Un retour aux racines

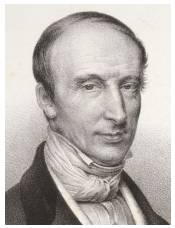
Le théorème fondamental de l'algèbre rendu effectif : une preuve algébrique réelle par des suites de Sturm

Michael Eisermann

Institut Fourier, Université Grenoble

Carl Friedrich Gauss (1777–1855)

Séminaire des Magistères
Jeudi 2 octobre 2008



Augustin Louis Cauchy (1789–1857)

Charles-François Sturm (1803–1855)

http://www-fourier.ujf-grenoble.fr/~eiserm/enseignement

Plan de l'exposé

Le théorème fondamental de l'algèbre

Théorème

Tout polynôme complexe de degré n admet n racines complexes.

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

Théorème

Pour tout polynôme

$$F = a_n Z^n + a_{n-1} Z^{n-1} + \cdots + a_1 Z + a_0$$

à coefficients complexes $a_0, a_1, \ldots, a_{n-1}, a_n \in \mathbb{C}$ vérifiant $a_n \neq 0$
il existe $z_1, z_2, \ldots, z_n \in \mathbb{C}$ telles que

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

Questions :

- L'énoncé serait faux sur $\mathbb{Q}$. Pourquoi est-il vrai sur $\mathbb{R}$?
- Quels outils sont utilisés dans les diverses démonstrations ?
- Pour quels corps ordonnés l'énoncé est-il vrai ?
- Peut-on améliorer l'énoncé ? le rendre effectif ?

Réflexions sur le théorème fondamental de l'algèbre

Ce théorème est un des résultats les plus classiques des mathématiques.
Il est fréquemment utilisé / enseigné et mérite donc une attention particulière.

Tout polynôme P de degré n sur un corps K a au plus n racines dans K.
En général il en a moins : $X^2 + 1$ sur $\mathbb{R}$ n'a aucune racine dans $\mathbb{R}$.
On rajoute donc une racine imaginaire : $\mathbb{C} = \mathbb{R}[i]$ où $i^2 + 1 = 0$.

A priori d'autres équations nécessiteraient encore d'autres adjonctions :
ainsi leurs racines se trouveraient dans une extension encore plus grande.

Ce qui est remarquable c'est qu'ici une seule adjonction suffit pour toutes :
Le corps $\mathbb{C} = \mathbb{R}[i]$ des nombres complexes est algébriquement clos.

Cet énoncé serait faux sur $\mathbb{Q}$. Pourquoi est-il vrai sur $\mathbb{R}$?
Quelles hypothèses sont nécessaires ? minimales ?
Pour quels corps ordonnés l'énoncé est-il vrai ?

L'énoncé ci-dessus n'est qu'une affirmation d'existence.
Il ne nous indique pas comment / où trouver les racines dans $\mathbb{C}$.
Comment les localiser / approcher par un algorithme ?

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

Aperçu historique

La résolution des équations quadratiques fut connue depuis l'antiquité.

Pour des avancées aux degrés supérieurs il fallait attendre le 16ème siècle, quand Del Ferro, Tartaglia et Cardano développèrent des solutions en degré 3, puis Ferrari en degré 4, exprimant les solutions en fonction des coefficients donnés à l'aide des opérations arithmétiques et des racines $n$ièmes.

Pendant trois siècles des formules semblables en degré ≥ 5 furent cherchées en vain. Au début du 19ème siècle Abel et Galois montrèrent que de telles formules n'existent pas en général. En absence de formule explicite il faut donc assurer par d'autres moyens au moins l'existence des racines.

L'existence des racines fut conjecturée par Girard, Descartes et Leibniz, mais ils n'affirmaient pas que toutes les racines étaient des nombres complexes. Avec l'expérience cette affirmation se précisait, et Euler, d'Alembert, Lagrange, Laplace publièrent des premières tentatives de démonstration. On considère que Gauss en donna la première démonstration rigoureuse.

Littérature : Ebbinghaus *et al.* : *Numbers*, chapitre 4, Springer 1990.

La démonstration que je présente ici combine l'idée géométrique de Gauss avec les techniques algébriques de Sturm et Cauchy. Elles mènent à une démonstration effective et réelle algébrique du théorème fondamental de l'algèbre. C'est une belle preuve qui semble peu connue de nos jours.

