

Le théorème fondamental de l'algèbre rendu effectif : une preuve réelle algébrique par les suites de Sturm

Michael Eisermann

Institut Fourier, Université Grenoble I
www-fourier.ujf-grenoble.fr/~eiserm

23 janvier 2009



Carl Friedrich Gauß (1777-1855)



Augustin Louis Cauchy (1799-1857)



Charles-François Sturm (1803-1855)

Séminaire de calcul formel et complexité, Université Rennes I

1/30

Prologue

Le théorème fondamental de l'algèbre, alias Gauß-d'Alembert, est un résultat classique des mathématiques du 19e siècle. Il est souvent utilisé, cité, enseigné, ... et mérite donc une attention appropriée. Il reste d'actualité, par exemple concernant ses aspects algorithmiques ou numériques.

Si de nos jours l'énoncé du théorème n'a plus rien de surprenant, la preuve réelle algébrique que je présente ici est très remarquable : elle est élégante, élémentaire, et effective. Cet exposé a pour objectif de la populariser.

La preuve réelle algébrique est basée sur des idées de Gauß (1799), Cauchy (1831/37), et surtout Sturm (1836), mais semble inconnue de nos jours. J'ai eu le plaisir de la découvrir en préparant un cours de calcul formel, et j'ai été ensuite très surpris de ne pas la trouver dans la littérature moderne.

Ainsi ma contribution consiste à remettre cette belle démonstration à la lumière du jour, après plus d'un siècle dans l'oubli, et de développer l'esquisse de Sturm en due rigueur.

2/30

Plan

- 1 Le théorème fondamental de l'algèbre
 - Le théorème et son histoire
 - Racines réelles de polynômes réels
 - Racines complexes de polynômes complexes
- 2 Sturm 1829/1835 : racines réelles de polynômes réels
 - L'indice de Cauchy pour les polynômes réels
 - La formule d'inversion de Cauchy
 - Suites de Sturm
- 3 Sturm 1836 : racines complexes de polynômes complexes
 - L'indice de Cauchy pour les polynômes complexes
 - La formule du produit
 - Invariance par homotopie
- 4 Conclusions et perspectives

Référence :

*The Fundamental Theorem of Algebra made effective :
an elementary real-algebraic proof via Sturm chains.*

www-fourier.ujf-grenoble.fr/~eiserm/publications.html#roots

3/30 1/1

Le théorème fondamental de l'algèbre

Théorème (version brève)

Tout polynôme complexe de degré n admet n racines complexes.

Théorème (version longue)

Soit \mathbb{R} le corps des nombres réels et soit $C = \mathbb{R}[i]$ où $i^2 = -1$.

Alors pour tout polynôme

$$F = Z^n + c_1 Z^{n-1} + \dots + c_{n-1} Z + c_n$$

à coefficients $c_1, \dots, c_{n-1}, c_n \in C$ il existe $z_1, z_2, \dots, z_n \in C$ tels que

$$F = (Z - z_1)(Z - z_2) \cdots (Z - z_n).$$

Questions naturelles :

- Existe-t-il une démonstration élémentaire ? qui capte la géométrie ?
- Peut-on affaiblir l'hypothèse ? à quels corps ordonnés au lieu de \mathbb{R} ?
- Peut-on renforcer la conclusion ? la rendre effective ?

4/30

Quelques protagonistes

Scipione del Ferro (1456-1526)
Niccolò Fontana Tartaglia (1500-1557)
Gerolamo Cardano (1501-1576)
Lodovico Ferrari (1522-1565)
...
Niels Henrik Abel (1802-1829)
Évariste Galois (1811-1832)

Albert Girard (1595-1632)
René Descartes (1596-1650)
Gottfried Leibniz (1646-1716)
...

Leonhard Euler (1707-1783)
Jean le Rond d'Alembert (1717-1783)
Joseph-Louis Lagrange (1736-1813)
Pierre-Simon Laplace (1749-1827)
...
Carl Friedrich Gauß (1777-1855)

Augustin Louis Cauchy (1789-1857)
Charles-François Sturm (1803-1855)

Tourisme mathématique



§1.1

§1.1

§1.1

Des nombres réels aux corps réels clos

Théorème (caractérisation des nombres réels)

Pour tout corps ordonné $(\mathbf{R}, +, \cdot, \leq)$ sont équivalents :

- 1 (\mathbf{R}, \leq) satisfait à l'axiome de la borne supérieure.
- 2 Tout intervalle $[a, b] \subset \mathbf{R}$ est compact.
- 3 Tout intervalle $[a, b] \subset \mathbf{R}$ est connexe.
- 4 Toute $f: \mathbf{R} \rightarrow \mathbf{R}$ continue a la propriété des valeurs intermédiaires :
 $a < b \wedge f(a) < 0 < f(b) \implies \exists x \in \mathbf{R} : a < x < b \wedge f(x) = 0$.

Deux tels corps sont isomorphes par un unique isomorphisme de corps.
Un tel objet existe : on l'appelle le corps des nombres réels, noté \mathbb{R} .

