Preuve du théorème fondamental de l'algèbre par les suites de Sturm

L. Dietrich et S. Billouet

15 décembre 2009

Table des matières

1 Introduction

Le théorème fondamental de l'algèbre a été énoncé pour la première fois par Jean d'Alembert au dix-huitième siècle :

Théorème :

Tout polynôme de degré $n \geq 1$ sur $\mathbb{C}$ admet n racines.

Diverses preuves sont connues à ce jour, utilisant :

- de l'analyse : preuve classique par compacité, par intégration, ou via les fonctions analytiques
- de l'algèbre : théorème des valeurs intermédiaires, théorie de Galois
- de la topologie algébrique.

FIGURE 1 – Jean d'Alembert – 1717-1783

La preuve que nous allons développer présente plusieurs intérêts :

- elle est valable sur l'extension algébrique de tout corps réel clos (un corps réel clos est un corps ordonné sur lequel tout polynôme vérifie le théorème des valeurs intermédiaires) ;
- elle est purement algébrique, au sens où elle ne nécessite pas de notions de logique du second ordre comme la compacité ;
- elle est constructive et implémentable par un algorithme qui permet de localiser les racines.

2 Préliminaires historiques – le cas réel

FIGURE 2 – Charles-François Sturm – 1803-1855

2.1 Suites de Sturm et indice de Cauchy

Définition : $(S_0, ..., S_n) \in R[X]^n$ est une *suite de Sturm* sur $[a, b] \subset R$ si elle vérifie : si $S_k(x) = 0$, $k \in [|1, n-1|]$, $x \in [a, b]$, alors $S_{k-1}(x)S_{k+1}(x) < 0$

Définition : si $(S_0, ..., S_n)$ est une suite de Sturm, on note

$$V_a(S_0, ..., S_n) = \sum_{k=1}^{n} \frac{1}{2}|sign(S_{k-1}(a)) - sign(S_k(a))|$$

On note aussi $V_a^b = V_a - V_b$.

Remarque : cela correspond simplement au nombre de changements de signe que l'on peut compter à la main, en prenant la convention $signe(0) = 0$, ce qui va permettre une additivité de V, propriété très importante que l'on utilisera souvent par la suite et dans l'algorithme.

Proposition : soient R et S des polynômes premiers entre eux, alors l'algorithme d'Euclide produit une suite de Sturm $S_0 = S, S_1 = R, ..., S_n = 1, S_{n+1} = 0$ avec $S_{k-1} = Q_k S_k - S_{k+1}$

Preuve : le signe moins placé devant le reste S_{k+1} assure que si $S_k(x) = 0$, alors $S_{k-1}(x)S_{k+1}(x) < 0$ (si l'un des deux était nul, on aurait que x est une racine commune de S et R, ce qui est exclu par la condition de primalité).

Proposition : caractérisation des suites de Sturm
Soit $(S_0, ..., S_n) \subset R[X]$ telle que :
- $A_k S_{k+1} + B_k S_k + C_k S_{k-1} = 0$ avec $A_k, B_k, C_k \in R[X]$ pour $0 < k < n$
- $A_k > 0$ et $C_k \geq 0$ sur $[a, b]$ pour $0 < k < n$

Alors $(S_0, ..., S_n)$ est une suite de Sturm sur $[a, b]$ ssi S_{n-1} et S_n n'ont pas de zéro commun.

Preuve : si S_{n-1} et S_n ont un zéro commun, il est clair que $(S_0, ..., S_n)$ n'est pas une suite de Sturm. Réciproquement, s'il n'y a pas de zéro commun, soit x un zéro de S_k. On a clairement $S_{k+1}(x) \neq 0$, donc $C_k(x)S_{k-1}(x) = -A_k(x)S_{k+1}(x) \neq 0$, donc $S_{k-1}(x)S_{k+1}(x) < 0$.

Proposition : division pseudo-euclidienne
Si A est un anneau intègre, si $S \in A[X]$ et $P \in A[X]^*$, $\exists !\ Q, R \in A[X]$ tels que $c^d S = PQ - R$ et $deg(R) < deg(P)$ avec c le coefficient dominant de P et $d = max\{0, 1 + deg(S) - deg(P)\}$.

