Autour de la méthode de Newton

Préparation à l'agrégation - option Calcul formel

Antoine Chambert-Loir

Résumé. — On expose dans ce petit cours la méthode de Newton et quelques avatars

1. La méthode de Newton en analyse

Soit I un intervalle de $\mathbf{R}$ et $f\colon I \to \mathbf{R}$ une application dérivable. Pour déterminer une approximation numérique des solutions de l'équation $f(t) = 0$, la méthode de Newton part d'une solution approchée x et remplace l'équation $f(t) = 0$ par l'équation approchée $f(x) + (t - x)f'(x) = 0$, d'où la solution

$$t = x - \frac{f(x)}{f'(x)}.$$

Bien entendu, cette formule n'a un sens que si $f'(x) \neq 0$. On espère que le nombre réel x_1 donné par cette équation est un peu plus proche d'une racine que ne l'était x. En tout cas, si t appartient encore à I, on peut recommencer en partant de t, etc., d'où une suite définie (pas forcément pour tout n) par la relation de récurrence :

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \qquad x_0 \in I.$$

Bien entendu, ce procédé d'approximation de l'équation vaut dans des contextes bien plus larges que celui d'une simple fonction de $\mathbf{R}$ dans $\mathbf{R}$. Pour l'analyse, on se place généralement dans le cas où f est une application dérivable d'un ouvert Ω d'un espace vectoriel normé E, disons complet, à valeurs dans E.[(1)]

On peut démontrer plusieurs types de résultats concernant la méthode de Newton :

(1) un énoncé de convergence locale : pour toute solution ξ de l'équation $f(\xi) = 0$ telle que $f'(\xi)$ soit inversible, il existe un voisinage U de ξ tel que pour tout $x_0 \in U$, la suite définie par la méthode de Newton soit définie pour tout entier et converge « quadratiquement » vers ξ.

(2) un énoncé de convergence un peu plus précis qui fournit, sous certaines hypothèses, une boule B telle que la suite donnée par la méthode de Newton soit définie pour tout premier terme dans B et converge (quadratiquement) vers un élément ξ de B qui est l'unique solution dans B de l'équation $f(x) = 0$.

(3) dans certains cas très particuliers, comme celui des fonctions convexes, ou la fonction $t \mapsto t^2 - a$, un énoncé de convergence globale, pour tout premier terme x_0.

[(1)] Certains théorèmes d'analyse comme celui de Nash-Moser requièrent même le cas de l'espace de Fréchet des fonctions $\mathscr{C}^\infty$...

Proposition 1.1. — *Soit E un espace vectoriel normé complet, soit Ω un ouvert de E et soit $f\colon \Omega \to E$ une application de classe $\mathscr{C}^2$. Soit ξ un élément de Ω tel que $f(\xi) = 0$ et $f'(\xi)$ soit inversible. Il existe alors un voisinage B de ξ tel que pour tout $x_0 \in B$, la suite (x_n) donnée par la relation de récurrence*

$$x_{n+1} = x_n - f'(x_n)^{-1}(f(x_n))$$

soit définie pour tout n et converge quadratiquement vers ξ au sens suivant : il existe un nombre réel $C > 1$ tel que pour tout $n \geqslant 0$,

$$\|x_n - \xi\| \leqslant C^{-2^n}.$$

Démonstration. — Comme f' est continue, que $f'(\xi)$ est inversible et que l'ensemble des éléments inversibles de $\mathscr{L}(E)$ est ouvert, il existe un nombre réel $r > 0$ tel que $f'(x)$ soit inversible pour tout $x \in B(\xi, r)$. Soit $F\colon B(\xi, r) \to E$ l'application définie par $F(x) = x - f'(x)^{-1}(f(x))$. Elle est de classe $\mathscr{C}^1$ sur B. Il se trouve que, quitte à diminuer r, F est contractante. Calculons en effet sa différentielle en ξ :

$$\begin{aligned} F(\xi + h) - F(\xi) &= F(\xi + h) - \xi = h - f'(\xi + h)^{-1}(f(\xi + h)) \\ &= h - (f'(\xi) + \mathrm{O}(\|h\|))^{-1}(f'(\xi)(h) + \mathrm{O}(\|h\|^2)) \\ &= h - f'(\xi)^{-1} \circ f'(\xi)(h) + \mathrm{O}(\|h\|^2) \\ &= \mathrm{O}(\|h\|^2). \end{aligned}$$