Ceci nécessite la logique de second ordre. Beaucoup moins suffira :

Définition (corps réel clos)

Un corps ordonné $(\mathbf{R}, +, \cdot, \leq)$ est dit *réel clos* si tout polynôme $P \in \mathbf{R}[X]$ satisfait à la propriété des valeurs intermédiaires sur \mathbf{R} .

Exemples : les nombres réels \mathbb{R} , les réels algébriques $\mathbb{Q}^c \subset \mathbb{R}$, ...

Tout corps ordonné admet une unique clôture réelle. Exemple : $\mathbb{R}(X)^c$.

§1.2

§1.2

§1.2

Stratégies de preuve

On connaît trois stratégies de preuve :

- 1 Analyse : compacité, fonctions analytiques, intégration, Stokes, ... (d'Alembert 1746, Argand 1814, Cauchy 1820) ;
- 2 Algèbre : TVI pour polynômes, fonctions symétriques / théorie de Galois (Euler 1749, Lagrange 1772, Laplace 1795, Gauß 1816) ;
- 3 Topologie algébrique : notion d'indice [winding number] (Gauß 1799/1816, Cauchy 1831, Sturm-Liouville 1836)

La preuve présentée ici est *réelle algébrique* et se situe entre 2 et 3.

Cette preuve réelle algébrique, qu'est-ce qu'elle offre d'intéressant ?

- ✓ Elle est élémentaire : arithmétique + TVI des polynômes réels.
- ✓ Tous les arguments sont valables sur un corps réel clos.
- ✓ La preuve est constructive : elle permet de localiser les racines.
- ✓ L'algorithme est facile à implémenter et suffisamment efficace.
- ✓ Démonstration formelle du théorème et de l'algorithme.

Sous des hypothèses minimales nous obtenons des conclusions maximales.

Compléments sur les corps réels clos

Remarque

Dans un corps réel clos l'ordre est déterminé par $a \geq 0 \Leftrightarrow \exists r \in \mathbf{R} : r^2 = a$.

Démonstration. Pour $a > 0$ le polynôme $X^2 - a$ a une racine dans $[0, 1 + a]$.

Théorème (clôture réelle)

Tout corps ordonné $(\mathbf{K}, +, \cdot, \leq)$ admet une clôture réelle, c'est-à-dire une extension algébrique $\mathbf{R} \supset \mathbf{K}$ qui soit réelle close.
Deux telles clôtures sont isomorphes par un unique isomorphisme de corps.

La clôture réelle est donc bien plus rigide que la clôture algébrique !

Théorème (Artin-Schreier 1927)

Soit \mathbf{R} un corps et soit $\mathbf{C} \supset \mathbf{R}$ un corps algébriquement clos.
Si $1 < \dim_{\mathbf{R}}(\mathbf{C}) < \infty$ alors \mathbf{R} est réel clos et $\mathbf{C} = \mathbf{R}[i]$.

Ainsi les corps réels clos nous fournissent l'hypothèse minimale.

Théorème (Tarski 1951, Seidenberg 1954)

Les corps réels clos ont tous la même théorie élémentaire.

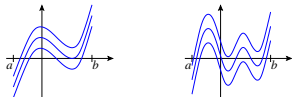
§1.2

§1.2

§1.2

Racines réelles de polynômes réels

Comment déterminer/majorer le nombre de racines de $P \in \mathbf{R}[X]$ dans $[a, b]$?



Réponses partielles par Descartes (1596-1650), Fourier (1768-1830), ...

Théorème de Sturm (1829/35)

Si \mathbf{R} est réel clos, alors $\#\{x \in [a, b] \mid P(x) = 0\} = V_a^b(S_0, S_1, \dots, S_n)$.

Ici la suite S_0, S_1, \dots, S_n est obtenue de $S_0 = P$ et $S_1 = P'$ par division euclidienne itérée, $S_{k-1} = Q_k S_k - S_{k+1}$, jusqu'à ce que $S_{n+1} = 0$.

Ce théorème permet de compter puis de localiser toutes les racines réelles :

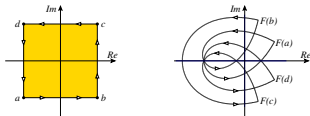


§1.2

§9.0 §1.3

L'indice complexe : motivation géométrique

Exemple : $F = Z^5 - 5Z^4 - 2Z^3 - 2Z^2 - 3Z - 12$ et $\Gamma = [-1, +1] \times [-1, +1]$.



Idée géométrique (Gauß 1799) :

On définit $\text{ind}(F|_{\partial\Gamma})$ comme le nombre des tours de $F|_{\partial\Gamma}$ autour de 0.

- Si Γ est grand, alors $\text{ind}(F|_{\partial\Gamma}) = \text{ind}(Z^n|_{\partial\Gamma}) = n$.
- Si Γ est petit, alors $\text{ind}(F|_{\partial\Gamma}) = \text{ind}(\text{const}|_{\partial\Gamma}) = 0$.
- L'indice ne change que si $F|_{\partial\Gamma}$ passe par 0.

Par conséquent, en degré $n \geq 1$, le polynôme F doit avoir une racine.

Problème technique : Comment définir rigoureusement cet indice ?

