COURS 15

Localisation de racines de polynômes[1]

Résumé

Dans le chapitre qui suit, le problème étudié est celui de la localisation des racines d'un polynôme, qu'il soit réel ou complexe. Selon les deux cas possibles, les méthodes ne seront pas les mêmes. Dans le cas complexe (plus général en un sens), on donnera d'abord un algorithme performant pour calculer la valeur d'une racine, puis nous présenterons quelques résultats théoriques de localisation. En revanche, lorsque le polynôme sera supposé réel, l'accent sera porté sur le nombre de racines du polynôme dans un intervalle prescrit, problème que circonscrit très convenablement le théorème de Sturm.

1. La méthode de Newton

On se donne un polynôme $f \in \mathbb{C}[X]$, et on définit, pour x n'annulant pas f', le nombre $N_f(x) = x - \frac{f(x)}{f'(x)}$. Modulo une hypothèse de stabilité, le résultat suivant, connu sous le nom de *méthode de Newton*, montre qu'il est facile d'approcher une racine simple de f grâce à N_f. Plus précisément :

THÉORÈME 22. *Si ζ est une racine simple de f, et $r > 0$ est tel que*

$$r \cdot \sup_{|x-\zeta|\leq r} |f''(x)| \leq \frac{1}{2}|f'(\zeta)|,$$

alors pour tout x_0 tel que $|x_0 - \zeta| \leq r$, la suite $(x_k)_{k\geq 0}$ définie par l'itération $x_{k+1} = N_f(x_k)$ converge vers ζ avec vitesse quadratique :

$$|x_k - \zeta| \leq \left(\frac{1}{2}\right)^{2^k-1} |x_0 - \zeta|.$$

DÉMONSTRATION. Il y a deux étapes :

1. Commençons par montrer que f' ne s'annule pas sur $D(\zeta, r)$, ce qui justifiera l'existence de la suite $(x_k)_{k\in\mathbb{N}}$. Pour cela, il suffit d'écrire, pour $x \in D(\zeta, r)$:

$$|f'(x) - f'(\zeta)| = \left|\int_\zeta^x f''(t)dt\right| \leq \frac{|x-\zeta|}{2r}|f'(\zeta)| \leq \frac{|f'(\zeta)|}{2}$$

   et donc $|f'(x)| \geq \frac{|f'(\zeta)|}{2} > 0$ ce qui règle le premier point.

[1] La première rédaction de ce chapitre est due à Henri Guenancia.

2. Quant à la convergence, on va procéder par récurrence sur k, l'initialisation étant bien sûr vérifiée. Alors, il s'agit seulement de bien exploiter la formule de Taylor (celle avec reste intégral). En effet, pour tout $x \in D(\zeta, r)$ on a
$$0 = f(\zeta) = f(x + (\zeta - x)) = f(x) + (\zeta - x)f'(x) + \int_x^\zeta (\zeta - t)f''(t)dt.$$
Cela se réécrit encore :
$$x - \frac{f(x)}{f'(x)} - \zeta = \frac{1}{f'(x)} \int_x^\zeta (\zeta - t)f''(t)dt$$
et donc
$$|N_f(x) - \zeta| \leq \frac{|f'(\zeta)|}{|f'(x)|} \frac{1}{2r} \frac{|\zeta - x|^2}{2} \leq \frac{1}{2r}|\zeta - x|^2.$$
Alors, on termine en écrivant :
$$|x_{k+1} - \zeta| \leq \frac{1}{2r} \left(\frac{1}{2}\right)^{2^{k+1}-2} \times 2|x_0 - \zeta|^2 \leq \left(\frac{1}{2}\right)^{2^{k+1}-1} |x_0 - \zeta|.$$

□

Cette méthode est donc performante (elle converge vite, et demande finalement assez peu de calculs), d'autant plus qu'elle est largement généralisable à des classes de fonctions bien plus larges que les polynômes. Cependant, un des inconvénients est qu'il faut trouver une borne *a priori* qui garantisse la validité des calculs. Cela requiert donc en particulier une localisation plus ou moins bonne de la racine étudiée ; c'est l'objet de la section qui suit.

