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

L. Dietrich et S. Billouet

13 mai 2009

Introduction

Énoncé

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

Preuves connues :

- Analyse (compacité, intégration, fonctions analytiques ...)
- Algèbre (TVI, théorie de Galois)
- Topologie algébrique (indice)

FIG.: Jean d'Alembert – 1717-1783

Intérêts de la preuve

- Valable sur tout $C = R[i]$ où R est un corps réel clos (existence du TVI pour les polynômes)
- Purement algébrique (ne nécessite pas de notions du second ordre comme la compacité)
- Constructive et implémentable par un algorithme qui permet de localiser les racines

Préliminaires historiques – le cas réel

FIG.: Charles-François Sturm – 1803-1855

Suites de Sturm et indice de Cauchy

Suite de Sturm

$(S_0, ..., S_n) \in R[X]^n$ 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 $S_{k-1}(x)S_{k+1}(x) < 0$

Suites de Sturm et indice de Cauchy

Suite de Sturm

$(S_0, ..., S_n) \in R[X]^n$ 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 $S_{k-1}(x)S_{k+1}(x) < 0$

Nombre de changements de signe

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 $V_a^b = V_a - V_b$.

Suites de Sturm et indice de Cauchy

Suite de Sturm

$(S_0, ..., S_n) \in R[X]^n$ 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 $S_{k-1}(x)S_{k+1}(x) < 0$

Nombre de changements de signe

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 $V_a^b = V_a - V_b$.

Algorithme d'Euclide

Si $pgcd(R, S) = 1$, l'algorithme d'Euclide produit une suite de Sturm
$S_0 = S, S_1 = R, ..., S_n = 1, S_{n+1} = 0$ avec $S_{k-1} = Q_k S_k - S_{k+1}$

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 $0 < k < n$
- $S_k(x) = 0$ implique $(A_k(x) \neq 0$ et $A_k(x)C_k(x) \geq 0)$ 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.

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 $0 < k < n$
- $S_k(x) = 0$ implique $(A_k(x) \neq 0$ et $A_k(x)C_k(x) \geq 0)$ 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.

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 $d = max\{0, 1 + deg(S) - deg(P)\}$.

Indice de Cauchy en un point

Soit $f \in R(X)^*$ et $a \in R$. L'*indice de Cauchy* de f en a est :
$Ind_a(f) = Ind_a^+(f) - Ind_a^-(f)$, avec

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

Indice de Cauchy en un point

Soit $f \in R(X)^*$ et $a \in R$. L'*indice de Cauchy* de f en a est :
$Ind_a(f) = Ind_a^+(f) - Ind_a^-(f)$, avec

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

Indice sur un intervalle

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 $Ind_a^b(\frac{R}{S}) = 0$

Indice de Cauchy en un point

Soit $f \in R(X)^*$ et $a \in R$. L'*indice de Cauchy* de f en a est :
$Ind_a(f) = Ind_a^+(f) - Ind_a^-(f)$, avec

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

Indice sur un intervalle

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 $Ind_a^b(\frac{R}{S}) = 0$

Propriété importante de l'indice

L'indice sur un intervalle est additif, i.e. $Ind_a^b(f) + Ind_b^c(f) = Ind_a^c(f)$

Les principaux résultats dans le cas réel

Dérivée logarithmique

Pour $f \in R(X)^*$, on a : $Ind_a(f'/f) = \begin{cases} +1 \text{ si a est une racine de f} \\ -1 \text{ si a est un pôle de f} \\ 0 \text{ sinon} \end{cases}$

Les principaux résultats dans le cas réel

Dérivée logarithmique

Pour $f \in R(X)^*$, on a : $Ind_a(f'/f) = \begin{cases} +1 \text{ si a est une racine de f} \\ -1 \text{ si a est un pôle de f} \\ 0 \text{ sinon} \end{cases}$

Comptage théorique de racines

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

La formule d'inversion

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

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

Ébauche de 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 $V_a(P, Q) = \frac{1}{2}$ et $V_b(P, Q) = 1$. On traite de même l'autre cas.

Corollaire de la formule d'inversion

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

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

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.

Corollaire de la formule d'inversion

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

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

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.

Théorème de Sturm

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

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

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 $R = P', S = P$.

