation P.

Au système linéaire résolu sur ( T , W , P ) correspond le système linéaire ( T.P , W ). L'ensemble des systèmes linéaires résolu sur L est un $\mathcal{P}$-ensemble.

Lorsqu'un système linéaire sur L est sous forme résolue, il est immédiat de déterminer s'il est impossible, s'il a une solution unique, ou si on peut choisir arbitrairement un certain nombre d'inconnues, et d'exprimer les autres en fonction de celles choisies.

**Définition C.c2** : On dira que les systèmes linéaires sont $\mathcal{P}$-résolubles dans L s'il existe une $\mathcal{P}$-opération qui transforme tout système linéaire en un système linéaire résolu équivalent.

**Proposition C.c2** : Pour tout $\mathcal{P}$-corps commutatif L , les propriétés suivantes sont équivalentes:

- les systèmes linéaires sont $\mathcal{P}$-résolubles dans L
- la dépendance linéaire est $\mathcal{P}$-calculable dans L
- les déterminants sont $\mathcal{P}$-calculables dans L
- les matrices inversibles sont $\mathcal{P}$-inversibles dans L

*preuve*> vu la prop. C.c1 les 3 dernières propriétés sont équivalentes.
Si L est **Dep-c$\mathcal{P}$c**, et si on veut résoudre un système linéaire ( A , V ) , on considère les vecteurs colonnes de la matrice A , on en extrait un système libre maximal. Si V est linéairement indépendant de ce système libre, le système linéaire de départ est impossible. Dans le cas contraire, les variables correspondant aux vecteurs colonnes ne faisant pas partie du système libre maximal choisi peuvent être choisies librement. Et il est immédiat de déterminer le "système linéaire résolu" équivalent au système de départ à partir de l'expression des vecteurs colonnes considérés en fonction des vecteurs du système libre maximal.
Si maintenant les systèmes linéaires sont $\mathcal{P}$-résolubles dans L, il est immédiat que L est **Dep-c$\mathcal{P}$c** : l'indépendance linéaire de k vecteurs revient à l'impossibilité d'exprimer linéairement chaque vecteur en fonction des précédents, c.-à-d. à l'impossibilité pour k systèmes linéaires etc ... ❑

## Géométrie des sous-espaces

On considère maintenant le calcul "géométrique" dans **Sv**(L) , **Fsv**(L) .

**Proposition C.c3** : Soit L un $\mathcal{P}$-corps commutatif **Det-c$\mathcal{P}$c**.

i. L'égalité dans **Sv**(L) et **Fsv**(L) est $\mathcal{P}$-décidable. Et il y a une $\mathcal{P}$-opération qui associe à tout élément de **Sv**(L) ou de **Fsv**(L) une base de cet espace.
ii. La somme et l'intersection sont complètement $\mathcal{P}$-calculables dans **Sv**(L) et **Fsv**(L)
iii. La fonction $f \longrightarrow Im(f)$ de **Mat**(L) vers **Fsv**(L) est $\mathcal{P}$-calculable
iv. La fonction $f \longrightarrow Ker(f)$ de **Mat**(L) vers **Fsv**(L) est $\mathcal{P}$-calculable

*preuve>* le i. est laissé au lecteur; le iii. est trivial; que la somme de sous-espaces soit complètement $\mathcal{P}$-calculable est
également trivial; pour l'intersection nous utiliserons un peu de dualité:

**Lemme:**

La fonction : E $\longrightarrow$ orthogonal de E
(sous-espace de $K^h$) (dans le dual de $K^h$ identifié à $K^h$ )
de **Fsv**(L) vers **Fsv**(L) est une $\mathcal{P}$-fonction.

Il est clair que ce lemme implique que l'intersection est complètement $\mathcal{P}$-calculable dans **Fsv**(L) , d'où ensuite dans **Sv**(L).
Par ailleurs le lemme est équivalent à l'affirmation que la fonction Ker de **Mat**(L) vers **Fsv**(L) est une $\mathcal{P}$-fonction. Mais calculer le noyau d'une application linéaire f de matrice A n'est rien d'autre que résoudre le système linéaire $A.X = 0$ (détails laissés au lecteur). ❑

## Calculs dans les anneaux de polynômes

**Proposition C.c4** : Soit L un $\mathcal{P}$-corps commutatif **Det-c$\mathcal{P}$c**. Alors:

- L'interpolation est une $\mathcal{P}$-fonction de **Lst**($L \times L$) vers L[X] .
- La relation de Bezout dans L[X] est $\mathcal{P}$-calculable.

*preuve>* L'interpolation est la fonction qui associe à une liste de couples $(x_0,y_0)$, $(x_1,y_1)$, ... , $(x_d,y_d)$ (avec les $x_i$ 2 à 2 distincts) l'unique polynôme P de L[X] qui vérifie $\deg(P) \leq d$ et $P(x_i) = y_i$ pour i de 0 à d . Il s'agit en fait de résoudre un système d'équations linéaires dont les coefficients sont les $x_{ij}$ (pour i et j de 0 à d ), avec les $y_i$ au second membre. C'est donc une $\mathcal{P}$-fonction.

Pour ce qui concerne la relation de Bezout dans L[X] , il s'agit encore de résolution de systèmes linéaires : cf. par exemple [**ALFA**] p 211-216. ❑

## Stabilité **Det-c$\mathcal{P}$c** pour des anneaux de polynômes

**Proposition C.c5** : Soit L un $\mathcal{P}$-corps commutatif **Det-c$\mathcal{P}$c**, corps des fractions de K , $\mathcal{P}$-divisible. Supposons de plus L $\mathcal{P}$-infini, (c.-à-d. : il existe une $\mathcal{P}$-fonction injective de $\mathbb{N}_1$ vers L ).
Alors les déterminants sont $\mathcal{P}$-calculables par interpolation dans $K[X_1, X_2, \ldots, X_n]$ et $L[X_1, X_2, \ldots, X_n]$ .

*preuve>* Voyons que K[X] est **Det-c$\mathcal{P}$c** . Cela permettra de conclure pour $K[X_1,X_2,\ldots,X_n]$ par récurrence : on rappelle que $K[X_1,X_2,\ldots,X_n]$ est $\mathcal{P}$-divisible (prop.

C.a4)).
Comme K[X] est une partie $\mathcal{P}$-détachable de L[X] , il suffit de traiter ce dernier cas.
Pour calculer le déterminant d'une matrice à coefficients dans L[X] , on donne une majoration d de son degré (par ex la somme des sup des degrés dans chaque colonne) puis on calcule le polynôme par interpolation en d + 1 points distincts, d'où la nécessité, pour rester dans la classe des $\mathcal{P}$-calculs, de disposer d'une $\mathcal{P}$-fonction injective $\mathbb{N}_1$ vers L . Pour calculer la valeur du polynôme "déterminant" en l'un de ces d + 1 points, il faut évaluer chaque coefficient de la matrice en ce point, puis calculer dans L le déterminant de la matrice obtenue. Tout ceci reste un $\mathcal{P}$-calcul. ❑

**Remarques:**

1) Si L est un corps fini, on peut voir que les déterminants sont $\mathcal{P}$-calculables dans $L[X_1,X_2,...,X_n]$ de différentes manières possibles:
– si $L = \mathbf{F_p}$, p premier, L[X] est un $\mathcal{P}$-quotient de $\mathbb{Z}[X]$ . Or $\mathbb{Z}[X]$ est **Det-c$\mathcal{P}$c** (par l'argument précédent, ou bien par le Th C.b1 appliqué à $\mathbb{Q}[X]$ , ou bien par la proposition C.b2).
– dans le cas général, on peut appliquer un argument du même type, en remplaçant $\mathbb{Z}$ par une extension finie . Le plus simple est d'invoquer la proposition C.b2 ou le Th C.d1 à venir.

2) Tout $\mathcal{P}$-corps commutatif L qui est **Det-c$\mathcal{P}$c** et de caractéristique nulle est $\mathcal{P}$-infini, et les calculs de polynômes par interpolation peuvent être faits pour des x entiers: en effet l'homomorphisme canonique $\mathbb{Z} \rightarrow L$ est une $\mathcal{P}$-fonction puisque L est **Mat-c$\mathcal{P}$c** et que $\mathbb{Z}$ est initial parmi les anneaux où addition et produit sont c-$\mathcal{P}$-c.

## Stabilité **Det-c$\mathcal{P}$c** pour les extensions finies

**Proposition C.c6** : Soit L un $\mathcal{P}$-corps commutatif **Det-c$\mathcal{P}$c**.
Soit L' une algèbre de dimension finie n sur L qui est de plus un corps, commutatif ou non. Alors L' est **Inv-c$\mathcal{P}$c** (pour sa présentation comme L-espace vectoriel de dimension finie n ). Si L' est commutatif, il est **Det-c$\mathcal{P}$c**

*preuve*> Le corps L' est $\mathcal{P}$-isomorphe à une sous-L-algèbre $\mathcal{P}$-détachable de $\mathbf{M_n}(L)$ (cf. la preuve de la proposition C.a2 viii). Il suffit donc d'appliquer la proposition C.a3.❑

# d) Evolution des coefficients dans la méthode du pivot (méthode de Bareiss)

Nous en venons à l'analyse concrète de la méthode du pivot.

Nous sommes intéressés par un critère permettant d'assurer que la méthode du pivot, économique en nombre d'opérations arithmétiques élémentaires (nettement plus efficace de ce point de vue que la méthode de Leverrier utilisée au Th C.b1) ne conduit pas à une explosion des coefficients.

L'analyse que nous faisons de la méthode du pivot est essentiellement la même que celle de Gantmacher dans **[Gan]** chapitre 2 . Nous en déduisons une présentation "pédagogique" de la méthode dite de Bareiss (cf. **[Bar]**). Il s'avère que la méthode de Bareiss était peut-être connue de Sylvester et sûrement de Aitken (cf. **[Ait]**).

Nous obtenons essentiellement la même condition suffisante pour la $\mathcal{P}$-calculabilité des déterminants dans un $\mathcal{P}$-anneau K, qu'au Théorème C.b2 .

Nous avons besoin dans les 2 cas de supposer la $\mathcal{P}$-réductibilité de K et la $\mathcal{P}$-majoration des déterminants en taille.
Alors que le Th C.b1 s'applique à toutes les $\mathcal{P}$-$\mathbb{Q}$-algèbres, la méthode du pivot ne s'applique qu'avec des anneaux intègres (mais sans hypothèse de caractéristique nulle).
D'autre part la méthode "améliorée" à la Bareiss (qui est nécessaire pour éviter l'explosion de la taille des coefficients), exige en plus que l'anneau soit $\mathcal{P}$-divisible.
En un sens, il est d'ailleurs peu surprenant qu'un gain dans la vitesse du calcul soit compensé par des hypothèses renforcées sur l'anneau K et la $\mathcal{P}$-calculabilité dans l'anneau K.

## L'exemple de $\mathbb{Z}$ : résolution (dans $\mathbb{Q}$) d'un système linéaire à coefficients dans $\mathbb{Z}$

Un système linéaire de Nlin équations à Ncol inconnues peut être considéré comme donné par une matrice C (de type Nlin × Ncol) et un vecteur colonne "second membre" B (à Nlin coefficients).

On peut désirer traiter plusieurs seconds membres simultanément, de sorte que B serait une matrice avec Nsm colonnes (par exemple, si C est une matrice carrée, on peut prendre pour B la matrice carrée I, ce qui conduit directement au calcul de l'inverse de C).
L'inconnue est alors une matrice X de type Ncol × Nsm pour laquelle on demande que soit réalisée l'égalité C.X = B .

### *Présentation des coefficients:*

Les coefficients sont des nombres rationnels. Décidons de représenter les systèmes linéaires successifs avec un dénominateur fixé dans chaque ligne.

Nous utiliserons pour cela une seule matrice A , à coefficients entiers, avec Nlin lignes et 1+Ncol+Nsm colonnes numérotées de 0 à Ncol+Nsm: dans la colonne 0 nous mettrons le dénominateur commun à la ligne correspondante.

### *Un exemple pour voir :*

Nous allons voir maintenant sur un exemple "comment les choses se passent". Cet exemple a été choisi "au hasard", et aucune permutation de lignes ou de colonnes n'intervient. Au départ, les coefficients sont entiers, c.-à-d. que dans la colonne "dénominateurs" il n'y a que des 1.

DEPART

| dénominateurs | matrice 5 x 5 | | | | | second membre |
|---|---|---|---|---|---|---|
| 1 | 9 | 7 | 8 | 5 | 6 | 4 |
| 1 | 12 | 3 | 56 | 84 | 75 | 10 |
| 1 | 35 | 62 | 14 | 18 | 23 | 11 |
| 1 | 20 | 3 | 6 | 5 | 4 | 7 |
| 1 | 51 | 23 | 51 | 42 | 2 | 57 |

après le 1er pivot

| | | | | | |
|---|---|---|---|---|---|
| 1 | 9 | 7 | 8 | 5 | 6 | 4 |
| 9 | 0 | -57 | 408 | 696 | 603 | 42 |
| 9 | 0 | 313 | -154 | -13 | -3 | -41 |
| 9 | 0 | -113 | -106 | -55 | -84 | -17 |
| 9 | 0 | -150 | 51 | 123 | -288 | 309 |

On a donc, normalement, des dénominateurs égaux à 9 sur les lignes 2, 3, 4, 5. On s'attend a priori à obtenir, après le 2ème pivot, des dénominateurs égaux à $9 \times 57$ ; mais on a l'agréable surprise de voir que tous les numérateurs (par exemple $57 \times 154 - 313 \times 408$) sont multiples de 9, et le dénominateur 57 suffit pour les lignes 3, 4, 5 :

après le 2ème pivot

| | | | | | |
|---|---|---|---|---|---|
| 1 | 9 | 7 | 8 | 5 | 6 | 4 |
| 9 | 0 | -57 | 408 | 696 | 603 | 42 |
| -57 | 0 | 0 | -13214 | -24123 | -20952 | -1201 |
| -57 | 0 | 0 | 5794 | 9087 | 8103 | 635 |
| -57 | 0 | 0 | 6477 | 10821 | 11874 | -1257 |

On s'attend a priori à obtenir, après le 3ème pivot, des dénominateurs égaux à $57 \times 13214$; mais on a l'agréable surprise de voir que tous les numérateurs (par exemple $13214 \times 9087 - 24123 \times 5794$) sont multiples de 57, et le dénominateur 13214 suffit pour les lignes 4, 5 .

De même après le 4ème pivot, tous les numérateurs de la 5ème ligne seront multiples de 13214, et le dénominateur 345492 suffira :

après le 3ème pivot

| | | | | | | |
|---|---|---|---|---|---|---|
| 1 | 9 | 7 | 8 | 5 | 6 | 4 |
| 9 | 0 | -57 | 408 | 696 | 603 | 42 |
| -57 | 0 | 0 | -13214 | -24123 | -20952 | -1201 |
| -13214 | 0 | 0 | 0 | -345492 | -251278 | 25128 |
| -13214 | 0 | 0 | 0 | -232561 | 371876 | -427875 |

après le 4ème pivot

| | | | | | | |
|---|---|---|---|---|---|---|
| 1 | 9 | 7 | 8 | 5 | 6 | 4 |
| 9 | 0 | -57 | 408 | 696 | 603 | 42 |
| -57 | 0 | 0 | -13214 | -24123 | -20952 | -1201 |
| -13214 | 0 | 0 | 0 | -345492 | -251278 | 25128 |
| -345492 | 0 | 0 | 0 | 0 | 14145425 | -11629422 |

Nous allons maintenant expliquer l'origine de ces simplifications automatiques.

Nous supposons tout d'abord que les pivots qui se présentent successivement en position (k,k) sont tous non nuls.

Nous notons $a_{k,ij}$ le coefficient rationnel dans la matrice $C_k$ transformée de la matrice C après le $k^{ème}$ pivot.

Pour i et j > k , nous notons $C_{k,ij}$ la matrice extraite de la matrice $C_k$ sur les lignes 1,2,...,k,i et sur les colonnes 1,2,...,k,j.

La matrice $C_{k,ij}$ est une matrice surtriangulaire, son déterminant est égal au produit de ses éléments diagonaux. Mais le produit des k premiers éléments sur la diagonale est aussi le déterminant de la matrice $C_{k-1,kk}$, et on obtient l'égalité : $a_{k,ij}.\det(C_{k-1,kk}) = \det(C_{k,ij})$.

Maintenant nous remarquons que les matrices $C_{k-1,kk}$ et $C_{k,ij}$ sont obtenues à partir des matrices correspondantes extraites de la matrice de départ C au moyen d'une succession de transformation élémentaires qui ne modifient pas les déterminants.

Ainsi $d_k = \det(C_{k-1,kk})$ est un entier, déterminant de la matrice extraite de C sur les lignes 1,2,...,k et sur les colonnes 1,2,...,k , et pourra servir de dénominateur commun pour les lignes k+1,...,Nlin de la matrice $C_k$.

Nous pouvons donc organiser le traitement algorithmique de manière à donner $d_k$ pour dénominateur commun aux coefficients des lignes k+1,..., Nlin de la matrice $C_k$ comme suit:

**Module de traitement du pivot n° Npiv**

Description des variables utilisées

```
matrices d'entiers
A        :   matrice contenant les coefficients des systèmes linéaires
             successifs. Les seconds membres sont stockés dans les
             colonnes Ncol+1 à Ncol+Nsm. Dans la colonne 0 on met le
             dénominateur commun à la ligne concernée. Le nombre de
             colonnes effectivement utilisées dans la matrice A est
             Ncol+Nsm+1.

compteurs
Npiv     :   numéro du pivot
Nlin     :   nombre de lignes effectivement utilisées dans A c.-à-d.
             nombre d'équations du système linéaire
Ncol     :   nombre d'inconnues du système linéaire
Nsm      :   nombre de colonnes dans le second membre
I        :   indices pour les lignes dans les boucles
J        :   indices pour les colonnes dans les boucles

entiers
Piv      :   numérateur du pivot
Coef     :   coefficient en situation I, Npiv  (avant traitement)
Denm     :   dénominateur du pivot
```

```
MODULE  TRAITERPIVOT
        annulation des coefficients en dessous du pivot par
        manipulations élémentaires de lignes
                Variables
              entrées      : Nlin, Ncol, Nsm
              itératives   : A, Npiv
              locales      : I, J, Piv, Coef, Denm
Début
Piv ← A(Npiv,Npiv) ;
Pour  I  de  Npiv + 1  à  Nlin  faire
   Début
   Coef ← A(I,Npiv) ;  Denm  ← A(Npiv,0) ;
   A(I,Npiv) ← 0      ;  A(I,0) ← Piv ;
   Pour  J  de  Npiv + 1  à  Ncol+Nsm  faire
       A(I,J)  ←  (Piv × A(I,J) - Coef × A(Npiv,J))/Denm ;
   fin ;
Npiv ← Npiv+1
fin
```

L'exemple que nous avons donné "pour voir" a été traité par l'algorithme ci-dessus. On avait $d_1 = 9$ , $d_2 = -57$ , $d_3 = -13124$ , $d_4 = -345492$ , $d_5 = 14145425$ qui est le déterminant de la matrice de départ.

A chaque étape, on a effectué les simplifications automatiques (la division par Denm dans l'algorithme, qui donne toujours un résultat entier) sans se préoccuper des simplifications éventuellement plus poussées dues "au hasard" (par exemple toutes les fractions de la 4ème ligne après le 3ème pivot auraient pu être simplifiées par 2).

L'algorithme est correct (si Piv est non nul) parce qu'il augmente Npiv de 1 tout en conservant l'affirmation suivante vraie: "les lignes Npiv, Npiv+1, ..., Nlin sont écrites avec pour dénominateur commun $d_k$ , où k = Npiv-1". (en convenant que $d_0 = 1$). Ainsi les divisions par Denm donnent bien à chaque fois un résultat entier.

Nous remarquons également que l'algorithme ne fait aucune différence entre la partie "1er membre C" et la partie "2ème membre B" de la matrice A ; les arguments concernant l'existence de simplifications automatiques s'appliquent donc aussi bien au second membre.

Insistons sur le fait que le traitement des coefficients successifs "avec divisions automatiques" fournit en position i , j , après traitement du kème pivot, (où k < i et k < j ) le coefficient $\det(C_{k,ij})$ extrait de la matrice de départ (et il y a, en colonne "dénominateurs", en position i , 0 le déterminant $\det(C_{k-1,kk})$ ); c.-à-d. que tous les coefficients qui apparaissent sont des déterminants extraits de la matrice de départ.

On peut remarquer enfin que $\det(C_{k-1,kk})$ n'avait pas besoin en fait d'être stocké en colonne 0 puisqu'il est déjà en position k-1,k-1 .

### *De l'influence éventuelle des permutations de lignes et colonnes:*

Nous avons raisonné jusqu'à présent en supposant que les pivots successifs qui se présentent sur la diagonale sont non nuls. Nous allons voir maintenant ce qui se passe lorsqu'on est obligé de permuter des lignes ou/et des colonnes pour ramener un pivot non nul en position convenable.

Nous supposons Nlin, Ncol, Nsm donnés et nous notons PivLin(A,Npiv) la transformation subie par A et Npiv lorsqu'on exécute le module TRAITERPIVOT . Nous notons EchLin(A,$i_1$,$i_2$) la transformation sur la matrice A consistant à échanger les lignes $i_1$ et $i_2$ , et EchCol(A,$j_1$,$j_2$) la transformation sur la matrice A consistant à échanger les colonnes $j_1$ et $j_2$.

On constate sans difficulté que: si $i_1$ et $i_2$ sont > Np alors les transformations PivLin(A,Np) et EchLin(A, $i_1$, $i_2$ ) commutent entre elles. De même: si $j_1$ et $j_2$ sont > Np alors les transformations PivLin(A,Np) et EchCol (A, $j_1$, $j_2$ ) commutent entre elles.

Or, lorsqu'on déplace un pivot non nul pour l'amener dans la position Npiv , Npiv les transformations PivLin déjà effectuées l'ont été avec des numéros de pivot Np < Npiv, et le pivot déplacé est en situation $i_2$ , $j_2$ avec $i_2$ et $j_2 \geq$ Npiv .

Ainsi, du point de vue des transformations subies par A, toutes les transformations EchLin et EchCol auraient pu avoir été effectuées avant les transformations PivLin , de manière à trouver systématiquement des pivots non nuls en bonne position sur la diagonale. Ceci montre que l'algorithme de triangulation par la méthode du pivot, avec le module TRAITERPIVOT décrit précédemment, est correct (les divisions par Denm sont toujours des divisions exactes en entiers), même si on effectue des échanges de lignes et/ou des échanges de colonnes entre les transformations PivLin successives, à condition que les échanges portent sur des lignes (ou colonnes) de numéros strictement supérieurs à ceux des pivots déjà traités.

De plus, le coefficient en position i , j , après traitement du $k^{ème}$ pivot, (où $k < i$ et $k < j$ ), est égal, au signe près, à un déterminant extrait de la matrice de départ, à savoir le déterminant extrait sur les lignes (resp. colonnes) du départ qui, après les échanges de lignes (resp. colonnes), se retrouvent (juste après le traitement du $k^{ème}$ pivot) en positions 1, 2, ..., k, i (resp. 1, 2, ..., k, j ); le signe étant donné par la somme des parités des permutations subies. (même remarque pour la colonne "dénominateurs")

Signalons pour terminer que la colonne 0 où nous avons placé les dénominateurs s'avère en fait superflue puisque les dénominateurs successifs sont 1 puis les coefficients diagonaux de la matrice A transformée: dans TRAITERPIVOT, on peut remplacer l'affectation Denm ⟵ A(Npiv,0) par Denm ⟵ A(Npiv-1,Npiv-1) (sauf Denm ⟵ 1 pour Npiv = 1).

***Résolution du système triangulé; de nouvelles simplifications automatiques:***

Discutons tout d'abord, pour simplifier, le cas où la matrice C de départ est de rang Nlin.

Nous commençons par "oublier" la colonne 0 de notre matrice A: dans la mesure où nous nous préoccupons seulement de résoudre le système linéaire, peu importe si nous multiplions les coefficients d'une ligne par leur dénominateur commun.

Si nous avions à résoudre dans $\mathbb{Q}$ un système triangulaire "arbitraire" à coefficients entiers, de la forme:

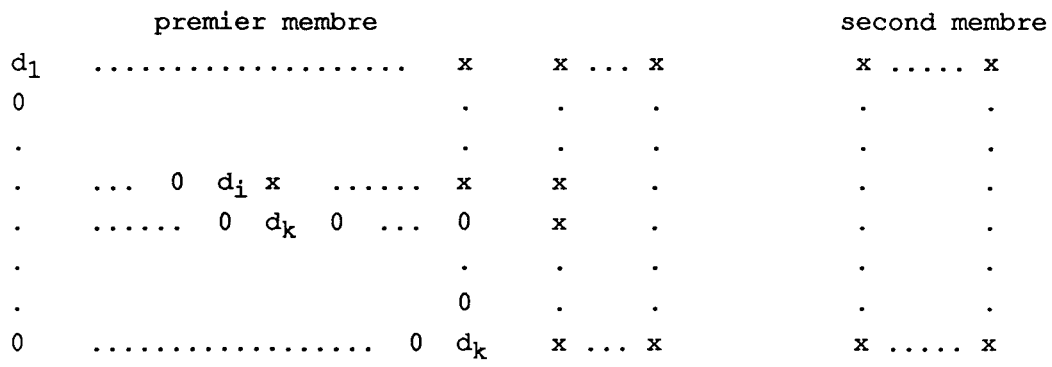
| | premier membre | | | | second membre |
|---|---|---|---|---|---|
| $d_1$ | x ..... | x | x ... x | | x ..... x |
| 0 | $d_2$ x ..... | x | x ... x | | x ..... x |
| . | 0 | . | . . | | . . |
| . | . | . | . . | | . . |
| . | . | x | . . | | . . |
| 0 | 0 ..... 0 | $d_k$ | x ... x | | x ..... x |

nous serions obligés de prévoir pour les solutions des dénominateurs égaux à $d_k$, $d_k d_{k-1}$, ... , $d_k d_{k-1} ... d_1$ sur les lignes k, k-1, ... ,1.

Mais le système triangulaire auquel nous avons affaire n'est pas n'importe quel système triangulaire. Nous pouvons profiter du fait que nous savons que les solutions de notre système triangulé peuvent être mises sous forme de quotient de déterminants entiers (des déterminants extraits de la matrice de départ C), le dénominateur étant justement égal, au signe près, au coefficient $d_k$.

Nous résolvons donc le système par substitution, en partant de la dernière ligne et en remontant, (ce qui revient à faire subir à A quelques manipulations élémentaires de lignes pour remplacer la partie triangulaire par une partie diagonale), en sachant qu'en fin de compte on pourra se ramener à des $d_k$ sur la diagonale avec une matrice à coefficients tous entiers.

Ainsi, si, avant le traitement de la $i^{ème}$ ligne, notre système a la forme:




nous aurions a priori, après traitement, un $d_i\, d_k$ sur la diagonale pour pouvoir conserver tous les coefficients de la $i^{ème}$ ligne entiers, mais nous savons que $d_k$ suffira, et donc tous les coefficients calculés en ligne i seront sûrement multiples de $d_i$.