Soit donc C tel que $\|F(x) - \xi\| \leqslant C\,\|x - \xi\|^2$ pour tout $x \in B$. De cette inégalité, il vient $\|F(x) - \xi\| \leqslant \|x - \xi\|$ si $\|x - \xi\| < 1/C$. Ainsi, remplaçant r par $\min(r, 1/C)$, il vient que la suite (x_n) définie par la relation $x_{n+1} = F(x_n)$ est définie pour tout n si x_0 appartient à $B(\xi, r)$. En outre, l'inégalité $\|x_n - \xi\| \leqslant C^{-1}C^{-2^n}$ entraîne que

$$\|x_{n+1} - \xi\| \leqslant C^{-1}C^{-2^{n+1}}.$$

Par suite, si $\|x_0 - \xi\| \leqslant C^{-2}$, on aura $\|x_n - \xi\| \leqslant C^{-1-2^n}$ pour tout n. Autrement dit, la méthode de Newton converge quadratiquement vers ξ pour tout premier terme x_0 qui appartient à la boule de centre ξ et de rayon $1/C^2$, ce qui démontre la proposition. □

Voici un énoncé d'apparence plus générale, quoiqu'assez impraticable.

Théorème 1.2. — *Soit E un espace vectoriel normé complet, soit Ω un ouvert de E et soit $f\colon \Omega \to E$ une application de classe $\mathscr{C}^1$ qui vérifie les hypothèses suivantes :*

(i) *f' est γ-lipschitzienne ;*
(ii) *pour tout $x \in \Omega$, $f'(x)$ est inversible et $\|f'(x)^{-1}\| \leqslant \beta$.*

Pour tout $x_0 \in \Omega$ tel que

$$\alpha = \|f(x_0)\| \leqslant \frac{2d(x_0, \complement\Omega)}{\beta(2 + \beta\gamma d(x_0, \complement\Omega)},$$

la méthode de Newton avec premier terme x_0 converge vers un élément $\xi \in \Omega$ tel que $f(\xi) = 0$. En outre, pour tout $n \geqslant 0$,

$$\|x_n - \xi\| \leqslant \alpha\beta \frac{h^{2^n-1}}{1-h^{2^n}}, \quad \text{avec } h = \alpha\beta^2\gamma/2 < 1.$$

Démonstration. — La démonstration est hélas un peu technique, elle se trouve un peu partout : Chambert-Loir/Fermigier (vol. 2, p. 206), Kolmogorov/Fomine, Dieudonné, etc. □

En une variable, les fonctions convexes se comportent plutôt bien.

Proposition 1.3. — *Soit I un intervalle compact de $\mathbf{R}$, soit $f\colon I \to \mathbf{R}$ une fonction strictement croissante, convexe, de classe $\mathscr{C}^1$. Si x_0 est un point de I tel que $f(x_0) > 0$, la méthode de Newton avec premier terme x_0 fournit une suite (x_n) décroissante qui, ou bien est définie pour tout n et converge vers l'unique racine de f dans I, ou bien n'est pas définie pour tout n et f ne s'annule pas dans I.*

Démonstration. — Comme f est convexe, f' est croissante ; comme f est strictement croissante, $f' \geqslant 0$ et ne s'annule identiquement sur aucun intervalle. Par suite, on a $f'(x) > 0$ pour tout $x \in I$, sauf peut-être sa borne inférieure.

Si f est positive dans I, il est clair que la suite (x_n) est décroissante tant qu'elle est définie. Si elle converge, sa limite a appartient à I. On a

$$f'(a)a = \lim f'(x_n)x_{n+1} = \lim f'(x_n)(x_n - \frac{f(x_n)}{f'(x_n)}) = f'(a)a - f(a),$$

d'où $f(a) = 0$.

Sinon, soit $a \in I$ tel que $f(a) = 0$. On a $x_0 > a$ car f est croissante et $f(x_0) > 0$. Remarquons que le graphe de f est au-dessus de sa tangente en x, donc

$$0 = f(a) \geqslant f(x) + (a-x)f'(x)$$

et

$$x > x - \frac{f(x)}{f'(x)} > a.$$

Il en résulte que la suite (x_n) est décroissante, minorée par a, donc converge. Sa limite est nécessairement égale à a. □

Notons que cette proposition s'applique notamment aux polynômes P de coefficient dominant positif, sur l'intervalle $[b, +\infty[$, où b est la plus grande racine de P'. En pratique, si x_0 est un nombre réel supérieur à la plus grande racine réelle de P, la méthode de Newton avec premier terme x_0, si elle fournit une suite strictement décroissante, converge vers cette racine de P. Il suffit pour cela de connaître un majorant de la plus grande racine réelle de P, ce qui est fourni par le lemme élémentaire suivant.