Remarque : quitte à multiplier S et P par c, on peut supposer que d est pair, et c'est ce qu'on fera dans la suite pour satisfaire à la positivité demandée dans la proposition précédente.
Remarque : on ne prouve pas cette dernière proposition, mais elle se traite de même que la division euclidienne classique, à laquelle on se ramène via multiplication par le coefficient en question.

Définition : indice de Cauchy
Soit $f \in R(X)^*$ et $a \in R$. L'*indice de Cauchy* de f en a est :

$$Ind_a(f) = Ind_a^+(f) - Ind_a^-(f), \text{ avec } Ind_a^\epsilon(f) = \begin{cases} +1/2 \text{ si } lim_a^\epsilon(f) = +\infty \\ -1/2 \text{ si } lim_a^\epsilon(f) = -\infty \\ 0 \text{ sinon} \end{cases}$$

Si $a < b$, on note $Ind_a^b(f) = Ind_a^+(f) + \sum_{x \in]a,b[} Ind_x(f) - Ind_b^-(f)$
Si $b < a$, on note $Ind_a^b(f) = -Ind_b^a(f)$

Si $R = 0$ ou $S = 0$, on pose $Ind_a^b(\frac{R}{S}) = 0$

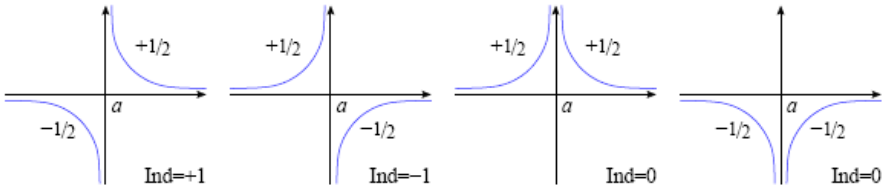

FIGURE 3 – Les 4 indices possibles

Remarque : cette définition peut paraître analytique, mais peut se voir de façon purement algébrique : en effet, il suffit d'écrire que $f = (X - a)^m g$ avec $m \in \mathbb{Z}$ et g qui n'a ni pôle ni racine en a, et alors l'indice est totalement déterminé par la parité de m et le signe de $g(a)$.

Proposition : l'indice sur un intervalle est additif, i.e.

$$Ind_a^b(f) + Ind_b^c(f) = Ind_a^c(f)$$

Preuve : $\sum_{x\in]a,c[} Ind_x(f) = \sum_{x\in]a,c[\setminus\{b\}} Ind_x(f) + Ind_b(f) = \sum_{x\in]a,b[} Ind_x(f) + \sum_{x\in]b,c[} Ind_x(f) + Ind_b^+(f) - Ind_b^-(f)$, d'où le résultat.

2.2 Les principaux résultats dans le cas réel

Proposition : pour $f \in R(X)^*$, on a :

$$Ind_a(f'/f) = \begin{cases} +1 \text{ si a est une racine de f} \\ -1 \text{ si a est un pôle de f} \\ 0 \text{ sinon} \end{cases}$$

Preuve : on écrit $f = (X-a)^m g$ avec $m \in \mathbb{Z}$ et g n'ayant ni pôle ni racine en a. Alors $\frac{f'}{f} = \frac{m}{X-a} + \frac{g'}{g}$, ce qui montre bien que l'indice est déterminé par le signe de m.

Corollaire : comptage théorique de racines

$$Card\,\{x \in [a,b], P(x) = 0\} = Ind_a^b(\tfrac{P'}{P})$$

Preuve : P étant un polynôme, il n'a pas de pôle, d'où le résultat.
Remarque : ce résultat est intéressant théoriquement mais l'indice est difficile à calculer *a priori*. On va donc maintenant s'intéresser à une méthode de calcul pratique de l'indice *via* les suites de Sturm.

Proposition : formule d'inversion
Si $P, Q \in R[X]$ n'ont pas de racine commune en a et b, alors

$$Ind_a^b(\tfrac{Q}{P}) + Ind_a^b(\tfrac{P}{Q}) = V_a^b(P,Q)$$

Preuve :
On peut supposer que $P, Q \neq 0$ et $pgcd(P,Q) = 1$.

