Astérisque

J. P. VAN DE WIELE

La complexité du calcul des polynômes

Astérisque, tome 38-39 (1976), p. 265-274

<http://www.numdam.org/item?id=AST_1976__38-39__265_0>

Société Mathématique de France
Astérisque 38-39 (1976) p.265-274

LA COMPLEXITÉ DU CALCUL DES POLYNÔMES

J.P. Van de Wiele
IRIA - LABORIA
Domaine de Voluceau
Rocquencourt
78150 Le Chesnay

°°°

0. RÉSUMÉ

Bien que la complexité du calcul des polynômes soit un "vieux" problème en complexité algébrique (les premiers résultats datent de 1954), il est loin d'être résolu complètement et suscite encore de l'intérêt parmi les chercheurs. Après avoir défini le modèle de calcul que nous utilisons, nous traitons dans cet article plus particulièrement du problème suivant : étant donné un polynôme

$$p = \sum_{i=0}^{n} \alpha_i X^i$$

à une variable et à coefficients dans un corps infini, quel est dans ce modèle le "meilleur" algorithme d'évaluation de p ? Nous nous intéressons essentiellement aux résultats de bornes inférieures et l'objet de ce papier est de tenter d'en donner une vue d'ensemble en les ordonnant autour de deux méthodes principales : une méthode algébrique et une méthode analytique.

Pour une étude approfondie d'un grand nombre des résultats qui sont présentés ici on se reportera utilement au livre de Borodin et Munro [75].

1. INTRODUCTION

Définissons d'abord brièvement notre modèle de calcul. Soit K un corps infini, et soit

$$p = \sum_{i=0}^{n} \alpha_i X^i$$

un polynôme de $K[X]$.

Un calcul (algorithme d'évaluation) β de p dans K(x) mod $K \cup \{x\}$ est une suite de fonctions de K(x), $f_1, f_2, \ldots, f_n = p$, chaque f_i étant ou bien dans $K \cup \{x\}$, l'ensemble de fonctions que l'on se donne au départ, ou bien de la forme $f_j \circ f_k$, f_j et f_k étant deux fonctions précédemment calculées, et $\circ$ l'une des quatre opérations arithmétiques de l'ensemble $\Omega = \{+,-,\cdot,\div\}$.

Le coût (longueur) de β est en général défini comme le nombre d'opérations arithmétiques qu'il contient. On le note : $L(\beta/K(x) \text{ mod. } K \cup \{x\})$ et plus simplement $L(\beta)$ quand il n'y a pas d'ambiguité. Les significations de $L(\pm/\beta)$ et de $L(\dot{\div}/\beta) \equiv L(*/\beta)$ sont évidentes.

La complexité de p est : $\min_{\beta \text{ calcule } p} L(\beta) \equiv L(p)$.

Démontrer l'optimalité d'un calcul A de p revient à montrer que tout calcul de p contient au moins autant d'opérations que A. Nous cherchons donc à borner inférieurement L(p).

Un calcul de p dans notre modèle est une bonne approximation d'un calcul effectif de p sur un ordinateur pour de nombreuses raisons que nous ne discuterons pas ici.

Remarque 1 : Il est possible de généraliser au cas de plusieurs variables la plupart des résultats que nous énonçons ici. Mais pour la simplicité de l'exposition, nous nous limitons au cas d'une variable. Par contre nous supposons le corps infini K quelconque bien que dans la plupart des applications : $K = Q, \mathbb{R}$ ou $\mathbb{C}$.

Remarque 2 : Il est essentiel de distinguer le problème de l'évaluation de polynômes particuliers, que nous traitons ici, de celui de l'évaluation du polynôme général de degré n. Ce dernier se formule de la façon suivante : étant donné un polynôme général de degré n

$$q = \sum_{i=0}^{n} a_i X^i \in K[X],$$

quel est le meilleur algorithme d'évaluation de q dans $K(X, a_0, a_1, a_2, \ldots, a_n)$ mod. $K \cup \{X, a_0, a_1, \ldots, a_n\}$? C'est à dire que nous considérons la classe des algorithmes qui évaluent tous les polynômes de degré n. Ce problème est complètement résolu. La règle de Horner (n additions-soustractions, n mul-

tiplications) est uniquement optimale. (Borodin [71]).