Les trois types de preuve du théorème fondamental

On connaît essentiellement trois types de démonstration :

1. analyse : compacité, exponentielle, intégration, Stokes, ...
2. algèbre : théorie de Galois / fonctions symétriques, valeurs intermédiaires
3. topologie algébrique : degré d'un lacet $\gamma\colon \mathbb{S}^1 \to \mathbb{C} \smallsetminus \{0\}$

Selon le contexte, chaque preuve a son intérêt et son propre charme.

La preuve que je présenterai ici est du dernier type mais *réelle algébrique*. Elle n'est pas la plus courte mais elle offre de nombreux avantages :

- La démonstration est élémentaire.
  - Arithmétique des polynômes réels à une variable,
  - Le théorème de valeurs intermédiaires.
- Ainsi tous les arguments sont valables sur un corps réel clos.
- La preuve est constructive : elle permet de localiser les racines.
- L'algorithme est relativement facile à implémenter sur ordinateur.
- La démarche mène à une preuve formelle du théorème.
- Parallèlement elle fournit une preuve formelle de l'implémentation.

Commentaires

Pour illustration, je discuterai brièvement une preuve typique de chacune des trois catégories. Pour une discussion plus détaillée consultez

1. Rudin, *Principles of mathematical analysis*, McGraw-Hill, New York 1976 (chapter 8)
2. Ebbinghaus *et al*, *Numbers*, Springer, New York 1990 (chapter 4, appendix)
3. Bredon, *Topology and geometry*, Springer, New York 1993 (chapter 3, §3)

Finalement, le livre suivant développe et met en parallèle plusieurs preuves :

4. Fine, Rosenberger : *The fundamental theorem of algebra*, Springer, New York 1997.

Analyse : la preuve d'Argand–Cauchy

Lemme (existence d'un minimum global)

La fonction $|F|\colon \mathbb{C} \to \mathbb{R}$ *atteint son minimum :*
il existe $z_0 \in \mathbb{C}$ *tel que* $|F(z_0)| \leq |F(z)|$ *pour tout* $z \in \mathbb{C}$.

Lemme (analyse d'un minimum local)

Si $|F(z_0)| > 0$ *alors* $|F(z_0)|$ *n'est pas minimal : il existe*
$z_1 \in \mathbb{C}$ *arbitrairement proche de* z_0 *tel que* $|F(z_1)| < |F(z_0)|$.

Ces deux lemmes prouvent qu'il existe $z_0 \in \mathbb{C}$ tel que $F(z_0) = 0$.
On factorise $F = (Z - z_0)G$, puis on conclut par récurrence sur le degré.

Outils utilisés :

- La boule $\bar{B}(r) = \{z \in \mathbb{C} \mid |z| \leq r\}$ est compacte. (Heine–Borel)
- Toute fonction continue $f\colon B(r) \to \mathbb{R}$ atteint son minimum.
- La fonction exponentielle $\exp\colon \mathbb{C} \to \mathbb{C}$ vérifie $\exp(nx) = \exp(x)^n$.
- Tout $z \in \mathbb{C}$ s'écrit comme $z = \rho \exp(i\theta)$ où $\rho, \theta \in \mathbb{R}$ et $\rho \geq 0$.

Inconvénient : Cette preuve n'est pas constructive, ni effective.
Elle ne nous indique pas comment trouver les racines.

Analyse : la preuve d'Argand–Cauchy (suite)

Lemme (existence d'un minimum global)

Si F *est un polynôme, alors la fonction* $|F|\colon \mathbb{C} \to \mathbb{R}$ *atteint son minimum :*
il existe $z_0 \in \mathbb{C}$ *tel que* $|F(z_0)| \leq |F(z)|$ *pour tout* $z \in \mathbb{C}$.

⚠ L'hypothèse que F soit un polynôme est cruciale :
la fonction $f\colon \mathbb{C} \to \mathbb{R}$, $f(z) = \frac{1}{1+|z|^2}$ n'atteint pas de minimum.