L'indice complexe : propriétés algébriques

Soit \mathbf{R} un corps réel clos et soit $\mathbf{C} = \mathbf{R}[i]$, $i^2 = -1$.

Soit $\Omega = \{\text{lacets } \gamma : [0, 1] \rightarrow \mathbf{C}^*, \gamma(0) = \gamma(1), \text{polynomiaux par morceaux}\}$.

Théorème

Il existe une application $\text{ind} : \Omega \rightarrow \mathbf{Z}$ ayant les propriétés suivantes :

- 1 Calculabilité : $\text{ind}(\gamma)$ se calcule par l'algorithme de Sturm sur \mathbf{R} .
- 2 Normalisation : Pour tout rectangle $\Gamma \subset \mathbf{C}$ on a

$$\text{ind}(\partial\Gamma) = \begin{cases} 1 & \text{si } 0 \in \text{Int } \Gamma, \\ 0 & \text{si } 0 \in \mathbf{C} \setminus \Gamma. \end{cases}$$

- 3 Multiplicativité : $\text{ind}(\gamma_1 \cdot \gamma_2) = \text{ind}(\gamma_1) + \text{ind}(\gamma_2)$.
- 4 Invariance par homotopie : $\text{ind}(\gamma_0) = \text{ind}(\gamma_1)$ si $\gamma_0 \sim \gamma_1$ dans \mathbf{C}^* .

La difficulté réside dans la construction ! Tous les moyens sont bons :

- Théorie des revêtements, appliquée à $\text{exp} : \mathbf{C} \rightarrow \mathbf{C}^*$ avec groupe \mathbb{Z} .
- Groupe fondamental, $\text{ind} : \pi_1(\mathbf{C}^*, 1) \xrightarrow{\sim} \mathbb{Z}$ via Seifert-van Kampen.
- Homologie, $\text{ind} : H_1(\mathbf{C}^*) \xrightarrow{\sim} \mathbb{Z}$ via les axiomes d'Eilenberg-Steenrod.
- Topologie différentielle, théorème de Sard et degré topologique.
- Analyse complexe, indice analytique $\text{ind}(\gamma) = \frac{1}{2\pi i} \int_{\gamma} \frac{dz}{z}$.
- Algèbre réelle, indice algébrique $\text{ind} : \Omega \rightarrow \mathbb{Z}$ via les suites de Sturm.

§1.3

§1.3 §1.3

Complément : le théorème fondamental par homotopie

On fixe $\Gamma = [-1, +1] \times [-1, +1] \subset \mathbf{C}$. Pour tout polynôme

$$F = Z^n + c_{n-1}Z^{n-1} + \dots + c_1Z + c_0$$

on construit une homotopie $H : [0, 1] \times \partial\Gamma \rightarrow \mathbf{C}$ comme suit.

Pour $t > 0$ nous posons

$$H_t(z) = t^n F(z(1-t)/t).$$

Ceci s'étend continûment en $t = 0$ par

$$H_t(z) = (1-t)^n z^n + c_{n-1}(1-t)^{n-1} t z^{n-1} + \dots + c_1(1-t)t^{n-1} z + c_0 t^n.$$

Nous obtenons une homotopie entre $H_0(z) = z^n$ et $H_1(z) = c_0$ dans \mathbf{C} .

Si F n'a pas de racines dans \mathbf{C} , alors l'homotopie est dans \mathbf{C}^* .

L'indice nous donne $n = \text{ind}_{\partial\Gamma}(H_0) = \text{ind}_{\partial\Gamma}(H_1) = 0$.

Par contraposé : si $n \geq 1$, alors F admet au moins une racine $z \in \mathbf{C}$.

On factorise $F = (Z - z)G$, puis on conclut par récurrence sur le degré n .

§2.0

Racines complexes de polynômes complexes

Soit \mathbf{R} un corps réel clos et soit $C = \mathbf{R}[i]$, $i^2 = -1$.

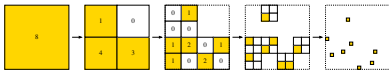
Nous savons construire l'indice, ayant les bonnes propriétés,

$$\text{ind} : \left\{ \begin{array}{l} \text{lacets } \gamma: [0, 1] \rightarrow C^* \\ \text{polynomiaux par morceaux} \end{array} \right\} \rightarrow \mathbb{Z}$$

Ceci donne une preuve effective du théorème fondamental de l'algèbre :

- $\text{ind}_{\partial\Gamma}(F)$ compte les racines de F dans Γ . (Sturm complexe)
- $\text{ind}_{\partial\Gamma}(F) = \deg(F)$ pour Γ assez grand. (Borne de Cauchy)

Ainsi l'indice permet de localiser toutes les racines de F dans C :



Une fois qu'on a bien séparé les racines, on passe à la méthode de Newton.