Lemme 1.4. — *Soit $P = a_0 + a_1X + \cdots + a_nX^n$ un polynôme à coefficients réels. Toute racine complexe x de P vérifie l'inégalité*

$$|x| \leqslant \max\big(1, |a_n|^{-1} \sum_{k=0}^{n-1} |a_k|\big).$$

Démonstration. — Soit x une racine de P dans $\mathbf{C}$. L'inégalité est évidente si $|x| \leqslant 1$. Supposons donc $|x| \geqslant 1$. Alors, $a_nx = -\big(\sum_{k=0}^{n-1} a_kx^k\big)x^{1-n}$, d'où par l'inégalité triangulaire

$$|a_nx| \leqslant \sum_{k=0}^{n-1} |a_k|.$$

Le lemme est démontré ! □

Exercice 1.5. — a) Lorsque $f(x) = x^2 - a$, avec $a > 0$, la méthode de Newton conduit à la suite définie par la récurrence

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right).$$

Montrer qu'elle converge pour tout premier terme x_0 non nul vers la racine carrée de a de même signe que x_0.

b) De même, lorsque $f(x) = \dfrac{1}{x} - a$, avec $a \neq 0$, la méthode de Newton conduit à la récurrence

$$x_{n+1} = x_n(2 - ax_n).$$

Montrer qu'elle converge vers $1/a$ dès que x_0 vérifie l'inégalité $0 < ax_0 < 2$.

2. La méthode de Newton en algèbre

Contrairement aux apparences, et plus encore que l'analyse, l'algèbre permet de donner de très nombreux sens au mot « limite ». Par exemple, on peut convenir — et c'est ce qu'on fait lorsqu'on étudie les développements limités — qu'un polynôme est très petit lorsqu'il est divisible par une très grande puissance de l'indéterminée t. L'exercice apparemment anecdotique de calcul d'inverses par la méthode de Newton devient intéressant quand il permet de diviser sans division !

Proposition 2.1. — *Soit A un anneau (commutatif unitaire), soit $f \in A[X]$ de terme constant 1 et soit (g_n) la suite d'éléments de $A[X]$ définie par*

$$g_{n+1} = g_n(2 - fg_n), \qquad \text{avec } g_0 = 1.$$

Alors, $fg_n \equiv 1 \pmod{X}^{2^n}$ pour tout entier n.

Démonstration. — On a $fg_0 = 1 \equiv 1 \pmod{X}$; l'assertion est donc vraie pour $n = 0$. Supposons-la vraie pour n et soit h_n le polynôme défini par $h_n = (fg_n - 1)/X^{2^n}$. On a ainsi $fg_n = 1 + X^{2^n}h_n$, donc

$$fg_{n+1} = fg_n(2 - fg_n) = (1 + X^{2^n}h_n)(1 - X^{2^n}h_n) = 1 - X^{2^{n+1}}h_n^2 \equiv 1 \pmod{X^{2^{n+1}}}.$$

Cela démontre la proposition par récurrence. □

Cela fournit une méthode pratique pour calculer l'inverse d'un polynôme f de terme constant 1 modulo X^n en $\lceil \log_2(n) \rceil$ étapes, chacune requiérant deux multiplications de polynômes. On notera en outre qu'il suffit, pour mener à bien l'étape n, de calculer g_n modulo X^{2^n}.

Remarques 2.2. — 1) Par une méthode naïve, le calcul de l'inverse modulo X^n d'un polynôme f de terme constant 1 utilise n étapes, la première requérant n multiplications dans l'anneau de base, la seconde $n-1$, etc., soit de l'ordre de n^2 opérations.

Soit $M(n)$ le nombre de multiplications dans l'anneau de base requises par la multiplication de deux polynômes de degrés n. Par la méthode naïve, $M(n) = n^2$; avec les algorithmes efficaces (Karatsuba ou transformation de Fourier rapide) on a $M(n) = \mathrm{O}(n^{1+\varepsilon})$, ce qui permet d'accelérer significativement les calculs. La méthode de Newton permet alors le calcul de l'inverse en $\mathrm{O}(M(n))$: la division est à peine plus difficile qu'une multiplication.

2) Faut-il faire observer que cette méthode, jointe à une multiplication, permet le calcul d'une division de deux polynômes modulo une puissance de X, résultat de ce qu'on appelle classiquement « division suivant les puissances croissantes » ? Rappelons que cette division intervient dans les calculs de développements limités.