Ceci nous conduit donc au module suivant pour résoudre le système triangulé:

## Module de résolution du système triangulé

Description des variables utilisées

```
matrice d'entiers
A   : matrice contenant les coefficients des systèmes linéaires successifs.

compteurs
Ncol : nombre d'inconnues du système linéaire
Nsm  : nombre de colonnes dans le second membre
I,I1 : indices pour les lignes dans les boucles
J    : indices pour les colonnes dans les boucles
Rang : rang de la matrice du premier membre

entiers
Di   : Ième élément sur la diagaonale (avant traitement)
Det  : déterminant de la matrice carrée extraite de rang  Rang
Asom : variable pour calculer une somme itérée
```

```
MODULE  TRIANGSOL
        la matrice est ramenée à une forme complètement résolue ,
        c.-à-d. diagonale pour les inconnues principales. La diagonale
        est remplie de Det et tous les coefficient sont entiers.
              Variables
              entrées      : Ncol, Nsm, Rang, Det
              itératives   : A
              locales      : I, I1, J, Di, Asom
Début
Pour  I  de  Rang-1  à  1  en descendant  faire
   Début
   Di ← A(I,I) ; A(I,I) ← Det ;
   Pour  J  de  Rang+1  à  Ncol+Nsm  faire
       Début
       Asom  ←  A(I,J) × Det  ;
       Pour  I1  de  I+1  à  Rang  faire
                                  Asom ← Asom-A(I1,J) × A(I,I1);
       A(I,J) ← Asom/Di ;
       fin;
   Pour  I1 de  I+1  à  Rang  faire  A(I,I1) ← 0 ;
   fin
fin
```

On pourra remarquer que si le système est de rang strictement inférieur à Nlin, le module TRIANGSOL le transforme également en un système équivalent dont les coefficients restent entiers, puisque l'on ne tient pas compte des lignes Rang+1 ... Nlin.

On notera enfin que tous les coefficients calculés pendant l'exécution de TRIANGSOL sont, encore une fois, égaux, au signe près, à des déterminants extraits de la matrice de départ, puisque la solution théorique donne des quotients de 2 déterminants dont le $2^{ème}$ est égal, au signe près, au coefficient sur la diagonale, qui sert justement de dénominateur.

## La méthode du pivot améliorée (à la Bareiss)

La méthode du pivot décrite ci-dessus dans $\mathbb{Z}$ est "améliorée" en ce sens que même si nous n'avions pas de $\mathcal{P}$-calcul de la réduite d'une fraction dans $\mathbb{Q}$, nous pourrions la pratiquer sans explosion de la taille des coefficients, alors que la méthode du pivot "ordinaire", directement dans $\mathbb{Q}$, et sans réduction systématique des fractions obtenues, conduirait, elle, à une explosion de la taille des coefficients.

Ici, nous n'avons pas pratiqué une réduction systématique de toutes les fractions obtenues, mais nous avons pratiqué toutes les réductions "automatiques" possibles (ce qui réclame seulement que la division exacte soit un $\mathcal{P}$-calcul dans $\mathbb{Z}$ ). Et cela a suffi pour éviter l'explosion des coefficients. Or il est en général bien plus facile d'avoir une division exacte en temps polynomial dans un anneau intègre, plutôt que le calcul en temps polynomial d'une forme réduite (presque) canonique pour les fractions dans le corps des fractions: le passage de l'anneau intègre à son corps des fractions ne conserve pas a priori le caractère $\mathcal{P}$-dénombrable ou $\mathcal{P}$-réductible.

Nous donnons maintenant une description précise de la méthode du pivot améliorée pour la résolution d'un système d'équations linéaires à coefficients dans K et inconnues dans L, lorsque K est un $\mathcal{P}$-anneau intègre, $\mathcal{P}$-divisible et $\mathcal{P}$-réductible.

### *Description détaillée de la méthode*

Nous explicitons le caractère $\mathcal{P}$-réductible de K par la donnée d'une $\mathcal{P}$-opération $Red : K \longrightarrow K$ , avec $x =_K Red(x)$ pour $x \in K$.

Nous procédons alors comme suit:

- D'abord on triangule au moyen du module TRAITERPIVOT, avec les précisions suivantes:

* Avant le traitement du $(k+1)^{ème}$ pivot, on cherche un coefficient non nul dans la partie utile restante de la matrice (c'est un $\mathcal{P}$-calcul parce que le test d'égalité à 0 est $\mathcal{P}$ . Puis on peut opérer à loisir une permutation quelconque sur les lignes de numéro > k , et pareil pour les colonnes, notamment dans le but de ramener en position k+1 , k+1 un coefficient non nul qui va servir de pivot. (il est recommandé de garder en mémoire la permutation subie par les lignes (resp. colonnes) depuis le départ.

* Dans TRAITERPIVOT lui-même, on remplace l'affectation:

```
A(I,J) ⟵ ( Piv ×A(I,J) − Coef ×A(Npiv,J) )/Denm ,
```

par l'affectation:

```
A(I,J) ⟵ Red( ( Piv ×A(I,J) − Coef ×A(Npiv,J) )/Denm )
```

- Après avoir triangulé jusqu'à épuisement des pivots non nuls, on exécute TRIANGSOL avec les précisions suivantes: on remplace l'affectation:

```
                Asom ⟵ Asom - A(I1,J) ×A(I,I1)  ,
```

par l'affectation:

```
                Asom ⟵ Red( Asom - A(I1,J) ×A(I,I1) ) ;
```

et l'affectation:

```
                A(I,J) ⟵ Asom/Di  ,
```

par l'affectation:

```
                A(I,J) ⟵ Red(Asom/Di)
```

### *Le Théorème concernant la méthode de Bareiss*

**Théorème C.d1 :**

Lorsqu'on traite un système d'équations linéaires par la méthode du pivot améliorée, tous les coefficients qui sont calculés sont égaux, au signe près, à des déterminants extraits de la matrice de départ.

Supposons que K soit un $\mathcal{P}$-anneau intègre, $\mathcal{P}$-divisible et $\mathcal{P}$-réductible, et soit L son corps de fractions . Alors la résolution dans le $\mathcal{P}$-corps L d'un système linéaire à coefficients dans K par la méthode du pivot améliorée est un $\mathcal{P}$-calcul si et seulement si la fonction "déterminant" est **RESP** dans K, c.-à-d. ssi les déterminants sont polynomialement majorés en taille.

En particulier le calcul du déterminant (et de l'inverse) d'une matrice par la méthode du pivot améliorée est un $\mathcal{P}$-calcul si et seulement si les déterminants sont polynomialement majorés en taille.

**Remarques:**

1) lorsque K est $\mathcal{P}$-réductible (en particulier s'il est $\mathcal{P}$-dénombrable) nous pouvons donc rajouter une nouvelle propriété équivalente à celles énoncées à la proposition C.c2 : les déterminants sont polynomialement majorés en taille.

2) on se reportera au § B.g pour obtenir un bon stock d'anneaux où la fonction déterminant est **RESP**

3) pour un système linéaire à coefficients dans L, on est facilement ramené au cas du théorème C.c1 par une réduction préalable de toutes les fractions au même dénominateur: la multiplication dans K est en effet c-$\mathcal{P}$-c, puisqu'elle est **RESP** (cf.: déterminant d'une matrice diagonale)

4) si L est lui-même $\mathcal{P}$-réductible, on déduit facilement du théorème C.c1 que la méthode du pivot "ordinaire", avec réduction systématique de toute fraction après calcul, est un $\mathcal{P}$-calcul si et seulement si la fonction déterminant est **RESP** (dans K , donc aussi dans L ici). Cependant, il est en général beaucoup plus économique de travailler avec K : par exemple si $K = \mathbb{Z}$ ou $\mathbb{Z}[X]$ , la division exacte dans K est nettement plus rapide que le calcul du pgcd dans K ; or la réduction d'une fraction nécessite un tel calcul de pgcd.

*preuve>* Tous les coefficients calculés, sont polynomialement majorés en taille à partir de la taille des données: en effet ils sont égaux, au signe près, et sous forme réduite, à des déterminants extraits de la matrice de départ.

Le nombre de coefficients calculés est également polynomialement majoré à partir de la taille des données.

Le seul problème est donc de vérifier que tous les éléments de K calculés au cours de TRAITERPIVOT et TRIANGSOL restent polynomialement majorés en taille: <u>après</u> TRAITERPIVOT ou TRIANGSOL, tout va bien. Que se passe-t-il <u>pendant</u> ? Pendant TRAITERPIVOT on effectue 2 multiplications, une soustraction et une division exacte à partir de coefficients sous forme réduite, puis on réduit le résultat obtenu. Ceci reste polynomialement majoré puisque K est un $\mathcal{P}$-anneau. Pendant TRIANGSOL, c'est le même raisonnement, mais cette fois-ci, il y a une addition itérée (au maximum Nlin fois) portant sur des produits de 2 éléments réduits égaux à des déterminants de matrices extraites de la matrice de départ. Or nous avons déjà établi le résultat adéquat pour l'addition itérée à la proposition C.b2 . ❑

# Notes

## n.1 §A.a : propriétés de la classe de constructions $\mathfrak{C}$, quelques précisions

**NB** : il est préférable de lire cette note après avoir lu le § A b)

Nous considérons 3 symboles spéciaux servant à écrire des listes : [ , ] , ; encore appelés **scl** (symboles constructeurs de listes).

Si A ne contient aucun **scl**, l'ensemble **Lst** (A*), des listes d'éléments de A* , peut être réalisé comme une partie du langage A°* (où A° est la réunion de A et des **scl**).
Si $X_1$ , $X_2$ ,..., $X_n$ sont des parties de A*, l'ensemble $X_1 \times X_2 \times ... \times X_n$ peut être réalisé comme une partie de **Lst**(A*) (listes convenables de n éléments).

Nous supposons qu'aucun alphabet désigné par A , B , C ne contient de **scl**.

Si X est un ensemble, **Lsp**(X) , ensemble des "lispes" d'éléments de X , est défini récursivement par :

**Lst'**(X) := **Lst** (X) ∪ X

**Lsp**(X) := **Lst'**(X) ∪ **Lst'**(**Lst'**(X)) ∪ **Lst'**(**Lst'**(**Lst'**(X))) ...

L'ensemble **Lsp**(A*) peut être réalisé comme une partie de A°* . Cette partie est **DT1**-détachable. Un mot de **Lsp**(A*) peut être lu sans ambiguïté comme liste de mots de **Lsp**(A*) , et sans ambiguïté comme lispe d'éléments de A* . En particulier, les opérations de **Lsp**(A*) vers **Lsp**(A*) : "1er mot de la liste" et "liste privée du 1er mot" sont bien définies, ce sont même des $\mathfrak{C}$-opérations.

Pour conserver cette propriété de lisibilité, nous allons restreindre notre "univers" . Nous posons d'abord la :

**Définition n.1 :** Un $\mathfrak{C}$-préensemble X est donné lorsque :
- on considère un alphabet A ne contenant pas de **scl**
- on considère une $\mathfrak{C}$-opération $P_X$ de **Lsp**(A*) vers {**oui**,**non**}
  Les éléments de X sont les mots m de **Lsp**(A*) pour lesquels
  $P_X(m)$ = oui . Nous dirons que X est un $\mathfrak{C}$-préensemble construit sur A[1] .

**Convention :** dans la suite, lorsque nous parlons d'une $\mathfrak{C}$-partie d'un langage L* , nous supposons toujours (implicitement ou explicitement) qu'elle est présentée comme un $\mathfrak{C}$-préensemble construit sur un alphabet A (donc en particulier L = A° pour un alphabet A, et les mots de X sont tous dans **Lsp**(A*) )

Nous sommes maintenant en mesure de préciser l'énoncé des propriétés de stabilité de la classe $\mathfrak{C}$ , et d'en tirer quelques conséquences (admises implicitement dans le texte) .

Nous commençons par préciser une condition concernant la mesure $\| \ \|_X$ pour la grandeur des éléments d'un $\mathfrak{C}$-préensemble X construit sur A . Notons $\| \ \|_{A^*}$ la mesure naturelle (longueur du mot). Nous supposons que nous avons toujours:

- l'identité $I : x \to x$ de $(X, \| \ \|_{A^*})$ vers $(X, \| \ \|_X)$ est une $\mathfrak{C}$-opération

La stabilité pour la composition des opérations et celle pour la définition par cas ne posent pas problème.

Nous donnons par contre des précisions en ce qui concerne la stabilité pour **Lst** . Nous devons tout d'abord énoncer la propriété de stabilité suivante:

---

[1] Nous disons $\mathfrak{C}$-préensemble, plutôt que $\mathfrak{C}$-ensemble, parce que ne sont définies, ni l'égalité de X , ni la mesure de la grandeur de ses éléments.

– si X est un $\mathfrak{C}$-préensemble construit sur A , alors **Lst**(X) est un $\mathfrak{C}$-préensemble
construit sur A .

Par ailleurs, lorsqu'on a défini une mesure $\| \ \|_X$ pour la grandeur des éléments de X , il faut considérer que :

– la mesure de l'élément $[x_1,x_2,...,x_n]$ de **Lst**(X) est définie par :
$$\| [x_1,x_2,...,x_n] \|_{\mathbf{Lst}(X)} := \| x_1 \|_X + \| x_2 \|_X + ... + \| x_n \|_X$$

Les propriétés de stabilité suivantes sont alors immédiates:

**$\mathfrak{C}$-parties**: si $Z_1$ et $Z_2$ sont des $\mathfrak{C}$-parties de X , il en est de même pour $Z_1 \cup Z_2$ , $Z_1 \cap Z_2$ et $X - Z_1$ .

**$\mathfrak{C}$-produits**: si X est un $\mathfrak{C}$-préensemble construit sur A , et Y est un $\mathfrak{C}$-préensemble construit sur B , alors $X \times Y$ est un $\mathfrak{C}$-préensemble construit sur $A \cup B$ .

**opérations élémentaires portant sur les listes**: les opérations élémentaires suivantes sont des $\mathfrak{C}$-opérations:

- les applications **Lst**($\mathbb{N}$) $\longrightarrow \mathbb{N}^k$ : k premiers éléments de la liste, éventuellement complétée par des 0
- les injections canoniques $\mathbb{N}^k \longrightarrow$ **Lst**($\mathbb{N}$)
- l'application **Lst**(N) × **Lst**($\mathbb{N}$) $\longrightarrow$ **Lst**($\mathbb{N}$) : concaténation de 2 listes
- l'application **Lst**($\mathbb{N}$) $\longrightarrow$ **Lst**($\mathbb{N}$) : retrait du premier élément d'une liste

**Remarque :** si ce n'est essentiellement pour des raisons de commodité, on pourrait d'ailleurs admettre une bonne fois pour toutes un seul alphabet A , ce qui ferait de tous les $\mathfrak{C}$-préensembles des $\mathfrak{C}$-parties de "l'univers" **Lsp**(A*).

## n.2 §A.b : $\mathfrak{C}$-équivalences entre les différents A*, et avec différentes présentations de $\mathbb{N}$

Les résultats affirmés dans le § "quelques $\mathfrak{C}$-équivalences" résultent essentiellement de l'existence d'algorithmes rapides (**DTNLG**) pour la multiplication des entiers en binaire. On pourra par exemple consulter [**DACA**] pour la preuve du caractère **DTNLG** d'un changement de base de numération.

Lorsque nous considérons un entier n écrit en base b , la longueur du mot qui le représente est à peu près proportionnelle à **lg**(n) (sa longueur lorsqu'il est écrit en base 2) dans le rapport log(2)/log(b). Le changement de base de numération est donc $\mathbf{DTNLG}_1$ (et, avec une modification convenable de la mesure de la taille pour les entiers en base b , il est $\mathbf{DTNLG}_0$ )

Par ailleurs, on constate facilement que "$\mathbb{N}$ en base b" est **DT0**-équivalent à A* , où A est un alphabet à b lettres. Considérons ces lettres comme représentant les chiffres 0, 1, ..., b−1 . Associons au mot m de longueur k : $a_1a_2...a_k$ l'entier u(m) qui s'écrit $1a_1a_2...a_k$ en base b . Soit alors num(m) le numéro de m lorsqu'on numérote les u(m) en ordre croissant.

On obtient par un petit calcul: $num(m) = u(m) + ( 1 + b + ... + b^{k-1} ) - b^k$ , et il est clair que le calcul de num(m) à partir de m , et vice versa, sont dans **DT0** .

### n.3 §A.b : produit de 2 $\mathfrak{C}$-ensemble-discrets comme produit au sens des catégories

Tout d'abord, les projections canoniques $X \times Y \longrightarrow X$ et $X \times Y \longrightarrow Y$ sont dans **DT0**. D'autre part, si $f : Z \longrightarrow X$ et $g : Z \longrightarrow Y$ sont dans la classe $\mathfrak{C}$, vue la stabilité de $\mathfrak{C}$ pour les listes, nous savons que $f \times g : Z \times Z \longrightarrow X \times Y$ est dans $\mathfrak{C}$.
Vue la stabilité de $\mathfrak{C}$ pour la composition, nous obtenons donc : si la classe $\mathfrak{C}$ contient les applications diagonales $Z \longrightarrow Z \times Z$ le produit de 2 $\mathfrak{C}$-ensemble-discrets est bien le produit au sens de la catégorie des $\mathfrak{C}$-ensemble-discrets Ce sera le cas lorsque $\mathfrak{C}$ contient **DT1**, mais pas pour la classe $\mathfrak{P}_0$ par exemple.

### n.4 §A.b : structures algébriques avec des axiomes "purement universels"

Nous considérons qu'une structure algébrique (sur un ensemble discret) est donnée par un certain nombre de relations (unaires, binaires...) décidables, par un certain nombre de constantes, et par un certain nombre de lois de compositions non nécessairement "partout définies": mais le domaine de définition doit être décidé par l'une des relations faisant partie de la structure.

Les axiomes liant ces éléments de structure sont dits purement universels, s'ils affirment que certaines égalités (écrites en utilisant exclusivement les constantes et lois de composition de la structure) sont vérifiées en tout point d'une partie décidée par l'une des relations de la structure.

Dans ce cas on a immédiatement une $\mathfrak{C}$-version de la structure considérée, en demandant que les relations de la structure soient $\mathfrak{C}$-décidables et que les lois de composition soient des $\mathfrak{C}$-fonctions.

Par exemple on a une formulation purement universelle de la notion de corps en utilisant les constantes 0 et 1, la partie $K-\{0\}$, les lois $+$ , $\times$ , $x \longrightarrow -x$ , et $x \longrightarrow 1/x$ , et des axiomes évidents.

L'une des principales difficultés pour un traitement constructif de l'algèbre discrète vient de l'impossibilité de formuler certaines notions classiques sous forme purement universelle (notamment la notion de noethériannité, qui n'est même pas formulable "au premier ordre"[1]), d'où la difficulté de donner une traduction "calculatoire" de certaines définitions classiques.

Notons que dans la formulation constructive de la notion d'anneau factoriel A , la décomposition en facteurs premiers est une "loi de composition" avec pour ensemble d'arrivée **Lst**(A) . Dans les premières versions du Moderne Algebra de Van der Waerden, un corps est appelé "factoriel" si l'anneau des polynômes à une indéterminée est factoriel au sens constructif.

### n.5 §A.b : les structures algébriques récursivement présentées dans [F-S] (article de Frölich et Shepherdson sur les procédures effectives en théorie des corps)

L'article en question est sans doute le premier texte systématique sur l'étude des extensions de corps du point de vue de la récursivité. Dans [F-S] une structure algébrique "effective" est donnée par un ensemble énuméré, avec une relation d'égalité récursivement

---

[1] C'est à dire au moyen d'une formule du premier ordre portant sur des variables dans la structure algébrique considérée.
Noter cependant que la notion de noetheriannité a néanmoins reçu un traitement constructif entièrement satisfaisant dans le cas des anneaux commutatifs discrets. cf. [CAL]

décidable (dans l'énumération considérée). Les lois de composition définissant la structure sont vues comme des relations (binaires, ternaires ... $z = x + y$ par exemple), de sorte qu'une loi de composition est dite "effective" si son graphe est récursivement décidable.

Dans la mesure où il s'agit de lois de composition partout définies, ou définies sur des parties récursivement décidables, on obtient bien la même notion que celle de structure algébrique récursivement présentée (c.-à-d. encore de **Rec**-structure) que nous avons définie.

Dans cet article sont définis 2 corps récursifs K et L avec les propriétés suivantes:

K[X] est récursivement factoriel, L[X] ne l'est pas

K et L sont isomorphes

Evidemment K et L ne sont pas récursivement isomorphes, et la preuve de l'existence d'un isomorphisme entre K et L est non constructive.

## n.6 §A.b : les structures algébriques de type fini sont "naturellement primitives récursives"

**NB :** il est préférable de lire cette note après avoir lu le § B a)

Une structure algébrique (X,Y,Z) est de type fini si tout objet peut être obtenu à partir des constantes et d'un nombre fini d'éléments $a_1,a_2,...,a_n$ en utilisant de manière répétée les lois de composition.

Considérons alors la partie T de **Calc**(X,Y,Z) (cf. §B.a) où seules interviennent les constantes et les éléments $a_1,a_2,...,a_n$ : on obtient, en prenant la relation d'égalité convenable, une présentation de $X \cup Y \cup Z$, pour laquelle les lois de composition sont $\mathbb{P}\mathbf{r}$, et même **DT0**. Mais il reste le problème de l'égalité dans T, qui peut n'être pas $\mathbb{P}\mathbf{r}$, ni même **Rec**. Et pareil pour les autres relations faisant partie de la structure.

Si (X,Y,Z) était au départ une $\mathbb{P}\mathbf{r}$-structure, alors il est clair que la relation d'égalité dans T (et les autres relations faisant partie de la structure) sont $\mathbb{P}\mathbf{r}$, et que la bijection naturelle de **Calc**(X,Y,Z) vers $X \cup Y \cup Z$ est une $\mathbb{P}\mathbf{r}$-fonction.

Cette présentation naturelle est $\mathbb{P}\mathbf{r}$-équivalente à la présentation de départ si et seulement si on peut retrouver, par un calcul $\mathbb{P}\mathbf{r}$, une manière d'écrire n'importe quel objet x à partir des constantes et des éléments en nombre fini considérés: c.-à-d. si la structure, avec sa présentation, est "de type fini de manière $\mathbb{P}\mathbf{r}$".

On remarquera qu'il semble difficile d'obtenir une structure algébrique qui serait "intrinsèquement $\mathbb{P}\mathbf{r}$" au sens suivant : 2 $\mathbb{P}\mathbf{r}$-présentations de la structure sont nécessairement $\mathbb{P}\mathbf{r}$-isomorphes. (sauf dans le cas des structures finies). Par contre les structures algébriques récursivement présentées de type fini sont "intrinsèquement récursives" (au même sens que ci-dessus).

## n.7 §A.c : rapports entre $\mathcal{P}$-dénombrabilité et $\mathcal{P} = \mathcal{NP}$ ?

On l'implication :

si $\mathcal{P} = \mathcal{NP}$, alors tout $\mathcal{P}$-ensemble est $\mathcal{P}$-dénombrable.

Soit en effet X un $\mathcal{P}$-ensemble et (f,r) une $\mathcal{P}$-énumération de X.

Pour $n, p \in \mathbb{N}$ définissons :

$$g(n,p) := \begin{cases} 0 & \text{si } f(n) = u \text{ ou } f(p) = u \text{ ou } f(n) \neq f(p) \\ n - p & \text{sinon} \end{cases}$$

Pour n fixé tel que $f(n) \neq u$, g(n,p) est maximum pour la 1ère valeur de p vérifiant $f(p) = f(n)$, et la connaissance de ce maximum permet de retrouver p.

Si donc on définit $h(n,p) := \sup\{g(n,q)\,,\, q \leq p\}$ on a le dénombrement (f,r') de X donné par : $r'(x) := f(\, n - h(n,n)\,)$, où $n := r(x)$.

Or $\mathcal{P} = \mathcal{N}\mathcal{P}$ équivaut au fait que la fonctionnelle **Sup** qui fait passer de g à h transforme toute $\mathcal{P}$-fonction en une $\mathcal{P}$-fonction.

Notons enfin qu'il est probable qu'on ait l'implication réciproque : si tout $\mathcal{P}$-ensemble est $\mathcal{P}$-dénombrable, alors $\mathcal{P} = \mathcal{N}\mathcal{P}$ .