2. Distance aux racines

2.1. Borne de Cauchy. Soit $P = X^n + a_{n-1}X^{n-1} + \cdots + a_0 \in \mathbb{C}[X]$. Le résultat que l'on commence par présenter montre que pour majorer le module des racines de P, il suffit de se ramener à un polynôme réel :

Théorème 23. *Si au moins un des a_i est non-nul, alors toutes les racines de P sont majorées en module par l'unique racine strictement positive de*
$$Q(x) = x^n - \sum_{i=0}^{n-1} |a_i|x^i.$$

Démonstration. Tout d'abord, l'écriture $\frac{Q(x)}{x^n} = 1 - f(x)$ où $f : \mathbb{R}_+^\star \to \mathbb{R}$ est une fonction continue positive décroissant strictement de $+\infty$ à 0, montre que Q admet en effet une seule racine strictement positive $r > 0$.

L'inégalité triangulaire entraîne $|P(z) - z^n| \leq \sum_{i=0}^{n-1} |a_i| \cdot |z|^i$ pour tout $z \in \mathbb{C}$. Il s'ensuit que pour toute racine $\zeta \in \mathbb{C}^\star$ de P on a $\sum_{i=0}^{n-1} |a_i| \cdot |\zeta|^i \geq |\zeta|^n$ et donc $Q(|\zeta|) \leq 0$, ou encore $f(|\zeta|) \geq 1 = f(r)$. Puisque f est strictement décroissante, cela permet de conclure que $|\zeta| \leq r$. □

Dans le cas où 0 n'est pas racine de P, on peut appliquer le résultat précédent au polynôme réciproque $X^nP(1/X)$ et obtenir ainsi une borne inférieure sur le module des racines de P.

Corollaire 1. *Si $P(0) \neq 0$, le module de la racine de P la plus proche de* 0 *est minoré par l'unique racine positive du polynôme $|a_0| - \sum_{i=1}^{n-1} |a_i| x^i - x^n$.*

COROLLAIRE 2 (Borne de Cauchy). *Toutes les racines du polynôme* $P(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_0 \in \mathbb{C}[X]$ *sont contenues dans le disque centré à l'origine et de rayon* $1 + \max_{1 \leq i \leq n}(|a_i|)$.

DÉMONSTRATION. Soit A le maximum des modules des coefficients a_i. Pour $s > 1 + A$ on a les majorations successives

$$|a_0| + \cdots + |a_{n-1}|s^{n-1} < 1 + A + sA + \cdots + s^{n-1}A = 1 + A \cdot \frac{s^n - 1}{s - 1} < s^n.$$

Cela entraîne que l'unique racine strictement positive $r > 0$ du polynôme $X^n - |a_{n-1}|X^{n-1} - \cdots - |a_0|$ est bornée par $1 + A$. Le Théorème 23 permet alors de conclure. □

COROLLAIRE 3 (Borne de Eneström et Kakeya). *Si tous les coefficients du polynôme* $R(X) = a_nX^n + a_{n-1}X^{n-1} + \cdots + a_0 \in \mathbb{R}[X]$ *sont strictement positifs, alors les racines de* R *sont toutes contenues dans la couronne* $\alpha \leq |z| \leq \beta$, *où*

$$\alpha := \min_{1 \leq k \leq n} \left(\frac{a_{k-1}}{a_k}\right), \qquad \beta := \max_{1 \leq k \leq n} \left(\frac{a_{k-1}}{a_k}\right).$$

DÉMONSTRATION. Le polynôme $P(X) = (X - \beta)/a_n \cdot R(X)$ s'écrit $P(X) = X^{n+1} - c_nX^n - \cdots - c_0$, où $c_k = (\beta a_k - a_{k-1})/a_n$ pour $k \geq 1$ et $c_0 = \beta a_0/a_n$. L'hypothèse entraîne que tous les c_k sont positifs et $c_0 > 0$. Le réel β étant l'unique racine réelle de P, le Théorème 23 montre alors que toutes les autres racines de P – c'est-à-dire les racines de R – ont un module au plus égal à β. La borne inférieure s'obtient en considérant le polynôme réciproque de P. □

2.2. Bornes a posteriori. Le résultat que nous allons présenter maintenant est tout à fait non banal en première approche, et pourtant, sa preuve est plutôt brève.