- Si P et Q ne s'annulent pas sur $[a,b]$, $Ind_a^b(\frac{Q}{P}) = Ind_a^b(\frac{P}{Q}) = 0$.
  D'après le théorème des valeurs intermédiaires, P et Q restent de signe constant, donc $V_a^b(P,Q) = 0$.
- Puisque l'indice est additif, quitte à dichotomiser l'intervalle, on peut supposer qu'il ne contient qu'une seule singularité, et quitte à recommencer, que c'est a. Par ailleurs on peut supposer (symétrie en P et Q) que $P(a) = 0$ et $Q(a) \neq 0$.
  Dans le cas où $Ind_a^b(\frac{P}{Q}) = -\frac{1}{2}$, on a bien $V_a(P,Q) = \frac{1}{2}$ (immédiat) et $V_b(P,Q) = 1$ (car P est non nul, et Q reste négatif, sinon il y aurait une autre singularité entre a et b). On traite de même l'autre cas.

Corollaire : si $(S_0, ..., S_n)$ est une suite de Sturm dans $R[X]$ sur $[a, b]$, alors :

$$Ind_a^b(\tfrac{S_1}{S_0}) + Ind_a^b(\tfrac{S_{n-1}}{S_n}) = V_a^b(S_0, ..., S_n)$$

Preuve : en effet, la formule d'inversion livre une somme télescopique : pour chaque zéro de S_k, S_{k-1} et S_{k+1} sont de signes opposés et leur contribution à la somme est opposée. D'autre part, $V_a^b(P, Q) + V_a^b(Q, R) = V_a^b(P, Q, R)$ donne le résultat voulu.

Théorème : (Sturm, 1835)

$$Ind_a^b(\tfrac{R}{S}) = V_a^b(S_0, ..., S_n)$$

avec $(S_0, ..., S_n)$ la suite de Sturm obtenue par l'algorithme d'Euclide pour $S_0 = S, S_1 = R$.

Preuve : $Ind_a^b(\tfrac{S_{n-1}}{S_n}) = 0$ par convention ($S_n = 0$).

Corollaire : comptage effectif de racines sur un intervalle

$$Card\,\{x \in [a, b], P(x) = 0\} = V_a^b(S_0, ..., S_n)$$

avec $(S_0, ..., S_n)$ la suite de Sturm obtenue par l'algorithme d'Euclide pour $R = P', S = P$.

Preuve : cela se déduit immédiatement du comptage théorique des racines et du théorème de Sturm.

3 Le cas complexe

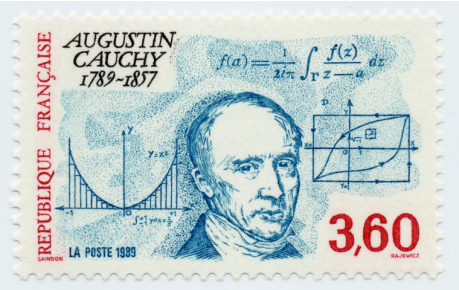

FIGURE 4 – Augustin-Louis Cauchy – 1789-1857

3.1 Indice complexe

Définition : indice de Cauchy pour un polynôme complexe
Pour $F \in C[Z]$ et a, $b \in C$ on définit

$$ind_{[a,b]}(F) = \frac{1}{2} Ind_0^1(\frac{re\hat{F}}{im\hat{F}}) \text{ où } \hat{F}(T) = F((b-a)T + a)$$

Définition : indice sur un rectangle
Si $\Gamma = [x_0, x_1] \times [y_0, y_1]$, en notant $a = (x_0, y_0)$, $b = (x_1, y_0)$, $c = (x_1, y_1)$, $d = (x_0, y_1)$, on définit l'indice sur un rectangle par :

$$ind_{\partial\Gamma}(F) = ind_{[a,b]}(F) + ind_{[b,c]}(F) + ind_{[c,d]}(F) + ind_{[d,a]}(F)$$

Proposition : cas du degré 1 – normalisation
Soit $z_0 \in C$, alors

$$ind_{\partial\Gamma}(Z - z_0) = \begin{cases} 1 \text{ si } z_0 \text{ est à l'intérieur de } \Gamma \\ 1/2 \text{ si } z_0 \text{ est sur une arête de } \Gamma \\ 1/4 \text{ si } z_0 \text{ est un sommet de } \Gamma \\ 0 \text{ si } z_0 \text{ est à l'extérieur de } \Gamma \end{cases}$$