## n.8 §B.a : $\mathbb{Q}$ est $\mathcal{P}$-isomorphe à sa présentation en fraction continue standard

Il reste à montrer que le passage du rationnel q = a/b à la liste des entiers constituant son **dfc** (développement en fraction continue) standard est une $\mathcal{P}$-fonction. Or le calcul du **dfc** se fait par divisions successives, et n'est rien d'autre que l'algorithme d'Euclide pour le pgcd du numérateur et du dénominateur. Si c est le sup de la valeur absolue du numérateur et du dénominateur, les entiers du **dfc** sont tous inférieurs à c (pour le premier: en valeur absolue) et leur nombre est majoré par 2.**lg**(c) : le pire cas est obtenu avec les **dfc** dont tous les termes sont égaux à 1 (cf. par ex D. Knuth **[ACP 4]** p 343)

# BIBLIOGRAPHIE


[Ait] On the evaluation of determinants, the formation of their adjugates, and the practical solution of simultaneous linear equations.
**Aitken A. C.**
Proc. Edinburgh Math. Soc. ser 2 III , 207-219 , (1932)

[ALFA] Algèbre (cours de mathématiques, tome 1)
**J. Lelong-Ferrand, J.-M. Arnaudiès**
(Dunod; 1974 )

[ACP4] The Art of Computer Programming (vol 2).
Chap 4: Arithmetic Seminumerical Algorithms. 2ème édition
**D. E. Knuth**
(Addison-Wesley; 1973)

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

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

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

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

[CASAC] Computer Algebra: Symbolic and algebraic computation
**Edited by Buchberger, Collins and Loos**
(Springer-Verlag; 1982)

[CFA] Constructive Functional Analysis
**D. Bridges**
(Pitman, London; 1979)

[DACA] The Design and Analysis of Computer Algorithms
**A.V. Aho, J. E. Hopcroft, J. D. Ullman**
(Addison-Wesley; 1974)

[F-S] Effective procedures in Field Theory
**A. Frölich, J. C. Shepherdson**
Philos. Trans. Roy. Soc. London (Ser A) 284 (1955) 407-432

[FCM] Foundations of Constructive Mathematics
**M. Beeson**
(Springer -Verlag; 1985)

[Gan] Théorie des Matrices
**Gantmacher F. R.**
DUNOD (1966) (traduit du russe)

[ITALC] Introduction to the Theory of Automata, Languages and Computability
**J. E. Hopcroft, J. D. Ullman**
(Addison-Wesley; 1979)

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

# INDEX


| | chapitre et page | explication rapide éventuelle |
|---|---|---|
| **Calc**(X) | : B 17 | : ensemble des écritures décrivant des calculs à effectuer dans la structure algébrique X |
| **Col**(K) | : C 44 | : ensemble des vecteurs colonnes à coefficients dans K |
| complètement $\mathfrak{C}$-calculable | : B 17 | : une structure algébrique, avec un nombre fini de lois de composition, donnée avec une $\mathfrak{P}$-présentation, est dite complètement $\mathfrak{P}$-calculable si l'évaluation des formules est un $\mathfrak{P}$-calcul |
| c-$\mathfrak{P}$-c | : B 18 | : complètement $\mathfrak{P}$-calculable |
| $\mathfrak{C}$-dénombrement | : A 13 | |
| $\mathfrak{C}$-détachable | : A 9 | |
| $\mathfrak{C}$-divisible ( $\mathfrak{C}$-monoïde) | : A 15 | |
| $\mathfrak{C}$-équivalence | : A 8 | |
| $\mathfrak{C}$-ensemble-discret | : A 8 | |
| $\mathfrak{C}$-fonction | : A 8 | |
| $\mathfrak{C}$-numérotation | : A 13 | |
| $\mathfrak{C}$-présentation | : A 10 | : d'un ensemble, d'une structure algébrique |
| $\mathfrak{C}$-quotient | : A 9 | |
| $\mathfrak{C}$-sous-structure | : A 11 | |
| $\mathfrak{C}$-structure-quotient | : A 11 | |
| $\mathfrak{C}$-surjective | : A 9 | |
| **Dep-c$\mathfrak{P}$c** | : C 44 | : un anneau commutatif K est **Dep-c$\mathfrak{P}$c** lorsque les relations de dépendance linéaires sont $\mathfrak{P}$-calculables (en un sens convenable) |
| **Det-c$\mathfrak{P}$c** | : C 40 | : un anneau commutatif K est **Det-c$\mathfrak{P}$c** lorsque la fonction "déterminant" est $\mathfrak{P}$-calculable |
| dénombrable, dénombrement | : A 13 | |
| discret (ensemble) | : intro 3 | |
| **DSP1** | : A 5 | : **DSPACE**(O(n)) |
| **DTIME$_1$**(f(n)) | : A 5 | : **RES1** $\cap$ **DTIME**(f(n)) |
| **DTNLG** | : A 5 | : $\cup_b$ **DTIME**(O(n.lg$^b$(n))) |
| **DTNLG$_0$** | : A 5 | : **DTNLG** $\cap$ **RES0** |
| **DT1** | : A 5 | : **LINTIME**, **DTIME**(O(n)) |
| énumérable, énumération | : intro 3 | |
| **Flin**(K) | : C 35 | : voir **Mat**(K) |
| **Fsv**(K) | : C 35 | |
| **Inv-c$\mathfrak{P}$c** | : C 39 | : un anneau K est **Inv-c$\mathfrak{P}$c** lorsque l'inversion des matrices carrées à coefficients dans K est un $\mathfrak{P}$-calcul |

| | | | |
|---|---|---|---|
| **Inv-1-c$\mathfrak{P}$c** | : C 40 | | |
| lispe | : note 1 | : | liste de listes de ... listes |
| **Lin**(K) | : C 35 | | |
| **Lst**(X) | : A 6 | : | ensemble des listes d'éléments de X |
| **Lsp**(X) | : note 1 | : | ensemble des lispes d'éléments de X |
| **Mat**(K) | : C 34 | : | ensemble des matrices n × h à coefficients dans K , (réunion disjointe sur les n × h ): la réunion emboitée est notée **Flin**(K) |
| **Mat-c$\mathfrak{P}$c** | : C 37 | : | un anneau K est **Mat-c$\mathfrak{P}$c** (pour une $\mathfrak{P}$-présentation donnée) si le produit des matrices est c-$\mathfrak{P}$-c dans Mat(K) |
| mesure (de la taille de ) | : A 6 | : | voir fin A a 2 les conditions supposées vérifiées par une telle "mesure" |
| méthode du pivot améliorée | : C 55 | | |
| $\mathbf{M_n}$(K) | : B 20 | : | matrices carrées n x n à coeffs dans K |
| $\mathbb{N}$ | : A 11 | : | entiers naturels présentés en binaire |
| $\mathbb{N}_1$ | : A 11 | : | entiers naturels présentés en unaire |
| $\mathbb{N}_b$ | : A 12 | : | entiers naturels présentés en bibase b |
| naturellement c-$\mathfrak{P}$-c | : B 19 | : | (structure algébrique) |
| naturellement de type $\mathfrak{C}$ | : A 10 | : | (structure algébrique) |
| naturellement primitive récursive | : A 10 | : | (structure algébrique) |
| numérotable, numérotation | : A 13 | | |
| $\mathfrak{P}$ | : A 5 | : | $\cup_b$ **DTIME**($O(n^b)$) |
| $\mathfrak{P}$-calculable | : | : | calculable en temps polynômial |
| $\mathfrak{P}$-dénombrable | : A 13 | | |
| $\mathfrak{P}$-divisible (anneau intègre) | : A 15 | | |
| $\mathfrak{P}$-ensemble, $\mathfrak{P}$-... | : | : | voir $\mathfrak{C}$-ensemble, $\mathfrak{C}$-... |
| $\mathfrak{P}$-libre ( $\mathfrak{P}$-espace) | : B 20 | | |
| $\mathfrak{P}$-réductible ( $\mathfrak{P}$-ensemble) | : A 14 | | |
| $\mathfrak{P}$-résoluble | : C 46 | : | (les systèmes linéaires dans L sont ...) |
| $\mathfrak{P}_0$ | : A 5 | : | **RES0** $\cap$ $\mathfrak{P}$ |
| $\mathfrak{P}_0$-anneau | : B 30 | | |
| $\mathfrak{P}_1$ | : A 5 | : | **RES1** $\cap$ $\mathfrak{P}$ |
| $\mathbb{P}$**r** | : A 5 | : | classe des opérations primitives récursives |
| présentation (d'un ensemble) | : A 7 | | |
| présentation c-$\mathfrak{P}$-c naturelle | : B 19 | : | d'une structure algébrique |
| présentation creuse | : B 28 | : | d'un anneau de polynôme |
| présentation en magma | : B 26 | : | de X sur un système générateur de X, d'une structure algébrique X |
| **PSPACE** | : A 5 | : | $\cup_b$ **DSPACE**($O(n^b)$) |
| $\mathbb{Q}$ | : A 16 | : | nombres rationnels présentés en binaire |

| | | | |
|---|---|---|---|
| **Rec** | : A 5 | : | classe des opérations récursives |
| **RES0** | : A 5 | : | $\cup_c$ **SPACERES**(n+c) |
| **RES1** | : A 5 | : | **SPACERES**(O(n)) |
| **RESP** | : A 5 | : | $\cup_b$ **SPACERES**($O(n^b)$) |
| séparation | : intro 3 | | |
| **SPACERES**(f(n)) | : A 5 | : | la mesure de la taille du résultat est majoré par f(n), où n est la mesure de la taille de l'entrée |
| **Sv**(L) | : C 35 | | |
| **Syslin**(L) | : C 45 | : | ensemble des systèmes d'équations linéaires à coefficients dans L |
| systéme linéaire résolu | : C 46 | | |
| **Trimat**(K) | : C 35 | : | ensemble des matrices carrées n x n supérieures, à coeffs dans K , avec des 1 sur la diagonale (réunion disjointe sur les n ) |
| $\mathbb{Z}$ | : A 15 | : | entiers relatifs présentés en binaire |
| $\mathbb{Z}_1$ | : A 15 | : | entiers relatifs présentés en unaire |
| $\mathbb{Z}[X]$ | : B 21 | : | polynômes en présentation "dense" |
| $\mathbb{Z}[X]_c$ | : B 21 | : | polynômes en présentation "creuse" |

# SOUS-RESULTANTS , SUITE DE STURM , SPECIALISATION

Introduction ..... 2

1) Matrice de Sylvester , polynômes sous-résultants et suite des restes: notations, premiers résultats 4

2) Un peu d'algèbre linéaire: conséquences pour la complexité 7

3) Sous-pgcd et vrais restes : des formules explicites 9

Le cas générique ..... 10
Le cas défectueux ..... 11
Discussion sur le temps de calcul de la suite des restes ..... 12

4) Spécialisation 14

5) Algorithmes de calcul de polynômes sous-résultants

Algorithme n°1 ..... 16
Algorithme n°2 ..... 17
Algorithme n°3 ..... 18
Algorithme n°4 ..... 19
Algorithme n°5 ..... 20
Conclusions ..... 21

6) Nombre de changements de signes dans la suite des restes signés

Vrais signes des restes ..... 22
Nombre de changements de signes dans la suite des restes signés ..... 23
Le théorème important ..... 24
Spécialisation de la suite de Habicht ..... 26
Permutation des deux polynômes de départ dans la suite de Habicht ..... 28

7) Suite de Sturm et spécialisation

Notations et définitions ..... 29
Spécialisation de la suite de Sturm-Habicht ..... 29
Algorithme pour calculer la suite de Sturm-Habicht ..... 30

8) Traitement de la matrice de Sylvester par la méthode de Bareiss 30

9) Relation de Bezout complète entre plusieurs polynômes

Position du problème ..... 34
Preuve du théorème ..... 35
Algorithme pour une relation de Bezout complète ..... 37

Bibliographie ..... 38

# SUBRESULTANT POLYNOMIALS, STURM SEQUENCE SPECIALISATION


## Abstract

We give a slight generalisation of subresultant polynomials of two polynomials in A[X] where A is an integral domain.

This allows to simplify proofs related to subresultant polynomials, specialisations of subresultant polynomials and algorithms to compute them.

We give explicit relations between the remainder sequence ans the subresultant sequence of two polynomials, and discuss the polynomial time computability of the remainder sequence.

We prove that the formal version of the Sturm sequence of two polynomials (here called the Sturm-Habicht sequence) is as good as the Sturm sequence for computing the number of real roots of a polynomial on an interval, but it is necessary to introduce a special counting rule for the number of changes of signs of the Sturm-Habicht sequence at a real root of a defective subresultant polynomial.

We study the polynomial time computability of a complete Bezout relation for a list of polynomials.



## Key-Words

Subresultant polynomials, subresultant algorithm, Sturm sequence, formal Sturm sequence, complete Bezout relation, Smith normal form, polynomial time computability

# SOUS-RESULTANTS , SUITE DE STURM , SPECIALISATION

Henri LOMBARDI Laboratoire de Mathématiques
UFR des Sciences et Techniques Besançon
Université de Franche Comté

Résumé

Nous donnons une légère généralisation de la notion de *polynôme sous-résultant* (ou *sous-pgcd*) de 2 polynômes de A[X] où A est un anneau intègre.

Ceci permet de simplifier les démonstrations concernant les sous-pgcd, de préciser les algorithmes pour les calculer, et de traiter de manière agréable les problèmes de *spécialisation* (c.-à-d. lorsqu'on transforme les coefficients par un homomorphisme d'anneaux), même lorsqu'il y a chute du degré (d'un ou même parfois des deux polynômes), en particulier dans le cas de la *suite de Sturm* et de ses généralisations.

Nous explicitons les relations entre suite des restes et suite des sous-pgcd, et discutons la question de la calculabilité en temps polynomial de la suite des restes.

Nous démontrons que la version formelle de la suite de Sturm, que nous appelons *suite de Sturm-Habicht*, fonctionne aussi bien que la suite de Sturm pour le comptage des racines réelles sur un intervalle, à condition d'introduire une règle particulière pour évaluer le nombre de changements de signes de la suite de Sturm-Habicht lorsqu'on est en un zéro d'un polynôme sous-résultant défectueux.

Nous étudions la calculabilité en temps polynomial d'une *relation de Bezout complète* entre plusieurs polynômes, qui est un cas particulier de la réduction de Smith d'une matrice.

Mots clé

Sous-résultants, sous-pgcd, algorithme des sous-résultants, suite de Sturm, suite de Sturm formelle, spécialisation, relation de Bezout, forme normale de Smith, calculabilité en temps polynomial

## Remerciements

Je remercie vivement Marie-Françoise Roy et Laureano Gonzalez. Sans les longues discussions que nous avons eues et les éclaircissements qu'elles m'ont apportés, cet article n'aurait pas vu le jour.

### *Introduction*

Nous donnons une légère généralisation de la notion de *polynômes sous-résultants* de 2 polynômes (cf **[Hab]**, **[Loos]**), polynômes que nous appelons également des *sous-pgcd*. L'utilité de cette généralisation s'avère lorsque nous étudions les problèmes liés à la spécialisation (§ 4 notamment) , en outre, les preuves de plusieurs résultats sont simplifiées.

La suite de polynômes sous-résultants est en quelque sorte une version formelle de la suite des restes, beaucoup plus facile à calculer, notamment pour les raisons suivantes:

– lorsque nous travaillons dans un anneau intègre **A** où les divisions exactes sont relativement aisées, on peut utiliser des algorithmes de calcul des sous-pgcd qui n'utilisent que des additions, multiplications et divisions exactes dans l'anneau **A** , et les coefficients obtenus restent polynomialement majorés si les déterminants sont polynomialement majorés dans **A** (par exemple avec $A = \mathbb{Z}[X_1, ..., X_n]$ )

– les sous-pgcd se spécialisent bien : si les divisions exactes dans un autre anneau **A'** ne sont pas aisées, on peut utiliser un algorithme de calcul des sous-pgcd dans **A** puis une *spécialisation* (i.e. un homomorphisme d'anneaux) de **A** vers **A'** (par exemple avec $A' = \mathbb{Z}[\xi_1, ..., \xi_n]$ où les $\xi_i$ sont des nombres algébriques).

– si on essaye de calculer la suite des restes directement dans le corps des fractions de **A** , on est confronté à l'alternative suivante: ou bien ne pas simplifier les fractions obtenues au fur et à mesure, mais alors la taille des coefficients explose presque à tout coup; ou bien simplifier les fractions obtenues, mais cela exige un calcul de pgcd dans **A** (en général nettement plus coûteux qu'une division exacte dans **A**), et on n'est même pas prémuni contre une possible explosion de la taille des fractions réduites (cf prop 8).

Supposons maintenant que le corps des fractions de **A** est muni d'un ordre.

Le nombre de changements de signes dans la suite des restes (convenablement modifiée) intervient dans le théorème de Sturm et dans des généralisations du théorème de Sturm (cf par exemple **[Syl]** ). Il s'avère en fait que la suite des sous-pgcd fait presque aussi bien l'affaire que la suite des restes pour calculer, à une constante près, le nombre de changements de signes. Ceci peut se déduire de résultats de Habicht, comme l'a montré Laureano Gonzalez (cf **[Gon]**). Nous en donnons dans cet article une preuve "directe".

Nous étudions dans le § 2 la suite des sous-pgcd par une méthode qui n'utilise pas de calculs de déterminants et se généralise au cas de n polynômes (cf § 9 ). Cela suffit à établir des résultats généraux concernant la calculabilité en temps polynomial de la suite des restes "à facteurs multiplicatifs non nuls près".

Nous établissons dans le § 3 les formules reliant explicitement la suite des restes à la suite des sous-pgcd. Nous retrouvons en les précisant les résultats du § précédent. Nous discutons la question de la calculabilité en temps polynomial de la suite des restes.

Dans le § 4 , étant donnée une spécialisation $Sp: A \to A'$, nous étudions la possibilité de calculer "facilement" les polynômes sous-résultants de Sp(P) et Sp(Q) lorsqu'on connaît les polynômes sous-résultants de P et Q .

Dans le § 5 , nous donnons différentes variantes de l'algorithme des sous-résultants de Habicht-Loos (donné dans **[Ha]** et modifié dans **[Loos]**). Nous sommes en désaccord avec **[Loos]** sur certains points de détail.

Dans le § 6 , nous définissons la *suite des restes signés* de 2 polynômes (qui est la suite des restes avec des signes modifiés selon une convention à la Sturm), puis la *suite de*

*Habicht* de 2 polynômes qui est une version formelle de la suite des restes signés. Nous démontrons par une méthode directe un résultat de Laureano Gonzalez qui montre que la suite de Habicht fait aussi bien l'affaire que la suite des restes signés dans le théorème de Sturm ou dans ses généralisations. Nous étudions les problèmes liés à la spécialisation d'une suite de Habicht.

Dans le § 7 , nous définissons *suite de Sturm-Habicht d'un polynôme* et indiquons comment elle se spécialise lorsque le degré du polynôme chute de 1 par spécialisation

Dans le § 8 , nous indiquons comment la suite de Habicht est calculée lors du traitement de la matrice de Sylvester (convenablement présentée) par la méthode du pivot améliorée à la Bareiss.

Dans le § 9 , nous étudions la calculabilité d'une *relation de Bezout complète entre plusieurs polynômes*, à titre de prolongement naturel des § 2 et 8 .

## Plan de l'article

1) Matrice de Sylvester , polynômes sous-résultants et suite des restes: notations, premiers résultats
2) Un peu d'algèbre linéaire: conséquences pour la complexité
3) Sous-pgcd et vrais restes: des formules explicites
4) Spécialisation
5) Algorithmes de calcul des polynômes sous-résultants
6) Nombre de changements de signes dans la suite des restes signés
7) Suite de Sturm et spécialisation
8) Traitement de la matrice de Sylvester par la méthode de Bareiss
9) Relation de Bezout complète entre plusieurs polynômes

# 1) Matrice de Sylvester , polynômes sous-résultants et suite des restes: notations, premiers résultats

Nous donnons dans ce § une légère généralisation de la notion de polynôme sous-résultant. L'utilité de cette généralisation s'avèrera lorsque nous étudierons les problèmes liés à la spécialisation.

Nous établissons en outre les relations liant polynômes sous-résultants "ordinaires" et "généralisés".

On considère un anneau intègre **A** et son corps de fractions **K** .

## Notations

*Polynômes, suite des restes*

Nous noterons $d(P)$ le degré d'un polynôme P ,
$cd(P)$ son coefficient dominant et
$\mathbf{cf}_j(P)$ son coefficient de degré j (égal à 0 si j est > d(P)).

Si R est le reste de la division de P par Q dans **K**[X], on note

$$\mathbf{Rst}(P,Q) := R ,$$

Nous considérons maintenant la **suite des restes de l'algorithme d'Euclide**, démarrant avec le numéro 0, et définie de manière récurrente par :

$$\mathbf{Rst}^0(P,Q) := P, \qquad \mathbf{Rst}^1(P,Q) := Q,$$
$$\mathbf{Rst}^{m+1}(P,Q) := \mathbf{Rst}(\mathbf{Rst}^{m-1}(P,Q) , \mathbf{Rst}^m(P,Q) )$$

On arrête la suite au premier reste nul. Le pgcd de P et Q est le dernier reste non nul.
En posant $t := \sup(d(P),d(Q)+1)$, nous noterons

$\mathbf{Rst}_t(P,Q) := P$ , $\mathbf{Rst}_{t-1}(P,Q) := Q$ et
(pour $-1 < j < t-1$ ) $\mathbf{Rst}_j(P,Q)$ le reste de plus fort degré inférieur ou égal à j dans la suite des restes $\mathbf{Rst}^m(P,Q)$ avec $m \geq 1$.

*Matrice de Sylvester*

Si P et Q sont dans **A**[X], p , q , et j des entiers avec $d(P) \leq p$, $d(Q) \leq q$ et $j < \inf(p,q)$ , nous notons $\mathbf{Sylv}_j(P,p, Q,q)$ la j-ème matrice extraite de "*la matrice de Sylvester de P et Q considérés comme étant de degrés p et q*" :
sur la base $X^{p+q-j-1},..., X^2, X, 1,$ les vecteurs lignes successifs de cette matrice sont : $P.X^{q-j-1},..., P.X, P, Q.X^{p-j-1},..., Q.X, Q$ .
Cette matrice possède $Nl = p+q-2j$ lignes et $Nc = p+q-j$ colonnes.

Nous noterons $\mathbf{Esylv}_j(P,p, Q,q)$ le sous espace de **K**[X] engendré par les polynômes $P.X^{q-j-1},..., P.X, P, Q.X^{p-j-1},..., Q.X, Q$ (lignes de la matrice $\mathbf{Sylv}_j(P,p, Q,q)$ ).

## Définitions

*Polynôme associé à une matice*

Par définition, le **polynôme associé à une matrice** possédant Nl lignes et Nc colonnes, où $Nc = Nl + j$ , est un polynôme de degré $\leq j$ : son coefficient de degré d est le déterminant extrait de cette matrice sur les colonnes 1,2,...,Nl-1,Nc-d .

*Polynômes sous-résultants (ou sous-pgcd) de deux polynômes*

Les **polynômes sous-résultants** (ou encore : **les sous-pgcd** ) (de P et Q considérés comme étant de degrés p et q) sont les polynômes associés aux matrices

$\mathbf{Sylv}_j(P,p,\ Q,q)$.

Ils seront notés $\mathbf{Sres}_j(P,p,\ Q,q)$ $\quad j < \inf(p,q)$

Il est clair que les polynômes sous-résultants sont à coefficients dans **A** et que $\mathbf{Sres}_j(P,p,\ S,s)$ est de degré inférieur ou égal à j . Si $\mathbf{Sres}_j(P,p,\ S,s)$ est de degré < j on dit qu'il est **défectueux**.

Si $p = d(P) \geq q = d(Q)$, on appelle le **pseudo-reste** de la division de P par Q le polynôme à coefficients dans **A** :

$$\mathbf{Prst}(P,Q) := \mathbf{Sres}_{q-1}(Q,q,\ P,p) = cd(Q)^{p-q+1}\ .\ \mathbf{Rst}(P,Q)$$

La **suite des polynômes sous-résultants** est la liste des $\mathbf{Sres}_j(P,p,\ Q,q)$ pour j descendant de $\inf(p,q)-1$ à 0 .

Nous appellerons **polynôme sous-résultant standard** un sous-pgcd $\mathbf{Sres}_j(P,p,\ Q,q)$ où $d(P) = p$ et $d(Q) = q \leq p$ .

*Coefficients sous-résultants de deux polynômes*

Les **coefficients sous-résultants** (ou encore les **sous-résultants** ) (de P et Q considérés comme étant de degrés p et q) sont les coefficients suivants:

$$\mathbf{sr}_j(P,p,\ Q,q) := \mathbf{cf}_j(\mathbf{Sres}_j(P,p,\ Q,q)) \qquad j < \inf(p,q)$$

**Remarques**

(a) Le sous-pgcd $\mathbf{Sres}_0(P,p,\ Q,q) = \mathbf{sr}_0(P,p,\ Q,q)$ est le résultant de P et Q si $p = d(P)$ et $q = d(Q)$ .

(b) On a les relations $\mathbf{Sres}_j(a.P,p,\ b.Q,q) = a^{q-j}.b^{p-j}.\mathbf{Sres}_j(P,p,\ Q,q)$

$\mathbf{Rst}(a.P,b.Q) = a.\mathbf{Rst}(P,Q)$ et

$\mathbf{Prst}(a.P,b.Q) = a.b^{p-q+1}.\mathbf{Prst}(P,Q)$.

(c) Les polynômes sous-résultants définis dans [**Loos**] p. 118 sont les sous-pgcd standards.

**Autres définitions et notations**

Nous donnons au § 5 p 18 une extension "raisonnable" de la suite des sous-pgcd en la faisant démarrer à $j = p$ , du moins lorsque $p > q = d(Q)$. Cette définition est utilisée au début du § 6 p. 23 dans la définition de la **suite de Habicht** de 2 polynômes. On trouve au début du § 6 p. 23 la définition de la **suite des restes signés** de 2 polynômes ainsi que les notations et conventions concernant le nombre de changements de signes dans une suite. Au début du § 7 p. 29 on trouve les définitions de la **suite de Sturm** et de la **suite de Sturm-Habicht** d'un polynôme

**Relations entre sous-pgcd généraux et sous-pgcd standards**

Ordinairement on calcule les sous-pgcd standards. Mais après spécialisation, il se peut que le degré de P ou celui de Q se retrouve diminué, aussi est-il intéressant d'étudier le comportement des sous-pgcd dans le cas où l'un des 2 degrés est plus petit que le degré annoncé. Si les 2 degrés sont trop petits, tous les sous-pgcd sont nuls.

Les autres sous-pgcd peuvent tous être facilement calculés à partir des sous-pgcd standards (ou vice-versa si l'autre sous-pgcd n'est pas identiquement nul), en appliquant la proposition suivante (pour plus de détails cf § 4) :

**Proposition 1 :**

Nous supposons $d(P) \le p$ , $d(Q) \le q$ , $j < \inf(p,q)$

a) Si $d(P) < p$ et $d(Q) < q$ , alors

$$\mathbf{Sres}_j(P,p,Q,q) = 0$$

b) $$\mathbf{Sres}_j(P,p,Q,q) = (-1)^{(p-j)(q-j)}\ \mathbf{Sres}_j(Q,q,P,p)$$

c) Si $q' \ge q$ et $d(P) = p$ alors

$$\mathbf{Sres}_j(P,p,Q,q') = cd(P)^{q'-q}\ .\ \mathbf{Sres}_j(P,p,Q,q)$$

$$\mathbf{Sres}_j(Q,q',P,p) = ((-1)^{p-j}\, cd(P))^{q'-q}\ .\ \mathbf{Sres}_j(Q,q,P,p)$$

*preuve>*

a) la première colonne de $\mathbf{Sylv}_j(P,p,Q,q)$ est nulle

b) cela revient à calculer le signe d'une permutation

c1) la nouvelle matrice est obtenue à partir de l'ancienne en rajoutant q'-q colonnes nulles à gauche, puis q'-q lignes au dessus, chacune portant le polynôme P décalé à chaque fois d'un cran. Les déterminants intervenant dans le calculs des $\mathbf{Sres}_j$ sont donc tous multipliés par $cd(P)^{q'-q}$.

c2) on applique c1) et 2 fois b) ❑

**NB :** Lorsque P est unitaire, la proposition 1 c1) montre que le polynôme sous-résultant $\mathbf{Sres}_j(P,p,Q,q)$ ne dépend pas du choix de $q \ge d(Q)$ .

**Proposition 2 :**

Soient P et Q des polynômes de degrés p et $q < p-1$, alors :

a) $$\mathbf{Sres}_j(P,p,Q,p-1) = 0 \quad \text{si } q < j < p-1$$

b) $$\mathbf{Sres}_q(P,p,Q,p-1) = (cd(P)\, cd(Q))^{p-q-1}\ Q$$

c) $$\mathbf{Sres}_j(P,p,Q,p-1) = cd(P)^{p-q-1}\ \mathbf{Sres}_j(P,p,Q,q) \qquad \text{pour } j < q$$

d) $$\mathbf{Sres}_{q-1}(P,p,Q,p-1) = (-cd(P))^{p-q-1}\ \mathbf{Prst}(P,Q)$$

*preuve>* a) et b) : faire un dessin de la matrice $\mathbf{Sylv}_j$ : par exemple

avec $p = 6$ , $q = 3$ , $j = 4$

```
x x x x x x x      (les . représentent les 0 de la matrice)
• • y y y y .      (les • sont les coeffs 0 de Q au dessus du degré)
. • • y y y y      (les x représentent les coeffs de P )
                   (les y représentent les coeffs de Q )
```

avec $p = 6$ , $q = 3$ , $j = 3$

```
x x x x x x x .
. x x x x x x x
• • y y y y . .
. • • y y y y .
. . • • y y y y
```

c) c'est la prop 1c

d) on applique c) avec $j = q-1$ et on remarque que

$$\mathbf{Prst}(P,Q) = \mathbf{Sres}_{q-1}(Q,q,P,p) = (-1)^{p-q-1}\ \mathbf{Sres}_{q-1}(P,p,Q,q)\ ,$$

(la 1ère égalité par définition, la 2ème en appliquant prop 1 b) ) .❑

# 2) Un peu d'algèbre linéaire: conséquences pour la complexité

Nous étudions dans ce § la suite des sous-pgcd *dans le cas standard* par une méthode utilisant uniquement des arguments de dimension, sans faire appel à des calculs de déterminants. Cela suffit à établir des résultats généraux concernant la calculabilité en temps polynomial de la suite des restes *à facteurs multiplicatifs non nuls près.*

On considère 2 polynômes P et Q , de degrés p et q avec $p \geq q$ , à coefficients dans l'anneau intègre **A** de corps des fractions **K** .

**Notations :**
On note pour abréger $Rst_j$ au lieu de $Rst_j(P,Q)$, $Sylv_j$ au lieu de $Sylv_j(P,p,Q,q)$, $Esylv_j$ au lieu de $Esylv_j(P,p,Q,q)$.

**Proposition 3 :**

Soient 2 polynômes P et Q , de degrés p et q avec $p \geq q$ .