THÉORÈME 24. *Soit* $P \in \mathbb{C}[X]$ *un polynôme unitaire de degré* n, *et* $z_0 \in \mathbb{C}$ *un nombre complexe quelconque. Alors :*

1. *Le disque fermé* $\overline{D}(z_0, |P(z_0)|^{1/n})$ *contient au moins une racine de* P.
2. *Si de plus* $P'(z_0) \neq 0$, *alors le disque* $\overline{D}\left(z_0, n\left|\frac{P(z_0)}{P'(z_0)}\right|\right)$ *aussi.*

REMARQUE 1. Si $P = X^n$, alors la première borne est atteinte pour tout z_0.

DÉMONSTRATION. On peut tout d'abord se ramener au cas où $z_0 = 0$, par exemple en considérant $P_{z_0}(X) = P(X - z_0)$. Adoptons quelques notations : soient ζ_i les racines de P, et $\rho = \min_i |\zeta_i|$. On peut tout de suite écarter le cas où $\rho = 0$ dans lequel les deux assertions sont bien vérifiées. Ainsi, le polynôme réciproque de P, égal à $X^nP(\frac{1}{X}) = a_0X^n + a_1X^{n-1} + \cdots + 1$ a pour racines les $\frac{1}{\zeta_i}$, et donc grâce aux relations entre coefficients et racines, on peut écrire :

$$\left|\frac{a_k}{a_0}\right| = \left|\sum_{1 \leq i_1 < \cdots < i_k \leq n} \frac{1}{\zeta_{i_1} \cdots \zeta_{i_k}}\right| \leq C_n^k \rho^{-k}.$$

Dans l'hypothèse où $a_k \neq 0$, on a donc $\rho^k \leq C_n^k \frac{|a_0|}{|a_k|}$. D'où les deux cas :

– Si $k = n$, alors $\rho^n \leq |a_0| = |P(0)|$.
– Si $k = 1$ et $P'(0) = a_1$ est non nul, alors $\rho \leq n\frac{|P(0)|}{|P'(0)|}$.

La preuve est donc maintenant complète. □

3. Comptage de racines

3.1. Indice de Cauchy. Soit f une fraction rationnelle dans $\mathbb{R}(X)$.

DÉFINITION 1. *L'indice de Cauchy de f en t, noté $I_{t_-}^{t^+}$ vaut :*
- 0 *si* $f(t_-) = f(t^+)$,
- 1 *si* $f(t_-) = -\infty$ *et* $f(t^+) = +\infty$,
- -1 *si* $f(t_-) = +\infty$ *et* $f(t^+) = -\infty$.

L'indice de Cauchy de f sur l'intervalle $[a, b]$ vaut

$$I_a^b := \sum_{t \in [a,b]} I_{t_-}^{t_+}$$

PROPOSITION 1. *Si $P \in \mathbb{R}[X]$ est tel que $P(a)P(b) \neq 0$, alors le nombre de racines de P dans l'intervalle $[a, b]$ vaut $I_a^b\left(\frac{P'}{P}\right)$.*

DÉMONSTRATION. La preuve repose sur l'égalité $\frac{P'(x)}{P(x)} = \sum_{P(x_i)=0} \frac{m_i}{x - x_i}$, où m_i est la multiplicité de la racine $x_i \in [a, b]$ de P. Alors, comme $m_i > 0$, en chaque x_i, on est dans le second point de la définition, et la conclusion s'ensuit. □

3.2. Changements de signe. Si $(a_0, \dots, a_n) \in \mathbb{R}^{n+1}$, on note $v(a_0, \dots, a_n)$ son nombre de changements de signe, c'est à dire le nombre d'indices i tels qu'il existe $k > 0$ vérifiant $a_i a_{i+k} < 0$ et $a_{i+j} = 0$ pour tout $j \in \{1, \dots, k-1\}$.

On se convaincra aisément que cela correspond bien à la notion intuitive de nombre de changements de signe (pour une famille finie de réels) qu'on peut avoir.