3) On passe de la division suivant les puissances croissantes à la division suivant les puissances décroissantes en changeant l'indéterminée en son inverse. Cette méthode devient alors utile pour calculer efficacement la division euclidienne de polynômes. Soit en effet a et b des éléments de $A[X]$, le polynôme b étant unitaire. Si $\deg a < \deg b$, la division euclidienne de a par b est triviale. Posons sinon $m = \deg(a) - \deg(b)$, ainsi que $a^* = X^{\deg(a)} a(1/X)$ et $b^* = X^{\deg(b)} b(1/X)$. Comme b est unitaire, on a $b^*(0) = 1$; soit f^* un polynôme de degré $\leqslant m$ tel que $f^* b^* \equiv 1 \pmod{X^{m+1}}$. Soit q^* le reste de la division euclidienne de $f^* a^*$ par X^{m+1} ; c'est un polynôme de degré au plus m. (Attention ! ce n'est pas une division euclidienne sérieuse, il s'agit juste de récupérer les termes de degré $\leqslant m$.) Alors, $f^* a^* \equiv q^* \pmod{X^{m+1}}$, d'où

$$a^* \equiv f^* a^* b^* \equiv q^* b^* \pmod{X^{m+1}}.$$

Par suite, $a^* - q^* b^*$ est divisible par X^{m+1}. Posons $r^* = (a^* - q^* b^*)/X^{m+1}$; c'est un polynôme de degré au plus $\deg(b) - 1$. Alors,

$$\begin{aligned} a(X) &= X^{\deg(a)} a^*(1/X) \\ &= X^{\deg(a)} \big(q^*(1/X) b^*(1/X) + X^{-m-1} r^*(1/X) \big) \\ &= X^m q^*(1/X) b(X) + X^{\deg(b)-1} r^*(1/X). \end{aligned}$$

Posons $q(X) = X^m q^*(1/X)$ $r(X) = X^{\deg(b)-1} r^*(1/X)$. Ce sont des polynômes de degrés au plus m et au plus $\deg(b) - 1$ respectivement tels que $a = bq + r$.

Ce qui précède concernait le calcul d'un inverse, c'est-à-dire la résolution de l'équation $ax = 1$, dans l'anneau $A[X]$ modulo une grande puissance de X. Cela s'applique

à d'autres équations, comme l'équation $x^2 = a$ et plus généralement toute équation polynomiale.

Exercice 2.3. — a) Écrire une procédure Maple (par exemple) qui, étant donnés un polynôme P et un entier n, renvoie un polynôme Q (de degré $\leqslant n/2$) tel que $Q^2 \equiv P \pmod{X^n}$.
b) Utiliser cette procédure pour retrouver le polynôme P, son carré étant donné.
c) Utiliser le changement de polynôme $P \leftrightarrow X^{\deg(P)}P(1/X)$ pour déterminer, un polynôme P étant donné, des polynômes Q et R de degrés $\leqslant \deg(P)/2$ tels que $P = Q^2 + R$.
d) Dans toutes ces questions, on pourra supposer que $P(0) = 1$. Expliquer néanmoins comment on pourrait se débarraser de cette hypothèse.

Exercice 2.4. — Voici une application arithmétique amusante. Soit P un polynôme à coefficients entiers, de terme dominant X^{2n}. On suppose que pour tout entier m assez grand, $P(m)$ est le carré d'un entier. Montrer que P est un carré. (Si $P = Q^2 + R$ comme ci-dessus, le terme dominant de Q étant X^n, montrer que $(Q(m) + 1)^2 > P(m) > (Q(m) - 1)^2$ pour m assez grand.)

Cela s'applique aussi dans d'autres anneaux, et notamment dans $\mathbf{Z}$. Il s'agit alors, une équation $P(x) = 0$ étant donnée, de trouver des nombres entiers n tels que $P(n)$ soit multiple d'une grande puissance d'un nombre premier fixé. Voyons-en un exemple.

Proposition 2.5. — *Soit p un nombre premier impair, soit a un nombre entier qui n'est pas multiple de p. Soit x_0 un entier tel que $x_0^2 \equiv a \pmod{p}$. Définissons une suite (x_n) par récurrence en posant*

$$x_{n+1} = \frac{1}{2}\big(x_n + y_n\big), \qquad \text{où } y_n \text{ est un entier tel que } y_n x_n \equiv a \pmod{p^{2^{n+1}}}.$$

Alors $x_{n+1} \equiv x_n \pmod{p^{2^n}}$ et $x_n^2 \equiv a \pmod{p^{2^n}}$ pour tout n.

Démonstration. — La démonstration est voisine de celle donnée pour la division. On a $x_0^2 \equiv a \pmod{p}$. Supposons que $x_n^2 \equiv a \pmod{p^{2^n}}$, soit r_n et s_n des entiers tels que $x_n^2 - a = p^{2^n} r_n$ et $x_n y_n - a = p^{2^{n+1}} s_n$. On a en particulier $x_n y_n \equiv x_n^2 \pmod{p^{2^n}}$. Par récurrence, $x_n \equiv x_0 \pmod{p}$, donc x_n est premier à p. On peut alors simplifier par x_n dans cette congruence, d'où $x_n \equiv y_n \pmod{p^{2n}}$. Par suite, $x_{n+1} \equiv x_n \pmod{p^{2^n}}$.

