Résolution numérique certifiée des systèmes algébriques

Joelle Saadé

Soutenance de stage du M2R

« Algèbre Appliquée à la Cryptographie et au Calcul Formel »

effectué du 1er mars au 30 juin

au Laboratoire d'informatique de l'École polytechnique

sous la direction de Grégoire Lecerf (CR CNRS)

La résolution des systèmes d'équations polynomiales :

- C'est un problème fondamental du Calcul Formel ;
- C'est un problème NP-Complet ;
- C'est un problème qui a beaucoup d'applications dans différents domaines ;
- L'homotopie : méthode de résolution numérique par déformation ;
- Certifier l'homotopie $\rightarrow$ Certifier les suivis de chemin $\rightarrow$ Opérateur de Newton ; [2, 3, 4]
- Critère de convergence de l'opérateur de Newton. [1]

Bibliographie

[1] M. Giusti, G. Lecerf, B. Salvy, et J.-C. Yakoubsohn. On location and approximation of clusters of zeros: case of embedding dimension one. *Foundations of Computational Mathematics*, 7(1):1–49, 2007.

[2] J. van der Hoeven. *Journées Nationales de Calcul Formel (2011)*, volume 2 de *Les cours du CIRM*, chapitre Calcul analytique. CEDRAM, 2011. Exp. No. 4, 85 pages, http://ccirm.cedram.org/ccirm-bin/fitem?id=CCIRM_2011__2_1_A4_0.

[3] J. van der Hoeven. Effective complex analysis. *JSC*, 39:433–449, 2005.

[4] J. van der Hoeven. Reliable homotopy continuation. Technical Report, LIX, Ecole polytechnique, 2015.

- But : résolution de n équations polynomiales à n variables

$$P(x) = (P_1(x), ..., P_n(x)) = 0$$

où $P_i \in \mathbb{C}[x] = \mathbb{C}[x_1, ...x_n]$ pour $i = 1, ..., n$.

- Départ : système de n équations à n variables

$$Q(x) = (Q_1(x), ..., Q_n(x)) = 0$$

facile à résoudre.

- Exemple :

$$\begin{cases} P_1(x,y) &= x^2 - y^2 + x + 3 = 0 \\ P_2(x,y) &= x^2 + 2\,x\,y + 7\,y^2 - 8\,y + 2 = 0 \end{cases} \quad \text{et} \quad \begin{cases} Q_1(x,y) &= x^2 - 1 = 0 \\ Q_2(x,y) &= y^2 - 1 = 0 \end{cases}$$

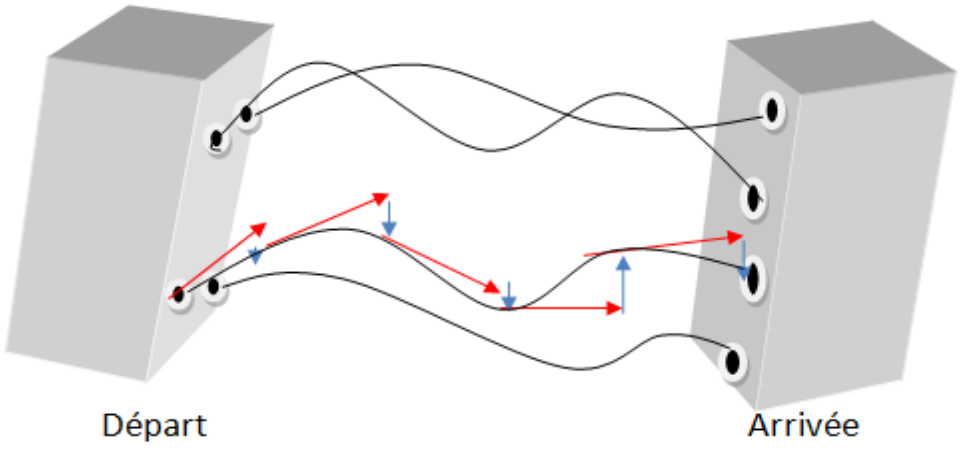
- Homotopie : on construit une homotopie linéaire par exemple

$$H(x,t) = (1-t)\,P(x) + t\,Q(x), t \in [0,1].$$

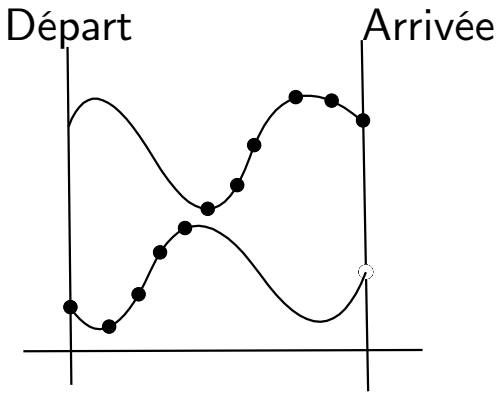
Solutions de $Q \longrightarrow$ Solutions de P.