2. LA MÉTHODE ALGÉBRIQUE

K est un corps quelconque infini.

Théorème 1 (Motzkin [55], Belaga [58]) (cf. Borodin-Munro [75]) :

Soit $p \in K[X]$, $\deg(p)=n$ $\quad p = \sum_{i=0}^{n} \alpha_i X^i$.

Si : $L(*/p \text{ dans } K(X) \text{ mod. } K \cup \{X\}) < \lceil (n+1)/2 \rceil$

ou bien : $L(\pm/p \text{ dans } K(X) \text{ mod. } K \cup \{X\}) < n$

alors il existe $H \in Q[y_0, y_1, \ldots, y_n]$, un polynôme à (n+1) variables et à coefficients rationnels, tels que :

$$H(\alpha_0, \alpha_1, \ldots, \alpha_n)=0 \ , \ H \neq 0$$

Corollaire : Presque tous les polynômes de $\mathbb{C}[X]$ (resp. $\mathbb{R}[X]$) requièrent dans $\mathbb{C}(X)$ mod. $\mathbb{C} \cup \{X\}$ (resp. $\mathbb{R}(X)$ mod. $\mathbb{R} \cup \{X\}$) $\geq \lceil (n+1)/2 \rceil$ multiplications divisions et $\geq n$ additions soustractions.

Preuve : Dans $\mathbb{C}^{n+1}$ (resp. $\mathbb{R}^{n+1}$) l'ensemble des tuplets qui annulent un polynôme à coefficients rationnels est de mesure nulle.

Ce résultat est "presque" optimal en ce sens que tout polynôme de degré n peut être calculé dans $\mathbb{C}(X)$ (resp. $\mathbb{R}(X)$) mod. $\mathbb{C} \cup \{X\}$ (resp. $\mathbb{R} \cup \{X\}$) avec $\lfloor n/2 \rfloor$ + 2 multiplications et n additions. (Knuth [69]).

Preuve du théorème 1 : Considérons uniquement le cas des mult.-div. Soit p un polynôme tel que : $L(*/p) = k$. Alors p est calculable par un calcul du type suivant :

$$\begin{aligned}
\beta \ : \ & s_0 \leftarrow X \\
& s_1 \leftarrow (c_1+m_{1,0}\, s_0) * (d_1+n_{1,0} s_0) \\
& \vdots \\
& s_k \leftarrow (c_k + \sum_{j<k} m_{k,j}\, s_j)(d_k + \sum_{j<k} n_{k,j} s_j) \\
& s_{k+1} \leftarrow C_{k+1} + \sum_{j \leq k} m_{k+1,j} s_j
\end{aligned}$$

avec : $c_i, d_i \in \mathbb{C}$, $m_{ij} \in \mathbb{Z}$, $1 \leq i \leq k+1$, $0 \leq j \leq k$.

On en déduit facilement qu'il existe des polynômes p_i en k+1 variables et à coefficients rationnels, $p_i \in Q[z_1, z_2, \ldots, z_{k+1}]$ tels que :

$$\alpha_i = p_i(c_1, c_2, \ldots, c_{k+1})$$

$$k+1 < n+1 \Longrightarrow \exists H \in Q[y_o, y_1, \ldots, y_n] \text{ tel que}$$

$$H(\alpha_o, \alpha_1, \ldots, \alpha_n) = 0 \ , \ H \neq 0$$

(résultat de géométrie algébrique élémentaire).

Malheureusement tous les polynômes de $Q[X]$ sont des exceptions du théorème de Motzkin et Belaga . Peut-on dire quelque chose sur les polynômes de $Q[X]$?