Démonstration. Le lemme est clair pour une fonction constante $F = a_0$.
Il reste à analyser le cas où $n \geq 1$ et $a_n \neq 0$. Pour tout $z \in \mathbb{C}$ on a

$$\begin{aligned} & a_n z^n = F(z) - a_{n-1} z^{n-1} - \cdots - a_1 z - a_0 \\ \Rightarrow \quad & |a_n z^n| \leq |F(z)| + |a_{n-1} z^{n-1}| + \cdots + |a_1 z| + |a_0| \\ \Rightarrow \quad & |F(z)| \geq |a_n||z|^n - |a_{n-1}||z|^{n-1} - \cdots - |a_1||z| - |a_0| \\ & \qquad = |z|^n \big(|a_n| - |a_{n-1}||z|^{-1} - \cdots - |a_1||z|^{1-n} - |a_0||z|^{-n}\big) \end{aligned}$$

Ce minorant tend vers $+\infty$ pour $|z| \to +\infty$.
Il existe alors $r \geq 0$ tel que $|F(z)| > |a_0|$ pour $|z| > r$.

Soit $\mu := \inf\{\ |F(z)| \mid z \in \mathbb{C}\ \}$. On a $\mu \leq |F(0)| = |a_0|$.
Ainsi $\mu = \inf\{\ |F(z)| \mid z \in \bar{B}(0,r)\ \}$ où $\bar{B}(0,r) = \{z \in \mathbb{C} \mid |z| \leq r\}$.

Puisque $|F|\colon \bar{B}(0,r) \to \mathbb{R}$ est continue et $\bar{B}(0,r)$ est compact,
le minimum est atteint : il existe $z_0 \in \bar{B}(0,r)$ tel que $|F(z_0)| = \mu$. □

Analyse : la preuve d'Argand–Cauchy (fin)

Lemme (analyse d'un minimum local)

Si $|F(z_0)| > 0$ *alors* $|F(z_0)|$ *n'est pas minimal : il existe*
$z_1 \in \mathbb{C}$ *arbitrairement proche de* z_0 *tel que* $|F(z_1)| < |F(z_0)|$.

Démonstration. On définit $f\colon \mathbb{C} \to \mathbb{C}$ par $f(z) = F(z + z_0)/F(z_0)$.
Ainsi $\min|F| = |F(z_0)|$ si et seulement si $\min|f| = |f(0)| = 1$.

La fonction f est polynomiale de même degré n, donc
$f(z) = 1 + b_k z^k + \cdots + b_n z^n$ où $b_k, \ldots, b_n \in \mathbb{C}$ et $b_k \neq 0$.

Il existe $\theta \in \mathbb{R}$ tel que $b_k e^{ik\theta} = -|b_k|$. Pour $z = re^{i\theta}$ on trouve
$f(re^{i\theta}) = 1 + b_k r^k e^{ik\theta} + \cdots + b_n r^n e^{in\theta} = 1 - |b_k| r^k + \cdots + b_n r^n e^{ik\theta}$.

Ainsi $|f(re^{i\theta})| \leq \big|\ 1 - |b_k| r^k\ \big| + \cdots + |b_n| r^n$. Pour $0 \leq r \leq |b_k|^{-1/k}$
ceci devient $|f(re^{i\theta})| \leq 1 - r^k\big(|b_k| - |b_{k+1}| r - \cdots - |b_n| r^{n-k}\big)$.

Pour $r > 0$ assez petit le terme en parenthèse est positif.
On conclut que $|f(z)| < 1$, donc $|f(0)| = 1$ n'est pas minimal. □

Ces deux lemmes prouvent le théorème fondamental de l'algèbre :
il existe $z_0 \in \mathbb{C}$ tel que $|F(z_0)| = \min|F|$, puis $F(z_0) = 0$.
On factorise $F = (Z - z_0)G$ puis on conclut par récurrence sur le degré.

Algèbre : la preuve de Laplace–Gauss

Réduction aux polynômes réels : Afin d'établir que tout $Q \in \mathbb{C}[X]$
est scindé sur $\mathbb{C}$ il suffit de prouver que Q admet une racine z_1 dans $\mathbb{C}$:
on factorise $Q = (X - z_1)Q_1$, puis on conclut par récurrence sur le degré.