3.3. Suites de Sturm.

DÉFINITION 2. *Une suite finie $(P_0, \dots, P_n)$ de polynômes réels est une suite de Sturm pour l'intervalle $[a, b]$ si :*

1. $P_0(a)P_0(b) \neq 0$,
2. $\forall x \in [a, b], P_n(x) \neq 0$,
3. *Si, pour $x \in [a, b]$, $P_k(x) = 0$, alors soit $P_{k-1}(x)P_{k+1}(x) < 0$ (si $k > 0$), soit $P_1(x) \neq 0$ (si $k = 0$).*

Le théorème suivant, démontré par Charles Sturm en 1829, s'énonce très clairement grâce aux notations introduites :

THÉORÈME 25. *Si $(P_0, \dots, P_n)$ est une suite de Sturm pour $[a, b]$, alors $I_a^b\left(\frac{P_1}{P_0}\right) = v((P_i(a))_i) - v((P_i(b))_i)$.*

DÉMONSTRATION. Tout d'abord, remarquons que $v((P_k(t))_k)$ ne peut changer qu'en un t racine d'un P_k.
D'autre part, si $P_k(t) = 0$, et $k > 0$, alors $P_{k-1}(t)P_{k+1}(t) < 0$, donc $v((P_k(t))_k)$ ne peut changer qu'en une racine de P_0.
Reste à étudier le cas de P_0. On se place donc autour d'une racine t de P_0. Si le signe ϵ_- de $P_0(t_-)$ est le même que ϵ_+ celui de $P_0(t_+)$, alors $v((P_k(x))_k)$ n'en change pas en $x = t$, et on a aussi $I_{t_-}^{t^+}\left(\frac{P_1}{P_0}\right) = 0$ grâce à la propriété 3 vérifiée par une suite de Sturm.

Si $(\epsilon_-, \epsilon_+) = (-, +)$ et $P_1(t) > 0$, alors $I_{t_-}^{t^+}\left(\frac{P_1}{P_0}\right) = 1$.

Si $(\epsilon_-, \epsilon_+) = (-, +)$ et $P_1(t) < 0$, alors $I_{t_-}^{t^+}\left(\frac{P_1}{P_0}\right) = -1$.
On obtient de même les résultats attendus dans les cas où $(\epsilon_-, \epsilon_+) = (+, -)$, ce qui conclut quant à la démonstration du théorème de Sturm : en effet, on a vérifié qu'en chaque point où pouvait changer la valeur de $v((P_k(t))_k)$, celle de $I_{t_-}^{t^+}\left(\frac{P_1}{P_0}\right)$ changeait dans les mêmes quantités, et donc le résultat global s'obtient en passant à la somme des indices de Cauchy en les racines de P_0. □

Voyons maintenant quelle application concrète on peut faire de ce théorème en ce qui concerne le nombre de racines d'un polynôme.

3.4. Racines d'un polynôme dans un intervalle. Rappelons d'abord ce que l'on entend par algorithme d'Euclide signé. Étant donnés deux polynômes P_0 et P_1, on peut appliquer l'algorithme d'Euclide habituel, mais en introduisant un signe dans les restes comme suit :

$$\begin{aligned} P_0 &= Q_1 P_1 - P_2 \\ P_1 &= Q_2 P_2 - P_3 \\ &\vdots \\ P_{m-2} &= Q_{m-1} P_{m-1} - P_m \end{aligned}$$

avec $P_m = \mathrm{pgcd}(P_0, P_1)$.
L'algorithme décrit ci-dessus fournissant la suite $(P_i)_{2 \leq i \leq m}$ à partir de la donnée (P_0, P_1) s'appelle *l'algorithme d'Euclide signé*. Voici un résultat, de vérification très élémentaire, nous donnant un algorithme (s'appuyant de manière essentielle sur celui qu'on vient de décrire) pour construire des suites de Sturm :

Théorème 26. *Soient P_0 un polynôme réel, $a < b$ deux réels en lesquels P_0 ne s'annule pas, et $P_1 := P_0'$. Alors les polynômes $(P_0, P_1, \ldots, P_m)$ construits par l'algorithme d'Euclide signé sont tels que la suite $(\frac{P_0}{P_m}, \frac{P_1}{P_m}, \ldots, 1)$ est une suite de Sturm pour l'intervalle $[a, b]$.*

En considérant conjointement la proposition 1 ainsi que les théorèmes 25 et 26, on obtient le

Corollaire 4. *Les suites de Sturm permettent de calculer le nombres de racines distinctes dans un intervalle.*

Insistons bien sur le fait qu'avec ce résultat, on ne prend pas en compte la multiplicité des racines.