Le cas complexe

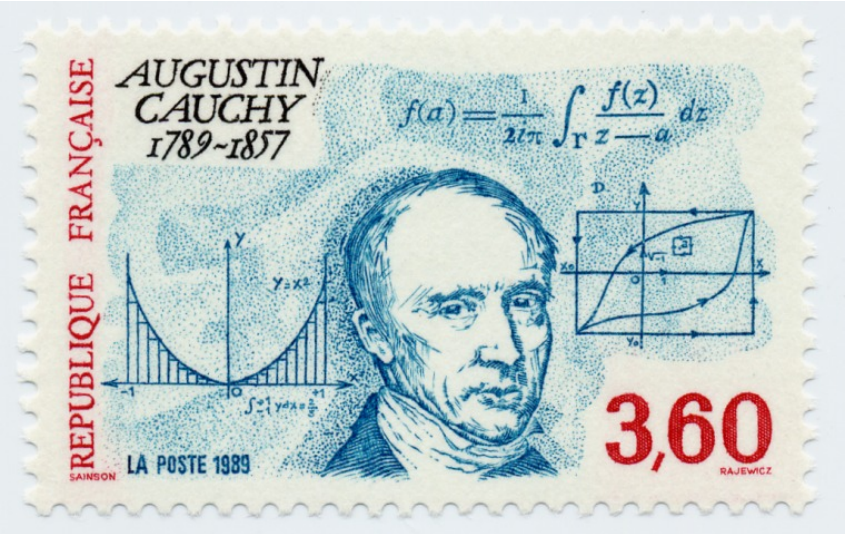

FIG.: Augustin-Louis Cauchy – 1789-1857

Indice complexe

Indice de Cauchy pour un polynôme complexe

Pour $F \in C[Z]$ et $a, b \in C$ on définit

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

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 :

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

Cas du degré 1 – normalisation

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

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 $z_0 = 0$.

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 $ind_{\partial\Gamma}(F_1) = 0$, $ind_{\partial\Gamma}(F_0) = 1/4$, $ind_{\partial\Gamma}(F_{-1}) = 1/2$.

Formule du produit

Si $\frac{P}{Q}, \frac{R}{S} \in R(X)^*$ sont telles que ni P et Q, ni R et S n'ont de racine commune en a ou b, on a

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

Formule du produit

Si $\frac{P}{Q}, \frac{R}{S} \in R(X)^*$ sont telles que ni P et Q, ni R et S n'ont de racine commune en a ou b, on a

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

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,

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

Formule du produit

Si $\frac{P}{Q}, \frac{R}{S} \in R(X)^*$ sont telles que ni P et Q, ni R et S n'ont de racine commune en a ou b, on a

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

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,

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

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.

Indice, racines, degré

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 : continuité.

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 $2ind_{\partial\Gamma}(F) = Ind_{x_0}^{x_1}(\frac{reF}{imF}|Y = y_0) + Ind_{y_0}^{y_1}(\frac{reF}{imF}|X = x_1) + Ind_{x_1}^{x_0}(\frac{reF}{imF}|Y = y_1) + Ind_{y_1}^{y_0}(\frac{reF}{imF}|X = y_0)$
  $= V_{x_0}^{x_1}(S_0, ..., S_n|Y = y_0) + V_{y_0}^{y_1}(S_0, ..., S_n|X = x_1) + V_{x_1}^{x_0}(S_0, ..., S_n|Y = y_1) + V_{y_1}^{y_0}(S_0, ..., S_n|X = x_0) = 0$
- 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.

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.

Lemme de localisation

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

Lemme de localisation

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

Lemme d'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 $ind_{\partial\Gamma}(F(0, Z)) = ind_{\partial\Gamma}(F(1, Z))$.

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.

Conclusion

Résumé de la progression

- Cas réel : l'indice compte les racines sur un intervalle, et on le calcule grâce aux suites de Sturm.
- Cas complexe :
  - l'indice compte les racines dans un rectangle, et se calcule encore grâce aux suites de Sturm ;
  - sur un rectangle assez grand, en particulier sur C, il vaut le degré.

Conclusion : on a prouvé le théorème fondamental de l'algèbre sur la clôture algébrique de tout corps réel clos, en particulier sur $\mathbb{C}$. De plus, on peut localiser ces racines algorithmiquement.

Références et remerciements


M. Eisermann, *The Fundamental Theorem of Algebra made effective : an elementary real-algebraic proof via Sturm chains*, 05/2009

Nous remercions vivement Marie-Françoise Roy pour son encadrement et Michael Eisermann pour avoir pris le temps de répondre à nos questions.