Afin de montrer que tout $Q \in \mathbb{C}[X]$ admet une racine dans $\mathbb{C}$ il suffit de
le montrer pour tout $P \in \mathbb{R}[X]$: si $Q \in \mathbb{C}[X]$, alors $Q\bar{Q} \in \mathbb{R}[X]$.

Notation : Dans $\mathbb{Z}[X_1, \ldots, X_n][X]$ développons le produit

$$(X + X_1) \cdots (X + X_n) = X^n + S_1 X^{n-1} + \cdots + S_n.$$

Les coefficients $S_1 = X_1 + \cdots + X_n$, …, $S_n = X_1 \cdots X_n$ sont appelés
polynômes symétriques élémentaires en $X_1, \ldots, X_n$.

Outils utilisés :

- Tout polynôme réel de degré impair admet au moins une racine réelle.
  (C'est un corollaire du théorème des valeurs intermédiaires.)
- Pour tout polynôme $P \in \mathbb{R}[X]$ il existe un corps de décomposition.
  (C'est-à-dire que P est scindé sur une certaine extension $E \supset \mathbb{R}$.)
- Tout polynôme symétrique $P \in \mathbb{R}[X_1, \ldots, X_n]$ est un polynôme réel
  dans les polynômes symétriques élémentaires $S_1, \ldots, S_n$.
- Tout polynôme $P \in \mathbb{C}[X]$ de degré 2 est scindé sur $\mathbb{C}$.

Algèbre : la preuve de Laplace–Gauss

Théorème

Tout polynôme $P \in \mathbb{R}[X]$ admet au moins une racine dans $\mathbb{C}$.

Démonstration. Soit $P = X^n + s_1 X^{n-1} + \cdots + s_n$ où $s_1, \ldots, s_n \in \mathbb{R}$.
Décomposons le degré $n = 2^k q$ où $k \geq 0$ et $q \in \mathbb{N}$ est impair.

On démontre le théorème par récurrence sur k.
Le cas $k = 0$ est assuré par le théorème des valeurs intermédiaires.

Supposons $k \geq 1$. Soit $E \supset \mathbb{R}$ un corps de décomposition de P, c'est-à-dire $P = (X + x_1) \cdots (X + x_n)$ où $x_1, \ldots, x_n \in E$. Pour $t \in \mathbb{R}$ nous définissons

$$L_t = \prod_{1 \leq j < k \leq n} (X - x_j - x_k - t x_j x_k) \in E[X].$$

Ce polynôme est symétrique en $x_1, \ldots, x_n$, donc un polynôme en $s_1, \ldots, s_n$.
Par conséquent, nous obtenons un polynôme réel $L_t \in \mathbb{R}[X]$ pour tout $t \in \mathbb{R}$.

Puisque L_t est de degré $\frac{n}{2}(n-1)$, notre hypothèse de récurrence s'applique : L_t admet au moins une racine complexe $z_t \in \mathbb{C}$.

Pour tout $t \in \mathbb{R}$ il existe une paire $j < k$ tels que $x_j + x_k + t x_j x_k \in \mathbb{C}$.
Il existe une infinité de nombres réels mais seulement $\frac{n(n-1)}{2}$ paires $j < k$, il existe $s \neq t$ et $j < k$ tels que $x_j + x_k + t x_j x_k \in \mathbb{C}$ et $x_j + x_k + s x_j x_k \in \mathbb{C}$.
Ceci implique que $u := x_j + x_k \in \mathbb{C}$ et $v := x_j x_k \in \mathbb{C}$.
Ainsi x_j, x_k sont les racines de $X^2 - vX + u \in \mathbb{C}[X]$, donc $x_j, x_k \in \mathbb{C}$ □

La preuve par la topologie algébrique : outils utilisés

On note $\mathbb{S}^1 = \{z \in \mathbb{C} \mid |z| = 1\}$ le cercle unité dans $\mathbb{C}$.
Un lacet dans $\mathbb{C}^* = \mathbb{C} \smallsetminus \{0\}$ est une application continue $\gamma \colon \mathbb{S}^1 \to \mathbb{C}^*$.