Preuve : on a encore l'additivité sur des rectangles bien choisis, car $ind_{[a,b]}(F) = -ind_{[b,a]}(F)$. On se ramène ainsi au cas d'un sommet, et on se ramène au rectangle unité avec $z_0 = 0$.
On a ensuite $ind_{[0,1]}(X) = \frac{1}{2} Ind_0^1(\frac{X}{0}) = 0$
$ind_{[0,1]}(1 + iX) = \frac{1}{2} Ind_0^1(\frac{1}{X}) = \frac{1}{4}$
$ind_{[0,1]}(1 + i - X) = \frac{1}{2} Ind_0^1(\frac{1-X}{1}) = 0$
$ind_{[0,1]}(i - iX) = \frac{1}{2} Ind_0^1(\frac{0}{1-X}) = 0$
En sommant, on a bien le résultat.

Remarque : attention, le fait que l'indice vale $1/4$ pour une racine située sur un sommet ne fonctionne que pour les polynômes de degré 1. En effet, dès le degré 2, $F_t = Z(Z - 2 - it)$ fournit un contre-exemple : sur le carré unité, 0 est un sommet et une racine de F_t pour tout t, mais $ind_{\partial\Gamma}(F_1) = 0$, $ind_{\partial\Gamma}(F_0) = 1/4$, $ind_{\partial\Gamma}(F_{-1}) = 1/2$.

Proposition : formule du produit
Si $\frac{P}{Q}, \frac{R}{S} \in R(X)^*$, on a

$$Ind_a^b(\tfrac{PR-QS}{PS+QR}) = Ind_a^b(\tfrac{P}{Q}) + Ind_a^b(\tfrac{R}{S}) - V_a^b(1, \tfrac{PS+QR}{QS})$$

Dans le cas $P = S$ et $Q = R$, c'est exactement la formule d'inversion.

Remarque : La preuve est technique et similaire à celle de la formule d'inversion.

Corollaire : formule du produit complexe sur un rectangle
Pour F, $G \in C[X, Y]$ n'admettant aucun sommet de $\Gamma \subset R^2$ pour racine,

$$ind_{\partial\Gamma}(FG) = ind_{\partial\Gamma}(F) + ind_{\partial\Gamma}(G)$$

Preuve : si $F = P + iQ$ et $G = R + iS$, $Re(FG) = PR - QS$ et $Im(FG) = PS + QR$. On applique alors la formule du produit. Les V s'annulent au final, puisqu'on fait la somme $V_a^b + V_b^c + V_c^d + V_d^a$, dont les termes une fois développés, s'annulent deux à deux.

Corollaire : comptage de racines pour un polynôme scindé
Si $F = c(Z - z_1)...(Z - z_n)$ n'admet aucun sommet de Γ pour racine, $ind_{\partial\Gamma}(F)$ compte le nombre de racines de F dans Γ dans le sens suivant :

- Une racine dans l'intérieur de Γ compte pour sa multiplicité.
- Une racine sur la frontière de Γ compte pour la moitié de sa multiplicité.

Preuve : cela découle de la formule du produit et de la proposition de normalisation.

3.2 Indice, racines, degré

Lemme : indice en l'absence de zéro – version locale
Si $F \in C[X,Y]$ vérifie $F(x_0,y_0) \neq 0$, alors il existe $\delta > 0$ tel que $ind_{\partial\Gamma}(F) = 0$ pour tout rectangle $\Gamma \subset [x_0-\delta, x_0+\delta] \times [y_0-\delta, y_0+\delta]$.