On a alors, modulo $p^{2^{n+1}}$,

$$\begin{aligned}
4x_{n+1}^2 x_n^2 &= x_n^4 + 2x_n^2 x_n y_n + x_n^2 y_n^2 \\
&= (a + p^{2^n} r_n)^2 + 2(a + p^{2^n} r_n)(a + p^{2^{n+1}} s_n) + (a + p^{2^{n+1}} s_n)^2 \\
&\equiv a^2 + 2p^{2^n} r_n a + 2a^2 + 2p^{2^n} r_n a + a^2 \\
&\equiv 4a^2 + 4p^{2^n} r_n a \\
&\equiv 4a(a + p^{2^n} r_n) = 4ax_n^2.
\end{aligned}$$

Comme x_n et p sont premiers entre eux, et que p est impair, on peut simplifier par $4x_n^2$ cette congruence et en déduire la relation $x_{n+1}^2 \equiv a \pmod{p^{2^{n+1}}}$ voulue. □

Remarque 2.6. — Lorsque $p = 2$, il faut partir de x_0 tel que $x_0^2 \equiv a \pmod 4$; un raisonnement analogue montre qu'alors $x_n^2 \equiv a \pmod{2p^{2^n}}$ pour tout n.

Toutefois, le calcul précédent est rendu malpratique du fait de la nécessité d'inverser x_n modulo $p^{2^{n+1}}$ à chaque étape. Quitte à perdre la convergence quadratique et à se rabattre vers une convergence linéaire, on peut supprimer la nécessité du calcul d'un inverse à chaque étape de la façon suivante. C'est le match Newton/Hensel.

Proposition 2.7. — *Soit f un polynôme à coefficients entiers, soit p un nombre premier et soit x_1 un entier tel que $f(x_1) \equiv 0 \pmod p$ mais $f'(x_1) \not\equiv 0 \pmod p$. Soit y_1 un entier tel que $f'(x_1)y_1 \equiv 1 \pmod p$. Alors, la suite (x_n) définie par la relation de récurrence*

$$x_{n+1} = x_n - y_1 f(x_n)$$

vérifie $x_{n+1} \equiv x_n \pmod{p^n}$ et $f(x_n) \equiv 0 \pmod{p^n}$ pour tout n.

On remarquera qu'il s'agit presque de la méthode de Newton, si ce n'est que le calcul de l'inverse de $f'(x)$ n'est fait qu'une fois avec une précision minimale.

Démonstration. — Commençons par observer que l'on a une formule de Taylor : si x et $h \in \mathbf{Z}$, alors $P(x+h) - P(x) - P'(x)h$ est multiple de h^2. Il suffit en effet de le vérifier lorsque $P(X)$ est un monôme X^n, auquel cas on a

$$(x+h)^n - x^n - nx^{n-1}h = \sum_{k=2}^{n} \binom{n}{k} x^{n-k}h^k = h^2 \sum_{k=2}^{n} \binom{n}{k} x^{n-k}h^{k-2}.$$

Montrons alors la proposition par récurrence sur n. On a bien $f(x_1) \equiv 0 \pmod p$; la définition de x_{n+1} montre que $x_{n+1} \equiv x_n \pmod{p^n}$ si $f(x_n) \equiv 0 \pmod{p^n}$. Supposons donc $f(x_n) \equiv 0 \pmod{p^n}$ et montrons que $f(x_{n+1}) \equiv 0 \pmod{p^{n+1}}$. Soit $r_n \in \mathbf{Z}$ tel que $f(x_n) = p^n r_n$. D'après la formule de Taylor,

$$f(x_{n+1}) \equiv f(x_n) - y_1 f(x_n) f'(x_n) \equiv f(x_n)(1 - y_1 f'(x_n)) \pmod{p^{2n}}.$$

Comme $x_n \equiv x_1 \pmod p$, on a aussi $f'(x_n) \equiv f'(x_1) \pmod p$ si bien que $y_1 f'(x_n) \equiv 1 \pmod p$. Comme $n \geqslant 1$, $2n \geqslant n+1$ et $f(x_{n+1})$ est multiple de p^{n+1}. □