Outil utilisé : Il existe une fonction $\mathrm{ind} \colon \{\text{lacets dans } \mathbb{C}^*\} \to \mathbb{Z}$ qui compte le nombre de tours autour de 0. Plus précisément :

1. $\mathrm{ind}(\mathrm{id}) = 1$.
2. $\mathrm{ind}(\gamma_1 \cdot \gamma_2) = \mathrm{ind}(\gamma_1) + \mathrm{ind}(\gamma_2)$.
3. $\mathrm{ind}(\gamma_0) = \mathrm{ind}(\gamma_1)$ si $\gamma_0 \sim \gamma_1$ sont homotopes dans $\mathbb{C}^*$.

La difficulté est de démontrer l'existence d'une telle application ind !
Il faut d'abord la construire puis établir toutes les propriétés requises.

Pour réaliser cette construction, tous les moyens sont bons :

- Théorie des revêtements, appliquée à $\exp \colon \mathbb{C} \twoheadrightarrow \mathbb{C}^*$.
- Groupe fondamental $\mathrm{ind} \colon \pi_1(\mathbb{C}^*, 1) \xrightarrow{\sim} \mathbb{Z}$ via Seifert–van Kampen.
- Homologie $\mathrm{ind} \colon H_1(\mathbb{C}^*) \xrightarrow{\sim} \mathbb{Z}$ via les axiomes d'Eilenberg–Steenrod.
- Analyse complexe, indice analytique $\mathrm{ind}(\gamma) = \frac{1}{2i\pi} \int_\gamma \frac{dz}{z}$.
- Algèbre réelle, indice de Cauchy via les suites de Sturm.

La preuve par la topologie algébrique : construction d'une homotopie

Preuve du théorème fondamental de l'algèbre :

On considère un polynôme $F = Z^n + a_{n-1} Z^{n-1} + \cdots + a_1 Z + a_0$.

On construit une homotopie $H \colon [0,1] \times \mathbb{S}^1 \to \mathbb{C}$ comme suit :

Pour $t > 0$ on pose

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

Ceci se développe en

$$H_t(z) = (1-t)^n z^n + a_{n-1}(1-t)^{n-1} t z^{n-1} + \cdots + a_1 (1-t) t^{n-1} z + a_0 t^n.$$

Cette dernière formule s'étend continûment en $t = 0$.

On obtient une homotopie entre $H_0(z) = z^n$ et $H_1(z) = a_0$ dans $\mathbb{C}$.

Si F n'a pas de racines dans $\mathbb{C}$, alors H est une homotopie dans $\mathbb{C}^*$.

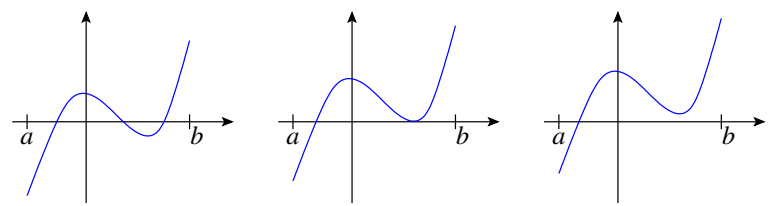
En utilisant l'indice on conclut que $n = \mathrm{ind}(H_0) = \mathrm{ind}(H_1) = 0$.

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

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

Objectif

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



Réponses partielles par René Descartes (1596-1650),
François Budan (1761-1840), Joseph Fourier (1768-1830), ...

La réponse complète fut donnée par Sturm en 1829/35 :

$$\#\{x \in [a,b] \mid P(x) = 0\} = V_a^b(S_0, S_1, \ldots, S_n).$$

Ici la suite de Sturm $S_0, S_1, \ldots, S_n \in \mathbf{R}[X]$ découle de l'algorithme d'Euclide.
La différence $V_a^b = V_a - V_b$ compare les variations des signes en a et en b.

Nombres réels et le théorème des valeurs intermédiaires

Théorème

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 $f : \mathbf{R} \to \mathbf{R}$ continue a la propriété des valeurs intermédiaires :*
   $a < b \ \wedge \ f(a) < 0 < f(b) \quad \Longrightarrow \quad \exists x \in \mathbf{R} : \ a < x < b \ \wedge \ f(x) = 0.$

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

Définition (corps réel clos)

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

Exemples : les réels $\mathbb{R}$, les réels algébriques $\mathbb{Q}^c \subset \mathbb{R}$, puis $\mathbb{R}(X)^c$, ...
Tout corps ordonné admet une unique clôture réelle.