§1.3

13/30

Quelques dates : aspects constructifs et algorithmiques

Indice réel algébrique

- Sturm 1829/35 : Mémoire sur la résolution des équations numériques
- Cauchy 1831/37 : Calcul des résidus et calcul des indices
- Sturm–Liouville 1836 : Démonstration d'un théorème de M. Cauchy

Réception dans des manuels

- Serret 1877 : Cours d'algèbre supérieure (Sturm réel et complexe)
- Weber 1898 : Lehrbuch der Algebra (Sturm réel, à peine complexe)
- Runge 1898 : Encyklopädie (Sturm réel et complexe)

Degré topologique

- Kronecker 1869 : Systeme von Functionen mehrer Variablen
- Brouwer 1912 : Abbildungen von Mannigfaltigkeiten
- Weyl 1924 : Fundamentalsatz der Algebra

Algorithmes et implémentations

- Lehmer 1969 : Search procedures for polynomial equation solving
- Wilf 1978 : Bisection algorithm for computing zeros of polynomials
- Schönhage 1982 : The fundamental theorem of algebra in terms of computational complexity

14/30

Changements de signe : conventions de comptage

On considère un corps ordonné $(\mathbf{R}, +, \cdot, \leq)$.

Nous comptons le nombre $V(s_0, s_1)$ des changements de signes :

$$V(+, -) = V(-, +) = 1,$$

$$V(+, +) = V(-, -) = V(0, 0) = 0,$$

$$V(+, 0) = V(0, +) = V(-, 0) = V(0, -) = \frac{1}{2}.$$

Définition

Pour une suite (s_0, \dots, s_n) d'éléments dans \mathbf{R} nous posons

$$V(s_0, \dots, s_n) := \sum_{k=1}^n V(s_{k-1}, s_k) = \sum_{k=1}^n \frac{1}{2} |\text{sign}(s_{k-1}) - \text{sign}(s_k)|.$$

Pour une suite (S_0, \dots, S_n) de polynômes dans $\mathbf{R}[X]$ nous posons

$$V_a(S_0, \dots, S_n) := V(S_0(a), \dots, S_n(a)).$$

Pour la différence en $a, b \in \mathbf{R}$ nous écrivons $V_a^b := V_a - V_b$.

⚠ Définition traditionnelle (Descartes, Fourier) : on forme d'abord la suite réduite \hat{s} en supprimant les éléments nuls de s , puis on définit $\hat{V}(s) := V(\hat{s})$.

§2.0

15/30

Les règles de Descartes et de Fourier

Comment déterminer le nombre de racines de $P \in \mathbf{R}[X]$ dans $]a, b[$?

Beaucoup de mathématiciens ont étudié cette question. . .

- La règle de Descartes majore le nombre des racines positives :

Théorème (règle de Descartes)

Pour tout polynôme $P = c_0 + c_1X + \dots + c_nX^n$ dans $\mathbf{R}[X]$ on a

$$\#_{\text{mult}} \{x \in \mathbf{R}_{>0} \mid P(x) = 0\} \leq \hat{V}(c_0, c_1, \dots, c_n).$$

- Fourier étendit cette majoration à tout intervalle réel :

Théorème (règle de Fourier)

Pour tout polynôme $P = c_0 + c_1X + \dots + c_nX^n$ dans $\mathbf{R}[X]$ on a

$$\#_{\text{mult}} \{x \in]a, b[\mid P(x) = 0\} \leq \hat{V}_a^b(P, P', \dots, P^{(n)}).$$

Si P a n racines dans \mathbf{R} , alors on a égalité pour tout intervalle $]a, b[\subset \mathbf{R}$.

Avantage : La majoration est facile à calculer.

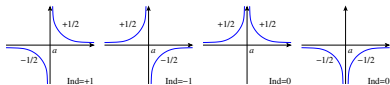
Inconvénient : La borne ainsi obtenue est souvent trop grossière.

C'était l'état de l'art avant la découverte spectaculaire de Sturm en 1829.

16/30

L'indice de Cauchy : comptage de pôles

On compte les pôles d'une fraction rationnelle $f \in \mathbf{R}(X)^*$ comme suit :



Définition (indice de Cauchy)

Pour $f \in \mathbf{R}(X)^*$ et $a \in \mathbf{R}$ on pose

$$\text{Ind}_a(f) := \text{Ind}_a^+(f) - \text{Ind}_a^-(f) \quad \text{où} \quad \text{Ind}_a^{\pm}(f) := \begin{cases} +\frac{1}{2} & \text{si } \lim_a^{\pm} f = +\infty, \\ -\frac{1}{2} & \text{si } \lim_a^{\pm} f = -\infty, \\ 0 & \text{sinon.} \end{cases}$$

Pour un intervalle $[a, b] \subset \mathbf{R}$ on pose

$$\text{Ind}_a^b(f) := \text{Ind}_a^+(f) + \sum_{x \in]a, b[} \text{Ind}_x(f) - \text{Ind}_b^-(f).$$

Propriétés : $\text{Ind}_a^b(f) + \text{Ind}_b^c(f) = \text{Ind}_a^c(f)$ et $\text{Ind}_a^b(f \circ \tau) = \text{Ind}_{\tau^{-1}(a)}^{\tau^{-1}(b)}(f)$.

12.1

17.10

L'indice de Cauchy : comptage de racines

Proposition (dérivée logarithmique)

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

Démonstration. On factorise $f = (X-a)^m g$ tel que $g(a) \in \mathbf{R}^*$.