Exercice 2.8. — a) Dans cet algorithme, x_n est en fait l'unique solution modulo p^n de l'équation $f(x) = 0$ qui est congrue à x_1 modulo p.
b) Cet algorithme, couplé avec la détermination des solutions modulo p, fournit une méthode pour déterminer les solutions *entières* d'une équation polynomiale $P(x) = 0$ à coefficients entiers. Il suffit en effet de calculer les solutions x modulo p^n où n est choisi de sorte que $p^n > 2M$, M étant un majorant des valeurs absolues des racines complexes. Soit le représentant $\tilde{x}$ de x dans l'intervalle $[-\frac{1}{2}p^n, \frac{1}{2}p^n]$ est racine de P, soit P n'a pas de racine congrue à x modulo p.

c) Étendre la méthode au cas où $f'(x_1)$ est multiple de p (mais non nul). Montrer que si $v_p(f(x_1)) \geqslant 1 + 2v_p(f'(x_1))$,[(2)] la relation de récurrence

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_1)}$$

définit une suite de nombre rationnels tels que pour tout n, $v_p(f(x_n)) \geqslant n + 2v_p(f'(x_1))$ et $v_p(x_{n+1} - x_n) \geqslant n + v_p(f'(x_1))$.

3. Résolution d'équations et factorisation de polynômes

J'expose dans ce dernier paragraphe la généralisation du précédent au cas d'équations polynomiales à plusieurs variables, avec en vue la factorisation des polynômes à coefficients entiers.

Soit en effet A un anneau, et considérons un polynôme de $A[X]$ de degré n, disons unitaire, que l'on souhaite factoriser en produit de deux polynômes unitaires P et Q de degrés p et q respectivement. On a donc $n = p + q$. Identifions un polynôme unitaire de degré p à la suite des p coefficients des termes de degrés 0, 1, …, $p - 1$. On s'intéresse ainsi à l'application

$$\mu\colon A^p \times A^q \to A^{p+q}, \quad (P, Q) \mapsto PQ. \tag{3.1}$$

C'est une application polynomiale ; sa *différentielle* — au sens formel, ou au sens usuel si $A = \mathbf{R}$ ou $\mathbf{C}$ — est l'application linéaire de $A^p \times A^q$ dans A^{p+q} donnée par la formule

$$d\mu(P, Q)\colon A^p \times A^q \to A^{p+q}, \quad (U, V) \mapsto UQ + VP. \tag{3.2}$$

En effet, on a

$$\mu(P + U, Q + V) = (P + U)(Q + V) = PQ + (UQ + VP) + UV$$

et le terme UV est quadratique en (U, V), donc $\mathrm{O}(\|U\| + \|V\|)$ dans le cas $A = \mathbf{R}$ ou $\mathbf{C}$, et négligé dans le cas général : ses dérivées partielles s'annulent en $(U, V) = (0, 0)$. La définition de la différentielle d'une application polynomiale fournit bien la formule voulue. La matrice de $d\mu(P, Q)$ dans les bases canoniques de $A^p \times A^q$ et A^{p+q} n'est autre que la matrice de Sylvester des polynômes P et Q (en degrés (p, q)). Autrement dit, le déterminant de cette application linéaire dans ces bases est le *résultant* en degrés (p, q) des polynômes P et Q ; notons-le $\mathrm{Res}_{p,q}(P, Q)$.

Par définition, il s'annule si et seulement s'il existe des polynômes non nuls U et V de degrés inférieurs à p et q respectivement tels que $UQ = VP$. Supposons que A soit un corps ou, plus généralement, un anneau factoriel ; alors $A[X]$ est un anneau factoriel et les polynômes P et Q ont un pgcd dans $A[X]$. Considérons deux tels polynômes et divisons cette égalité par le pgcd D de P et Q ; on obtient $UQ_1 = VP_1$, où $P = DP_1$ et $Q = DQ_1$. D'après le lemme de Gauss, P_1 divise U et Q_1 divise V ;

(2) Rappelons que $v_p(x)$ désigne la *valuation p-adique* d'un nombre rationnel non nul x ; c'est l'unique entier relatif n tel que l'on puisse écrire $x = p^n a/b$, où a et b sont des entiers relatifs premiers à p.

on écrit ainsi $U = P_1 W$ et $V = Q_1 W$. Si $\deg(D) > 0$, c'est-à-dire si P et Q ne sont pas premiers entre eux, $(U, V) = (P_1, Q_1)$ est solution. Si $\deg(D) = 0$, la relation $U = P_1 W$ entraîne que $\deg(U) \geqslant p$ ou $U = 0$. Autrement dit, $\mathrm{Res}_{p,q}(P, Q) = 0$ si et seulement si P et Q ont un facteur commun dans $A[X]$.