Complément sur les corps réels clos

Remarque

Si $\mathbf{R}$ est réel clos, alors tout $a \in \mathbf{R}_{\geq 0}$ admet une unique racine $r \in \mathbf{R}_{\geq 0}$ telle que $r^2 = a$: c'est la racine de $P = X^2 - a$ sur $[0, 1 + a]$.
Ainsi l'ordre est caractérisé algébriquement par $a \geq 0 \Leftrightarrow \exists r \in \mathbf{R} : r^2 = a$.

En particulier, un corps réel clos n'admet qu'un seul ordre.

Théorème (clôture réelle)

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

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

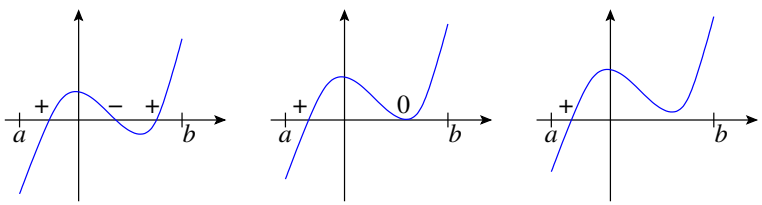
Théorème (Artin–Schreier 1927)

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 $\mathbf{C} = \mathbf{R}[i]$.

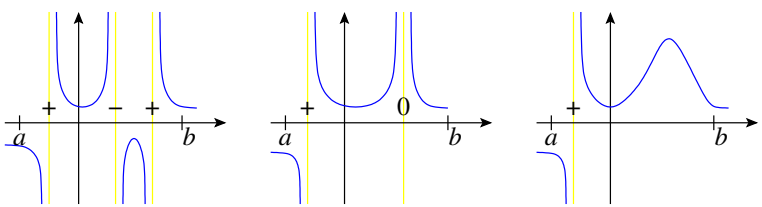
Ainsi les corps réels clos fournissent l'hypothèse minimale cherchée.

Zéros et pôles

On souhaite compter les zéros d'un polynôme $P \in \mathbf{R}[X]$ dans $[a, b]$:



De manière équivalente on peut compter les pôles de $1/P$:



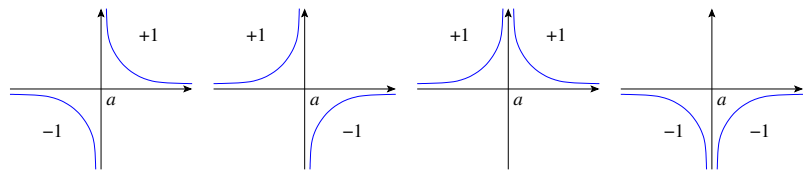
Il sera avantageux de considérer plus généralement des fractions Q/P.

Indice de Cauchy en un point

Pour $R, S \in \mathbf{R}[X]^*$ on définit $f : \mathbf{R} \smallsetminus \mathscr{Z}(S) \to \mathbf{R}$ par $f(x) = R(x)/S(x)$.
Ici $\mathscr{Z}(S) = \{x \in \mathbf{R} \mid S(x) = 0\}$ est l'ensemble des racines de S.

On note $\lim_a^- f$ et $\lim_a^+ f$ les limites à gauche et à droite en $a \in \mathbf{R}$.

$$\mathrm{Ind}_a^\varepsilon(f) := \begin{cases} +1 & \text{si } \lim_a^\varepsilon f = +\infty, \\ -1 & \text{si } \lim_a^\varepsilon f = -\infty, \\ 0 & \text{dans tous les autres cas.} \end{cases}$$



On définit l'indice de Cauchy de f en a par

$$\mathrm{Ind}_a(f) := \tfrac{1}{2}\big[\mathrm{Ind}_a^+(f) - \mathrm{Ind}_a^-(f)\big].$$

Indice de Cauchy sur un intervalle

Définition

Pour $a < b$ dans $\mathbf{R}$ on définit l'indice de Cauchy de f sur $[a,b]$ par

$$\operatorname{Ind}_a^b(f) := \tfrac{1}{2}\operatorname{Ind}_a^+(f) + \sum_{x\in]a,b[} \operatorname{Ind}_x(f) - \tfrac{1}{2}\operatorname{Ind}_b^-(f).$$

Pour $b < a$ on pose $\operatorname{Ind}_a^b(f) := -\operatorname{Ind}_b^a(f)$, et on pose $\operatorname{Ind}_a^a(f) := 0$.