Les solutions du premier système sont obtenues en suivant les racines de $Q(x)$ de $t = 1$ à $t = 0$.

Le cas générique : système choisi au hasard $\longrightarrow^{\sim}$ racines simples



Risque d'erreur :



Donc il est necéssaire de certifier les chemins pas à pas !

Pourquoi une arithmétique de boules

→ Approximations intervenant dans un calcul complexe : recommandation d'une arithmétique de boule

Implémentation de la classe d'arithmétique de boule :

→ Précision limitée

→ Mode d'arrondi selon la norme IEEE 754 :

- Au plus près ; (`FE_TONEAREST`)
- Vers zéro ; (`FE_TOWARDZERO`)
- Vers plus l'infini ; (`FE_UPWARD`)
- Vers moins l'infini ; (`FE_DOWNWARD`)

→ Implémentation portable : librairie standard C++, math.h.

→ Implémentation avec la technologie **S**treaming **S**IMD **E**xtensions. SSE2/SSE3

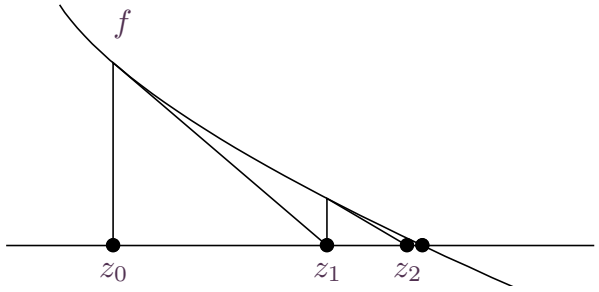
Avantage : opérer sur plusieures données en une seule instruction.

	Implémentation portable	Utilisation de la technologie SSE
Addition	448	64
Multiplication	3390	164
Division	7759	257

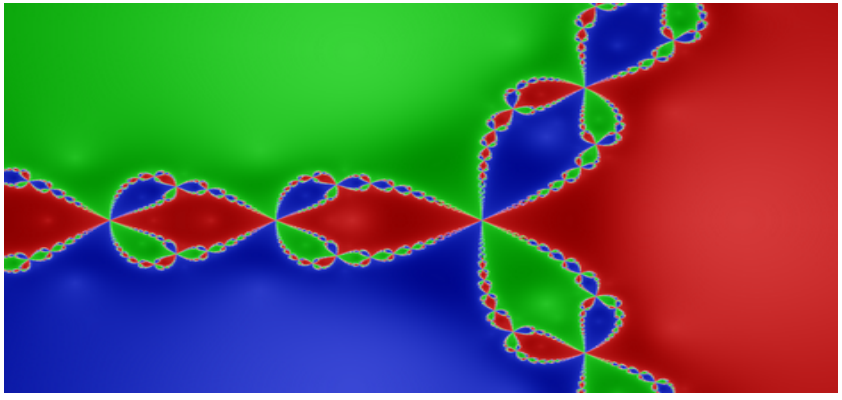
Comparaison sur le nombre de cycles du processeur :Intel(R) Core(TM) i3 CPU M330 @ 2.13 GHz, pour quelques opérations sur les boules $\mathfrak{B}(2+i,3)$ et $\mathfrak{B}(1,4)$.

La suite de Newton pour une fonction f et une valeur initiale z_0 :

$$z_{k+1} = z_k - \frac{f(z_k)}{f'(z_k)}$$



Il est difficile de supposer qu'un point va converger vers un zéro ou non. Dans la pratique on a besoin de critères rapides.



Théorème 1. *Soit $f \in \mathcal{C}^1(\Omega, \mathbb{Y})$, et $x_0 \in \Omega$ est tel que $Df(x_0)$ est inversible. Soit la constante β telle que, $\beta \geqslant \|Df(x_0)^{-1} f(x_0)\|$, et une fonction continue croissante $L: [0, R] \to \mathbb{R}_{\geqslant 0}$ qui satisfait :*

$$\|Df(x_0)^{-1}(Df(b) - Df(a))\| \leqslant L(r)\|b - a\|, \forall r \in [0, R] \text{ et } \forall a, b \in \bar{B}(x_0, r) \cap \Omega$$

et $B(x_0, R) \subseteq \Omega$. Soit la fonction ϕ, définie par $\phi(r) = \beta - r + \int_0^r L(s)(r - s)\,\mathrm{d}s$, on suppose qu'elle admet un unique zéro r_- dans $[0, R)$, et que $\phi(R) \leqslant 0$.