Si A est un anneau intègre, on peut raisonner dans le corps des fractions K de A ; on voit ainsi que $\mathrm{Res}(P, Q) = 0$ si et seulement si P et Q ont un facteur commun dans $K[X]$.

Cela équivaut à ce que les polynômes P et Q aient une racine commune dans une extension de K.

La première conséquence de ce calcul est l'applicabilité du théorème des fonctions implicites lorsque $A = \mathbf{R}$ ou $\mathbf{C}$ et que P et Q n'ont pas de racine commune dans $\mathbf{C}$.

Proposition 3.3. — *Soit P et Q des polynômes unitaires à coefficients complexes de degrés p et q, sans racine commune dans $\mathbf{C}$. Il existe un voisinage U de P dans $\mathbf{C}^p$, un voisinage de Q dans $\mathbf{C}^q$ et un voisinage W de PQ dans $\mathbf{C}^{p+q}$ tels que μ induise un difféomorphisme (holomorphe) de $U \times V$ dans W.*

En particulier, tout polynôme unitaire R_1 proche de $R = PQ$ a une unique factorisation $R_1 = P_1 Q_1$ où (P_1, Q_1) est proche de (P, Q).

Revenons à l'algèbre. Voici la variante du théorème d'inversion locale dont nous avons besoin.

Théorème 3.4. — *Soit $F = (F_1, \ldots, F_n)$ une famille de n polynômes en n indéterminées $X_1, \ldots, X_n$. Soit I un idéal de A et soit $(a_1, \ldots, a_n) \in A^n$ tel que*

$$F(a_1, \ldots, a_n) \equiv 0 \pmod{I}, \qquad dF(a_1, \ldots, a_n) \in \mathrm{GL}_n(A/I).$$

Pour tout k, il existe alors un unique élément $(b_1, \ldots, b_n) \in (A/I^k)$ tel que $a_i \equiv b_i \pmod{I^k}$ et $F(b_1, \ldots, b_n) \equiv 0 \pmod{I^k}$.

Démonstration. — Fixons une matrice $U \in \mathrm{M}_n(A)$ telle que $U \cdot dF(a_1, \ldots, a_n) = dF(a_1, \ldots, a_n) \cdot U = \mathrm{I}_n \pmod{I}$. Démontrons alors le théorème par récurrence sur k ; il n'y a rien a faire pour $k = 1$. Il suffit de montrer que si $(a_1, \ldots, a_n) \in A^n$ vérifie $F(a_1, \ldots, a_n) \equiv 0 \pmod{I}^k$ et $dF(a_1, \ldots, a_n)$ est inversible dans $\mathrm{M}_n(A/I)$, alors il existe un unique élément $b = (b_1, \ldots, b_n) \in (A/I^k)^n$ congru à $(a_1, \ldots, a_n)$ modulo I^k tel que $F(b_1, \ldots, b_n) \equiv 0 \pmod{I^{k+1}}$.

La formule de Taylor, pour $h = (h_1, \ldots, h_n)$, s'écrit

$$F(X_1 + Y_1, \ldots, X_n + Y_n) = F(X_1, \ldots, X_n) + dF(X_1, \ldots, X_n) \cdot (Y_1, \ldots, Y_n) + G(X_1, \ldots, X_n; Y_1, \ldots, Y_n)$$

où G est un polynôme en $2n$ variables $X_1, \ldots, X_n, Y_1, \ldots, Y_n$ à coefficients dans A, dont chaque monôme est de degré au moins 2 en $Y_1, \ldots, Y_n$.

Cherchons b sous la forme $a + h$, avec $h \in I^k$; on a ainsi

$$F(a + h) = F(a) + dF(a)(h) \pmod{I^{2k}}.$$

Multiplions cette relation par U. On obtient ainsi $h = -UF(a) \pmod{I}^{k+1}$, d'où l'existence et l'unicité de h modulo I^{k+1}. □

Appliqué au problème de la factorisation modulo une puissance d'un nombre premier, le théorème 3.4 devient :

Corollaire 3.5. — *Soit T un polynôme unitaire à coefficients entiers, soit P_1 et Q_1 des polynômes unitaires de degrés p et q respectivement, à coefficients entiers, et unitaires tels que $T = P_1Q_1 \pmod{\ell}$ et $\mathrm{Res}(P_1, Q_1) \neq 0 \pmod{\ell}$. Pour tout entier $k \geqslant 1$, il existe alors un couple (P_k, Q_k) de polynômes unitaires à coefficients entiers, de degrés p et q, unique modulo ℓ^k, congru à (P_1, Q_1) modulo ℓ tels que $T = P_kQ_k \pmod{\ell^k}$.*

Grâce aux bornes de Mignotte (prop. 4.2), cela permet ou bien de trouver une factorisation de T dans $\mathbf{Z}[X]$, ou bien de prouver qu'il n'en existe pas de factorisation $T = PQ$ telle que $(P, Q) \equiv (P_1, Q_1) \pmod{\ell}$.