Soient $j_1 \leq q$ et $j_2 \geq 0$ les degrés de 2 restes successifs (dans l'algorithme des divisions successives démarrant avec P et Q). Alors:

a) la matrice $\mathbf{Sylv}_{j_1-1}$ est de rang maximum (c.-à-d.: égal au nombre de ses lignes)

b) l'espace $\mathbf{Esylv}_{j_1-1}$ possède une base qui commence par $\mathbf{Rst}_{j_2}$ , $\mathbf{Rst}_{j_1}$ , et se poursuit par des polynômes dont les degrés augmentent régulièrement de 1 , tous de la forme $X^i.\mathbf{Rst}_j$ (avec $j \geq j_1$ )

c) si $j_1 > j_2 + 1$, la matrice $\mathbf{Sylv}_{j_2}$ est de rang maximum (c.-à-d.: égal au nombre de ses lignes) et: l'espace $\mathbf{Esylv}_{j_2}$ possède une base qui commence par $\mathbf{Rst}_{j_2}$ et se poursuit par des polynômes dont les degrés augmentent régulièrement de 1 , tous de la forme $X^i.\mathbf{Rst}_j$ (avec $j \geq j_2$ )

Avec les hypothèses $j_1 > 0$ et $\mathbf{Rst}_{j_2} = 0$ (degré = -1), on obtient :

a') la matrice $\mathbf{Sylv}_{j_1-1}$ est de rang égal au nombre de ses lignes moins 1

b') l'espace $\mathbf{Esylv}_{j_1-1}$ possède une base qui commence par $\mathbf{Rst}_{j_1}$ , et se poursuit par des polynômes dont les degrés augmentent régulièrement de 1 , tous de la forme $X^i.\mathbf{Rst}_j$ (avec $j \geq j_1$ )

*preuve>*

- On amorce la pompe avec $j_1 = q$ , les résultats a) et b) (ou a') et b') ) se montrent sans difficulté (diviser P par Q )

- On a la relation de récurrence $\mathbf{Esylv}_{m-1} = \mathbf{Esylv}_m + X.\mathbf{Esylv}_m$.

Par ailleurs $\mathbf{Sylv}_{m-1}$ possède 2 lignes de plus que $\mathbf{Sylv}_m$.

- Soient $j > k > h$ les degrés de 3 restes successifs. Supposons les résultats a) et b) démontrés avec $j_1 = j$ et $j_2 = k$ . Supposons pour fixer les idées $j = k+3$ .

Soit $\mathbf{Rst}_k$, $\mathbf{Rst}_j$, $R_1$, $R_2$,... $R_n$ une base de $\mathbf{Esylv}_{j-1}$ comme décrite en b). Le nombre de lignes de $\mathbf{Sylv}_{j-1}$ est donc égal à $n + 2$ .

- Vue la relation de récurrence l'espace $\mathbf{Esylv}_{j-2}$ contient les $n + 4$ polynômes $\mathbf{Rst}_k$, $X.\mathbf{Rst}_k$, $\mathbf{Rst}_j$, $X.\mathbf{Rst}_j$, $X.R_1$, $X.R_2$,... , $X.R_n$ dont les degrés sont strictement

croissants. Cela implique que la matrice $\mathbf{Sylv}_{j-2}$ est de rang maximum, (elle possède en effet $n+4$ lignes) et que les polynômes ci-dessus sont une base de $\mathbf{Esylv}_{j-2}$.
- De même, l'espace $\mathbf{Esylv}_k = \mathbf{Esylv}_{j-3}$ contient les $n+6$ polynômes

$\mathbf{Rst}_k$, $X.\mathbf{Rst}_k$, $X^2.\mathbf{Rst}_k$, $\mathbf{Rst}_j$, $X.\mathbf{Rst}_j$, $X^2.\mathbf{Rst}_j$, $X^2.R_1$, $X^2.R_2$,... $X^2.R_n$ dont les degrés croissent régulièrement de 1 en 1 . Cela implique que la matrice $\mathbf{Sylv}_k$ est de rang maximum, et que les polynômes ci-dessus sont une base de $\mathbf{Esylv}_k$. Ceci montre le c) pour $j_1 = j$ et $j_2 = k$.
- enfin l'espace $\mathbf{Esylv}_{k-1}$ contient les $n+8$ polynômes
$\mathbf{Rst}_k$, $X.\mathbf{Rst}_k$, $X^2.\mathbf{Rst}_k$, $X^3.\mathbf{Rst}_k$, $\mathbf{Rst}_j$, $X.\mathbf{Rst}_j$, $X^2.\mathbf{Rst}_j$, $X^3.\mathbf{Rst}_j$, $X^3.R_1$, $X^3.R_2$, $X^3.R_n$. Les 5 premiers polynômes engendrent le même espace que les polynômes $\mathbf{Rst}_h$, $\mathbf{Rst}_k$, $X.\mathbf{Rst}_k$, $X^2.\mathbf{Rst}_k$, $X^3.\mathbf{Rst}_k$: diviser $\mathbf{Rst}_j$ par $\mathbf{Rst}_k$. Et $\mathbf{Esylv}_{k-1}$ contient donc les polynômes $\mathbf{Rst}_h$, $\mathbf{Rst}_k$, $X.\mathbf{Rst}_k$, $X^2.\mathbf{Rst}_k$, $X^3.\mathbf{Rst}_k$, $X.\mathbf{Rst}_j$, $X^2.\mathbf{Rst}_j$, $X^3.\mathbf{Rst}_j$, $X^3.R_1$, $X^3.R_2$,... $X^3.R_n$ dont les degrés sont strictement croissants.
Si $\mathbf{Rst}_h$ est non nul, cela prouve donc a) et b) pour $j_1 = k$ et $j_2 = h$ .
Si $\mathbf{Rst}_h$ est nul avec $k > 0$, nous voulons montrer a') et b') pour $j_1 = k$ et $j_2 = h$ : nous connaissons déjà $n+7$ polynômes linéairement indépendants convenables dans le sous-espace $\mathbf{Esylv}_{k-1}$ ; il suffit donc de voir que la matrice $\mathbf{Sylv}_{k-1}$ ne peut pas être de rang maximum, et ceci se déduit du fait que tous les polynômes de $\mathbf{Esylv}_{k-1}$ sont multiples de $\mathbf{Rst}_k$ (pgcd de P et Q) et ont leur degré convenablement majoré. ❑

On en déduit immédiatement:

### Théorème 1

a) Pour $j < q$ le sous-pgcd $\mathbf{Sres}_j$ est proportionnel au reste $\mathbf{Rst}_j$
b) Si $j_1 \le q$ et $j_2$ sont les degrés de 2 restes successifs , les polynômes $\mathbf{Sres}_{j_1-1}$ et $\mathbf{Sres}_{j_2}$ sont multiples de $\mathbf{Rst}_{j_2}$ avec des facteurs non nuls
c) Si $j_1-1 > j > j_2$ , alors $\mathbf{Sres}_j$ est identiquement nul

*preuve>*

* Si la matrice $\mathbf{Sylv}_j$ n'est pas de rang maximum, le sous-pgcd est identiquement nul. Ceci se produit lorsque $\mathbf{Rst}_{j_1}$ est le dernier reste non nul, avec $j < j_1$ . En effet, les lignes de la matrice représentent des polynômes tous multiples de $\mathbf{Rst}_{j_1}$ , ce qui donne la dimension de $\mathbf{Sylv}_j$ : $p - q - j - j_1$. Or la matrice possède plus de lignes: $p - q - 2j$ . Ceci prouve a) lorsque j est plus petit que le degré du dernier reste non nul, ainsi que b) et c) lorsque $j_2 = -1$

* Voyons b) et c) lorsque $j_2 \ge 0$ et $j_1-1 \ge j \ge j_2$ : la matrice $\mathbf{Sylv}_j$ est alors de rang maximum. Une base de $\mathbf{Esylv}_j$ est donnée dans la proposition 3 (ou dans sa preuve). Or, lorsqu'on remplace les vecteurs lignes, supposés indépendants, d'une matrice, par une autre base de l'espace engendré, les polynômes associés aux deux matrices sont proportionnels avec pour rapport le déterminant d'une matrice de passage. Nous pouvons donc raisonner avec la matrice ayant pour vecteurs lignes la base de $\mathbf{Esylv}_j$ fournie à la proposition 3 (ou dans sa preuve). Cette matrice est sur-triangulaire et son polynôme associé est immédiatement calculé: c'est un multiple de $\mathbf{Rst}_{j_2}$, le facteur étant non nul juste pour $j = j_1-1$ ou $j = j_2$. ❑

### Corollaire

a) Si **A** est un anneau intègre où les déterminants sont calculables en temps polynomial[1], on peut calculer en temps polynomial une suite de

[1] Voir la remarque qui suit

polynômes de A[X] égaux, à des facteurs non nuls près, aux polynômes de la suite des restes de P et Q

b) Si en outre le corps des fractions **K** de **A** est muni d'un ordre tel que le signe d'un élément de **A** soit calculable en temps polynomial, alors on peut calculer en temps polynomial une suite de polynômes de **A**[X] égaux, à des facteurs strictement positifs près, aux polynômes de la suite des restes de P et Q

*preuve>*

a) immédiat

b) Il s'agit de multiplier chacun des sous-pgcd par le facteur +1 ou -1. Il faut voir que ces facteurs peuvent être déterminés en temps polynomial. On procède de proche en proche, en calculant, pour 2 sous-pgcd successifs, leur pseudo-reste, ce qui permet en utilisant les signes des différents coefficients dominants obtenus, de décider le "vrai signe" du reste correspondant. ❑

**Remarque :**

La phrase "les déterminants sont calculables en temps polynomial dans **A**" signifie précisément ceci: **A** est dénombrable et est codé de manière que l'on ait à la fois :

- l'égalité (dans **A** de 2 mots du code) est testable en temps polynomial, et
- les déterminants sont calculables en temps polynomial.

Dans l'anneau $\mathbf{A} = \mathbb{Z}[X_1, \ldots, X_n]$ les déterminants sont calculables en temps polynomial. Les méthodes les plus performantes semblent être actuellement les méthodes modulaires (cf [**Col**]).

Le plus souvent , on peut obtenir la calculablilité en temps polynomial des déterminants par utilisation de la méthode de Bareiss : grosso modo : si dans **A** (convenablement codé) les additions, multiplications, divisions exactes, tests d'égalité sont en temps polynomial et si la taille des déterminants est polynomialement majorée, alors le calcul des déterminants est en temps polynomial. Une étude générale de la calculabilité des déterminants en temps polynomial est donnée dans [**Lom**].

## 3) Sous-pgcd et vrais restes : des formules explicites

Nous établissons dans ce § les formules reliant explicitement la suite des restes à la suite des sous-pgcd standards. Nous retrouvons en les précisant les résultats du § précédent. Nous discutons la question de la calculabilité en temps polynomial de la suite des restes.

Dans tout le § nous utiliserons souvent l'hypothèse suivante, notée (H) :

| (H) $\quad p = d(P) \geq q = d(Q)$ , $R = \mathbf{Rst}(P,Q)$ , et $r = d(R)$ |
|---|

Nous commençons par une proposition qui sert de base aux calculs qui suivent:

**Proposition 4 :** Supposons (H) et $j < q$, alors:

$\mathbf{Sres}_j(P,p,Q,q) = \mathbf{Sres}_j(R,p,Q,q)$ et $\mathbf{Sres}_j(Q,q,P,p) = \mathbf{Sres}_j(Q,q\,R,p)$

*preuve>* Chaque ligne $P.X^k$ de la matrice $\mathbf{Sylv}_j(P,p,Q,q)$ peut être remplacée par la ligne $R.X^k$ en lui rajoutant des lignes $-c_m.Q.X^{k+m}$, en choisissant pour $c_m$ les coefficients du polynôme B dans l'identité de la division euclidienne: $P = B.Q + R$ . Ces manipulations

élémentaires ne modifient pas les déterminants extraits.
Or, la nouvelle matrice obtenue n'est autre que $\mathbf{Sylv}_j(R,p,Q,q)$
**NB**: cette dernière matrice, à coefficients dans **K**, admet donc un polynôme associé à coefficients dans **A** ❑

***Le cas générique*** *( les degrés dans la suite des restes décroissent de 1 en 1 )*

**Proposition 5 :** Supposons (H) et $p = q+1$. Alors nous avons:

a) $\mathbf{Sres}_{q-1}(P,p,Q,q) = cd(Q)^2\ R = \mathbf{Prst}(P,Q)$

b) $\mathbf{Sres}_j(P,p,Q,q) = cd(Q)^2\ \mathbf{Sres}_j(Q,q,R,q-1)$ pour $j < q-1$

**Théorème 2 :** Supposons (H), et que les degrés dans la suite des restes décroissent de 1 en 1 (en commençant au polynôme P).
Posons $c(q) := cd(Q)$ et, pour $j < q$, $c(j) := cd(\mathbf{Rst}_j(P,Q))$. Alors:

$$\mathbf{Sres}_j(P,p,Q,q) = (\,c(q).c(q-1)...c(j+1)\,)^2\,\mathbf{Rst}_j(P,Q) \quad \text{pour } j < q$$

En particulier, si on est dans un corps ordonné, chaque sous-pgcd a "même signe" que le reste correspondant.

*preuve>* Le théorème résulte immédiatement de la proposition 5. La prop 5a résulte de l'égalité pour le pseudo-reste et de la prop 1c. La prop 5b résulte de la prop 4 et des prop 1c et 1b. ❑

Nous redémontrons maintenant le "théorème de Habicht" dans [**Loos**] par un calcul direct.

**Théorème de Habicht :**
Nous supposons $d(P) \le p = q+1$, $d(Q) \le q$ [1].
Nous posons $S_p := P$, $S_q := Q$, $S_j := \mathbf{Sres}_j(P,p,Q,q)$ pour $j < q$, $C(j) := \mathbf{cf}_j(S_j)$ pour $j \le q$, $C(p) := 1$. Alors, pour $0 \le h < j \le q$, on a :

$$C(j+1)^{2(j-h)}\ S_h = \mathbf{Sres}_h(\,S_{j+1}\,,j+1,S_j\,,j\,)$$

En particulier, lorsque $d(S_{j+1}) = j+1$ et $d(S_j) = j$, on obtient:

$$C(j+1)^2\ S_{j-1} = \mathbf{Prst}(\,S_{j+1}\,,S_j\,)$$

*preuve>* Le cas particulier résulte de l'égalité générale, avec $h = j-1$, et de la définition du pseudo-reste. Les égalités générales à démontrer sont des identités algébriques. On peut donc supposer que les coefficients de P et Q sont des *variables indépendantes*. On applique alors les résultats du théorème 2. Les 2 membres de l'égalité à établir sont des multiples de $\mathbf{Rst}_h(P,Q)$. Les calculs sont simples. Nous les explicitons en reprenant les notations du théorème 2.

Nous posons $R_j := \mathbf{Rst}_j(P,Q)$, $\gamma(j) := (\,c(q).c(q-1)...c(j+1)\,)^2 = C(j)/c(j)$.
On a donc $C(j+1)^2 = \gamma(j).\gamma(j+1)$, $S_j = \gamma(j).R_j$.
Par ailleurs
$$\begin{aligned}\mathbf{Sres}_h(S_{j+1},j+1,S_j,j) &= \gamma(j+1)^{j-h}.\gamma(j)^{j-h+1}\mathbf{Sres}_h(R_{j+1},j+1,R_j,j)\\ &= \gamma(j+1)^{j-h}.\gamma(j)^{j-h+1}.(\,c(j).c(j-1)...c(h+1)\,)^2R_h\\ &= \gamma(j+1)^{j-h}.\gamma(j)^{j-h}.\gamma(h).R_h\end{aligned}$$
et $S_h = \gamma(h).R_h$ ❑

[1] Pour que le théorème affirme autre chose que des égalités 0 = 0, il faut que l'on ait $d(P) = p$ ou $d(Q) = q$.

**Théorème de Habicht : (2ème version)**

Nous supposons $d(P) \leq p$ , $d(Q) = q$ , $p > q$.

Nous posons $T_q := cd(Q)^{p-1-q} Q$ , $T_j := \mathbf{Sres}_j(P,p,Q,q)$ pour $j < q$ . Alors, pour $0 \leq h < j < q$ , on a :

$$\mathbf{cf}_j(T_j)^{2(j-h)} T_h = \mathbf{Sres}_h(T_{j+1}, j+1, T_j, j)$$

En particulier, lorsque $d(T_{j+1}) = j+1$ et $d(T_j) = j$ , on obtient :

$$\mathbf{cf}_{j+1}(T_{j+1})^2 T_{j-1} = \mathbf{Prst}(T_{j+1}, T_j)$$

*preuve>* Le cas particulier résulte de l'égalité générale, avec $h = j - 1$ , et de la définition du pseudo-reste. Les égalités générales à démontrer sont des identités algébriques. On peut donc supposer que les coefficients de P et Q sont des *variables indépendantes.*
Si $q = p - 1$ , on retrouve la 1ère version du théorème de Habicht. Si $q < p - 1$ , on pose $S_j := \mathbf{Sres}_j(P,p,Q,p-1)$ pour $j \leq q$ , de sorte qu'on a pour tout $j \leq q$ , $S_j = cd(P)^{p-1-q} T_j$ (en appliquant les prop 1 c1) et 2 b) ), et on termine en remarquant que l'égalité du théorème de Habicht est homogène, donc passe des $S_j$ aux $T_j$ . ❑

***Le cas défectueux***

**Proposition 6 :** Supposons (H) . On a:

a) $\mathbf{Sres}_{q-1}(P,p,Q,q) = (-cd(Q))^{p-q+1} R = (-1)^{p-q+1} \mathbf{Prst}(P,Q)$

$\mathbf{Sres}_j(P,p,Q,q) = ((-1)^{q-j} cd(Q))^{p-q+1} \mathbf{Sres}_j(Q,q,R,q-1)$ pour $j < q - 1$

b) On en déduit

$\mathbf{Sres}_j(P,p,Q,q) = 0$ si $r < j < q - 1$

$\mathbf{Sres}_r(P,p,Q,q) = ((-1)^{p-q-1} cd(Q).cd(R))^{q-r-1} \mathbf{Sres}_{q-1}(P,p,Q,q)$

$\mathbf{Sres}_j(P,p,Q,q) = (-1)^{(p-q-1)(q-j)} cd(Q)^{p-r} \mathbf{Sres}_j(Q,q,R,r)$ pour $j < r$

**Corollaire :**

a) Supposons (H). Le polynôme $\mathbf{Sres}_j(P,p,Q,q)$ est ou bien nul, ou bien égal, à un facteur non nul près dans **K** , à $\mathbf{Rst}_j(P,Q)$ .
Le cas "facteur non nul" se produit lorsque j ou j+1 est le degré d'un reste, et pour $j = d(Q) - 1$.

b) Ce résultat reste vrai si $d(P) \leq p$ , $d(Q) \leq q$ , l'une des 2 inégalités étant une égalité, et $j < \inf(d(P),d(Q))$

**Théorème 3 :**

Supposons (H) , et définissons $R_{-1} := P$ , $R_0 := Q$ , $R_i := \mathbf{Rst}^{i+1}(P,Q)$

$d_i = d(R_i)$ , $e_i = d_{i-1} - d_i + 1$ , $f_i = d_{i-1} - d_{i+1}$ , $c_i = cd(R_i)$

alors, pour tout degré $d_i < q$ , on a:

$$\mathbf{Sres}_{d_i - 1}(P,p,Q,q) = \varepsilon_i \cdot c_0^{f_0} \cdot c_1^{f_1} \ldots c_{i-1}^{f_{i-1}} \cdot c_i^{e_i} \cdot R_{i+1}$$

où $\varepsilon_i = 1$ si $\Sigma_{0 \leq k \leq i} (1 + d_k - d_i).e_k$ est pair , $-1$ sinon.

**Corollaire**

Si **A** est un anneau intègre où les déterminants sont calculables en temps polynomial, on peut calculer en temps polynomial une suite de polynômes de

A[X] égaux, à des facteurs carrés non nuls près dans **A** , aux polynômes de la suite des restes de P et Q

preuves>

*Proposition 6*

a1) en effet : $\mathbf{Sres}_{q-1}(Q,q,\ P,p) = cd(Q)^{p-q+1}\ R = \mathbf{Prst}(P,Q)$ et
$\mathbf{Sres}_{q-1}(P,p,\ Q,q) = (-1)^{p-q+1}\ \mathbf{Sres}_{q-1}(Q,q,\ P,p)$

a2) on a $\mathbf{Sres}_j(P,p,\ Q,q) = \mathbf{Sres}_j(R,p,\ Q,q)$ (prop 4)
$= (-1)^{(p-j)(q-j)}\ \mathbf{Sres}_j(Q,q,\ R,p)$ (prop 1b)
$= (-1)^{(p-j)(q-j)}\ cd(Q)^{p-q+1}\ \mathbf{Sres}_j(Q,q,\ R,q-1)$ (prop 1 c1)

b3) si $r = q - 1$ c'est simplement a2), sinon on applique la prop 2c) à $\mathbf{Sres}_j(Q,q,\ R,q-1)$

b1) on applique a2) puis la prop 2a) à $\mathbf{Sres}_j(Q,q,\ R,q-1)$

b2) si $r = q - 1$ c'est trivial, sinon on applique a2) puis la prop 2 b) à $\mathbf{Sres}_j(Q,q,\ R,q-1)$ et on utilise a1) .

*Corollaire de la proposition 6*

*a)* c'est vérifié pour $j = q - 1, \ldots, r$ d'après la proposition 6 , alinéas b1) et b2) . Pour $j < r$ on utilise l'alinéa b3) qui nous ramène au cas de la suite des restes démarrant avec Q et R . (preuve par induction sur le degré de P donc ) .

*b)* la proposition 1c montre que $\mathbf{Sres}_j(P,p,\ Q,q)$ et $\mathbf{Sres}_j(P,d(P),\ Q,d(Q))$ sont proportionnels dans un facteur non nul pour $j < \inf(d(P),d(Q))$

*Théorème 3*

Se démontre par récurrence sur i en utilisant la proposition 6 . On amorce la pompe avec a1) et la recurrence fonctionne grâce à b3).

*Corollaire du théorème*

Posons $s_{i+1} := cd(\mathbf{Sres}_{d_i-1}(P,p,\ Q,q))$

La formule du théorème appliquée aux coefficients dominants donne:

$$s_{i+1} := \varepsilon_i \cdot {c_0}^{f_0} \cdot {c_1}^{f_1} \ldots\ldots {c_{i-1}}^{f_{i-1}} \cdot {c_i}^{e_i} \cdot c_{i+1}$$

et on voit qu'on a un polynôme $V_{i+1}$ égal à $R_{i+1}$ *à un facteur carré non nul près* qui est de la forme :

$$V_{i+1} := \alpha_i \cdot {s_0}^{g_{0,i}} \cdot {s_1}^{g_{1,i}} \ldots\ldots {s_{i-1}}^{g_{i-1,i}} \cdot {s_i}^{g_{i,i}} \cdot \mathbf{Sres}_{d_i-1}(P,p,\ Q,q)$$

avec les $g_{k,i}$ égaux à 0 ou 1 , et $\alpha_i = \pm 1$ qui peuvent être calculés de proche en proche

❑

### *Discussion sur le temps de calcul de la suite des restes*

Nous établissons maintenant des formules plus explicites lorsque les degrés descendent de 1 en 1 et lorsqu'ils descendent de 3 en 3. Dans le premier cas, on conclut que la suite des restes (pour deux polynômes de départ à coefficients entiers) est calculable en temps polynomial. Le deuxième cas a été choisi parce qu'apparaît alors dans les formules la possibilité d'une explosion exponentielle de la taille des coefficients de la suite $\mathbf{Rst}_j$ . Nous n'avons cependant pas d'exemple explicite d'explosion, et nous laissons donc la question ouverte.

**Proposition 7 :** Nous reprenons les hypothèses du théorème 2, et nous définissons: $c(j) := cd(\mathbf{Rst}_j)$ , $C(j) := cf_j(\mathbf{Sres}_j)$ $C(q) := c(q) := cd(Q)$

Alors les c(j) et les C(j) sont liés par les relations: (pour $0 \le j < q$)

$C(j) = (\ c(q).c(q-1)...c(j+1)\ )^2\ c(j)$

$C(j)\ /\ C(j+1) = c(j+1).c(j)$

$c(q).c(q-1)...c(q-2i) = C(q).C(q-2)....C(q-2i) / C(q-1).C(q-3)....C(q-2i+1)$

$c(q).c(q-1)...c(q-2i-1) = C(q-1).C(q-3)...C(q-2i-1) / C(q).C(q-2)...C(q-2i)$

En particulier, si les déterminants sont calculables en temps polynomial dans **A**, les $\mathbf{Rst}_j$ peuvent être calculés en temps polynomial.

**Proposition 8 :** Nous supposons (H), q = p-1, et que les degrés successifs dans la suite des restes descendent de 3 en 3 (à partir de q). Nous reprenons les notations du théorème 3 et nous définissons en outre:

$S_0 = R_0 = Q$, $S_{i+1} := \mathbf{Sres}_{d_i - 1}$, $C_i := cd(S_i)$.

On obtient alors les relations:

$S_0 = R_0$, $S_1 = c_0^2.R_1$, $S_2 = c_0^4.c_1^4.R_2$

$S_{i+1} = c_0^4.(c_1....c_{i-1})^6 . c_i^4 . R_{i+1}$ $(i \geq 2)$

$C_{i+1} = c_0^4.(c_1....c_{i-1})^6 . c_i^4 . c_{i+1}$ $(i \geq 2)$

$C_0 = c_0$, $C_1 = c_0^2.c_1$

$C_{i+1} / C_i = c_{i-1}^2 c_i^3 c_{i+1}$ $(i \geq 1)$

$c_i . c_{i+1} = C_{i+1} / (C_i . (c_{i-1} c_i)^2)$ $(i \geq 1)$

$c_{i+1} . c_{i+2} = (c_{i-1} c_i)^4 . C_{i+2}. C_i^2 / C_{i+1}^3$ $(i \geq 1)$

$c_0. c_1 = C_1 / C_0$

$c_2. c_3 = C_3.C_1 / C_2^3.C_0^4$

$c_4. c_5 = C_5.C_3^6.C_1^{24} / C_4^3.C_2^{12}.C_0^{16}$ et plus généralement

$$c_{2i}. c_{2i+1} = C_{2i+1} . (\prod_{t=0}^{i-1} C_{2i-2t-1}^{6.4^t}) / (\prod_{t=0}^{i-1} C_{2i-2t}^{3.4^t}) C_0^{4^i}$$

*preuve>* on applique le théorème 2 pour la proposition 7 et le théorème 3 pour la proposition 8 ❑

**Question ouverte :** La proposition 8 incline en faveur de l'affirmation suivante:
Lorsque $\mathbf{A} = \mathbb{Z}$, les coefficients des polynômes de la suite des restes de 2 polynômes P et Q ne sont pas polynomialement majorés en taille à partir de la taille des données (qui sont les listes des coefficients de P et Q)
Soit en effet v une valuation p-adique (ou le logarithme de la valeur absolue archimédienne); on considère l'égalité :

$c_{i+1} . c_{i+2} = (c_{i-1} c_i)^4 . C_{i+2}. C_i^2 / C_{i+1}^3$

on pose $i = 2k+1$, $\gamma_k = c_{2k}.c_{2k+1}$ et on obtient

$v(\gamma_{k+1}) = 4.v(\gamma_k) + v(C_{2k+3}) + 2.v(C_{2k+1}) - 3.v(C_{2k+2})$

On en déduit que :

si $v(\gamma_0) \geq 1$, et pour tout k $3.v(C_{2k+2}) \leq 2^k + v(C_{2k+3}) + 2.v(C_{2k+1})$

alors pour tout k $v(\gamma_k) \geq 2^k$

Si on arrive à réaliser les inégalités ci-dessus, ainsi que les annulations de déterminants qui forcent la suite des degrés des restes à descendre 3 en 3, avec des polynômes P et Q de degrés arbitrairement grands et ayant des coefficients majorés en taille par une fonction polynôme du degré, on aura établi l'affirmation en italique ci-dessus.
Si cette conjecture est vraie, cela signifie que la méthode des divisions successives ne doit pas être appliquée "telle quelle" dans les calculs de PGCD dans $\mathbb{Q}[X]$, et a fortiori dans les calculs

de base standard pour des idéaux de $\mathbb{Q}[X,Y]$ (par exemple), ou dans le calcul d'une forme réduite de Smith d'une matrice à coefficients dans $\mathbb{Q}[X]$.