Alors la suite de Newton $r_0 = 0$, $r_{k+1} = r_k - \frac{\phi(r_k)}{\phi'(r_k)}$ est bien définie dans $[0, r_-]$, et converge vers r_-. La suite $x_{k+1} = x_k - Df(x_k)^{-1} f(x_k)$ est bien définie dans $\bar{B}(x_0, r_-)$, et converge vers l'unique zéro ζ de f dans $B(x_0, R)$. De plus, on a $\|\zeta - x_k\| \leqslant r_- - r_k$ et $\|x_{k+1} - x_k\| \leqslant r_{k+1} - r_k$.

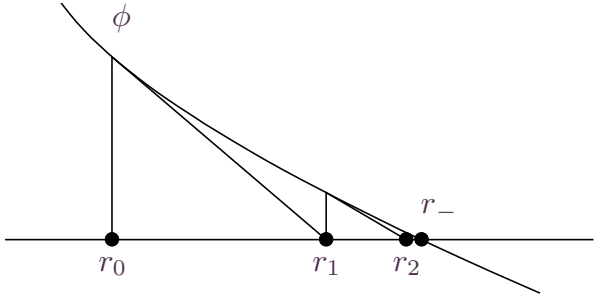
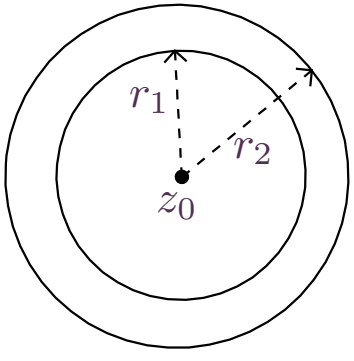


Figure 1. Graphe de ϕ et les premiers itérations de Newton.

Soit,

$$\gamma = \gamma(f, z_0) = \sup_{k \geqslant 2} \left\| Df(z_0)^{-1} \frac{D^k f(z_0)}{k!} \right\|^{1/(k-1)}, \ \beta = \beta(f, z_0) = \| Df(z_0)^{-1} f(z_0) \|.$$

Corollaire 2. *Si* $\alpha(f, z_0) = \gamma \beta < 3 - 2\sqrt{2}$, *alors* z_0 *est bien un zéro approché de* f *et va converger vers une unique solution exacte* ζ *dans* $\bar{\mathfrak{B}}\left(z_0, \left(1 + \frac{\sqrt{2}}{2}\right)\beta\right)$ *et* $\bar{\mathfrak{B}}\left(z_0, \left(1 - \frac{\sqrt{2}}{2}\right)\gamma^{-1}\right)$.

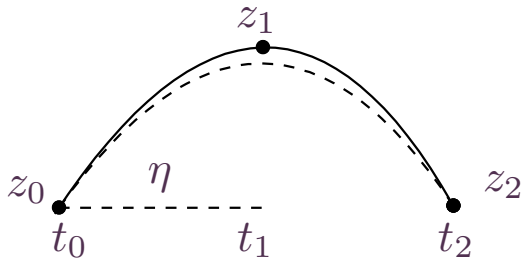


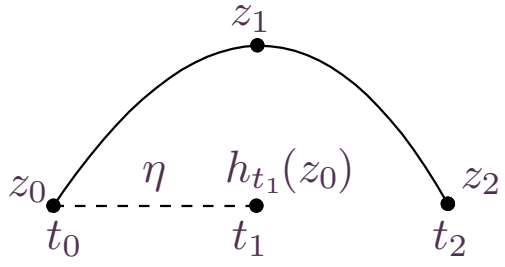
$$r_1 = \left(1 + \frac{\sqrt{2}}{2}\right)\beta(z_0)$$

$$r_2 = \left(1 - \frac{\sqrt{2}}{2}\right)\gamma^{-1}(z_0)$$

Démonstration. $L(r) = \frac{2\gamma}{(1-\gamma r)^3}$ *et* $\phi(r) = \beta - r + \frac{\gamma r^2}{1-\gamma r}$ □