On obtient $\frac{f'}{f} = \frac{m}{X-a} + \frac{g'}{g}$. Ainsi $\text{Ind}_a(\frac{f'}{f}) = \text{sign}(m)$. \square

Corollaire (racines réelles de polynômes réels)

L'indice $\text{Ind}_a^b(P'/P)$ compte les racines de $P \in \mathbf{R}[X]^*$ dans $]a, b[$:

$$\#\{x \in]a, b[\mid P(x) = 0\} = \text{Ind}_a^b\left(\frac{P'}{P}\right).$$

D'éventuelles racines sur le bord $\{a, b\}$ comptent pour un demi.

Problème : Comment calculer l'indice sans connaître les pôles ?

Exemple : le TVI se reformule comme $\text{Ind}_a^b(\frac{1}{P}) = V_a^b(1, P)$.

Ceci transforme le comptage sur $]a, b[$ en un comptage sur $\{a, b\}$.

Solution générale : Suite de Sturm pour calculer $\text{Ind}_a^b(\frac{Q}{P})$.

18.10

La formule d'inversion sur un corps réel clos

Formule d'inversion (Cauchy 1837)

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

$$\text{Ind}_a^b\left(\frac{Q}{P}\right) + \text{Ind}_a^b\left(\frac{P}{Q}\right) = V_a^b(P, Q).$$

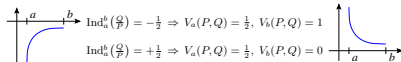
Démonstration. On peut supposer que $P \neq 0$ et $Q \neq 0$ et $\text{pgcd}(P, Q) = 1$.

① Supposons d'abord que $]a, b[$ ne contient pas de racine de P ni de Q .

- En absence de pôles, les indices $\text{Ind}_a^b(\frac{Q}{P})$ et $\text{Ind}_a^b(\frac{P}{Q})$ sont nuls.
- Le TVI assure que P et Q ne changent pas de signe, donc $V_a^b(P, Q) = 0$.

② La formule est additive par rapport à bisection de l'intervalle.

Il ne reste qu'à regarder le cas d'un seul pôle : $P(a) = 0$ et $Q(a) \neq 0$.



⚠ Pour l'argument local ②, sur des petits intervalles $[a, a + \delta]$ et $[a - \delta, a]$ autour d'un pôle, la continuité des polynômes P, Q suffit. Ceci est valable sur tout corps ordonné. La conclusion globale sur $]a, b[$ nécessite le théorème des valeurs intermédiaires ②.

12.2

19.10

Suites de Sturm

Définition (suite de Sturm)

Une suite (S_0, \dots, S_n) dans $\mathbf{R}[X]$ est dite **de Sturm** sur $]a, b[\subset \mathbf{R}$ si elle vérifie : Si $S_k(x) = 0$ pour $0 < k < n$ et $x \in]a, b[$, alors $S_{k-1}(x)S_{k+1}(x) < 0$.

Corollaire (de la formule d'inversion)

$$\text{Ind}_a^b\left(\frac{S_1}{S_0}\right) + \text{Ind}_a^b\left(\frac{S_{n-1}}{S_n}\right) = V_a^b(S_0, S_1, \dots, S_{n-1}, S_n).$$

Démonstration. La formule d'inversion est télescopique : pour $n = 2$ on a

$$\text{Ind}_a^b\left(\frac{S_1}{S_0}\right) + \text{Ind}_a^b\left(\frac{S_0}{S_1}\right) + \text{Ind}_a^b\left(\frac{S_2}{S_1}\right) + \text{Ind}_a^b\left(\frac{S_1}{S_2}\right) = V_a^b(S_0, S_1, S_2).$$

Proposition (fraction continue selon l'algorithme d'Euclide)

Pour $\frac{R}{S}$ où $\text{pgcd}(R, S) = 1$ l'algorithme d'Euclide produit une suite de Sturm $S_0 = S, S_1 = R, \dots, S_n = 1, S_{n+1} = 0$ telle que $S_{k-1} = Q_k S_k - S_{k+1}$. \square

Corollaire : le théorème de Sturm

Pour tout polynôme $P \in \mathbf{R}[X]$ sur un corps \mathbf{R} réel clos on a

$$\#\{x \in]a, b[\mid P(x) = 0\} = \text{Ind}_a^b\left(\frac{P'}{P}\right) = V_a^b(S_0, S_1, \dots, S_n).$$

20.10

L'indice complexe : rappel des propriétés algébriques

Soit \mathbf{R} un corps réel clos et soit $\mathbf{C} = \mathbf{R}[i], i^2 = -1$.