Preuve : On écrit $F(x_0+s_0, y_0+t_0) = F(x_0,y_0) + \sum_{j+k\geq 1} a_{jk} s^j t^k$. On pose $M = max^{j+k}\sqrt{|\frac{a_{jk}}{F(x_0,y_0)}|}$; on a donc $|a_{jk}| \leq |F(x_0,y_0)|M^{j+k}$. On pose également $\delta = \frac{1}{4M}$. Si $|s|$, $|t| < \delta$, on a
$|\sum_{j+k\leq 1} a_{jk} s^j t^k| \leq \sum_{n\geq 1} \sum_{j+k=n} |F(x_0,y_0)| M^n |s^j||t^k|$
$\leq |F(x_0,y_0)| \sum_{n\geq 1} (n+1)\frac{1}{4^n} \leq \frac{7}{9}|F(x_0,y_0)|$.
F ne s'annule donc pas sur $U = [x_0-\delta, x_0+\delta] \times [y_0-\delta, y_0+\delta]$. On suppose $F(x_0,y_0) = i$ (quitte à multiplier F par $\frac{i}{a}$, ce qui ne change pas l'indice par additivité, on peut s'y ramener). On a donc que $Im(F) > 0$ sur U, et donc que $ind_{\partial\Gamma}(F) = 0$ pour $\Gamma \subset U$.

Théorème : indice en l'absence de zéro – version globale

Si $F \in C[X,Y]$ ne s'annule pas sur $\Gamma = [x_0,x_1] \times [y_0,y_1]$, alors $ind_{\partial\Gamma}(F) = 0$.

Preuve : on construit la suite $(S_0, ..., S_n)$ associée à F par pseudo-divisions euclidiennes successives dans $R[Y][X]$.
On pose $\frac{Re(F)}{Im(F)} = \frac{S_0}{S_1}$ et on a donc $S_{k+1} = Q_k S_k - c_k^2 S_{k-1}$ avec $Q_k \in R[Y][X]$ et $c_k \in R[Y]^*$. On a $deg_X(S_k) < deg_X(S_k)$, si bien qu'il existe n tel que $S_n \in R[Y]^*$ et $S_{n+1} = 0$.

Cas 1 : S_n ne s'annule pas sur $[y_0,y_1]$. D'après la caractérisation des suites de Sturm, pour $y \in [y_0,y_1]$, $(S_0(y), ..., S_n(y))$ est une suite de Sturm dans $R[X]$. Idem pour x. On a donc
$2ind_{\partial\Gamma}(F) = Ind_{x_0}^{x_1}(\frac{reF}{imF}|Y = y_0) + Ind_{y_0}^{y_1}(\frac{reF}{imF}|X = x_1) + Ind_{x_1}^{x_0}(\frac{reF}{imF}|Y = y_1) + Ind_{y_1}^{y_0}(\frac{reF}{imF}|X = y_0)$
$= V_{x_0}^{x_1}(S_0, ..., S_n|Y = y_0) + V_{y_0}^{y_1}(S_0, ..., S_n|X = x_1) + V_{x_1}^{x_0}(S_0, ..., S_n|Y = y_1) + V_{y_1}^{y_0}(S_0, ..., S_n|X = x_0) = 0$
(la première égalité est due au théorème de Sturm)

Cas 2 : on isole les racines de S_n. Par symétrie et subdivision, on se ramène au cas où (x_0,y_0) est l'unique racine de S_n. Puisque F ne s'annule pas en (x_0,y_0), le lemme local nous donne l'existence d'un rectangle assez petit où l'indice est nul. Sur les trois autres rectangles, on applique le cas 1, ce qui achève la preuve.

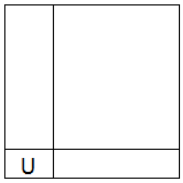

FIGURE 5 – Cas 2

Proposition : l'indice compte les racines complexes
Soit $F \in C[Z]^*$ et un rectangle $\Gamma \subset C$ tel que F ne s'annule pas sur ses sommets, alors $ind_{\partial\Gamma}(F)$ compte les racines de F dans Γ au sens déjà vu.

Preuve : on écrit $F = (Z - z_1)...(Z - z_k)G$ avec $G(z) \neq 0$ pour tout $z \in \Gamma$. D'après la multiplicativité de l'indice et la propriété précédente, on a le résultat voulu.

Lemme : localisation des racines
Soit $F = Z^n + c_{n-1}Z^{n-1} + ... + c_1 Z + c_0$, $M = max(|c_0|, ..., |c_{n-1}|)$ et $\rho_F = 1 + M$, alors toutes les racines de F sont dans $B(0, \rho_F)$.