## 4) Spécialisation

Etant donnée une spécialisation (i.e. un homomorphisme d'anneaux) Sp: $\mathbf{A} \to \mathbf{A'}$, nous étudions la possibilité de calculer "facilement" les polynômes sous-résultants standards de Sp(P) et Sp(Q) lorsqu'on connaît les polynômes sous-résultants standards de P et Q (polynômes de A[X]). Ceci est important car il est fréquent d'avoir un algorithme de division exacte facile dans **A** , et beaucoup plus difficile (voire impossible) dans **A'**. Or les algorithmes rapides de calculs des sous-pgcd utilisent des divisions exactes (cf § 5 et § 8 ).

**1er cas** : *les degrés de P et Q sont conservés au cours d'une spécialisation*

Les polynômes sous-résultants standards se spécialisent en les polynômes sous résultants standards.

**2ème cas** : *un seul des 2 degrés de P ou Q s'abaisse au cours d'une spécialisation*

Nous supposons que les divisions exactes se font "facilement" dans **A**.

Supposons que nous connaissions, par exemple par l'algorithme n°1 donné au § 5 , les polynômes sous-résultants $\mathbf{Sres}_j(P,p, Q,p-1)$.

Si $d(Sp(P)) = d(P)$, on obtient en spécialisant ces sous-pgcd des sous-pgcd non nuls, même si $d(Sp(Q)) < d(Q)$.

Par contre, si $d(Sp(Q)) = d(Q) = q < p-1$ et $d(Sp(P)) < d(P)$, on a $Sp(\mathbf{Sres}_j(P,p, Q,p-1)) = 0$. Il suffit cependant de calculer $\mathbf{Sres}_j(P,p, Q,q)$ à partir de $\mathbf{Sres}_j(P,p, Q,p-1)$ pour obtenir par spécialisation des sous-pgcd non nuls. De manière générale, il est utile de savoir calculer les sous-pgcd standards à partir d'autres sous-pgcd. Ceci est possible si les divisions exactes dans l'anneau considéré sont "faciles". On utilisera les résultats suivants, conséquences immédiates de la proposition 1.

**Proposition 9 :** Nous supposons $d(P) = p$ , $d(Q) = q$ , $j < \inf(p,q)$

a) Si $q \geq p$ alors

$\mathbf{Sres}_j(Q,q, P,q) = (-1)^{q-j}\, \mathbf{Sres}_j(P,q, Q,q) = \mathbf{Sres}_j(P,q+1, Q,q) / cd(Q)$

b) Si $p \geq q' > q$ alors

$\mathbf{Sres}_j(P,p, Q,q) = \mathbf{Sres}_j(P,p, Q,q') / cd(P)^{q'-q}$

$= \mathbf{Sres}_j(P,p, Q,p-1) / cd(P)^{p-q-1}$

c) Si $p' > p \geq q$ alors

$\mathbf{Sres}_j(P,p, Q,q) = \mathbf{Sres}_j(P,p', Q,q) / ( (-1)^{q-j}\, cd(Q) )^{p'-p}$

d) Si $p' \geq q \geq p$ alors

$\mathbf{Sres}_j(Q,q, P,p) = (-1)^{(p'-j)(q-j)}\, \mathbf{Sres}_j(P,p', Q,q) / cd(Q)^{p'-p}$

**3ème cas** : *les degrés de P et Q s'abaissent de 1 pour une raison commune*

Nous supposons que cd(P) et cd(Q) s'écrivent respectivement: $cd(P) = a.c_p$ et $cd(Q) = a.c_q$ avec $Sp(a) = 0$. Plus précisément nous écrivons:

$P = a.c_p\, X^p + a_{p-1}\, X^{p-1} + \ldots$ , $Q = a.c_q\, X^q + b_{q-1}\, X^{q-1} + \ldots$ et nous supposons que le déterminant $\boxed{det = c_p\, b_{q-1} - c_q\, a_{p-1}}$ se spécialise non nul.

**Proposition 10 :** Avec les hypothèses ci-dessus, et $p \geq q$

a) $Sp(\mathbf{Sres}_{q-1}(P,p,Q,q) / a) = Sp(\det . b_{q-1}^{p-q-1} . Q)$.

b) $Sp(\mathbf{Sres}_j(P,p,Q,q) / a) = (-1)^{q-j+1}.Sp(\det).\mathbf{Sres}_j(Sp(P),p-1,Sp(Q),q-1)$

pour $j < q-1$

*preuve>* Nous notons $P_1$ et $Q_1$ les polynômes P et Q tronqués de leur coefficient dominant. On a évidemment $Sp(P) = Sp(P_1)$ et $Sp(Q) = Sp(Q_1)$. Nous posons $P_2 := c_p.X^p + P_1$ , $Q_2 := c_q.X^q + Q_1$ .

Le polynôme $\mathbf{Sres}_j(P,p,Q,q)/a$ est le polynôme associé à la matrice $M_j$ dont les vecteurs lignes successifs sont $P_2.X^{q-j-1}$, $P.X^{q-j-2}$,..., P.X, P, $Q_2.X^{p-j-1}$, $Q.X^{p-j-2}$,..., Q.X, Q

*cas j= q-1 :* Après spécialisation la matrice $M_j$ est de la forme:

| | |
|---|---|
| $Sp(c_p)$ $Sp(a_{p-1})$<br>$Sp(c_q)$ $Sp(b_{q-1})$ | |
| 0 | matrice surtriangulaire dont les vecteurs lignes sont des polynômes $X^i\ Q_1$ |

D'où le résultat a)

*cas j < q-1 :* Si on regroupe en haut les 2 lignes portant $P_2.X^{q-j-1}$ et $Q_2.X^{p-j-1}$, et si on spécialise, on obtient une matrice de la forme

| | |
|---|---|
| $Sp(c_p)$ $Sp(a_{p-1})$<br>$Sp(c_q)$ $Sp(b_{q-1})$ | |
| 0 | $\mathbf{Sylv}_j(Sp(P_1),p-1,Sp(Q_1),q-1)$ |

D'où le résultat b) si on tient compte de la parité de la permutation de lignes éffectuée. ❑

**4ème cas** : *les degrés de P et Q s'abaissent de manière "incontrôlée"*

On n'obtient rien par spécialisation "directe". Néanmoins, si les divisions exactes sont nettement plus faciles dans A que dans A', on aura intérêt à poser

$Q_q := Q$ tronqué au dessus du degré de Sp(Q),

$P_p := P$ tronqué au dessus du degré de Sp(P),

à calculer les sous-pgcd de $P_p$ et $Q_q$, et spécialiser pour terminer.

## 5) Algorithmes de calcul de polynômes sous-résultants

Nous présentons 5 algorithmes de calculs des sous-pgcd de 2 polynômes. Les relations entre ces algorithmes seront discutées à la fin du paragraphe

**Algorithme n°1** (1)

Nous supposons $d(P) = p = n+1$ , $d(Q) = q \leq n$.

Nous posons $S_p := P$ , $S_n := Q$ , et $S_j := \mathbf{Sres}_j(P,n+1, Q,n)$ pour $j < n$.

Les $S_j$ peuvent alors être calculés par la méthode suivante (utilisant uniquement des calculs de pseudo-restes et des divisions exactes):

**entrées** : les polynômes P et Q

**sortie** : la suite des sous-pgcd $S_j$ $\quad (0 \leq j \leq n-1)$

**initialisation** :

| | | | |
|---|---|---|---|
| – si $q = n$ | $S_{q-1} := \mathbf{Prst}(P,Q)$ | (0) | |
| – si $q < n$ | $S_q := (cd(P)\, cd(Q))^{n-q}\, Q$ | (1) | |
| | $S_{q-1} := (-cd(P))^{n-q} . \mathbf{Prst}(P,Q)$ | (2) | |
| | en outre si $q < n-1$ et $q < k < n$ : $S_k := 0$ | (3) | |
| – | $j := q-1$ | | |

**étape suivante** : {$1 \leq j \leq q-1$ , $S_{j+1}$ et $S_j$ sont supposés déjà calculés, avec $d(S_{j+1}) = j+1$ et $s = d(S_j)$. On va calculer les $S_k$ manquants jusqu'à $S_{s-1}$ }

| | | | |
|---|---|---|---|
| – | $s := d(S_j)$ | | |
| – si $j = s$ | $S_{s-1} := \mathbf{Prst}(S_{j+1}, S_j) / cd(S_{j+1})^2$ | (4) | |
| – si $s < j$ | $S_s := S_j . cd(S_j)^{j-s} / cd(S_{j+1})^{j-s}$ | (5) | (*) |
| | $S_{s-1} := \mathbf{Prst}(S_{j+1}, S_j) / (-cd(S_{j+1}))^{j-s+2}$ | (6) | (*) |
| | en outre si $s < j-1$ et $s < k < j$ : $S_k := 0$ | (7) | |
| – | $j := s-1$ | | |

**fin** : l'algorithme se termine lorsqu'on a calculé $S_0$ cad lorsque $j \leq 0$

(*) (5) n'est pas exécuté si $s = -1$ (6) n'est pas exécuté si $s \leq 0$

*preuve>*

(0) par la définition du pseudo-reste

(1) par la prop 2 b)

(2) par la prop 2 d)

(3) par la prop 2 a)

(4) par le th de Habicht puisque $cd(S_{j+1}) = cf_{j+1}(S_{j+1})$

(5) le th de Habicht nous donne : $cd(S_{j+1})^{2(j-s)} . S_s = \mathbf{Sres}_s(S_{j+1}, j+1, S_j, j)$ , et la prop 2b : $\mathbf{Sres}_s(S_{j+1}, j+1, S_j, j) = (cd(S_j).cd(S_{j+1}))^{j-s} S_j$

(6) le th de Habicht nous donne : $cd(S_{j+1})^{2(j-s+1)} . S_{s-1} = \mathbf{Sres}_{s-1}(S_{j+1}, j+1, S_j, j)$ , et la

---

[1] Cet algorithme calcule les sous-pgcd $\mathbf{Sres}_j(P,n+1, Q,n)$ lorsque $d(P) = n+1 > d(Q)$ .

Le Subresultant Theorem p 122 de [**Loos**] semble, en première lecture, concerner ces sous-pgcd, puisque p 121, ce sont ces sous-pgcd (obtenus par spécialisation d'une suite où P et Q sont formellement de degrés n+1 et n ) qui sont considérés ... En fait le Subresultant Theorem est correct avec les $\mathbf{Sres}_j(P,n+1, Q,q)$ lorsque $n = p-1 \geq q$ , il est par contre incorrect lorsque $p \leq q$ . (Cf la dernière note bas de page au sujet de l'algorithme n°3)

prop 2d : $\mathbf{Sres}_{s-1}(S_{j+1}, j+1, S_j, j) = (-cd(S_{j+1}))^{j-s}\ \mathbf{Prst}(S_{j+1},S_j)$
(7) on applique le th de Habicht comme ci dessus et on conclut par la prop 2a. ❑

On remarque maintenant que les formules récurrentes (4) (5) (6) (7) sont homogènes. Si, en dessous d'un certain degré $k$, on sait que les $S_j$ sont tous multiples d'une constante $c$ de $\mathbf{A}$, les formules sont encore valables si on remplace les polynômes $S_j$ par les $S_j / c$. Nous en déduisons, lorsque $p = d(P)$, $q = d(Q) \le n = p - 1$, un algorithme pour calculer les sous-résultants standards $\mathbf{Sres}_j(P,p,Q,q) = \mathbf{Sres}_j(P,p,Q,n) / cd(P)^{n-q}$ (cf proposition 1 c)). On notera que l'algorithme ne diffère du précédent que lorsque $q < n$, et seulement dans la partie "initialisation". Nous n'avons réécrit *étape suivante* et *fin* qu'à cause du changement de notation ($T_j$ au lieu de $S_j$).

**Algorithme n°2 :**

**(algorithme des polynômes sous-résultants standards dans le cas d(Q) < d(P))**

Nous supposons $d(P) = p = n+1$, $d(Q) = q \le n$.

Nous posons $T_p := P$, $T_n := Q$, $T_q := cd(Q)^{n-q}.Q$, et $T_j := \mathbf{Sres}_j(P,p,Q,q)$ pour $j < q$. Les $T_j$ peuvent alors être calculés par la méthode suivante (utilisant uniquement des calculs de pseudo-restes et des divisions exactes):

**entrées** : les polynômes $P$ et $Q$,

**sortie** : la suite des sous-pgcds standards $T_j$ $\quad (0 \le j \le q)$

**initialisation** :

| | | | |
|---|---|---|---|
| – | $p := d(P)$ , $q := d(Q)$ , $n := p-1$ | | |
| – | $T_q := cd(Q)^{n-q}\ Q$ | (1) | |
| – | $T_{q-1} := (-1)^{n-q}\ .\ \mathbf{Prst}(P,Q)$ | (2) | |
| – | $j := q-1$ | | |

**étape suivante** : { $1 \le j \le q-1$, $T_{j+1}$ et $T_j$ sont supposés déjà calculés, avec $j+1 = d(T_{j+1})$ et $s = d(T_j)$. On va calculer les $T_k$ manquants jusqu'à $T_{s-1}$ }

| | | | | |
|---|---|---|---|---|
| – | $s := d(T_j)$ | | | |
| – si $j = s$ | $T_{s-1} := \mathbf{Prst}(T_{j+1}, T_j) / cd(T_{j+1})^2$ | (4) | | |
| – si $s < j$ | $T_s := T_j.\ cd(T_j)^{j-s} / cd(T_{j+1})^{j-s}$ | (5) | (*) | |
| | $T_{s-1} := \mathbf{Prst}(T_{j+1}, T_j) / (-cd(T_{j+1}))^{j-s+2}$ | (6) | (*) | |
| | en outre si $s < j-1$ et $s < k < j$ : $T_k := 0$ | (7) | | |
| – | $j := s-1$ | | | |

**fin** : l'algorithme se termine lorsqu'on a calculé $T_0$ cad lorsque $j \le 0$

(*) (5) n'est pas exécuté si $s = -1$ (6) n'est pas exécuté si $s \le 0$

**Remarques:**

1) si $j = s$ l'affectation (5) donnerait $T_j := T_j$. Et l'affectation (6) produirait le même effet que la (4)

2) on peut essayer de faire rentrer les affectations (1) et (2) dans le moule: (5) et (6). C'est possible en prenant $j = n$, $s = q$, et en faisant l'affectation $cd(T_{n+1}) := 1$ (qui est "fausse"). Avec cette philosophie, la suite des sous-pgcd commence à $T_{n+1} = P$ et il faut poser $T_k := 0$ si $q < k < p-1$. L'avantage est que les seules initialisations sont : $T_{n+1} := P$, $T_n := Q$, "$cd(T_{n+1}) := 1$". Et on passe directement à "étape suivante". Aussi ferons nous désormais la convention suivante:

**Définition (convention) :** Si $p \geq d(P)$ , $q = d(Q)$ et $p > q$ , on pose:

$\mathbf{Sres}_p(P,p,Q,q) := P$ , $\mathbf{Sres}_{p-1}(P,p,Q,q) := Q$ ,

$\mathbf{Sres}_q(P,p,Q,q) := cd(Q)^{p-1-q} . Q$ ,

$\mathbf{Sres}_k(P,p,Q,q) := 0$ si $q < k < p-1$

$\mathbf{sr}_p(P,p,Q,q) := 1$ , $\mathbf{sr}_j(P,p,Q,q) := \mathbf{cf}_j(\mathbf{Sres}_j(P,p,Q,q))$ si $j < p$

On notera qu'avec cette convention, de nombreux "cas distincts" dans les propositions établies précédemment "fusionnent" :

- prop 2 : a et b sont des cas particuliers de c ( avec $j < p-1$ )
- prop 5 : a est un cas particulier de b ( avec $j < q$ )
- théorème de Habicht : définition "uniforme" pour les $S_j$ et les C(j)
- prop 6 : a1 est un cas particulier de a2 , b1 et b2 sont des cas particuliers de b3
- prop 10: a est un cas particulier de b

En outre remarquons que

- la prop 1 c1 reste vraie dans les cas $j = q = d(Q) < p$ et $d(Q) = q < j < \inf(q', p-1)$
- la prop 1 c2 reste vraie dans le cas $j = p < q$ mais serait fausse pour $p < j = q < q'$ ou $p < j = q - 1 < q < q'$ [1]
- la proposition 9 c reste vraie pour $j = q$
- la proposition 9 d reste vraie pour $j = p < q \leq q'$

Par ailleurs, quoique bien utile, cette convention présente le défaut de n'être pas symétrique en P et Q , et il semble impossible de remédier à ce défaut lorsque $p = q$ .

Nous donnons maintenant une généralisation de l'algorithme précédent, conformément à la définition-convention ci-dessus.

## Algorithme n°3

**(algorithme généralisé des polynômes sous-résultants [2] )**

Nous supposons $p \geq d(P)$ , $q = d(Q)$ et $p > q$

Nous posons pour $j \leq p$ : $T_j := \mathbf{Sres}_j(P,p,Q,q)$ , $t_j := \mathbf{sr}_j(P,p,Q,q)$ (cf convention avant l'algorithme pour $q \leq j \leq p$ )

Les $T_j$ peuvent alors être calculés par la méthode suivante (utilisant uniquement des calculs de pseudo-restes et des divisions exactes):

**entrées** : les polynômes P et Q , l'entier $p \geq \sup(d(P), 1+d(Q))$

**sortie** : la suite des sous-pgcd standards $T_j$ ( $0 \leq j \leq p$ )

---

[1] Ainsi la convention concernant $\mathbf{Sres}_q(P,p,Q,q)$ pour $q = d(Q) < p$ tient correctement la route par rapport aux égalités générales données dans la prop 1 . Il en va de même avec les sous-pgcds identiquement nuls pour $q < j < p-1$ . La lecture des propositions 1c et 2 a) et b) pouvait d'ailleurs inciter à poser ces conventions au tout début de l'article. Il en va tout différemment en ce qui concerne la convention $\mathbf{Sres}_{p-1}(P,p,Q,q) = Q$ . Supposons en effet $p = d(P)$ , $q = d(Q)$ , $p > q+1$ : l'égalité $\mathbf{Sres}_{p-1}(P,p,Q,p) = cd(P)\ Q$ inciterait à poser, vue la proposition 1 c1) , $\mathbf{Sres}_{p-1}(P,p,Q,q) := Q/cd(P)^{p-q-1}$ tandis que l'égalité $\mathbf{Sres}_{p-1}(P,p+1,Q,q) = 0$ inciterait à poser, elle, vue 1 c2) , $\mathbf{Sres}_{p-1}(P,p,Q,q) := 0$ . La même critique vaut pour la convention concernant le sous-pgcd $\mathbf{Sres}_p(P,p,Q,q)$ .

[2] Le Subresultant Chain Algorithm dans [**Loos**] est celui-ci lorsque $p = d(P) > d(Q)$

**initialisation :**

- $T_p := P \; ; \quad t_p := 1$ (1)
- $T_{p-1} := Q$ (2)
- $j := p-1$

**étape suivante :** { $1 \le j \le n$, $T_{j+1}$, $t_{j+1}$ et $T_j$ sont supposés déjà calculés, $T_{j+1}$, $t_{j+1}$ non nuls, avec $s = d(T_j)$. On va calculer les $T_k$ manquants jusqu'à $T_{s-1}$ }

- $s := d(T_j)$
- si $j = s$ $\quad T_{s-1} := \mathbf{Sres}_{s-1}(T_j, s, T_{j+1}, j+1) \,/\, t_{j+1}^2$; (4)
- si $s < j$ $\quad T_s := T_j \cdot \mathbf{cf}_s(T_j)^{j-s} \,/\, t_{j+1}^{j-s}$ ; (5) (*)

  $T_{s-1} := \mathbf{Sres}_{s-1}(T_j, s, T_{j+1}, j+1) \,/\, (-t_{j+1})^{j-s+2}$ (6) (*)

  en outre si $s < j-1$ et $s < k < j$ : $T_k := 0$ (7)
- $j := s-1$ ; $\quad t_{j+1} := \mathbf{cf}_{j+1}(T_{j+1})$ [1]

**fin :** l'algorithme se termine lorsqu'on a calculé $T_0$ cad lorsque $j \le 0$

(*) (5) n'est pas exécuté si $s = -1$ (6) n'est pas exécuté si $s \le 0$

**NB:** On notera que dans (4) et (6) on peut toujours remplacer $\mathbf{Sres}_{s-1}(T_j, s, T_{j+1}, j+1)$ par $\mathbf{Prst}(T_{j+1}, T_j)$ sauf lors du premier passage[2] si $d(P) < p$. En outre, on a toujours l'égalité:

$$\mathbf{Sres}_{s-1}(T_j, s, T_{j+1}, j+1) = \mathbf{cf}_s(T_j)^{j-s+2}\, \mathbf{Rst}(T_{j+1}, T_j)$$

*preuve>* Tout d'abord, si $p = d(P) > q = d(Q)$, il s'agit d'une simple reformulation de l'algorithme précédent, tenu compte de la définition-convention. Ensuite on remarque que, pour une suite donnée de degrés dans la suite des restes (à partir de Q), l'algorithme écrit sous cette forme généralisée peut être vu comme une suite d'identités algébriques sous conditions (les conditions sont celles qui forcent les égalités $d(T_j) = s$, cad l'annulation de certains déterminants extraits de la matrice $\mathbf{Sylv}_0(P,p,Q,q)$). Il est donc encore valable si $d(P) < p$ (le coefficient dominant de P n'intervient pas dans l'algorithme). ❑

Nous donnons maintenant un algorithme pour les sous-pgcd standards dans le cas où $d(P) = d(Q)$. Il ne diffère de celui donné pour le cas $d(P) > d(Q)$ que dans la partie "initialisation".

**Algorithme n°4 :**

**(algorithme des polynômes sous-résultants standards dans le cas $d(Q) = d(P)$ )**

Nous supposons $d(P) = q$, $d(Q) = q$.

Nous posons $T_j := \mathbf{Sres}_j(P,q,Q,q)$ pour $j < q$ .

Les $T_j$ peuvent alors être calculés par la méthode suivante (utilisant uniquement des calculs de pseudo-restes et des divisions exactes):

**entrées** : les polynômes P et Q

**sortie** : la suite des sous-pgcd standards $T_j$ $\quad (0 \le j < q)$

---

[1] Cette affectation peut aussi s'écrire : $t_s := \mathbf{cf}_s(T_s)$

[2] Dans le Subresultant Theorem et le Subresultant Chain Algorithm de [**Loos**], c'est toujours $\mathbf{Prst}(T_{j+1}, T_j)$ qui intervient. Or ce polynôme n'est défini que pour $d(T_{j+1}) \ge d(T_j)$. En conséquence le Subresultant Theorem et le Subresultant Chain Algorithm sont "illisibles" pour $d(P) < d(Q)$ et incorrects pour $d(P) = d(Q)$. On notera également que l'on ne trouve pas dans [**Loos**] de définition explicite des $\mathbf{Sres}_j(P,n+1,Q,q)$ lorsque $n > j \ge q$ .

**initialisation :**

– $T_{q-1} := \mathbf{Prst}(Q,P)$ , $t := d(T_{q-1})$ (1)

– si $t = q-1$ $T_{q-2} := \mathbf{Prst}(P, T_{q-1}) / cd(P)$ (2)

– si $t < q-1$ on calcule $T_t$ et $T_{t-1}$ comme suit

$T_t := cd(T_{q-1})^{q-1-t} . T_{q-1}$ (1 bis)

$T_{t-1} := (-1)^{q-1-t} . \mathbf{Prst}(P, T_{q-1}) / cd(P)$ (2 bis)

en outre si $t < q-2$ et $t < k < q-1$ : $T_k := 0$ (3)

– $j := t-1$

**étape suivante** et **fin** : comme dans l'algorithme n°2

*preuve>* On pose pour $j < q$ : $S_j := \mathbf{Sres}_j(Q,q+1, P,q)$. Par la prop 1 on obtient : $S_j = cd(P).T_j$ .

(1) par l'égalité définissant le pseudo-reste

(2) Par le th de Habicht on a: $S_{q-2} = \mathbf{Prst}(P,S_{q-1})/cd(P)^2$ .

Par ailleurs $\mathbf{Prst}(P,T_{q-1}) = \mathbf{Prst}(P,S_{q-1})/cd(P)^2$ parce que $T_{q-1} = S_{q-1}/cd(P)$

(1 bis) On remplace dans le th de Habicht j par q–1 et h par t , on obtient:

$cd(P)^{2(q-1-t)} . S_t = \mathbf{Sres}_t(P,q, S_{q-1},q-1)$.

Par ailleurs $\mathbf{Sres}_t(P,q, S_{q-1},q-1) = \mathbf{Sres}_t(P,q, T_{q-1},q-1).cd(P)^{q-t}$ (il y a q–t lignes "portant $T_{q-1}$" dans la matrice correspondante). La prop 2b donne de plus:

$\mathbf{Sres}_t(P,q, T_{q-1},q-1) = (cd(P).cd(T_{q-1}))^{q-1-t}. T_{q-1}$ . Fin du calcul: élémentaire.

(2 bis) Par le th de Habicht on a: $cd(P)^{2(q-t)} . S_{t-1} = \mathbf{Sres}_{t-1}(P,q, S_{q-1},q-1)$

Par ailleurs $\mathbf{Sres}_{t-1}(P,q, S_{q-1},q-1) = \mathbf{Sres}_{t-1}(P,q, T_{q-1},q-1).cd(P)^{q-t+1}$ (même argument que ci-dessus). La prop 2d donne de plus:

$\mathbf{Sres}_{t-1}(P,q, T_{q-1},q-1) = (-cd(P))^{q-t-1} \mathbf{Prst}(P, T_{q-1})$ . Fin du calcul: élémentaire.

*étape suivante et fin :* (4) , (5) , (6) , (7) : vu le caractère homogène de ces formules, elles se déduisent des formules analogues pour les $S_j$ , obtenues par l'algorithme généralisé des polynômes sous-résultants. ❑

Nous présentons enfin un algorithme qui constitue une amélioration de l'algorithme n°2 lorsque cd(P) et cd(Q) sont divisibles par un même élément c de **A** .

**Algorithme n°5 :**

**( cas d(Q) < d(P) , cd(P) et cd(Q) divisibles par un même élément c )**

Nous supposons $d(P) = p = n+1$ , $d(Q) = q \le n$ , cd(P) et cd(Q) divisibles par un même élément c de **A** : $cd(P) = c.\gamma$ , $cd(Q) = c.\chi$

Nous posons $T_j := \mathbf{Sres}_j(P,p, Q,q)/c$ pour $j < n$.