Théorème 2 - (Paterson-Stockmeyer [73], Borodin-Cook [74])

$\exists f$, $f \neq 0$, $f \in \mathbb{Z}[y_o, y_1, \ldots, y_n]$ tel que tout polynôme de $K[X]$ calculable avec moins de $\sqrt{n}$ multiplication-divisions non scalaires ou moins de $\sqrt{n+1} - 2$ add.-soustr. satisfait : $f(\alpha_o, \alpha_1, \ldots, \alpha_n) = 0$.

Corollaire : Presque tous les polynômes de $Q[X]$ requièrent dans $\mathbb{C}(X)$ mod. $\mathbb{C} \cup \{X\}$ $\geq \sqrt{n}$ k-opérations et $\geq \sqrt{n+1} - 2$ ± opérations.

Ce résultat est presque optimal pour les mult.-div. non scalaires (Paterson-Scotkmeyer [73]), mais on ne sait pas en général évaluer un polynôme de degré n à coefficients rationnels avec moins de n add.-soustr., et ceci quel que soit le nombre de mult.-div. utilisé ; est-il possible de "sauver" des additions avec des multiplications comme il est possible de sauver des multiplications avec des additions ? La question est ouverte.

Preuve du théorème 2 :

Nous nous limiterons encore au cas des multiplications-divisions. L'idée est de considérer dans le programme β ci-dessus aussi bien les $m_{ij} \in \mathbb{Z}$ que les $c_i, d_i \in \mathbb{C}$ comme des paramètres. On définit ainsi un calcul unique (canonique) qui évalue tous les polynômes calculables

avec moins de k additions. La preuve est ensuite analogue à celle du théorème 1.

Le résultat précédent n'est pas constructif. On sait qu'il existe des polynômes de degré n à coefficients rationnels qui sont "difficiles à calculer" (de l'ordre de $2\sqrt{n}$ opérations) et que la plupart le sont, mais on ne sait toujours rien dire sur la complexité d'un polynôme déterminé de $Q[X]$ pris au hasard.

Le résultat suivant résoud partiellement cette question.

Théorème 3 - (Strassen [74]) :

Soit β un calcul dans $\mathbb{C}(X)$ mod. $\mathbb{C} \cup \{X\}$ du polynôme $p(X) = \sum_{\delta=0}^{n} 2^{2^{\delta n^3}} X^{\delta}$.

Soit $u = L(\pm/\beta)$, $v = L(*,\beta)$

Pour n assez grand : $m = \min(u,2v) \geq n-4$

ou bien : $s = u+v > \frac{n^2}{\log_2 n}$

En d'autres termes si β contient moins de $\frac{n^2}{\log n}$ opérations, il comprend au moins n-4 add.-soustr. et n/2-2 mult.-div. ; donc $L(p) \geq 3n/2-6$.

Preuve du théorème 3 : Le point de départ est le théorème de Motzkin et Belaga . Celui-ci dit en substance : si β calcule un polynôme q de degré n et contient "peu" d'opérations alors les coefficients de q qui s'expriment rationnellement en fonction d'un "petit" nombre de paramètres sont algébriquement dépendants sur Q.

L'idée de Strassen consiste à examiner, en reprenant les notations de la preuve du théorème 1, le degré et le poids (somme des valeurs absolues des coefficients) des polynômes p_i et à en déduire des majorations sur le degré et la hauteur du polynôme H.

Lemme 1 : Soit K un corps infini, β un calcul de $\sum_{\delta=0}^{n} \alpha_\delta X^\delta$, $\alpha_\delta \in K$, dans $K(X)$ mod. $K \cup \{X\}$.

Soit : $L(\pm/\beta) = u$, $L(*/\beta) = v > 0$

$m = \min(u,2v)$ $\quad s = u+v$

Alors il existe $P_i \in \mathbb{Z}[z_1,z_2,\ldots,z_m]$, $1 \leq i \leq n$ et $\lambda, \gamma_1, \gamma_2, \ldots, \gamma_m \in K$ tels que :

$$\alpha_i = \lambda P_i(\gamma_1,\gamma_2,\ldots,\gamma_m)$$

$$\text{et de plus : } \max_{1\le i\le n} \deg(P_i) \le n^s \ , \ \max_{1\le i\le n} \text{poids}(P_i) \le 2^{n^s}$$