Exercice 3.6. — On se place sous les hypothèses du théorème 3.4.
a) Montrer que $dF(a)$ est inversible modulo I^k.
b) Construire cet inverse par la méthode de Newton.
c) Expliciter la méthode de Newton dans le cas de l'équation $F(x) = 0$ et montrer la convergence quadratique attendue.

4. Les bornes de Mignotte

Si $P = a_0X^d + \cdots + a_d$ est un polynôme à coefficients complexes, de racines $z_1, \ldots, z_d$, répétées suivant leur multiplicité, on pose

$$M(P) = |a_0| \prod_{j=1}^{d} \max(1, |z_j|) \quad \text{et} \quad \|P\|_p = \left(\sum_{j=0}^{d} |a_j|^p \right)^{1/p},$$

pour $1 \leqslant p \leqslant \infty$.

Lemme 4.1. — *On a les inégalités*

$$2^{-d} \|P\|_1 \leqslant M(P) \leqslant \|P\|_2 .$$

Démonstration. — Les formules coefficients racines entraînent

$$a_j = (-1)^{d-j} a_0 \sum_{i_1 < \cdots < i_j} z_{i_1} \ldots z_{i_j},$$

d'où

$$|a_j| \leqslant |a_0| \sum_{i_1 < \cdots < i_j} |z_{i_1}| \ldots \left|z_{i_j}\right|,$$

d'où finalement

$$\|P\|_1 = \sum_{j=0}^{d} |a_j| \leqslant |a_0| \prod_{j=1}^{d} (1 + |z_j|) \leqslant |a_0| \prod_{j=1}^{d} (2 \max(1, |z_j|)) = 2^d M(P).$$

Démontrons l'autre inégalité. Notons $z_1, \dots, z_k$ les racines de P de valeur absolue au moins 1, et $z_{k+1}, \dots, z_d$ les autres. Soit Q le polynôme défini par

$$Q(X) = a_0 \prod_{j=1}^{k} (\bar{z}_j X - 1) \prod_{j=k+1}^{d} (X - z_j).$$

D'après la formule de Parseval,

$$\|P\|_2^2 = \frac{1}{2\pi} \int_0^{2\pi} \left|P(e^{i\theta})\right|^2 \, d\theta.$$

Observons que pour $\theta \in \mathbf{R}$, $\bar{z}_j e^{i\theta} - 1$ et $e^{i\theta} - z_j$ ont même module, puisque

$$\overline{e^{i\theta} - z_j} = e^{-i\theta} - \bar{z}_j = -e^{-i\theta}(\bar{z}_j e^{i\theta} - 1).$$

Par suite, $\left|Q(e^{i\theta})\right| = \left|P(e^{i\theta})\right|$ pour tout $\theta \in \mathbf{R}$; en particulier, $\|Q\|_2 = \|P\|_2$.

Posons $Q = b_0 X^d + \cdots + b_d$. On a

$$b_0 = a_0 \prod_{j=1}^{k} \bar{z}_j,$$

donc $|b_0| = M(P)$. Comme $|b_0| \leqslant \|Q\|_2$, on en déduit l'inégalité $M(P) \leqslant \|P\|_2$. □

Proposition 4.2. — *Soit $P = a_0 X^d + \cdots + a_d$ un polynôme à coefficients entiers et soit $Q = b_0 X^q + \cdots + b_q$ un diviseur de Q dans $\mathbf{Z}[X]$. Alors, $\|Q\|_1 \leqslant 2^q \|P\|_2$.*

Démonstration. — L'entier b_0 divise a_0, donc $|a_0| \leqslant |b_0|$. En outre, toute racine de Q étant une racine de P, $M(Q)/|b_0| \leqslant M(P)/|a_0|$, d'où $M(Q) \leqslant M(P)$. Alors,

$$\|Q\|_1 \leqslant 2^q M(Q) \leqslant 2^q M(P) \leqslant 2^q \|P\|_2 .$$

□

Exercice 4.3. — a) Calculer l'intégrale

$$I(a) = \int_0^{2\pi} \log \left|e^{i\theta} - a\right| \, d\theta$$

en fonction du nombre complexe a. (On trouve 0 si $|a| < 1$, $2\pi \log |a|$ sinon.)
b) En déduire la formule

$$\log M(P) = \frac{1}{2\pi} \int_0^{2\pi} \log \left|P(e^{i\theta})\right| \, d\theta.$$