Nous voulons construire l'indice algébrique

$$\text{ind} : \left\{ \begin{array}{l} \text{lacets } \gamma : [0, 1] \rightarrow \mathbf{C}^* \\ \text{polynomiaux par morceaux} \end{array} \right\} \rightarrow \mathbf{Z}$$

On exige les propriétés suivantes :

1 Normalisation : Pour tout rectangle $\Gamma \subset \mathbf{C}$ on a

$$\text{ind}(\partial\Gamma) = \begin{cases} 1 & \text{si } 0 \in \text{Int } \Gamma, \\ 0 & \text{si } 0 \in \mathbf{C} \setminus \Gamma. \end{cases}$$

2 Multiplicativité : $\text{ind}(\gamma_1 \cdot \gamma_2) = \text{ind}(\gamma_1) + \text{ind}(\gamma_2)$.

3 Invariance par homotopie : $\text{ind}(\gamma_0) = \text{ind}(\gamma_1)$ si $\gamma_0 \sim \gamma_1$ dans \mathbf{C}^* .

Bénéfice algorithmique : calcul par les suites de Sturm.

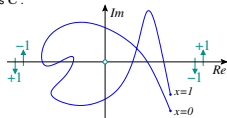
Calcul formel au lieu du numérique : tous les calculs sont exacts.

§3.1

21/30

L'indice complexe : motivation géométrique et définition algébrique

Pour $F \in \mathbf{C}[X]$ la restriction $\gamma : [0, 1] \rightarrow \mathbf{C}$, $\gamma(x) = F(x)$, décrit un chemin dans \mathbf{C} :



Observation

L'indice $\text{ind}_{[0,1]}^{\mathbf{C}}(F) := \frac{1}{2} \text{Ind}_0^{\mathbf{C}}\left(\frac{F(x)}{i\pi F(x)}\right)$ compte les tours autour de 0.

Plus généralement : pour $a, b \in \mathbf{C}$ on considère $\gamma(x) = F(a + (b-a)x)$.

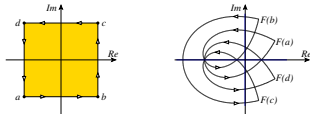
Définition

Pour $F \in \mathbf{C}[Z]$ et $a, b \in \mathbf{C}$ on pose $\text{ind}_{[a,b]}^{\mathbf{C}}(F) = \text{ind}_{[0,1]}^{\mathbf{C}} F(a + (b-a)X)$.

22/30

L'indice par rapport à un rectangle

Exemple : $F = Z^5 - 5Z^4 - 2Z^3 - 2Z^2 - 3Z - 12$ et $\Gamma = [-1, +1] \times [-1, +1]$.



Définition

Pour tout polynôme $F \in \mathbf{C}[Z]$ et tout rectangle $\Gamma \subset \mathbf{C}$ on pose

$$\text{ind}_{\partial\Gamma}^{\mathbf{C}}(F) := \text{ind}_{[a,b]}^{\mathbf{C}}(F) + \text{ind}_{[b,c]}^{\mathbf{C}}(F) + \text{ind}_{[c,d]}^{\mathbf{C}}(F) + \text{ind}_{[d,a]}^{\mathbf{C}}(F).$$

Proposition (normalisation)

$$\text{On a } \text{ind}_{\partial\Gamma}^{\mathbf{C}}(Z - z_0) = \begin{cases} 1 & \text{si } z_0 \text{ est dans l'intérieur de } \Gamma, \\ \frac{1}{2} & \text{si } z_0 \text{ est dans une arête de } \Gamma, \\ \frac{-1}{2} & \text{si } z_0 \text{ est un sommet de } \Gamma, \\ 0 & \text{si } z_0 \text{ est à l'extérieur de } \Gamma. \end{cases}$$

§3.1

23/30

La formule du produit

Pour $F = P + iQ$ et $G = R + iS$ on a $FG = (PR - QS) + i(PS + QR)$.

Lemme (formule du produit)

Pour toute paire de fractions rationnelles $\frac{P}{Q}, \frac{R}{S} \in \mathbf{R}(X)^*$ nous avons

$$\text{Ind}_a^{\mathbf{C}}\left(\frac{PR - QS}{PS + QR}\right) = \text{Ind}_a^{\mathbf{C}}\left(\frac{P}{Q}\right) + \text{Ind}_a^{\mathbf{C}}\left(\frac{R}{S}\right) - V_a^{\mathbf{C}}\left(1, \frac{P}{Q} + \frac{R}{S}\right).$$

Cas particulier : pour $P = S$ et $Q = R$ on retrouve la formule d'inversion. Le cas général se démontre exactement comme le cas particulier.

Théorème (multiplicativité)

Si $F, G \in \mathbf{C}[Z]$ n'ont pas de racines sur les sommets de $\Gamma \subset \mathbf{R}^2$, alors

$$\text{ind}_{\partial\Gamma}^{\mathbf{C}}(F \cdot G) = \text{ind}_{\partial\Gamma}^{\mathbf{C}}(F) + \text{ind}_{\partial\Gamma}^{\mathbf{C}}(G).$$

Corollaire (comptage de racines complexes, cas scindé)

Soit $F \in \mathbf{C}[Z]$ scindé, $F = c(Z - z_1) \cdots (Z - z_n)$ sur \mathbf{C} , sans racines sur les sommets de Γ . Alors l'indice $\text{ind}_{\partial\Gamma}^{\mathbf{C}}(F)$ compte les racines dans Γ .

⚠ Nous devons encore montrer que sur \mathbf{C} tout polynôme est scindé sur \mathbf{C} .