Preuve : supposons $M > 0$ (sinon, cela veut dire que $F = Z^n$ et le résultat est clair). Si $|z| \geq \rho_F$, on a alors
$|F(z) - z^n| = |c_0 + c_1 z + ... + c_{n-1}z^{n-1}| \leq |c_0| + ... + |c_{n-1}||z^{n-1}|$
$\leq M\frac{|z|^n - 1}{|z| - 1} \leq |z|^n - 1$, car $|z| - 1 \geq \rho_F - 1 = M$
On en tire $|F(z)| = |(z^n - F(z)) - z^n| \geq |z^n| - |F(z) - z^n| \geq |z^n| - |z^n| + 1 = 1 > 0$.

Lemme : invariance par homotopie
Soit $F \in C[T, Z]$. On suppose que pour tout $t \in [0, 1]$, $F(t, Z) \in C[Z]$ ne s'annule pas sur $\partial\Gamma$. Alors $ind_{\partial\Gamma}(F(0, Z)) = ind_{\partial\Gamma}(F(1, Z))$.

Preuve : F est ici vue comme une fonction de $C \times R$ dans C. L'hypothèse est que F n'a pas de zéro sur $\partial\Gamma \times [0, 1]$. Sur chaque arête de Γ (par exemple $[a, b]$), on obtient un rectangle $\tilde{\Gamma} = [a, b] \times [0, 1]$. D'après le théorème sur l'indice en l'absence de zéros, on a $ind_{\delta\tilde{\Gamma}}(F) = 0$, i.e. $ind_{[(a,0),(b,0)]}(F) - ind_{[(a,1),(b,1)]}(F) = ind_{[(a,0),(a,1)]}(F) - ind_{[(b,0),(b,1)]}(F)$. En sommant sur les quatre arêtes, on obtient finalement $ind_{\partial\Gamma}(F(0, Z)) = ind_{\partial\Gamma}(F(1, Z))$.

Proposition : l'indice sur un rectangle assez grand vaut le degré
Pour $F \in C[Z]^*$ et $\Gamma \supset B(0, \rho_F)$, on a

$$ind_{\partial\Gamma}(F) = deg(F)$$

Preuve : $F(t, Z) = Z^n + t(c_{n-1}Z^{n-1} + ... + c_0)$ déforme $F = F(1, Z)$ en $F(0, Z) = Z^n$ dont on connaît les racines.

4 Conclusion

Résumons notre progression :

Dans le cas réel, on a montré que l'indice comptait les racines sur un intervalle, et qu'on pouvait le calculer grâce aux suites de Sturm.
Dans le cas complexe :

- on a montré d'une part que l'indice comptait les racines dans un rectangle, et se calculait encore grâce aux suites de Sturm ;
- d'autre part, on a vu que sur un rectangle assez grand, en particulier sur C, l'indice valait le degré du polynôme considéré.

On a donc prouvé le théorème fondamental de l'algèbre sur la clôture algébrique de tout corps réel clos, en particulier sur $\mathbb{C}$.

On peut de plus localiser ces racines algorithmiquement : partant d'un rectangle assez grand (contenant toutes les racines), on va le découper en quatre rectangles semblables. On vérifie la présence de racine ou non en chaque sommet. S'il y en a, on les met dans la pile. On vérifie ensuite la présence de racines sur chaque arête et dans l'intérieur de chaque rectangle grâce au calcul d'indice, et on met dans la pile les arêtes et les rectangles contenant des racines. On réitère le procédé de la sorte sur les éléments de la pile :

- On ne touche pas à un point ;
- On divise une arête en deux, on la sort de la pile, et on y rentre celles qui contiennent des racines (ou éventuellement le point où l'on a coupé) ;
- On divise un rectangle en quatre rectangles semblables, on le sort de la pile, on y rentre ceux contenant des racines.

On finit par avoir des objets qui contiennent une et une seule racine, puisque les diamètres de ces parties décroissent strictement. On peut éventuellement par la suite, appliquer sur les arêtes une méthode de type Newton, pour converger rapidement vers la racine qu'elle contient.

Références

[1] Michael Eisermann, *The fundamental theorem of algebra made effective : an elementary real-algebraic proof via Sturm chains*, `http://www.igt.uni-stuttgart.de/eiserm/publications/`

Merci à Michael Eisermann pour son article, et à Marie-Françoise Roy pour nous avoir encadré et répondu à nos nombreuses questions.