**entrées** : les polynômes P et Q

**sortie** : la suite des $T_j$ définis ci-dessus ( $0 \le j \le n-1$)

**initialisation :**

**1er cas : d(Q) + 1 < d(P)** (cad $p > q+1$)

– $T_q := \chi^{n-q} . c^{n-q-1} . Q$ (1)

– $T_{q-1} := (-1)^{n-q} . \mathbf{Prst}(P,Q) / c$ (2)

– si $q < p-2$ et $q < k < p-1$ : $T_k := 0$ (3)

– $j := q-1$

**2ème cas : d(Q) + 1 = d(P)** (cad $p = q+1$)

- $T_q := Q$
- $T_{q-1} := \mathbf{Prst}(P,Q) / c$ (1)
- $t := d(T_{q-1})$
- si $t = q-1$ $T_{q-2} := \mathbf{Prst}(Q, T_{q-1}) / (c.\chi^2)$ (2)
- si $t < q-1$ on calcule $T_t$ et $T_{t-1}$ comme suit

  $T_t := cd(T_{q-1})^{q-1-t} . T_{q-1} / \chi^{q-1-t}$ (1 bis)

  $T_{t-1} := (-1)^{p-t} . \mathbf{Prst}(Q, T_{q-1}) / (c.\chi^{p-t})$ (2 bis)

  en outre si $t < q-2$ et $t < k < q-1$ : $T_k := 0$ (3)
- $j := t-1$

**étape suivante** et **fin** : comme dans l'algorithme n°2

*preuve*> On pose pour $j < q$ : $S_j := \mathbf{Sres}_j(P,p, Q,q)$. On a $S_j = c.T_j$ pour $j < p-1$. On regarde comment se modifie l'algorithme n°2 lorsqu'il traite les $T_j$ au lieu des $S_j$, ce qui donne l'initialisation. Dès qu'on a obtenu $T_{j+1}$ de degré $j+1 < p-1$ et $T_j$ on peut se brancher sur "étape suivante" en raison du caractère homogène des formules ❑

### *Conclusions*

*Comparaison des différents algorithmes proposés*

Les algorithmes n°2 et 4 sont ceux qu'on utilisera en pratique pour calculer les sous-pgcd: en effet les sous-pgcd non standards sont des multiples des sous-pgcd standards. Ils occupent en général plus de place et occasionnent des calculs plus longs. Notons cependant que l'algorithme n°3 avec $p = d(P) > q = d(Q)$ peut remplacer le n°2 (il effectue exactement les mêmes calculs) et est plus facile à écrire. L'algorithme n°5 est une amélioration de l'algorithme n°2, pour les 2 raisons suivantes : primo, si le facteur c qu'on connait dans cd(P) et cd(Q) est "grand" (du point de vue de la taille occupée), les calculs seront a priori plus rapides avec les coefficients divisés par ce facteur commun; secundo, si le facteur c s'annule par spécialisation, la suite calculée par l'algorithme n°5 peut s'avérer utile tandis que celle calculée par l'algorithme n°2 serait inutilisable (cf la proposition 10 )

L'algorithme n°1 est une sorte d'algorithme intermédiaire qui permet de démontrer facilement les algorithmes n°2 et 4 à partir du théorème de Habicht.

L'algorithme n°3 est sans doute celui qui éclaire le mieux la question des sous-pgcd : les polynômes P et Q font partie ici de la suite des sous-pgcd de manière tout à fait naturelle, et l'étape d'initialisation est réduite au strict minimum. Il n'est donc pas étonnant de voir dans la suite certaines démonstrations (celle du théorème 4 § 6 par exemple) reposer sur la correction de cet algorithme n°3.

*Philosophie des sous-pgcd standards*

La philosophie des sous-pgcd standards est que ce sont sans doute les meilleurs polynômes à coefficients dans **A** possibles pour remplacer les polynômes de la suite des restes (qui a priori sont seulement à coefficients dans **K** ). Il ne pourrait alors pas y avoir d'algorithme plus performant que les n°2 et 4 dans la mesure où les divisions exactes qui y sont opérées seraient toujours les meilleures possibles. Ceci sert de base à la *question ouverte* suivante:

*Considérons une liste d'entiers* $[d_1, \dots, d_k]$ *vérifiant* $d_1 \geq d_2 > d_3 > \dots > d_k$ *qui sera la liste des degrés d'une suite des restes.*
*Soit* $A = \mathbb{Z}[a_0, \dots, a_{d_1}, b_0, \dots, b_{d_2}]$ *où les* $a_i$ *et les* $b_j$ *sont des variables indépendantes. Soient* P *et* Q *les polynômes de* A[X] *ayant respectivement pour coefficients les* $a_i$ *et les* $b_j$. *Soit* $\mathfrak{I}$ *l'idéal de* A *engendré par les déterminants dont les annulations forcent les degrés dans la suite des restes de* P *et* Q *à être les degrés prévus dans la liste* $[d_1, \dots, d_k]$. *Soit* B *l'anneau* $A/\mathfrak{I}$.
*Question ouverte : Alors dans l'anneau* B[X] *, les sous-pgcd standards* $\mathbf{Sres}_{d_i-1}(P,d_1, Q,d_2)$ *sont-ils irréductibles ?*

*Théorème des sous-résultants*

Nous pouvons déduire de la correction de l'algorithme n°3 la version suivante du théorème des sous-résultants :

**Théorème des sous-résultants**

Soient P et Q des polynômes à coefficient dans un anneau commutatif.
Nous supposons $p \geq d(P)$ , $q = d(Q)$ et $p > q$
Nous posons pour $j \leq p$ : $T_j := \mathbf{Sres}_j(P,p, Q,q)$ , $t_j := \mathbf{sr}_j(P,p, Q,q)$ ;
$d_p := p$ , et pour $j < p$ , $d_j := d(T_j)$ .
Pour tout $j < p$ tel que $d_{j+1} = j+1$, on a, en posant $s := d_j$

- si $j = s$ $\quad t_{j+1}^{\,2}\, T_{s-1} = \mathbf{Sres}_{s-1}(T_j, s, T_{j+1}, j+1)$ ;
- si $s < j$ $\quad t_{j+1}^{\,j-s}\, T_s = T_j \cdot \mathbf{cf}_s(T_j)^{j-s}$ ;
  $\quad (-t_{j+1})^{j-s+2}\, T_{s-1} = \mathbf{Sres}_{s-1}(T_j, s, T_{j+1}, j+1)$
  en outre si $s < j-1$ et $s < k < j$ : $T_k = 0$

Lorsque l'anneau est intègre, on a de plus dans les égalités ci-dessus:
$\mathbf{Sres}_{s-1}(T_j, s, T_{j+1}, j+1) = \mathbf{cf}_s(T_j)^{j-s+2}\, \mathbf{Rst}(T_{j+1}, T_j)$
et, sauf dans le cas où $j = p-1$ et $d(P) < p$ :
$\mathbf{Sres}_{s-1}(T_j, s, T_{j+1}, j+1) = \mathbf{Prst}(T_{j+1}, T_j)$

## 6) Nombre de changements de signes dans la suite des restes signés

***Vrais signes des restes***

Une suite de sous-pgcd est beaucoup plus facile à calculer que la suite des restes.

Si on désire vraiment avoir les restes sans facteur multiplicatif, le mieux sera en général de calculer la suite des sous-pgcd puis de retrouver les restes en utilisant le théorème 3

Lorsque le corps des fractions de A est un corps ordonné, on a souvent besoin de calculer, à défaut des restes eux-mêmes, des polynômes qui leur sont proportionnels dans des rapports positifs. Nous supposons connaître une suite de polynômes $V_q, V_{q-1}, \dots V_0$ proportionnels, dans des rapports positifs , aux sous-pgcd standards. Cette suite peut avoir été obtenue par une spécialisation, et non directement par l'algorithme donné ci-dessus. Nous

pouvons alors calculer, de proche en proche, une suite de signes sn(q), sn(q−1), ... , sn(0), telle que sn(j).$V_j$ soit proportionnel dans un rapport $>0$ à $\mathbf{Rst}_j$ , sauf si $V_j = 0$.

Pour cela, nous avons 2 méthodes possibles:

a) nous référer au théorème 3 ou à son corollaire

b) suivre pas à pas l'algorithme des sous-pgcd standards, sans l'exécuter, et en déduire les sn(j) de proche en proche.

Dans les 2 cas, nous utiliserons, pour calculer sn(j) : les signes sn(k) pour $k > j$, les signes des coefficients dominants de P et des $T_k$ pour $k \geq j$, les degrés des $T_k$ pour $k \geq j$.

### *Nombre de changements de signes dans la suite des restes signés*

Nous supposons le corps des fractions de **A** muni d'un ordre.

Le nombre de changements de signes dans la suite des restes (convenablement modifiée) intervient dans le théorème de Sturm et dans des généralisations du théorème de Sturm[1] . Il s'avère en fait que la suite des sous-pgcd fait aussi bien l'affaire que la suite des restes pour calculer, à une constante près , le nombre de changements de signes. Ceci peut se déduire de résultats de Habicht, comme l'a montré Laureano Gonzalez (cf [**Gon**]). Nous en donnons ici une preuve "directe".

**Notations et définitions**

Introduisons tout d'abord quelques notations pour nous placer dans les conditions d'application du Théorème de Sturm (ou de ses généralisations).

*La suite des restes signés de P et Q*

Etant donnés 2 polynômes P et Q nous appellerons:

**suite des restes signés de P et Q**

la suite des restes de l'algorithme d'Euclide (démarrant avec P et Q ) avec des modifications de signes convenables comme suit :

$$\boxed{\mathbf{Rss}^m(P,Q) := (-1)^{\frac{m(m-1)}{2}} \mathbf{Rst}^m(P,Q)}$$

de sorte qu'on ait la relation de récurrence :

$$\mathbf{Rss}^{m+1}(P,Q) = -\,\mathbf{Rst}(\,\mathbf{Rss}^{m-1}(P,Q)\,,\,\mathbf{Rss}^m(P,Q)\,).$$

En posant $t = \sup(d(P),d(Q)+1)$, nous notons

$\mathbf{Rss}_t(P,Q) := P$ , $\mathbf{Rss}_{t-1}(P,Q) := Q$ et

(pour $-1 < j < t-1$ ) $\mathbf{Rss}_j(P,Q)$ le reste signé de plus fort degré inférieur ou égal à j dans la suite des $\mathbf{Rss}^m(P,Q)$ avec $m \geq 1$.

*La suite de Habicht de P et Q*

Nous donnons par ailleurs des définitions et notations analogues, concernant cette fois-ci la suite des sous-pgcd, avec les signes modifiés de façon convenable. Nous définissons la :

**suite de Habicht de P et Q**

par: $q := d(Q)$, $t := \sup(q+1,d(P))$

$$\boxed{\mathbf{Ha}_{t-m}(P,Q) := (-1)^{\frac{m(m-1)}{2}} \mathbf{Sres}_{t-m}(P,t,Q,q)} \qquad (0 \leq m \leq t)$$

(pour $t \geq t-m \geq q$ nous faisons appel à la convention donnée au § 5 )

---

[1] Une généralisation importante du théorème de Sturm est obtenue par Sylvester ([**Syl**]) en 1853 (cf également [**G-L-R-R**]). Une généralisation similaire est utilisée par Tarski ([**Tar**]) pour les problèmes de décision concernant les systèmes d'équations et inéquations dans un corps réel clos.

Notez que les $\mathbf{Sres}_j(P,t,Q,q)$ ci-dessus peuvent être calculés par l'algorithme généralisé des polynômes sous-résultants.

*Le nombre de changements de signes*

**Définition:**

Soient **K** un corps ordonné, $a \in \mathbf{K} \cup \{+\infty\} \cup \{-\infty\}$, $f := [f_0,f_1,...,f_n]$ une liste de polynômes de **K**[X]. On note $\mathbf{Ncs}(f_0,f_1,...,f_n ; a)$ ou $\mathbf{Ncs}(f ; a)$ le nombre entier défini comme suit :

- on extrait tout d'abord la suite $[g_0,g_1,...,g_m] = [f_{j0},f_{j1},...,f_{jm}]$ formée des polynômes non identiquement nuls
- on compte ensuite le nombre de changements de signes dans la suite $[g_0(a),g_1(a),...,g_m(a)]$ *en adoptant les conventions suivantes concernant les 0 :*
  - * comptent pour 1 changement de signe les segments suivants
    $-,0,+$ ou $+,0,-$ ou $+,0,0,-$ ou $-,0,0,+$
  - * comptent pour 2 changements de signe les segments suivants
    $+,0,0,+$ ou $-,0,0,-$

On définit enfin : $\mathbf{Ncs}(f ; a,b) := \mathbf{Ncs}(f ; a) - \mathbf{Ncs}(f ; b)$

Le nombre $\mathbf{Ncs}(f_0,f_1,...,f_n ; a)$ reste donc non défini pour des suites comportant des segments avec des 0 non couverts par la convention ci-dessus. Il est cependant clair qu'il est défini lorsque $f_0,f_1,...,f_n$ est une suite de restes signés (un 0 est toujours isolé et entouré de 2 signes opposés) et nous verrons qu'il est défini pour une suite de Habicht (les 0 isolés sont entourés de 2 signes opposés, et il n'y a pas de 0 triples)

Soient a, b sont des éléments de $\mathbf{K} \cup \{+\infty\} \cup \{-\infty\}$, P et Q des polynômes de **K**[X], $q := d(Q)$, $t := \sup(q+1,d(P))$

on note
$$\mathbf{W_{Rss}}(P,Q;a) := \mathbf{Ncs}([\mathbf{Rss}^m(P,Q)]_{m=0,1,...} ; a)$$
$$\mathbf{W_{Rss}}(P,Q;a,b) := \mathbf{W_{Rss}}(P,Q;a) - \mathbf{W_{Rss}}(P,Q;b)$$

et
$$\mathbf{W_{Ha}}(P,Q;a) := \mathbf{Ncs}([\mathbf{Ha}_j(P,Q)]_{j=t,t-1,...,0} ; a)$$
$$\mathbf{W_{Ha}}(P,Q;a,b) := \mathbf{W_{Ha}}(P,Q;a) - \mathbf{W_{Ha}}(P,Q;b)$$

On notera que pour une suite de restes signés, en un point a non racine de PGCD(P,Q), le nombre de changements de signes s'obtient simplement en ne tenant pas compte des 0 éventuels, ceci parce qu'il n'y a jamais deux 0 consécutifs, et que tout 0 est entouré par 2 signes opposés, vue la relation de récurrence entre les $\mathbf{Rss}^m(P,Q)$.

***Le théorème important***

La suite de Habicht est la *version formelle* de la suite des restes signés. Dans le cas "ordinaire", où les degrés descendent de 1 en 1 dans la suite des restes, les deux suites sont formées des mêmes polynômes, à des facteurs carrés près (cf Théorème 2). Dans les autres cas, la suite des restes signés de P et Q possède moins de termes que la suite de Habicht[1].

---

[1] Si on note $\mathbf{W_{Rss'}}(P,Q;a,b) := \mathbf{Ncs}([\mathbf{Rss}_j(P,Q)]_{j=t,t-1,...,0} ; a,b)$ il est clair qu'on obtient $\mathbf{W_{Rss'}}(P,Q;a,b) = \mathbf{W_{Rss}}(P,Q;a,b)$, mais les facteurs de proportionnalité entre $\mathbf{Rss}_j(P,Q)$ et $\mathbf{Ha}_j(P,Q)$ ne sont pas en général des carrés

La suite de Habicht est beaucoup plus facile à calculer que la suite des restes signés. Aussi le théorème qui suit s'avère fort utile.

**Théorème 4 :**

Le théorème de Sturm (et ses généralisations) est encore valable si on remplace la suite des restes signés par la suite de Habicht . Précisément :

(i) Si a et b sont deux points non racines de Pgcd(P,Q) , on a:

$$\mathbf{W_{Rss}}(P,Q;a,b) = \mathbf{W_{Ha}}(P,Q;a,b)$$

(ii) Plus généralement, si $v \geq \sup(d(P),d(Q)+1)$ , et a et b sont deux points non racines de Pgcd(P,Q) , on a :

$$\mathbf{W_{Rss}}(P,Q;a,b) = \mathbf{Ncs}(\,[(-1)^{\frac{m(m-1)}{2}}\,\mathbf{Sres_j}(P,t,Q,q)\,)]_{j=v,v-1,\ldots,0}\,;a,b\,) \quad (j+m=v)$$

Le théorème 4 (ii) se démontre exactement comme le (i). Nous raisonnons avec la suite de Habicht pour simplifier les notations. Le théorème 4 (i) est une conséquence immédiate du lemme suivant

**Lemme 1 :** Pour 2 polynômes P et Q fixés, il existe une constante c telle que : $\mathbf{W_{Ha}}(P,Q,a) = \mathbf{W_{Rss}}(P,Q,a) + c$ en tout point a non racine de PGCD(P,Q)

La preuve du lemme 1 utilise le lemme 2 suivant:

**Lemme 2 :** Si $\sup(d(P),d(Q)+1) \geq s = d(\mathbf{Rss_s}(P,Q) > 0$ , alors

$$\frac{\mathbf{Ha_{s-1}}(P,Q)}{\mathbf{Rss_{s-1}}(P,Q)} \cdot \frac{\mathbf{Ha_s}(P,Q)}{\mathbf{Rss_s}(P,Q)} \quad \text{est un carré dans } \mathbf{K}$$

preuves>

*Notations:* $t = \sup(d(P),d(Q)+1)$, $q = d(Q)$ , $T_j = \mathbf{Sres_j}(P,t,Q,q)$, $R_j = \mathbf{Rst_j}(P,Q)$, $\mathbf{Rss_j} = \mathbf{Rss_j}(P,Q)$, $\mathbf{Ha_j} = \mathbf{Ha_j}(P,Q)$, $T_j / R_j = r_j$ .

*Montrons tout d'abord que le lemme 1 résulte du lemme 2.*

Si $s = t$ ou est le degré d'un reste $\mathbf{Rss^m}$ , alors $\mathbf{Rss_s}$ et $\mathbf{Rss_{s-1}}$ sont 2 polynômes successifs dans la suite des restes signés (les $\mathbf{Rss^m}$ ). Par ailleurs, tout polynôme non identiquement nul dans la suite de Habicht est de la forme $\mathbf{Ha_s}$ ou $\mathbf{Ha_{s-1}}$ avec s comme ci-avant. D'après le lemme 2, en un point a non racine de $\mathbf{Ha_s}$ ni $\mathbf{Ha_{s-1}}$ , il y a changement de signe entre $\mathbf{Rss_s}$ et $\mathbf{Rss_{s-1}}$ si et seulement si il y a changement de signe entre $\mathbf{Ha_s}$ et $\mathbf{Ha_{s-1}}$ . Dans la suite de Habicht, s'ajoutent d'éventuels changements de signes entre $\mathbf{Ha_{s-1}}$ et $\mathbf{Ha_k}$ si $k = d(\mathbf{Ha_{s-1}}) < s-1$ : mais les 2 polynômes étant proportionnels, ce changement de signe "supplémentaire" éventuel a lieu indépendamment du point a où sont évalués les polynômes. Ceci démontre le lemme 1 dans le cas où tous les $\mathbf{Rss^j}(P,Q)(a)$ sont non nuls.

Voyons maintenant le cas où l'un des polynômes, non défectueux dans la suite de Habicht, s'annule en a : par exemple $d(\mathbf{Ha_s}) = s$, $d(\mathbf{Ha_{s-1}}) = s-1$, et $\mathbf{Ha_{s-1}}(a) = 0$ . On sait alors que $\mathbf{Rss_s}(a).\mathbf{Rss_{s-2}}(a) < 0$ , ce qui compte pour un changement de signe dans la suite des restes signés.

En outre, pour a' suffisamment proche de a et distinct de a , on a $\mathbf{W_{Rss}}(P,Q,a) = \mathbf{W_{Rss}}(P,Q,a')$ et tous les $\mathbf{Rss^j}(P,Q)(a')$ sont non nuls .

En appliquant 2 fois le lemme 2 on voit que $\frac{\mathbf{Ha_{s-2}}}{\mathbf{Rss_{s-2}}} \cdot \frac{\mathbf{Ha_s}}{\mathbf{Rss_s}}$ est un carré dans $\mathbf{K}$ , et on obtient donc également un changement de signe dans la suite de Habicht. Et pour a'

suffisamment proche de a , on a $\mathbf{W_{Ha}}(P,Q,a) = \mathbf{W_{Ha}}(P,Q,a')$.

Donc $\mathbf{W_{Ha}}(P,Q,a) = \mathbf{W_{Rss}}(P,Q,a) + c$ avec la même valeur de c en a qu'en a'.

Voyons enfin le cas où l'un des polynômes défectueux dans la suite de Habicht, s'annule en a . Soit donc j avec $\mathbf{Ha_{j+1}}$ non défectueux (de degré j+1 ), $\mathbf{Ha_j}$ défectueux de degré $h < j$ et tel que $\mathbf{Ha_j}(a) = 0$ . D'après le lemme 2

$\frac{\mathbf{Ha_j}}{\mathbf{Rss_j}} \cdot \frac{\mathbf{Ha_{j+1}}}{\mathbf{Rss_{j+1}}}$ est un carré dans $\mathbf{K}$ et

$\frac{\mathbf{Ha_h}}{\mathbf{Rss_h}} \cdot \frac{\mathbf{Ha_{h-1}}}{\mathbf{Rss_{h-1}}}$ est un carré dans $\mathbf{K}$ .

Ceci signifie que les polynômes $\mathbf{Ha_{j+1}} \cdot \mathbf{Ha_j} \cdot \mathbf{Ha_h} \cdot \mathbf{Ha_{h-1}}$ et $\mathbf{Rss_{j+1}} \cdot \mathbf{Rss_j} \cdot \mathbf{Rss_h} \cdot \mathbf{Rss_{h-1}}$ sont de même signe en tout point a' non racine de $\mathbf{Ha_j}$ . Or $\mathbf{Rss_j} = \mathbf{Rss_h}$ et $\mathbf{Rss_{j+1}}$ et $\mathbf{Rss_{h-1}}$ sont de signe opposé en a . Si on considère donc un point a' non racine de $\mathbf{Ha_j}$ et suffisamment proche de a (tel qu'il n'y ait aucune racine d'un polynôme de la suite des restes signés de P et Q entre a et a' ), le polynôme $\mathbf{Ha_{j+1}} \cdot \mathbf{Ha_j} \cdot \mathbf{Ha_h} \cdot \mathbf{Ha_{h-1}}$ est négatif en a' et le nombre des changements de signe dans la suite $\mathbf{Ha_{j+1}}$ , $\mathbf{Ha_j}$ , $\mathbf{Ha_h}$ , $\mathbf{Ha_{h-1}}$ en a' vaut 2 si $\mathbf{Ha_{j+1}} \cdot \mathbf{Ha_{h-1}} > 0$ , 1 si $\mathbf{Ha_{j+1}} \cdot \mathbf{Ha_{h-1}} < 0$ .

On a donc $\mathbf{W_{Ha}}(P,Q; a) = \mathbf{W_{Ha}}(P,Q; a')$

*Voyons maintenant la preuve du lemme 2.*

Lorsque $s = t$ , le lemme 2 est trivial :

$$P = T_s = R_s = \mathbf{Rss_s} = \mathbf{Ha_s} \quad \text{et} \quad Q = T_{s-1} = R_{s-1} = \mathbf{Rss_{s-1}} = \mathbf{Ha_{s-1}}.$$

On utilise ensuite l'algorithme généralisé des polynômes sous-résultants et on regarde comment les choses évoluent lors de "étape suivante" , lorsqu'on passe de j+1, j à s, s−1. Si on pose $c_{j+1} := \mathbf{sr_{j+1}}(P,t,Q,q)$ et $c_j := \mathbf{cd}(T_j)$, on trouve:

$$\frac{r_s}{r_{s-1}} = \frac{r_j}{r_{j+1}} \cdot \left( \frac{c_{j+1}}{c_j} \right)^2 \cdot (-1)^{j-s}$$

Comme $\mathbf{Rss_{j+1}}$ , $\mathbf{Rss_j} = \mathbf{Rss_s}$ et $\mathbf{Rss_{s-1}}$ sont 3 restes successifs on a :

$$(\mathbf{Rss_{j+1}}/R_{j+1}).(\mathbf{Rss_j}/R_j).(\mathbf{Rss_s}/R_s).(\mathbf{Rss_{s-1}}/R_{s-1}) = -1$$

Enfin, on a:

$$(\mathbf{Ha_{j+1}}/T_{j+1}).(\mathbf{Ha_j}/T_j).(\mathbf{Ha_s}/T_s).(\mathbf{Ha_{s-1}}/T_{s-1}) = (-1)^{j-s+1}$$

Ce qui montre le lemme 2 en s, s−1 s'il était vrai en j+1, j . ❑

### *Spécialisation de la suite de Habicht*

On considère 2 polynômes P et Q de $\mathbf{A}[X]$ , et un homomorphisme $Sp : \mathbf{A} \to \mathbf{A'}$.

*Notations:* On note $P_1 = Sp(P)$ , $Q_1 = Sp(Q)$ , $p = d(P)$ , $q = d(Q)$ , $p_1 = d(P_1)$ , $q_1 = d(Q_1)$ , $t = \sup(p,q+1)$ , $t_1 = \sup(p_1,q_1+1)$ .

Le problème est le suivant : *est-ce que les polynômes* $Sp(\mathbf{Ha_j}(P,Q))$ *peuvent être utilisés à la place des polynômes de la suite de Habicht de* $P_1$ *et* $Q_1$ *dans le théorème de Sturm et ses généralisations ?* (le cas du théorème de Sturm lui-même est traité en détail au § 7 )

Si $P_1$ et $Q_1$ ont mêmes degrés que P et Q , tout se passe évidemment bien puisqu'alors : $Sp(\mathbf{Ha_j}(P,Q)) = \mathbf{Ha_j}(P_1,Q_1)$

Dans les autres cas, on peut se référer aux résultats du § 4 pour calculer la suite de Habicht de $P_1$ et $Q_1$ . Néanmoins, on n'a pas en général à se fatiguer à ce point. Nous allons examiner quelques cas où les choses se passent bien : notamment lorsque un seul des 2 degrés chute par spécialisation . Dans la suite, les points a et b sont des éléments de $\mathbf{K'}$, corps des fractions de $\mathbf{A'}$ , avec $a < b$ .

**1er cas** : $p_1 = p > q > q_1$

Pour $j \leq q_1$ on obtient $Sp(\mathbf{Ha}_j(P,Q)) = cd(P_1)^{q-q_1}.\mathbf{Ha}_j(P_1,Q_1)$ . La seule différence éventuelle dans le décompte du nombre de changements de signes se situe donc au niveau du dédoublement du polynôme $Q$. Si $a$ est non racine de $Q_1$ , cette différence est constante, indépendante de $a$. Si $a$ est racine de $Q_1$ , nous comparons à ce qui se passe en un point $a'$ très proche de $a$ , et pour cela nous examinons en détail le début des 2 suites $Sp(\mathbf{Ha}_j(P,Q))$ et $\mathbf{Ha}_j(P_1,Q_1)$ , pour $j = p, p-1, ..., q_1, q_1-1$ : dans la 1ère suite, on a : $P_1, Q_1, 0, ..., 0, cd(P_1)^{q-q_1}\,\mathbf{Ha}_{q_1-1}(P_1,Q_1)$ [1], $cd(P_1)^{q-q_1}\,\mathbf{Ha}_{q_1-1}(P_1,Q_1)$ et dans la 2ème, on a : $P_1, Q_1, 0, ..., 0, \mathbf{Ha}_{q_1-1}(P_1,Q_1), \mathbf{Ha}_{q_1-1}(P_1,Q_1)$ . Dans la 2ème suite, le produit des 4 polynômes non identiquement nuls, est $< 0$ en $a'$ très proche de $a$ , (cf preuve du Théorème 4 cas défectueux), donc également dans la 1ère suite. Et le nombre de changement de signes compté en $a$ et en $a'$ sur le début de la suite est donc le même ( 2 ou 1 selon que les polynômes extrêmes ont même signe ou non), aussi bien pour la 1ère que pour la 2ème.

**NB**: Si $q > q_1$ et $q \geq p$ alors on a $Sp(\mathbf{Ha}_j(P,Q)) = 0$ pour $j < q$ .

**2ème cas** : $p > p_1$ , $q = q_1$

Si $t_1 = t$ on a $Sp(\mathbf{Ha}_j(P,Q)) = \mathbf{Ha}_j(P_1,Q_1)$ et tout est OK.