24/30

Comptage de racines complexes

Nous devons montrer : si $\text{ind}_{\mathbb{R}^2}^{\mathbb{C}}(F) > 0$, alors $F(z) = 0$ pour un $z \in \Gamma$.

Par contraposé : si $F(z) \neq 0$ pour tout $z \in \Gamma$, alors $\text{ind}_{\mathbb{R}^2}^{\mathbb{C}}(F) = 0$.

Nous plongeons $\mathbb{C}[Z] \subset \mathbb{C}[X, Y]$ via $Z = X + iY$.

Lemme (version locale)

Si $F \in \mathbb{C}[X, Y]$ ne s'annule pas en $(x, y) \in \mathbb{R}^2$, alors il existe $\delta > 0$ tel que $\text{ind}_{\mathbb{R}^2}^{\mathbb{C}}(F) = 0$ pour tout $\Gamma \subset [x - \delta, x + \delta] \times [y - \delta, y + \delta]$.

Démonstration. Continuité (δ explicite, sur tout corps ordonné). \square

Théorème (version globale)

Si $F \in \mathbb{C}[X, Y]$ ne s'annule pas sur $\Gamma \subset \mathbb{R}^2$, alors $\text{ind}_{\mathbb{R}^2}^{\mathbb{C}}(F) = 0$.

Démonstration. Sur les nombres réels \mathbb{R} : compacité.

Sur un corps \mathbb{R} réel clos quelconque : suites de Sturm. \square

Corollaire (comptage de racines complexes, cas général)

Pour tout $F \in \mathbb{C}[Z]$ l'indice $\text{ind}_{\mathbb{R}^2}^{\mathbb{C}}(F)$ compte les racines de F dans Γ .

Démonstration. On suppose que F n'a s'annule pas sur les sommets de Γ .

Soit $F = (Z - z_1) \cdots (Z - z_m)G$ tel que G n'ait pas de racines dans \mathbb{C} .

On applique la multiplicativité de l'indice à F et le théorème ci-dessus à G .

3/3

25/30

Complément : démonstration par les suites de Sturm

On suppose que $F \in \mathbb{C}[X, Y]$ ne s'annule pas sur $\Gamma = [x_0, x_1] \times [y_0, y_1]$.

On considère $S_0 = \text{im } F$ et $S_1 = \text{re } F$ dans $\mathbb{C}[X, Y] = \mathbb{C}[Y][X]$.

On construit $S_2, \dots, S_n \in \mathbb{C}[Y][X]$ par division pseudo-euclidienne :

$$\begin{aligned} c_k^2 S_{k-1} &= Q_k S_k - S_{k+1}, & Q_k &\in \mathbb{C}[Y][X], & c_k &\in \mathbb{C}[Y], \\ \deg_X S_{k+1} < \deg_X S_k, & & \deg_X S_n &= 0, & S_n &\in \mathbb{C}[Y]. \end{aligned}$$

● Si S_n ne s'annule pas sur $[y_0, y_1]$, on peut appliquer Sturm :

À noter : si $S_k(x, y) = 0$ en $(x, y) \in \Gamma$, alors $S_{k-1}(x, y)S_{k+1}(x, y) < 0$.

$$\begin{aligned} 2 \text{ind}_{\mathbb{R}^2}^{\mathbb{C}}(F) &= + \text{Ind}_{x_0}^{\mathbb{C}} \left(\frac{\text{re } F}{\text{im } F} \mid Y = y_0 \right) + \text{Ind}_{y_0}^{\mathbb{R}} \left(\frac{\text{re } F}{\text{im } F} \mid X = x_1 \right) \\ &\quad + \text{Ind}_{x_1}^{\mathbb{C}} \left(\frac{\text{re } F}{\text{im } F} \mid Y = y_1 \right) + \text{Ind}_{y_1}^{\mathbb{R}} \left(\frac{\text{re } F}{\text{im } F} \mid X = x_0 \right) \\ &= + V_{x_0}^{\mathbb{C}}(S_0, \dots, S_n \mid Y = y_0) + V_{y_0}^{\mathbb{R}}(S_0, \dots, S_n \mid X = x_1) \\ &\quad + V_{x_1}^{\mathbb{C}}(S_0, \dots, S_n \mid Y = y_1) + V_{y_1}^{\mathbb{R}}(S_0, \dots, S_n \mid X = x_0) = 0. \end{aligned}$$

● Au pire il existe un nombre fini de paramètres critiques $y \in [y_0, y_1]$.

En raisonnant dans $\mathbb{C}[X][Y]$: paramètres critiques $x \in [x_0, x_1]$.

Autor d'un point critique (x, y) on applique le lemme local.

En dehors des points critiques on applique le cas ●.

26/30

Localisation grossière des racines

Définition (borne de Cauchy)

Soit $F = Z^n + c_{n-1}Z^{n-1} + \dots + c_1Z + c_0$ dans $\mathbb{C}[Z]$.

On pose $M := \max\{|c_0|, \dots, |c_{n-1}|\}$ et $\rho_F := 1 + M$.