On considère l'homotopie de Gauss : $h_t(z) = f(z) - t\,f(z_0)$

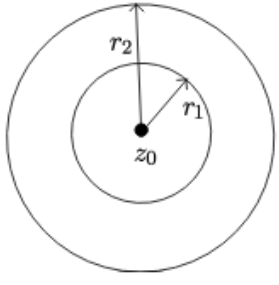


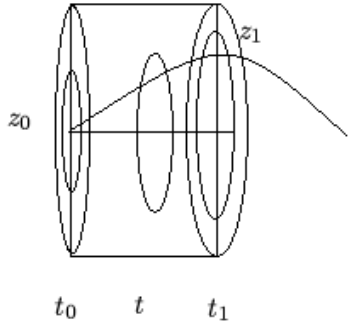


On choisit un pas η, on calcule $h_{t_1}(z_0)$, si $\alpha(h_{t_1}, z_0) \geqslant \alpha_0$ on diminue le pas. Si $\alpha(h_{t_1}, z_0) < \alpha_0$, la courbe approchée passe par un seul point $z_1 = N_f(z_0)$. Afin de certifier le chemin $[t_0, t_1]$, on vérifie les conditions du α-théorème pour tout $t \in [t_0, t_1]$. Pour une condition : $\max_{t_1 \leqslant t \leqslant t_0} \left(1 + \frac{\sqrt{2}}{2}\right) \beta_t(z_0) < \min_{t_1 \leqslant t \leqslant t_0} \left(\frac{1 - \frac{\sqrt{2}}{2}}{\gamma_t(z_0)}\right)$ on aura la certification suivante :

$$r_1 = \left(1 + \tfrac{\sqrt{2}}{2}\right)\beta_t(z_0)$$

$$r_2 = \left(1 - \tfrac{\sqrt{2}}{2}\right)\gamma^{-1}(z_0)$$



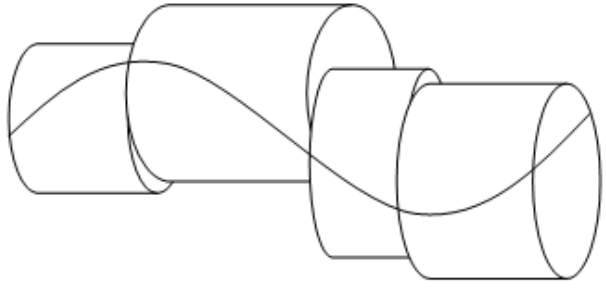


Un pas optimal η peut être calculé, garanti par la vérification des conditions du α-théorème, alors l'unique chemin dans chaque intervale $[t, t+\eta]$.

Algorithme 1

Entrée : un polynôme p et $z_0 \in \mathbb{C}$, t=1.
Sortie : z_{final} la valeur de z_t quand $t \to 0$.

1. Initialiser $z = z_0$, $h = p(z) - t\,p(z_0)$;
2. Tant que t>0
   Faire:
   i. On calcule $|\eta| \leqslant \left(\frac{\alpha_0}{\gamma} - \beta_t\right) / \|Dp(z)^{-1}\,p(z)\|$
   ii. $t = \max(t - \eta, 0)$
   iii. Calculer $h = p(z) - t\,p(z_0)$
   iv. $z = z - \frac{h(z)}{D\,h(z)}$
3. Retourner z



La certification obtenue :

- C'est un critère de convergence de l'opérateur de Newton.
- Apparu pour la 1ère fois en 1948 avec Kantorovich : sur des fonctions de classe $\mathcal{C}^2$ et des conditions sur $\|D^2 f\|$.
- 1968 $\rightarrow$ 1974, Ortega, Gragg et Tapia ont introduit la condition lipschitzienne pour des fonctions de classe $\mathcal{C}^1$ qui remplace les conditions sur les dérivées secondes, la version du théorème la plus connue.
- En 1993, Zheng Da Huang a introduit une extension de la version précédente sur des fonctions de classe $\mathcal{C}^2$, une version optimale pour les polynômes de degré 3.
- En 1999, Wang a introduit une forme générale de la constante lipschitzienne de la verison de Gragg ; alors pour des fonctions de classe $\mathcal{C}^1$.
- En 2013, Ezquerro, González et Hernández donnent une forme pour des ordres élevés; pour des fonctions de classe $\mathcal{C}^l, l \geqslant 2$.
- Dans « A short survey on Kantorovich-like theorems for Newton's method », on donne l'énoncé unificateur et une version générale pour des ordres élevés, puis, on montre que le cas général pour des fonctions de classe $\mathcal{C}^l, l \geqslant 2$ n'est qu'un cas particulier des classe $\mathcal{C}^1$ et alors n'existe pas une version plus générale. https://hal.archives-ouvertes.fr/hal-01196890
- On extrait l'α-théorie du cas général en introduisant la notion de série majorante.

MERCI POUR VOTRE ATTENTION