Sinon, on a quand même : $Sp(\mathbf{Ha}_j(P,Q)) = (-1)^{\frac{m(m-1)}{2}}\,\mathbf{Sres}_j(P_1,t,Q_1,q)$ $\quad (j+m = t)$

et le théorème 4 (ii) permet de conclure.

Nous récapitulons les résultats obtenus :

**Proposition 11** :

On considère 2 polynômes $P$ et $Q$ de $\mathbf{A}[X]$ , de degrés $p$ et $q$ , et un homomorphisme $Sp : \mathbf{A} \to \mathbf{A'}$. On note $P_1$ et $Q_1$ les polynômes $Sp(P)$ et $Sp(Q)$ , $p_1 = d(P_1)$ , $q_1 = d(Q_1)$ , $t = \sup(p,q+1)$, $t_1 = \sup(p_1,q_1+1)$.

On suppose que $p_1 = p$ , $q_1 = q$ ou $p_1 = p > q > q_1$ ou $p > p_1$ , $q = q_1$

Soient $a$ et $b$ des éléments de $\mathbf{K'}$ (corps des fractions de $\mathbf{A'}$) non racines de $Pgcd(P_1,Q_1)$ . Alors on a :

$$\mathbf{Ncs}([\,Sp(\,\mathbf{Ha}_j(P,Q)\,)\,]_{j=t,t-1,...,0}\,;a,b) = \mathbf{Ncs}([\,\mathbf{Ha}_j(P_1,Q_1)\,]_{j=t_1,t_1-1,...,0}\,;a,b)^2$$

et donc $\mathbf{W_{Rss}}(P_1,Q_1;a,b) = \mathbf{Ncs}(\,[Sp(\mathbf{Ha}_j(P,Q))]_{j=t,t-1,...,0}\,;a,b\,)$

**3ème cas** : *les degrés de P et Q s'abaissent de 1 pour une raison commune*

On suppose $p > q$ . On va appliquer la propostion 10. On note $c$ le facteur commun dans $cd(P)$ et $cd(Q)$ , et $\quad det := \dfrac{\mathbf{cf}_p(P).\mathbf{cf}_{q-1}(Q) - \mathbf{cf}_{p-1}(P).\mathbf{cf}_q(Q)}{c}$

Pour $j < q_1 = q-1$, $m+j = p$ , $m_1+j = p_1 = p-1$ , on obtient :

$$Sp(\mathbf{Ha}_j(P,Q)/c) = (-1)^{\theta(j)}\,.\,Sp(det)\,.\,\mathbf{Ha}_j(P_1,Q_1)$$

$$\text{avec } \theta(j) = \frac{m(m-1)}{2} + \frac{m_1(m_1-1)}{2} + q-j+1 \equiv p-q \pmod 2$$

D'où :

---

[1] $\mathbf{Ha}_{q_1-1}(P_1,Q_1) = cd(Q_1)^{p-1-q_1}\,Q_1$

[2] On remarquera que le second membre est $\mathbf{W_{Ha}}(P_1,Q_1;\,a,b)$

**Proposition 12 :**

On considère 2 polynômes P et Q de A[X], de degrés p et $q < p$, et un homomorphisme $Sp : A \to A'$. On note $P_1$ et $Q_1$ les polynômes Sp(P) et Sp(Q). On suppose cd(P) et cd(Q) divisibles par un même élément c de A, avec $Sp(c) = 0$, $p_1 = d(P_1) = p - 1$, $q_1 = d(Q_1) = q - 1$.

Pour $j < q - 1$ on obtient

$$Sp(\mathbf{Ha}_j(P,Q)/c) = (-1)^{p-q} . Sp\left(\frac{cf_p(P).cf_{q-1}(Q) - cf_{p-1}(P).cf_q(Q)}{c}\right). \mathbf{Ha}_j(P_1,Q_1)$$

***Permutation des deux polynômes de départ dans la suite de Habicht***

On considère 2 polynômes P et Q de A[X] de degrés p et $q \leq p$.

*On veut comparer les comportements de la suite de Habicht de P et Q et de celle de Q et P*. On note $t = \sup(p,q+1)$, $t' = \sup(q,p+1)$

**1er cas :** $p > q$

On a donc $t = p$, $t' = p+1$. Les indices j, m et m' sont liés par $j+m = t$, $j+m' = t'$.

*Pour* $j \leq q$, un calcul simple montre que :

$$\mathbf{Ha}_j(-Q,P) = (-1)^{m'+1}.\mathbf{Ha}_j(Q,P) = cd(P)^{p+1-q}.\mathbf{Ha}_j(P,Q)$$

*Comparons maintenant le début des deux suites de Habicht.*

Si $p = q+1$, le début de la suite $\mathbf{Ha}_j(P,Q)$ $(j = p, p-1)$ est donné par P, Q et celui de la suite $\mathbf{Ha}_j(-Q,P)$ $(j = p+1, p, p-1)$ est donné par $-Q$, P, $cd(P)^2.Q$

Si $p > q+1$, le début de la suite $\mathbf{Ha}_j(P,Q)$ $(j = p, p-1, \ldots, q+1)$ est formé par les polynômes P, Q, 0, ..., 0 ; et celui de la suite $\mathbf{Ha}_j(-Q,P)$ $(j = p+1, p, p-1, \ldots, q+1)$ est formé par les polynômes $-Q$, P, $cd(P)^2.Q$, 0, ..., 0. La différence entre le nombre de changements de signes compté dans la 1ère suite et celui compté dans la 2ème (à partir du polynôme P) se situe donc uniquement au niveau du dédoublement du polynôme Q. Or cette différence est constante pour les points a non racines de Q, et le cas des point a racines de Q se traite comme dans la spécialisation de la suite de Habicht, 1er cas.

On obtient donc

**Proposition 13 :**

On considère 2 polynômes P et Q de A[X], de degrés p et $q < p$.

On suppose que a et b ne sont pas racines de Q. Alors :

$$\mathbf{Ncs}([\mathbf{Ha}_j(P,Q)]_{j=p,p-1,\ldots,0}; a,b) = \mathbf{Ncs}([\mathbf{Ha}_j(-Q,P)]_{j=p,p-1,\ldots,0}; a,b)^1$$

**2ème cas :** $p = q$

On a donc $t = t' = p+1$. On obtient pour $j < q$ la relation

$$cd(P).\mathbf{Ha}_j(P,Q) = cd(Q).\mathbf{Ha}_j(-Q,P)$$

Mais les deux suites commencent respectivement par P, Q et $-Q$, P ce qui empêche d'avoir un résultat analogue à la proposition 13.

---

[1] On remarquera que le premier membre est $\mathbf{W_{Ha}}(P,Q; a,b)$ mais que le second n'est pas $\mathbf{W_{Ha}}(-Q,P; a,b)$ puisque la suite considérée démarre avec $j = p$ et non pas $j = p+1$

# 7) Suite de Sturm et spécialisation

Nous introduisons la notion de "suite de Sturm-Habicht d'un polynôme", indiquons l'algorithme pour la calculer et étudions les problèmes liés à la spécialisation . Nous supposons le corps des fractions de **A** muni d'un ordre. Tous les résultats sont obtenus par application immédiate de ceux des § précédents. Une référence classique pour le théorème de Sturm est **[VdW]**.

### *Notations et définitions*

Soit P un polynôme de A[X], de degré p. Rappelons que la **suite de Sturm de P** est égale à la suite des restes signés de P et P' . Nous définissons la **suite de Sturm-Habicht de P** comme étant "à peu près" la suite de Habicht de P et P' :

Nous définissons : | $\mathbf{StHa}_j(P) := \mathbf{Ha}_j(P,P')/cd(P)$ pour $j < p-1$
$\mathbf{StHa}_p(P) := P.cd(P)$, $\mathbf{StHa}_{p-1}(P) := P'.cd(P)$

Ce sont des polynômes à coefficients dans **A** .

L'intérêt essentiel de cette définition est donné par le théorème 4 du § précédent , qui donne dans le contexte actuel:

> Le théorème de Sturm , qui permet de calculer le nombre de racines de P sur un intervalle ]a,b[ (avec $P(a).P(b) \neq 0$ ) dans une clôture réelle de **K** , s'applique avec la suite de Sturm-Habicht (au lieu de la suite de Sturm) à condition d'utiliser la nouvelle convention pour le compte du nombre de changements de signes (cf § 6)

Un autre aspect interessant est le "bon" comportement par spécialisation.

### *Spécialisation de la suite de Sturm-Habicht*

**1er cas** : $d(Sp(P)) = d(P)$ : on a $Sp(\mathbf{StHa}_j(P)) = \mathbf{StHa}_j(Sp(P))$ .

**2ème cas** : $d(Sp(P)) < d(P) - 1$ : on a $Sp(\mathbf{StHa}_j(P)) = 0$ . Si on veut calculer la suite de Sturm avant spécialisation, on considère le polynôme $P_t$ := "P tronqué au delà du degré $d(Sp(P))$", on a $Sp(P) = Sp(P_t)$ et $d(P_t) = d(Sp(P_t))$. On calcule alors les $\mathbf{StHa}_j(P_t)$.

**3ème cas** : $d(Sp(P)) = d(P) - 1$ : on déduit immédiatement de la proposition 10 :

**Proposition 14** : Nous supposons $d(Sp(P)) = d(P) - 1$ , on a alors l'égalité:
$Sp(\mathbf{StHa}_j(P)) = Sp(\mathbf{cf}_{p-1}(P))^2 . \mathbf{StHa}_j(Sp(P))$ si $j < p-2$ .

*preuve*> Par application de la proposition 12 ❑

On notera que le signe obtenu après spécialisation est correct pour $j < p-2$ , mais que les 3 premiers polynômes se spécialisent mal: ils doivent être remplacés par les 2 polynômes $Sp(\mathbf{cf}_{p-1}(P).P)$ et $Sp(\mathbf{cf}_{p-1}(P).P')$

***Algorithme pour calculer la suite de Sturm-Habicht***

Nous présentons maintenant un algorithme pour calculer la suite de Sturm de P . Il est obtenu directement à partir de l'algorithme n°5.

**Algorithme pour la suite de Sturm-Habicht d'un polynôme P**

Nous supposons $d(P) = p$.

Nous posons $q := p - 1$ , $T_j := \mathbf{Sres}_j(P,p,P',q)/cd(P)$ pour $j < q$.

Les $T_j$ peuvent alors être calculés par la méthode suivante (utilisant uniquement des calculs de pseudo-restes et des divisions exactes):

**entrée** : le polynôme P

**sortie** : la suite de Sturm-Habicht de P

**initialisation** :

| | | | |
|---|---|---|---|
| – | $T_{q-1}$ | $:= \mathbf{Prst}(P,P') / cd(P)$ , $t := d(T_{q-1})$ | (1) |
| – si $t = q-1$ | $T_{q-2}$ | $:= \mathbf{Prst}(P', T_{q-1}) / (p^2.cd(P))$ | (2) |
| – si $t < q-1$ | on calcule $T_t$ et $T_{t-1}$ comme suit | | |
| | $T_t$ | $:= cd(T_{q-1})^{q-1-t} . T_{q-1} / p^{q-1-t}$ | (1 bis) |
| | $T_{t-1}$ | $:= (-1)^{p-t} . \mathbf{Prst}(P, T_{q-1}) / (p^{p-t}.cd(P))$ | (2 bis) |
| | en outre si $t < q-2$ et $t < k < q-1$ : | $T_k := 0$ | (3) |
| – | $j$ | $:= t - 1$ | |

**étape suivante** et **fin** comme l'algorithme n°2

**NB:** Pour obtenir la suite de Sturm-Habicht, il faut encore changer les signes des $T_j$ 2 fois sur 4 , et rajouter les 2 premiers polynômes.

## 8) Traitement de la matrice de Sylvester par la méthode de Bareiss

Lorsque $p > q = d(Q)$ , $p \geq d(P) \geq q$ , il est possible de réarranger les lignes de la matrice $\mathbf{Sylv}_0(P,p,Q,q)$ de manière que le traitement de la matrice obtenue par la *méthode du pivot améliorée à la Bareiss* [1] fournisse en cours de route tous les polynômes $\mathbf{Ha}_j(P,Q)$ .

Il faut noter cependant que l'algorithme des sous-pgcd standards est plus économe en nombre d'opérations arithmétiques que le traitement de $\mathbf{Hab}_0(P,p,Q,q)$ par la méthode de Bareiss: l'algorithme des sous-pgcd standards peut être vu comme une légère amélioration de la

---

[1] La méthode du pivot améliorée à la Bareiss (cf [**Bar**] ou [**Lom**]) est basée sur l'étude des valeurs des coefficients successifs obtenus lors d'une triangulation par la méthode du pivot: tout coefficient obtenu est le quotient de 2 déterminants extraits de la matrice de départ. Dans [**Gan**] tome 1 chap 2, Gantmacher, met clairement en évidence comment l'analyse détaillée de la méthode du pivot permet une démonstration simple des identités de Sylvester concernant les déterminants, identités qui garantissent la possibilité d'opérer les divisions exactes dans **A** qui interviennent dans la méthode de Bareiss. Notons qu'en 1932 , Aitken ([**Ait**]) signale "en passant" comment obtenir une triangulation entièrement dans $\mathbb{Z}$ en utilisant des divisions exactes (produit en croix divisé par le pivot précédent) ... cad par la méthode de Bareiss .

méthode de Bareiss (qui, elle, est une amélioration décisive de la méthode du pivot), due à la connaissance de la forme particulière de la matrice de Sylvester (qui est une matrice contenant beaucoup de 0 et beaucoup de coefficients égaux).

Pour comprendre le réarrangement de lignes necessaire dans la matrice de Sylvester, le mieux est de faire quelques dessins. Comme précédemment, les x représenteront les coefficients de Q et les y ceux de P, tandis que les 0 sont représentés par des points .

p = 6 q = 5

les polynômes $\mathbf{Ha}_j(P,Q)$ pour j = 4, 3, 2, 1, 0 sont les polynômes associés aux matrices suivantes:

j = 4

```
x x x x x x .
y y y y y y y
. x x x x x x
```

j = 3

```
x x x x x x . .
y y y y y y y .
. x x x x x x .
. y y y y y y y
. . x x x x x x
```

j = 2

```
x x x x x x . . .
y y y y y y y . .
. x x x x x x . .
. y y y y y y y .
. . x x x x x x .
. . y y y y y y y
. . . x x x x x x
```

j = 1

```
x x x x x x . . . .
y y y y y y y . . .
. x x x x x x . . .
. y y y y y y y . .
. . x x x x x x . .
. . y y y y y y y .
. . . x x x x x x .
. . . y y y y y y y
. . . . x x x x x x
```

j = 0

```
x x x x x x . . . . .
y y y y y y y . . . .
. x x x x x x . . . .
. y y y y y y y . . .
. . x x x x x x . . .
. . y y y y y y y . .
. . . x x x x x x . .
. . . y y y y y y y .
. . . . x x x x x x .
. . . . y y y y y y y
. . . . . x x x x x x
```

On voit que ces matrices, que nous noterons $\mathbf{Hab}_j(P,p,Q,q)$, sont toutes extraites (dans le coin supérieur gauche) de la plus grande d'entre elles, qui est la matrice de Sylvester avec un réarrangement des lignes.

Donnons les dessins pour des écarts plus grands entre p et q :

p = 7 q = 5

les polynômes $\mathbf{Ha}_j(P,Q)$ pour j = 4, 2, 0 sont les polynômes associés aux matrices suivantes:

j = 4

```
. x x x x x x .
x x x x x x . .
y y y y y y y y
. . x x x x x x
```

j = 2

```
. x x x x x x . . .
x x x x x x . . . .
y y y y y y y y . .
. . x x x x x x . .
. y y y y y y y y .
. . . x x x x x x .
. . y y y y y y y y
. . . . x x x x x x
```

j = 0

```
. x x x x x x . . . . .
x x x x x x . . . . . .
y y y y y y y y . . . .
. . x x x x x x . . . .
. y y y y y y y y . . .
. . . x x x x x x . . .
. . y y y y y y y y . .
. . . . x x x x x x . .
. . . y y y y y y y y .
. . . . . x x x x x x .
. . . . y y y y y y y y
. . . . . . x x x x x x
```

p = 8 q = 5

les polynômes $\mathbf{Ha}_j(P,Q)$ pour j = 5, 4, 2, 0 sont les polynômes associés aux matrices suivantes:

j = 5

```
. . x x x x x x
. x x x x x x .
x x x x x x . .
```

j = 4

```
. . x x x x x x .
. x x x x x x . .
x x x x x x . . .
y y y y y y y y y
. . . x x x x x x
```

j = 2

```
. . x x x x x x . . .
. x x x x x x . . . .
x x x x x x . . . . .
y y y y y y y y y . .
. . . x x x x x x . .
. y y y y y y y y y .
. . . . x x x x x x .
. . y y y y y y y y y
. . . . . x x x x x x
```

```
j = 0    . . x x x x x x . . . . .
         . x x x x x x . . . . . .
         x x x x x x . . . . . . .
         y y y y y y y y y . . . .
         . . . x x x x x x . . . .
         . y y y y y y y y y . . .
         . . . . x x x x x x . . .
         . . y y y y y y y y y . .
         . . . . . x x x x x x . .
         . . . y y y y y y y y y .
         . . . . . . x x x x x x .
         . . . . y y y y y y y y y
         . . . . . . . x x x x x x
```

De manière générale nous pouvons définir les matrices $\mathbf{Hab}_j(P,p,Q,q)$ pour $p > q \geq j$ comme suit :

– les lignes de $\mathbf{Hab}_q(P,p,Q,q)$ sont les polynômes Q, Q.X, ..., $Q.X^{p-q-1}$ sur la base $X^{p-1}$, ..., X, 1.

– la matrice $\mathbf{Hab}_{j-1}(P,p,Q,q)$ est obtenue à partir de la matrice $\mathbf{Hab}_j(P,p,Q,q)$ en rajoutant une colonne de 0 à droite, puis 2 lignes portant P et Q en dessous.

En traitant la matrice $\mathbf{Hab}_0(P,p,Q,q)$ par la méthode de Bareiss, avec des conventions convenables, on obtient en cours de route tous les polynômes $\mathbf{Ha}_j(P,Q)$. Les conventions sont les suivantes:

– si on tombe sur un coefficient nul en position de pivot, on cherche le pivot dans la colonne, ligne après ligne, et on effectue l'échange de lignes

– on conserve en mémoire la parité de la permutation éffectuée sur les lignes.

Comme le polynôme $\mathbf{Ha}_j(P,Q)$ est le polynôme associé à la matrice $\mathbf{Hab}_j(P,p,Q,q)$ on obtient, lorsque $d(\mathbf{Ha}_{j+1}(P,Q)) = j+1$ :

– les matrices $\mathbf{Hab}_{j+1}(P,p,Q,q)$ et $\mathbf{Hab}_j(P,p,Q,q)$ sont traitées "en cours de route" lors du traitement de $\mathbf{Hab}_0(P,p,Q,q)$ par la méthode de Bareiss. Plus précisément :

– les échanges de lignes correspondant au traitement jusqu'au pivot sur la ligne $p+q-2(j+1)$ sont tous "internes" à la matrice $\mathbf{Hab}_{j+1}(P,p,Q,q)$

– les échanges de lignes correspondant au traitement jusqu'au pivot sur la ligne $p+q-2j-1$ sont tous "internes" à la matrice $\mathbf{Hab}_j(P,p,Q,q)$ , et, lorsque ce pivot a été traité, on trouve sur la dernière ligne de la matrice $\mathbf{Hab}_j(P,p,Q,q)$, au signe près (donné par la parité de la permutation des lignes) le polynôme $\mathbf{Ha}_j(P,Q)$ .

# 9) Relation de Bezout complète entre plusieurs polynômes

***Position du problème***

On considère une liste de polynômes $\mathbf{P} = [P_1, P_2, ..., P_r]$ de $\mathbf{A}[X]$. On appelle **relation de Bezout complète entre ces polynômes** une matrice $\mathbf{C} \in \mathbf{M_r}(\mathbf{A}[X])$ qui vérifie : $\det(\mathbf{C}) \in \mathbf{A}-\{0\}$ , et $\mathbf{P.C} = [G, 0, ..., 0]$ .

On a alors nécessairement : G est un pgcd de la liste **P** dans **K**[X] (on rappelle qu'on suppose que **A** est un anneau intègre de corps de fraction **K** ). Les polynômes de la première colonne de **C** fournissent, eux, une relation de Bezout ordinaire entre les polynômes de la liste.

Le but du § est de montrer le théorème suivant :

**Théorème 5 :**

On suppose que les déterminants sont calculables en temps polynomial dans **A** . Alors on peut calculer en temps polynomial une relation de Bezout complète entre polynômes de **A**[X] . (les polynômes étant donnés en présentation dense).

**Questions ouvertes :**

1) Le problème connexe suivant semble actuellement ouvert : sous la même hypothèse, est-il vrai que les diviseurs d'un polynôme de **A**[X] sont polynomialement majorés en taille ? Plus précisément peut-on construire une opération calculable en temps polynomial F de $\mathbf{A}[X] \times \mathbf{A}[X]$ vers **A**[X] telle que :

si G divise P dans **K**[X] alors F(P,G) est associé à G dans **K**[X]
et sa taille est polynomialement majorée à partir de celle de P

Ce résultat constituerait une extension du lemme 4 ci-après.

2) Autre problème connexe qui semble actuellement ouvert : sous la même hypothèse, la réduction de Smith des matrices à coefficients dans **A**[X] est-elle calculable en temps polynomial ? Plus précisément, peut-on construire une opération calculable en temps polynomial qui, à partir d'une matrice **M** à coefficients dans **A**[X] (tout ceci est en présentation dense) , calcule 2 matrices carrées **C** et **L** et une liste de polynômes $\mathbf{G} = [G_1, G_2, ..., G_s]$ telles que :

- $\det(\mathbf{C})$ et $\det(\mathbf{L}) \in \mathbf{A}-\{0\}$
- **L.M.C** a des coefficients tous nuls, sauf ceux sur la diagonale qui sont égaux successivement à $G_1$, $G_1.G_2$, ..., $G_1.G_2...G_s$ .

**Notations :**

- $\mathbf{P} = [P_1, P_2, ..., P_r]$, $d_i = d(P_i)$ , $d = \sup(d_i)$ , $d' = \inf(d_i)$ ,
- G est un pgcd de **P** dans **K**[X]
- $\mathbf{Esylv^m(P)}$ est le sous-espace vectoriel de **K**[X] engendré par les polynômes $P_i.X^j$ de degré inférieur ou égal à m .
  Si $m \geq d$ , on a la relation de récurrence :
  $$\mathbf{Esylv^{m+1}(P)} = \mathbf{Esylv^m(P)} + X.\mathbf{Esylv^m(P)}$$
  Nous abrègerons $\mathbf{Esylv^m(P)}$ en $\mathbf{E^m}$
- $\mathbf{Sylv^m(P)}$ est une matrice dont les vecteurs lignes sur la base $X^m$, $X^{m-1}$, ..., $X^2$, X, 1 sont les polynômes $P_i.X^j$ de degré inférieur ou égal à m : nous préciserons l'ordre des lignes en indiquant que $\mathbf{Sylv^{j+1}(P)}$ est obtenue à partir de $\mathbf{Sylv^j(P)}$ en rajoutant une

colonne de 0 à droite, puis en dessous, des lignes portant les $P_i$ de degré $\leq m$ dans l'ordre des i croissants.

***Preuve du théorème***

Cette preuve comporte 4 lemmes préparatoires, puis la description d'un algorithme en temps polynomial qui réalise le théorème 5 .

**Lemme 1 :** Notons $G^m$ un polynôme de degré minimum dans $\mathbf{Esylv^m(P)}$. Si $m \geq d = \sup(d_i)$ on a l'équivalence:

$$\boxed{\dim(E^m) = m + 1 - d(G)} \Leftrightarrow \boxed{\dim(\mathbf{Esylv^{m+1}(P)}) = \dim(\mathbf{Esylv^m(P)}) + 1}$$

Lorsque ces conditions sont vérifiées $G^m$ est un pgcd de P

*preuve>*
*Voyons l'implication directe* :. Comme $E^m$ contient uniquement des polynômes de degré $\leq m$ et multiples de G , on a $\dim(E^m) \leq m + 1 - d(G)$ , et l'égalité implique que $E^m$ possède une base $G, G.X, \ldots, G.X^{m+1-d(G)}$.
*Voyons la réciproque.* Supposons pour simplifier que $d(P_1) = d$. Considérons une base $B^m$ de $E^m$ formée de polynômes de degrés strictement croissants. Comme $\dim(E^{m+1}) = \dim(E^m) + 1$, et que $X^{m+1-d_1}.P_1$ est dans $E^{m+1}$ mais pas dans $E^m$ , on a :

$$E^{m+1} = \langle X^{m+1-d_1}.P_1 \rangle + E^m \quad \text{et :}$$
$$X.E^{m+1} = \langle X^{m+2-d_1}.P_1 \rangle + X.E^m$$
$$X.E^{m+1} + E^{m+1} = \langle X^{m+2-d_1}.P_1 \rangle + X.E^m + \langle X^{m+1-d_1}.P_1 \rangle + E^m$$
$$= \langle X^{m+2-d_1}.P_1 \rangle + \langle X^{m+1-d_1}.P_1 \rangle + E^{m+1}$$
$$= \langle X^{m+2-d_1}.P_1 \rangle + E^{m+1}$$

cad $$E^{m+2} = \langle X^{m+2-d_1}.P_1 \rangle + E^{m+1}$$

Ceci montre que $\dim(E^{m+t}) = t + \dim(E^m)$ pour tout entier $t \geq 0$ . Or, pour t assez grand, on a $\dim(E^{m+t}) = m + t + 1 - d(G)$ .
Cela implique que $\dim(E^m) = m + 1 - d(G)$ . ❑

**Lemme 2 :** Si $m \geq d + d' - (1 + d(G))$ , alors :

$$\dim(\mathbf{Esylv^{m+1}(P)}) = \dim(\mathbf{Esylv^m(P)}) + 1$$

*preuve>* Pour $m = d$ , si le plus bas degré dans $B^m$ est $d'$ , cela signifie que les $P_i$ sont tous multiples de l'un d'entre eux, de degré $d'$ et on a $\dim(E^m) = m + 1 - d'$ . Sinon on a: $\dim(E^d) \geq d + 2 - d'$ . Si $\dim(E^{d+s}) \neq \dim(E^{d+s-1}) + 1$ , cela signifie que pour m de d à $d + s - 1$ on a :

$\dim(E^{m+1}) \geq \dim(E^m) + 2$ , donc :
$d + s + 1 - d(G) > \dim(E^{d+s}) \geq d + 2 - d' + 2s$ d'où :
$d + s < d + d' - 1 - d(G)$ ❑

**Lemme 3 :** Soit E un espace vectoriel de polynômes engendré par un nombre fini de générateurs et soit S une matrice dont les lignes sont ces générateurs, exprimés sur une base $X^m, X^{m-1}, \ldots, X, 1$. Alors la triangulation de S par la méthode du pivot sans échange de colonne fournit sur la dernière ligne non nulle de la triangulée un polynôme de degré minimum de E

**Remarque :** La triangulation sans échange de colonne fournit par exemple une matrice de la forme suivante, où les • représentent des 0 , les x représentent des coefficients non nuls et les # des coefficients "arbitraires":

```
X  #  #  #  #  #  #  #   ...   #
•  X  #  #  #  #  #  #   ...   #
•  •  •  •  X  #  #  #   ...   #
•  •  •  •  •  X  #  #   ...   #
              etc
```

La preuve du lemme est immédiate. Il faut aussi noter que 2 polynômes de degré minimum dans un sous-espace sont toujours proportionnels.