Voyons tout de suite un exemple pour comprendre la puissance de ces résultats :

Exemple 1. On prend $P = 2x^3 - 7x^2 + 3x - 2$, alors on dispose de la suite

$$\begin{aligned} P_0 &= 2x^3 - 7x^2 + 3x - 2 \\ P_1 &= 6x^2 - 14x + 3 \\ P_2 &= 62x + 15 \\ P_3 &= -1. \end{aligned}$$

On choisit $a = 0$ et $b = +\infty$ (ce n'est pas gênant, il suffit en fait de prendre b plus grand que tous les zéros des polynômes intervenant dans la suite). Alors, v vaut 2 en 0 et 1 en l'infini. On en déduit que P ne possède qu'une seule racine réelle positive.

Il ne faudrait cependant pas croire que la théorie s'arrête là : il y a beaucoup d'autres configurations (c'est-à-dire différentes de la recherche de racines de polynômes réels sur un intervalle) où on peut appliquer les résultats de ce chapitre. Nous en proposons une pour élargir un peu notre horizon, c'est l'objet de la partie qui suit.

3.5. Nombre de racines dans un demi-plan. On se donne $P \in \mathbb{C}[X]$ unitaire, et on décompose P en $P = R + iS$ avec $R, S \in \mathbb{R}[X]$. On peut alors montrer qu'il existe un lien entre le nombre de racines réelles et celui de racines dans le demi-plan supérieur. L'énoncé suivant rend ce lien explicite :

THÉORÈME 27. *Soient k le nombre de racines réelles de P, m le nombre de racines de P dans $\mathcal{H} = \{z \in \mathbb{C}; \Im(z) > 0\}$, et n le degré de P. Alors,*

$$m = \frac{1}{2}\left(n - k - I_{-\infty}^{+\infty}\left(\frac{S}{R}\right)\right).$$

3.6. Théorème de Sylvester-Hermite.

THÉORÈME 28. *Soit $P \in \mathbb{R}[X]$ de degré n ayant les racines $\zeta_1, \ldots, \zeta_n$ supposées deux à deux distinctes. Pour $k \geq 0$ on définit les sommes de Newton $N_k = \zeta_1^k + \cdots + \zeta_n^k$. Alors, le nombre de racines réelles de P est égal à la signature de la forme quadratique associée à la matrice symétrique réelle*

$$M = \begin{bmatrix} N_0 & N_1 & \cdots & N_{n-1} \\ N_1 & N_2 & \cdots & N_n \\ \vdots & \vdots & \cdots & \vdots \\ N_{n-1} & N_n & \cdots & N_{2n-2} \end{bmatrix}$$

DÉMONSTRATION. Soit $F(x_1, \ldots, x_n)$ la forme quadratique $y_1^2 + \cdots + y_n^2$, où $y_r = x_1 + \zeta_r x_2 + \cdots + \zeta_r^{n-1} x_n$ pour $1 \leq r \leq n$. Les coefficients de F étant symétriques en les racines ζ_i, ils sont réels. La forme F peut donc se représenter comme

$$h_1^2 + \cdots + h_p^2 - h_{p+1}^2 - \cdots - h_n^2, \tag{1}$$

où les h_i sont des formes linéaires en les ζ_i, à coefficients réels. La matrice de F est exactement la matrice M.

Aux racines réelles ζ_j de P correspondent des formes réelles y_j, aux racines purement imaginaires conjuguées $\zeta_k, \overline{\zeta}_k$ correspondent des formes complexes conjuguées :

$$y_k^2 + \overline{y}_k^2 = (\lambda_k + i\mu_k)^2 + (\lambda_k - i\mu_k)^2 = 2\lambda_k^2 - 2\mu_k^2.$$

Donc, le nombre $n - p$ de carrés portant un signe négatif dans (1) est égal au nombre de différents couples de racines purement imaginaires conjuguées de P. Il s'ensuit que le nombre de racines réelles de P vaut $n - 2(n - p)$, ce qui est égal à la signature de (1). $\square$