Lemme 2 : Soit $m \ge 1$, $c \ge 2$, $f \ge 4$, $q \ge 5 \in \mathbb{N}$ et soient $P_1,\ldots,Pq$ dans $\mathbb{Z}[z_1,z_2,\ldots,z_m]$ q polynômes en m variables à coefficients entiers tels que

$$\sup_{1\le i\le q} P_i \le c \qquad \sup_{1\le i\le q} \text{poids}(P_i) \le f$$

Si $g \in \mathbb{N}$ est tel que : $g^{p-m-2} > c^m q^q \log f$
il existe une forme non triviale $H \in \mathbb{Z}[y_1,\ldots,y_q]$ de degré g et de hauteur ≤ 3 telle que :

$$H(P_1,\ldots,P_q) \neq 0$$

[Hauteur (H) = Sup. des valeurs absolues des coefficients]

Si l'on applique le lemme 2 à la situation du lemme 1 nous obtenons

$$\text{Si } g \in \mathbb{N} \text{ est tel que : } g^{n-m-2} > n^{s(m+1)} n^n \qquad (1)$$

il existe une forme H non triviale $\in \mathbb{Z}[z_1,z_2,\ldots,z_m]$ de degré g et de hauteur ≤ 3 telle que : $H(\alpha_1,\alpha_2,\ldots,\alpha_n) = 0$.

Lemme 3 : Soient $\alpha_1,\alpha_2,\ldots,\alpha_n \in \mathbb{Z}$, $n \ge 5$, $g \in \mathbb{N}$, $g \ge 3$ et :

$$|\alpha_1| > 4 \ , \ |\alpha_k| > |n\alpha_{k-1}|^g \text{ pour } k > 1$$

Alors il n'existe pas de forme H non triviale dans $\mathbb{Z}[z_1,\ldots,z_n]$ de degré g et de hauteur ≤ 3 telle que :

$$H(\alpha_1,\alpha_2,\ldots,\alpha_n) = 0$$

On peut alors démontrer le théorème 3. Soit $\alpha_\delta = 2^{2^{\delta n^3}}$ et soit $g = 2^{n^3-n^2}$; g et α_δ satisfont les conditions du lemme 3.

Donc : $g^{n-m-2} \le n^{s(m+1)}\ n^n$ (lemme 2 + (1)).

$$\Rightarrow \quad n^4-n^3-\log n \le (s \log n+n^3-n^4)\ (m+2) \qquad (g=2^{n^3-n^2})$$

$$\Rightarrow \quad \left(s < \frac{n^2}{\log n} \Rightarrow m \ge n-4\right).$$

Il existe donc des polynômes à coefficients entiers qui requièrent à peu

près $\frac{3n}{2}$ opérations arithmétiques dans $\mathbb{C}(X)$ mod. $\mathbb{C} \cup \{X\}$ soit autant que les plus complexes des polynômes de $\mathbb{C}[X]$, ceux dont les coefficients sont algébriquement indépendants sur Q. (Théorème 1).

Mais on remarque que les polynômes de Strassen ont des coefficients qui croissent extrêmement rapidement. Existe-t-il des polynômes simples qui ont la même complexité ?

Problème ouvert : Construire un polynôme de degré n à coefficients 0-1 dont la complexité soit de l'ordre de n dans $\mathbb{C}(X)$ mod. $\mathbb{C} \cup \{X\}$.

Lipton a récemment obtenu le résultat partiel et non constructif suivant :

Théorème 4 (Lipton [75]).

Il existe des polynômes de degré n à coefficients 0-1 qui requièrent dans $\mathbb{C}(X)$ mod. $\mathbb{C} \cup \{X\}$ $\approx \frac{n^{1/4}}{\log n}$ mult.-div. non scalaires.

La preuve utilise essentiellement les résultats de Strassen [74].

Récemment l'auteur a obtenu un résultat un peu plus fort.