**Lemme 4 :** Si **A** est un anneau intègre où les déterminants sont calculables en temps polynomial, on peut calculer en temps polynomial un pgcd d'une liste de polynômes de **A**[X].
En particulier, il existe une majoration polynômiale de la taille de tous les pgcds de listes extraites de **P** .

*preuve>* Cela résulte des lemmes 1, 2, 3 et du fait que la triangulation "améliorée à la Bareiss" n'utilise comme coefficients que des déterminants extraits de la matrice de départ. On notera que même si les divisions exactes ne sont pas faciles dans **A** , la triangulation à la Bareiss peut être "mimée" à partir du moment où les détermùinants sont calculables en temps polynomial. ❑

**Remarque :** On peut , plutôt que trianguler directement la matrice $\mathbf{Sylv}^m(\mathbf{P})$ avec $m = d + d' - 1$ , trianguler successivement des matrices $\mathbf{Sylv}^m(\mathbf{P})$ avec m croissant de 1 en 1 à partir de d'. On obtiendra comme triangulées successives (en supprimant les lignes nulles du bas), des matrices par exemple de la forme:

- après la 1ère triangulation

```
X  #  #  #  #  #  #  #  #  #  #   ...   #
•  X  #  #  #  #  #  #  #  #  #   ...   #
•  •  •  •  X  #  #  #  #  #  #   ...   #
```

- après la 2ème triangulation

```
X  #  #  #  #  #  #  #  #  #  #   ...   #
•  X  #  #  #  #  #  #  #  #  #   ...   #
•  •  •  •  X  #  #  #  #  #  #   ...   #
•  •  X  #  •  #  #  #  #  #  #   ...   #
•  •  •  X  •  #  #  #  #  #  #   ...   #
•  •  •  •  •  X  #  #  #  #  #   ...   #
•  •  •  •  •  •  •  X  #  #  #   ...   #
```

- après la 3ème triangulation

```
X  #  #  #  #  #  #  #  #  #  #   ...   #
•  X  #  #  #  #  #  #  #  #  #   ...   #
•  •  •  •  X  #  #  #  #  #  #   ...   #
•  •  X  #  •  #  #  #  #  #  #   ...   #
•  •  •  X  •  #  #  #  #  #  #   ...   #
•  •  •  •  •  X  #  #  #  #  #   ...   #
•  •  •  •  •  •  •  X  #  #  #   ...   #
•  •  •  •  •  •  X  •  #  #  #   ...   #
•  •  •  •  •  •  •  •  X  #  #   ...   #
•  •  •  •  •  •  •  •  •  X  #   ...   #
```

etc ...

On peut noter en outre la simplification suivante dans l'algorithme: si, à une certaine étape, on a introduit un polynôme $P_i$ et que celui-ci a été tué (réduit à 0) par la triangulation, il est inutile de l'introduire aux étapes suivantes.

Enfin, **l'algorithme peut être complété par le calcul d'une relation de Bezout "ordinaire", à coefficients dans A, entre les $P_i$** de la manière suivante. Lorsqu'on a trouvé le degré d'un pgcd, on repère la première fois qu'il est apparu dans l'algorithme un polynôme G ayant ce degré. On repère de quels $X^j.P_i$ le polynôme G est combinaison linéaire de manière unique (il suffit pour cela d'avoir gardé en mémoire, au fur et à mesure des triangulations, de quels polynômes $X^j.P_i$ proviennent les lignes des triangulées). Le coefficient dominant c de G est alors (au signe près) le déterminant d'une matrice extraite de la matrice de Sylvester, obtenue en ne conservant que les lignes $X^j.P_i$ précédemment repérées et les colonnes des pivots, qui correspondent à des coefficients de degrés $r_1 > r_2 > ... > r_k = d(G)$. Enfin, on résout le système linéaire qui exprime qu'une certaine combinaison linéaire des $X^j.P_i$ en question a pour coefficients de degrés $r_1, ..., r_k$ les éléments $0, 0, ..., 0, c$. Il est clair que la solution obtenue est à coefficients dans A. Par ailleurs, on peut remplacer, dans la partie finale de l'algorithme, le coefficient dominant de G par tout autre coefficient non nul (on a sans doute intérêt à choisir un coefficient de taille minimale).

***Algorithme pour une relation de Bezout complète*** *(fin de la preuve du théorème 5)*

**Entre 2 polynômes P et Q**

On peut suivre la méthode ci-dessus pour la relation de Bezout ordinaire. Il faut alors compléter la matrice $\begin{bmatrix} U \\ V \end{bmatrix}$ en $\begin{bmatrix} U & -Q.cd(G)/G \\ V & P.cd(G)/G \end{bmatrix}$

**Entre $2^k$ polynômes** (ce qui règle naturellement le cas général)

On procède par recurrence sur k, en divisant la liste en 2, calculant une relation de Bezout complète pour chaque moitié, puis une relation de Bezout complète entre les 2 pgcds obtenus. Par exemple, la relation de Bezout complète entre P1, Q1, P2, Q2 sera obtenue en faisant un produit de matrices :

$$\begin{bmatrix} U1 & R1 & 0 & 0 \\ V1 & S1 & 0 & 0 \\ 0 & 0 & U2 & R2 \\ 0 & 0 & V2 & S2 \end{bmatrix} \cdot \begin{bmatrix} U3 & 0 & R3 & 0 \\ 0 & 1 & 0 & 0 \\ V3 & 0 & S3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Le produit d'une liste de matrices dans A est calculable en temps polynomial parce que les déterminants le sont[1]. Il reste à vérifier que l'on peut majorer polynomialement a priori la taille des matrices utilisées. Pour cela, il nous faut modifier un tout petit peu la récurrence, car on n'est pas sûr que les pgcds calculés en cascade sont de taille convenablement majorée: il risquerait d'y avoir explosion de la taille pour k variable, par introduction de facteurs

[1] cf par exemple [**Lom**], dans le cas présent on peut vérifer que les coefficients du produit sont seulement des produits (et non des sommes de produits) des coefficients des matrices de départ

multiplicatifs parasites. Pour plus de sûreté, nous calculons donc séparément les pgcds impliqués, à partir de listes extraites de **P** en appliquant le lemme 4 . Nous faisons ensuite un calcul à coefficients dans le corps des fractions **K** (sans réduction des fractions): nous calculons les relations de Bezout complètes $\begin{bmatrix} U & R \\ V & S \end{bmatrix}$ entre couples de pgcds précédemment calculés, (relations de Bezout qui vont se retrouver dans le produit de k matrices qui donne une relation de Bezout complète pour la liste entière), et nous multiplions la 1ère colonne par le coefficient nécessaire pour que le pgcd obtenu soit exactement celui qui avait été calculé a priori: il est clair que la taille de ce coefficient est bien maîtrisée. Nous obtenons en fin de compte une relation de Bezout complète à coefficients dans **K** . Il suffit alors de réduire dans chaque colonne les fractions de la matrice obtenue à un même dénominateur (qu'on chasse en suite comme un vilain) pour obtenir une relation de Bezout complète à coefficients dans **A** .

## Bibliographie


[Ait] Aitken A. C. : On the evaluation of determinants, the formation of their adjugates, and the practical solution of simultaneous linear equations. Proc. Edinburgh Math. Soc. ser 2 III , 207-219 , (1932)

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

[Col] Collins G. E. Computer Algebra of Polynomials and Rational Functions. Am. Math. Mon. 80 p 725-755 (1973)

[???] Référence plus récente pour une majoration du temps de calcul des déterminants dans un anneau de polynômes à coeffs entiers

[Gan] Gantmacher F. R. Théorie des Matrices DUNOD (1966) (traduit du russe)

[Gon] Gonzalez Laureano. The proof of the Sylvester Theorem through Habicht's sequence. prépublication . Université de Santander (Espagne). 1988

[G-L-R-R] Gonzalez L., Lombardi H., Recio T., Roy M.F. .Spécialisation de la suite de Sturm et sous-résultants. 1988

[Hab] Habicht .W. Eine Verallgemeinerung des Sturmschen Wurzelzählverfarhens. Comment. Math. Helvetici 21 p 96-116 (1948)..

[Lom] Lombardi Henri. : Calculabilité dans les structures algébriques dénombrables. Prépublication. Besançon. Juil 88

[Loos] Loos R.:. Generalized poynomial reaminder sequences. p 115-138 dans Computer Algebra, Symbolic and Algebraic Computation edité par Buchberger, Collins, Loos . Springer Verlag 1982 ref

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

[Syl] Sylvester J.J. On a theory of syzygetic relations of two integral functions, comprising an application to the theory of Sturm's function. Philos. Trans. Roy. Soc. London 143 (1853).
reprint dans : Sylvester : Collected Math Papers. Chelsea Pub. Comp. NY 1983 vol 1 429-586

[Tar] Tarski A. A decision method for elementary algebra and geometry. Prepared for publication by J.C.C. Mac Kinsey, Berkeley (1951)

[VdW] Van der Waerden B. L. Modern Algebra vol II
New York Frederick Ungar Publ Co 1953

# NOMBRES ALGEBRIQUES ET APPROXIMATIONS

Introduction ..... 2

A) LE CORPS DES NOMBRES REELS ALGEBRIQUES : PRESENTATION NAIVE

a) Présentation de $\mathbb{R}_{alg}$ ..... 5
b) Ralg comme $\mathcal{P}$-structure ..... 13
c) Situation des racines réelles d'un polynome de $\mathbb{Q}[X]$ ..... 16
d) Deux mots sur $\mathbb{C}_{alg}$ ..... 18
e) Une généralisation ..... 20

B) DISCUSSION A PROPOS DE DIFFERENTES PRESENTATIONS DES NOMBRES ALGEBRIQUES

a) Position du problème ..... 22
b) Systèmes d'équations en cascade, avant la levée de l'ambiguité ..... 27
c) Systèmes d'équations en cascade, après une levée de l'ambiguité à la Newton ..... 35

C) METHODES APPROXIMATIVES 43

a) Le théorème fondamental de l'algèbre est en temps polynomial ..... 44
b) Méthode des tableaux de signes approchés ..... 48
c) Fonctions approchables en temps polynomial par des fonctions polynômes ..... 54
d) Extensions possibles ..... 63


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
©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 $\xi$ est le polynôme minimum de $\xi$ 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 $\mathcal{P}$-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 $\mathcal{P}$-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 $\mathcal{P}$-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 $\geq 0$ . 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 $\alpha$ , sachant que $d(P) \leq p$ et **sup**(| coeffs de P |) $\leq H$ , $\mathbf{lg}(H) = h$ , il suffit de connaître $\alpha$ 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 $\alpha$ , 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 $| P(x_0) | \leq \inf(| P'(x_0) | / 2 M , r_0 / 2 )$ , 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 $| x_0 - u | < \inf( s / 4 m M , r / 3 m )$ , alors on a $r_0 > 2 r / 3$ , $P(x_0) < r_0 / 2$ , $| P'(x_0) | > s/2$ 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 $\mathbb{Q}[X,Y] \times \mathbb{R}_{alg} \times \mathbb{R}_{alg} \rightarrow \mathbb{R}_{alg}$

**Théorème** A.b3 :

Il existe une $\mathcal{P}$-opération Ev2 : $\mathbb{Q}[X,Y] \times \mathbb{R}_{alg} \times \mathbb{R}_{alg} \rightarrow \mathbb{R}_{alg}$ 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 $\zeta_i$ 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 $\zeta$ 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 $\alpha$ est une racine de T et $\beta$ 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 $\zeta_i$ et $\xi_i$ 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 $\geq 1$ (théorème A.a11 ),
- le fait que les produits dans $\mathbb{Z}$ sont en $O( n \log(n) \log\log(n) )$ (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 $\alpha$ 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 $\alpha$ est $\mathfrak{P}$-transcendant au sens suivant :

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

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

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 $\alpha$ 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 : $\quad | P(\alpha) | \geq 1/2^{c\, d^h\, ls(P)^k}$

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 $\xi_i$ 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 $\xi_i$ 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 $\xi_i$ 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 $\xi_i$ 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 $\xi_k$ 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* $\alpha$ *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 $\mathbf{lg(P)}$ 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 $\mathbf{k\ d\ lg(P)}$ .

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 $\mathbf{A}_{\mathbf{P}}$

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 $\mathbf{A}_{\mathbf{P}}$ 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 : $|\ M\ |_\infty \le |\ M\ |_2 \le |\ M\ |_1 \le \mathbf{lg}(\dim(M)) + |\ M\ |_\infty$ 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 $\qquad |\ MN\ |_1 \le |\ M\ |_1 + |\ N\ |_1$ .

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

Nous noterons $\alpha$ , $\beta$ , $\gamma$ ... des éléments de $\mathbf{A}_{\mathbf{P}}$ . Nous noterons $\lambda_1,\lambda_2,\ldots,\lambda_k$ les variables $X_1,X_2,\ldots,X_k$ vues comme éléments de $\mathbf{A}_{\mathbf{P}}$ .

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

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

et donc également :

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

Le produit dans $\mathbf{A}_{\mathbf{P}}$ est à peine plus compliqué. Notons $P_\alpha$ et $P_\beta$ les polynômes (de degrés $\le d_i - 1$ en $X_i$ (i=1,...,k) ) qui correspondent à $\alpha$ et $\beta$ . Le produit $\alpha.\beta$ dans $\mathbf{A}_{\mathbf{P}}$ 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 $\mathcal{B}_\mathbf{P}$ définie ci-après, et à garder les coordonnées utiles. La base $\mathcal{B}_\mathbf{P}$ 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 $\mathcal{B}_\mathbf{P}$ . En d'autres termes, la base $\mathcal{B}_\mathbf{P}$ 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 $\mathcal{B}_\mathbf{P}$ , 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 $\mathcal{B}_\mathbf{P}$ . 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 :

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

*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

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

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

b) On a la majoration

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

*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 $\alpha$ de $\mathbb{A}_P$ est en temps uniformément polynomial par rapport à $|\ \alpha\ |_1$ , **d** et **lg(P)** .

*preuve>* On calcule la liste $[\alpha^i]_{i=0,...,d-1}$ dans $\mathbb{A}_P$ (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 **d** × **d** dont les colonnes sont les $\alpha^i$ écrits sur la base canonique de $\mathbb{A}_P$ convenablement ordonnée. ❑

**Remarque** : Un élément $\alpha$ de $\mathbb{A}_P$ est inversible (dans $\mathbb{A}_P \otimes \mathbb{Q}$ ) 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* $\mathbb{A}_P \otimes \mathbb{Q}$ *en temps uniformément polynomial par rapport à* $|\ \alpha\ |_1$ , **d** *et* **lg(P)** .
De plus, lorsque $\alpha$ est inversible, $T(0)\ \alpha^{-1}$ s'exprime comme polynôme en $\alpha$ 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 $\beta$ ( $\beta \in \mathbb{A}_P \otimes \mathbb{Q}$ ) : $\alpha\,\beta = 1$ , comme un système de **d** équations à **d** inconnues dans $\mathbb{Q}$ (les coefficients de $\beta$ sur la base canonique de $\mathbb{A}_P$ ). Ce système d'équations est facile à écrire une fois qu'on a construit la table de multiplication de $\mathbb{A}_P$ .

**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 **P** est uniformément majorable par un polynôme en **d** et **lg(P)** .

*preuve>* Soit $\xi_1$ , ..., $\xi_k$ une solution du système. On applique la proposition précédente à $\xi_1$ , ..., $\xi_k$ vus comme éléments de $\mathbb{A}_P$ . 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 $\lambda_i$ 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 $\lambda_i$ 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 $dg(T) = d_1^{k}.d_2^{k-1}...d_k$. Le polynôme $P_j(X_1,X_2,...,X_j)$ définit un élément $\pi_j$ de $\mathbb{A}_T$ 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 $\Delta$ et $\Delta'$ 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* ( $\Delta$ et $\Delta'$ 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 $\Gamma$ 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 $\Gamma$ 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 $\alpha_\xi$ 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* $\alpha_\xi$ *dans* $\mathbb{C}_{sae,N}$ .
Nous disons que $\mathbf{d} = d_1...d_k$ est *le degré a priori* de l'entier algébrique $\alpha_\xi$ .
Nous notons $\mathbf{lg}(\alpha_\xi)$ la taille de $(\mathbf{P},[x_1,...,x_k],\alpha)$ ( **P** et $\alpha$ 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 $\rho$ . On note $\Gamma_0$ 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 $\Gamma_0$ , on note $\| \Gamma_0 \| := \sup_i ( \sum_j | a_{ij} | )$ .

On suppose que les majorations suivantes sont vérifiées :

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

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

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

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

Alors on est assuré que le processus itératif converge vers un point $\xi$ 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 $\Gamma_0$ 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 **d** et **lg(P)** pour la taille de $[x_1,x_2,\ldots,x_k]$ où $(P,[x_1,\ldots,x_k])$ 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 $\xi_i$ 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 $\alpha_j$ 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 $\alpha_j$ 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 $\xi$ 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 $\alpha_\xi$ dans $\mathbb{C}_{sae,N}$ .

a) Le calcul d'une approximation de $\alpha_\xi$ 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 $\alpha_\xi$ 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 $\xi$ 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 $\alpha_\xi$ : 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 $\alpha_\xi$ dans $\mathbb{C}_{sae,N}$ .

Alors le test d'égalité à 0 pour $\alpha_\xi$ , le calcul du signe de $\alpha_\xi$ dans le cas réel et, lorsque $\alpha_\xi \neq 0$ , le calcul de l'inverse de $\alpha_\xi$ (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 $\alpha$ . 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 $\alpha_\xi$ pour être assuré de son signe. En pratique, on a intérêt à mener en parallèle le calcul de plus en plus approché de $\alpha_\xi$ 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 $\alpha_\xi$ est la suivante :

- on majore les modules des conjugués des $\xi_i$ (ce qui peut être fait une fois pour toutes
- on en déduit une majoration m des modules des conjugués de $\alpha_\xi$
- comme $\alpha_\xi$ 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 $P_\alpha$ de $\alpha$ (sauf si on sait par un argument quelconque que $\alpha$ est inversible dans $\mathcal{A}_\mathbf{P} \otimes \mathbb{Q}$ ). A partir de $P_\alpha$ 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 $\alpha_\xi$ sous forme $\beta_\xi / n$ , 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 $(\mathbf{P},[x_1,...,x_k])$ une présentation dans $\mathbb{C}_{sae,N}$ de la solution $\xi_1,\xi_2,...,\xi_k$ du système d'équations algébriques emboitées $\mathbf{P}$ . 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 $\mathcal{B}_\xi$ cette clôture intégrale.
Ceci donnerait une version améliorée de $\mathbb{C}_{sae,N}$ où on représenterait tous les éléments de $\mathcal{B}_\xi$ 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 $\mathcal{B}_\xi$ , alors si $\alpha/n_P = \beta/n$ irréductible, avec $\alpha_\xi/n_P \in \mathcal{B}_\xi$ , on peut noter

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

$$| \beta_\xi/n |_{\mathcal{B}_\xi} \le | \beta |_1 + | n_P |_1$$

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

$$| \beta_\xi/n \,.\, \gamma_\xi/m |_{\mathcal{B}_\xi} \le | \beta_\xi/n |_{\mathcal{B}_\xi} + | \gamma_\xi/m |_{\mathcal{B}_\xi} + m_P \quad \text{(cf § b pour } m_P\text{)}$$

### 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 $\lambda$ 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 $\varepsilon$ , 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 $\varepsilon$ 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 $(\mathbf{P},[x_1,...,x_k])$ .

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 $\lambda$ 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 $\lambda$ 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 $\xi_i$ 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 $\alpha$ est un élément de $A_P$ nous dirons que $(P, [(x_1, s_1),...,(x_k,s_k)], \alpha)$ *est une présentation de l'entier algébrique* $\alpha_\xi$ *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 $\mathfrak{P}$-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 $\alpha_i$ de f et les zéros $\beta_i$ 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 :

$d([x_1,...,x_n],[y_1,...,y_n]) = \mathbf{inf}\,(\Sigma_i \mid x_i - y_{\sigma(i)} \mid)$ où $\sigma$ 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 $\alpha$ 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 $\varepsilon$ un rationnel positif. On appellera $\varepsilon$ - 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
* $a_0 = a < a_1 < ... < a_{k-1} < a_k = b$
* $a_1$ ... $a_{k-1}$ sont des dyadiques

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

Si $f$ est continument dérivable, un $\varepsilon$-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 $\varepsilon$ de taille raisonnable, la connaissance d'un $\varepsilon$-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 $\varepsilon$-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 $|b - a| \leq 2^k$ , $\| f' \| < M \leq 2^{n_1}$ ( $a , b \in \mathbb{Q}$ , $k , n_1 \in \mathbb{Z}$ ) , $\varepsilon = 1/2^n$ , et que $a$ et $b$ sont des dyadiques. On veut calculer un $\varepsilon$-tableau de signes pour $f$.

* *préliminaires* : on pose $\varepsilon' = 1/2^{n+k+3}$ , $\varepsilon'' = 1/2^{n+n_1+2}$ . On calcule un $\varepsilon'$-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 $f(a_i) \in [ (g_i - 1)/2^{n+3} , (g_i+1)/2^{n+3} ]$

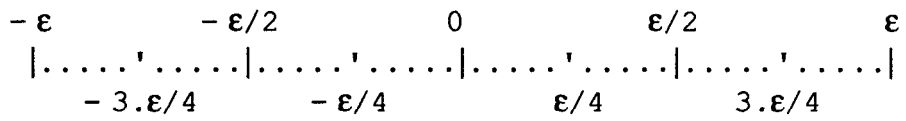
* *plan du travail :*

– 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 :

$T(n,r) \leq c_0$

$T(n,r-1) \leq c_0 + 2\,O((n+c)\,P(n+c))$

$T(n,r-2) \leq T(n+c,r-1) + 4\,O((n+c)\,P(n+c))$

$\leq c_0 + 2\,O((n+2c)\,P(n+2c)) + 4\,O((n+c)\,P(n+c))$

...

$T(n,r) \leq c_0 + 2\,O((n+rc)\,P(n+rc)) + 4\,O((n+(r-1)c)\,P(n+(r-1)c)) + ...$
$+ 2r\,O((n+c)\,P(n+c))$

$\leq c_0 + 2\,(1+2+...+r)\,O((n+rc)\,P(n+rc))$

(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 $\mathcal{P}$-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 $\varepsilon$-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 $\varepsilon$-tableau de signes pour la fraction rationnelle en calculant un $\varepsilon$-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 $2^m$ :

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

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 $d_\infty$.
$\mathbf{C}^{(k)}$ est l'espace des fonctions k fois continument dérivables sur $[-1\,,\,1\,]$.
$\mathbf{C}^{(\infty)}$ est l'espace des fonctions indéfiniment dérivables sur $[-1\,,\,1\,]$.
$\mathfrak{P}_n$ 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

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

$E_n(f) = d_\infty(f, \mathfrak{P}_n)$ 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 $\mathfrak{P}_n$ 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 $\mathfrak{P}_n$ 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

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

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

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

Les propriétés suivantes sont équivalentes

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

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

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

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

(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 $| A_n(f) | \leq S_n(f) + S_{n-1}(f)$

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

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

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

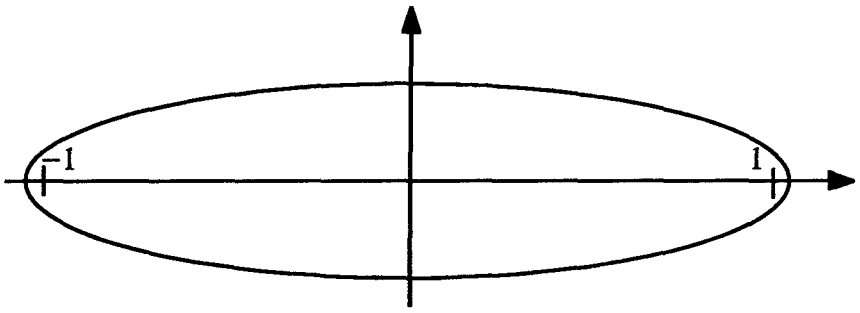
(iv) $\exists M>0 , r<1 \ \forall n \quad \| u_n^{[n]} - f \|_\infty \leq M r^n$

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




**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

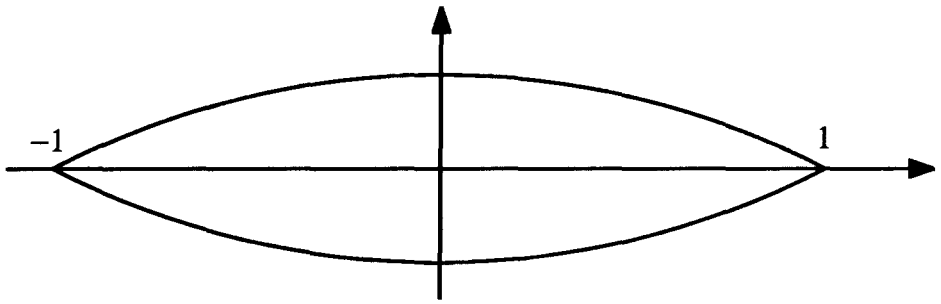
---

[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 $\mathcal{P}$-calculable pour un polynôme f (avec la présentation dense) : plus précisément, la fonction $f \rightarrow \|\,f\,\|_\infty$ , $\mathbb{Q}[X] \rightarrow \mathbb{R}$ est $\mathcal{P}$-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) | $\exists M>0, r<1, \gamma>0 \quad \forall n$ | $E_n(f) \leq M r^{n^\gamma}$ |
| ii) | $\exists M>0, r<1, \gamma>0 \quad \forall n$ | $S_n(f) \leq M r^{n^\gamma}$ |
| iii) | $\exists M>0, r<1, \gamma>0 \quad \forall n$ | $\mid A_n(f) \mid \leq M r^{n^\gamma}$ |
| iv) | $\exists M>0, r<1, \gamma>0 \quad \forall n$ | $\| u_n^{[n]} - f \|_\infty \leq M r^{n^\gamma}$ |
| j) | $\exists c, \beta>0 \quad \forall m \geq c\, n^\beta$ | $E_m(f) \leq 1/2^n$ |
| jj) | $\exists c, \beta>0 \quad \forall m \geq c\, n^\beta$ | $S_m(f) \leq 1/2^n$ |
| jjj) | $\exists c, \beta>0 \quad \forall m \geq c\, n^\beta$ | $\mid A_m(f) \mid \leq 1/2^n$ |
| jw) | $\exists c, \beta>0 \quad \forall m \geq c\, n^\beta$ | $\| u_m^{[m]} - f \|_\infty \leq 1/2^n$ |

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 $\gamma$ possibles est la même dans i), ii), iii) et iv) , la limite inférieure des $\beta$ 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 $\beta$ 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 $\alpha$ .

---

[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 $\alpha$ . 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 $\beta$ 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 :

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

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 $2\beta$ . ❑

**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 $\beta$ au moyen d'un calcul de majoration plus sophistiqué.

**Théorème** C.c5 :

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

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

b) f est polynomialement $\mathbf{C}^\infty$ c.-à-d.

$$\exists \, M > 0 \, , \, r < 1, \, k > 0 \text{ tels que : } \forall \, n \quad E_n(f) \leq M \, r^{n^\gamma}$$

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

d) la suite $f^{(n)}$ est une $\mathfrak{P}$-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 $\mathfrak{P}$-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 $\mathfrak{P}$-calculables

*preuve>* f est un $\mathfrak{P}$-point de $\mathbf{C}_W$ donné par une suite $m \to P_m$ , $\mathfrak{P}$-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 $\mathfrak{P}$-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 $\gamma$ ) . 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 $\gamma$ , 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') ≠ 0 et **Res**(P,Q) ≠ 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) | > ε , 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) | > ε/2 } du plan réel. Enfin s'il existe un chemin joignant 2 points dans la partie {(x,y); | f(x,y) | > ε } 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) | > ε/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*