Théorème (localisation grossière des racines)

Pour tout $z \in \mathbb{C}$ tel que $|z| \geq \rho_F$ on a $|F(z)| \geq 1$.

⇒ Toutes les racines de F dans \mathbb{C} sont dans $B(\rho_F) = \{z \in \mathbb{C} \mid |z| < \rho_F\}$.

Démonstration. L'énoncé est vrai pour $F = Z^n$: ici $M = 0$ et $\rho_F = 1$.

Dans la suite nous pouvons supposer $M > 0$ et $\rho_F > 1$.

Soit $z \in \mathbb{C}$ tel que $|z| \geq \rho_F$, c'est-à-dire $|z| - 1 \geq M$. On trouve

$$\begin{aligned} |F(z) - z^n| &= |c_0 + c_1z + \dots + c_{n-1}z^{n-1}| \leq |c_0| + |c_1||z| + \dots + |c_{n-1}||z|^{n-1} \\ &\leq M + M|z| + \dots + M|z|^{n-1} = M \frac{|z|^n - 1}{|z| - 1} \leq |z|^n - 1. \end{aligned}$$

Ainsi nous obtenons

$$\begin{aligned} |z^n| &= |z^n - F(z) + F(z)| \leq |z^n - F(z)| + |F(z)|, \quad \text{d'où} \\ |F(z)| &\geq |z^n| - |F(z) - z^n| \geq |z|^n - (|z|^n - 1) = 1. \end{aligned}$$

3/3

27/30

Invariance par homotopie

Théorème (invariance par homotopie)

Soit $F \in \mathbb{C}[T, Z]$. Supposons que pour tout $t \in [0, 1]$ le polynôme $F_t \in \mathbb{C}[Z]$ n'a pas de racine sur $\partial\Gamma$. Alors $\text{ind}_{\mathbb{R}^2}^{\mathbb{C}}(F_0) = \text{ind}_{\mathbb{R}^2}^{\mathbb{C}}(F_1)$.

Démonstration. L'absence de racines sur $[0, 1] \times [a, b]$ garantit que

$$\begin{aligned} \text{ind}_{[a,b]}^{\mathbb{C}}(F \mid T=0) - \text{ind}_{[a,b]}^{\mathbb{C}}(F \mid T=1) \\ = \text{ind}_{[0,1]}^{\mathbb{C}}(F \mid Z=a) - \text{ind}_{[0,1]}^{\mathbb{C}}(F \mid Z=b). \end{aligned}$$

La somme sur les quatre cotés de Γ donne $\text{ind}_{\mathbb{R}^2}^{\mathbb{C}}(F_0) - \text{ind}_{\mathbb{R}^2}^{\mathbb{C}}(F_1) = 0$. \square

Corollaire

Pour $F \in \mathbb{C}[Z]^*$ et $\Gamma \supset B(\rho_F)$ on a $\text{ind}_{\mathbb{R}^2}^{\mathbb{C}}(F) = \deg F$.

Démonstration. Soit $F = Z^n + c_{n-1}Z^{n-1} + \dots + c_0$.

Alors $F_t = Z^n + t(c_{n-1}Z^{n-1} + \dots + c_0)$ déforme $F_1 = F$ en $F_0 = Z^n$.

La borne de Cauchy $\rho_t = 1 + tM$ diminue de $\rho_1 = \rho_F$ à $\rho_0 = 1$.

Ainsi F_t n'a pas de racines sur $\partial\Gamma$, d'où $\text{ind}_{\mathbb{R}^2}^{\mathbb{C}}(F_1) = \text{ind}_{\mathbb{R}^2}^{\mathbb{C}}(F_0) = n$. \square

Ceci prouve le théorème fondamental : Γ contient n racines de F .

28/30

Conclusions et perspectives

Sur tout corps réel clos nous savons construire l'indice algébrique

$$\text{ind} : \left\{ \begin{array}{l} \text{lacets } \gamma: [0, 1] \rightarrow \mathbf{C}^* \\ \text{polynômes par morceaux} \end{array} \right\} \rightarrow \mathbb{Z}.$$

⇒ Preuve élémentaire et effective du théorème fondamental de l'algèbre.

Contrôle sur les degrés algébriques

A-t-on l'équivalence « \mathbf{R} réel n -clos $\iff \mathbf{R}[i]$ algébriquement n -clos » ?

L'implication « \Rightarrow » est valable en petit degré. La réciproque « \Leftarrow » est claire.

Optimisation algorithmique : complexité asymptotique

Localisation des racines à b bits d'un polynôme de degré n :
algébrique : $\tilde{O}(n^4 b^2)$ vs numérique : $\tilde{O}(n^2(n+b))$, Schönhage 1982.

Comment rendre la version algébrique plus efficace ?

Indice algébrique en dimension supérieure

Théorème du point fixe de Brouwer sur des corps réels clos.

Contrôle sur les degrés algébriques ?



Je vous remercie de votre attention !

Michael.Eisermann@ujf-grenoble.fr
www-fourier.ujf-grenoble.fr/eiserm

*The Fundamental Theorem of Algebra made effective :
an elementary real-algebraic proof via Sturm chains*

Vos remarques et suggestions seront les bienvenues !