Théorème 5 : Il existe des polynômes de degré n à coefficients 0-1 qui requièrent dans $\mathbb{C}(X)$ mod. $\mathbb{C} \cup \{X\}$ $\approx \frac{n^{1/3}}{\log n}$ mult.-div. non scalaires et de même il existe des polynômes de degré n à coefficients 0-1 qui requièrent dans $\mathbb{C}(X)$ mod. $\mathbb{C} \cup \{X\}$ $\approx \frac{n^{1/3}}{\log n}$ additions-soustractions.

La preuve de ce résultat apparaîtra dans un autre papier (Van de Wiele [76]). Notons enfin que Lipton [75] a utilisé des résultats de la théorie des langages pour mettre en évidence quelques propriétés de la structure des polynômes 0-1 difficiles à calculer.

Nous en venons maintenant à une deuxième catégorie de résultats ; ces résultats sont relatifs au nombre minimum d'additions-soustractions et contrairement aux précédents il utilisent des propriétés topologiques de $\mathbb{R}$. C'est pourquoi il nous semble raisonnable de parler de méthode analytique.

3. LA MÉTHODE ANALYTIQUE :

1) Théorème 6 - Borodin et Cook [74].

Soit E(k) l'ensemble des polynômes de $\mathbb{R}[X]$ calculables dans $\mathbb{R}(X)$ mod. $\mathbb{R} \cup X$ avec $\leq k$ additions-soustractions. (Le nombre de mult.-div. n'est pas borné). Le nombre de zéros réels distincts de tout polynôme de E(k) est borné par une certaine constante M(k) qui ne dépend que de k.

La preuve est par induction sur k et utilise fondamentalement le théorème de Rolle; Borodin et Cook mettent en évidence une borne supérieure de M(k) probablement très grossière : $M(k) \leq 2^{2^{\cdot^{\cdot^{\cdot^{2^k}}}}}$ (k)

Problème ouvert : $\exists c \in \mathbb{R}$, $C > 0$ tel que : $M(k) \leq c^k$.

La démonstration de cette conjecture impliquerait pour le nombre d'additions-soustractions une borne inférieure de l'ordre de log n analogue à celle que l'on obtient pour les mult.-div. par l'argument de croissance du degré ($L(*/x^n) \gtrsim \log n$).

Il serait extrêmement intéressant de tenter ensuite de généraliser ce résultat au cas de plusieurs variables.

Problème ouvert : Trouver $p \in E(k)$ dont le nombre de racines réelles distinctes soit supérieur à 3^k.

2) Nous ne connaissons aucune borne-inférieure non triviale sur le nombre d'add.-soustr. nécessaires pour évaluer le polynôme $\sum_{i=0}^{n} X^i$ dans l'anneau $\mathbb{R}[X]$ mod. $\mathbb{R} \cup \{X\}$ (sans division).

Motivés par ce problème, L. Hyafil et l'auteur ont montré le résultat suivant :

Théorème 7 : (Hyafil, Van de Wiele [75])

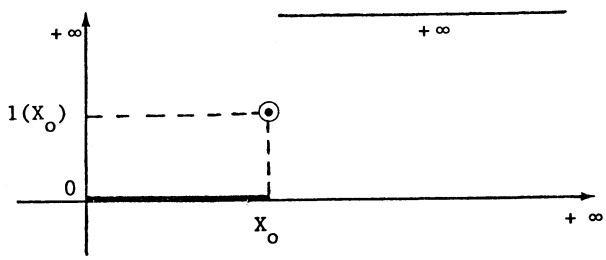
Soit F(k) l'ensemble des polynômes de $R_+[X]$ (à coefficients ≥ 0) qui requièrent moins de k add.-soustr. dans $\mathbb{R}_+[X]$ mod. $\mathbb{R}_+ \cup \{X\}$ (sans division et sans soustraction). Pour tout $k \in \mathbb{N}$, il existe $M'(k) > 0$ tel que :

$$\forall n > M'(k) \ , \ \sum_{i=0}^{n} X^i \notin F(k) \ , \text{ ou encore}$$

$$L(\pm/ \sum_{i=0}^{n} X^i \text{ dans } \mathbb{R}_+[X] \text{ mod. } \mathbb{R}_+ \cup \{X\}) > k.$$

En d'autres termes, dans ce modèle, la complexité additive de $\sum_{i=0}^{n} X^i$ diverge quand n tend vers $+\infty$.

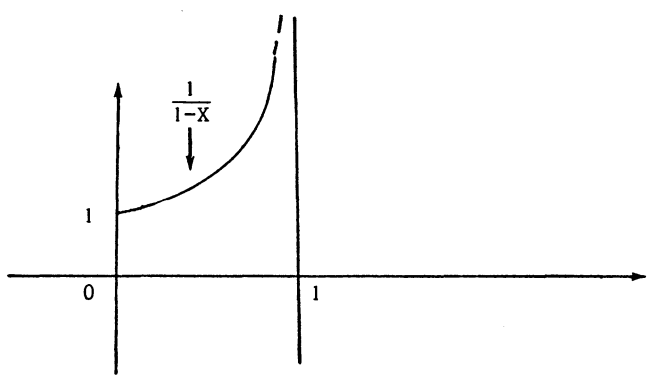
L'idée de la preuve est essentiellement la suivante. Nous considérons le comportement des limites de suites de fonctions de F(k). Toute suite de fonctions de F(0) qui "a une limite" est essentiellement de l'un des deux types suivants : ou bien la suite converge vers un monôme cx^γ, $0 \leq c$, $\gamma < +\infty$ ou bien, si le nombre de multiplications diverge, la suite se comporte comme $\{x^n\}$, c'est-à-dire que la limite a l'allure suivante :



La situation n'est pas essentiellement différente pour $k > 0$. Plus précisément nous montrons :

Théorème 8 : Soit $\{f_n\}$ une suite de fonctions de F(k) , $k \geq 0$. Et soit 1 la "limite" de f_n ; 1 n'est définie qu'en les points X de $\mathbb{R}_+$ où $\lim_{n \to +\infty} f_n(X)$ existe et est fini. Alors 1 est bornée sur tout borné de $\mathbb{R}_+$.

La limite de $\left\{ \sum_{i=0}^{n} X^i \right\}$ est $\frac{1}{1-X} = 1(X)$ sur $[0,1[$, qui a l'allure suivante :



Il est clair que l(X) n'est pas bornée supérieurement sur l'intervalle $[0,1[$; ce qui démontre le Théorème 7.

Problème ouvert :

Est-ce que le théorème 8 reste vrai lorsqu'on s'autorise les soustractions ?

4. CONCLUSION

Nous espérons avoir montré par cet exemple que des problèmes initialement "informatiques" peuvent conduire à des problèmes mathématiques intéressants dont la solution aurait un intérêt à la fois pratique et théorique.

5. BIBLIOGRAPHIE

1) BORODIN A. [71]. "Horner's rule is uniquely optimal", in Theory of Machines and Computation, Academic Press, New-York.

2) BORODIN A., COOK S. [74]. "On the number of additions to compute specific polynomials". Proc. 6th Annual ACM Symp. on Theory of Computing. pp. 342-347.

3) BORODIN A., MUNRO I. [75]."The computational complexity of algebraic and numeric problems". Elsevier Computer Science Library.

4) HYAFIL L., VAN DE WIELE [75]. "On the additive complexity of specific polynomials". Information Processing Letters. Vol. 4 N° 2, Nov. 75.

5) KNUTH D.E. [69]. "The art of computer programming : semi-numerical algorithms, vol. II, Addison-Wesley, Reading, Mass.

6) LIPTON [75]. "Polynomials with 0-1 coefficients that are hard to evaluate". IBM Research Report, RC 5612.

7) PATERSON-STOCKMEYER [73]. "On the number of non scalar multiplications necessary to evaluate polynomials", SIAM J. Computing 2 : 60-66, 1973.

8) STRASSEN V. [74]. "Polynomials with rational coefficients which are hard to compute". SIAM J. Computing 3 : 128-148, 1974.

9) VAN DE WIELE [76]. Rapport LABORIA à paraître.