La somme est bien définie : seul un nombre fini de points contribuent.

L'indice jouit des propriétés suivantes (similaire à l'intégrale) :

découpage: $\operatorname{Ind}_a^b(f) + \operatorname{Ind}_b^c(f) = \operatorname{Ind}_a^c(f)$ pour tout $a,b,c \in \mathbf{R}$.
invariance: $\operatorname{Ind}_a^b(f\circ\tau) = \operatorname{Ind}_{\tau(a)}^{\tau(b)}(f)$ pour tout fonction $\tau(x) = ux + v$.
somme: $\operatorname{Ind}_a^b(f+g) = \operatorname{Ind}_a^b(f) + \operatorname{Ind}_a^b(g)$ s'il n'y pas de pôle commun.
produit: $\operatorname{Ind}_a^b(gf) = \sigma \operatorname{Ind}_a^b(f)$ si $g|_{[a,b]}$ est de signe constant $\sigma \in \{\pm 1\}$.
réduction: $\operatorname{Ind}_a^b\big(QR/QS\big) = \operatorname{Ind}_a^b\big(R/S\big)$ pour tout $Q,R,S \in \mathbf{R}[X]^*$.

On peut donc définir $\operatorname{Ind}_a^b(\frac{R}{S}) := \operatorname{Ind}_a^b(f)$ pour $\frac{R}{S} \in \mathbf{R}(X)^*$.

Cas exceptionnels : si ou $R = 0$ ou $S = 0$, on pose $\operatorname{Ind}_a^b(\frac{R}{S}) := 0$.
Ceci définit l'indice sur toute la droite projective $\mathbf{PR}(X)$.

Comptage des racines

Proposition (dérivée logarithmique)

Nous avons $\operatorname{Ind}_a(f'/f) = \begin{cases} +1 & \text{si } a \text{ est un zéro de } f, \\ -1 & \text{si } a \text{ est un pôle de } f, \\ 0 & \text{sinon.} \end{cases}$

Démonstration.
On a $f = (X-a)^m g$ où $m \in \mathbb{Z}$ et $g \in \mathbf{R}(X)^*$ tel que $g(a) \in \mathbf{R}^*$.
D'après Leibniz on a $f' = m(X-a)^{m-1}g + (X-a)^m g'$ donc

$$\frac{f'}{f} = \frac{m}{X-a} + \frac{g'}{g}.$$

Ici g'/g n'a pas de pôle en a. On conclut que $\operatorname{Ind}_a(f'/f) = \operatorname{sign}(m)$. □

Corollaire

Pour tout polynôme $P \in \mathbf{R}[X]^*$ *et pour tout* $a < b$ *dans* $\mathbf{R}$
l'indice $\operatorname{Ind}_a^b(P'/P)$ *compte le nombre des racines de* P *dans* $[a,b]$.
D'éventuelles racines sur le bord $\{a,b\}$ *comptent pour un demi.*

Inconvénient : a priori $\operatorname{Ind}_a^b(f)$ nécessite la connaissance de tous les pôles.
Solution : Cette difficulté sera surmontée par le théorème de Sturm !

Variation des signes

On compte les changements de signes entre $s_0, s_1 \in \mathbf{R}$ par

$$V(s_0,s_1) := \tfrac{1}{2}\big|\operatorname{sign}(s_0) - \operatorname{sign}(s_1)\big|.$$

Ceci résume les 9 cas suivants :

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

Définition

La variation des signes d'une suite $s = (s_0,\dots,s_n)$ dans $\mathbf{R}$ est définie par

$$V(s) := \sum_{k=1}^{n} \tfrac{1}{2}\big|\operatorname{sign}(s_{k-1}) - \operatorname{sign}(s_k)\big|.$$

La variation des signes d'une suite $(S_0,\dots,S_n)$ dans $\mathbf{R}[X]$ en $a \in \mathbf{R}$ est

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

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

Le théorème des valeurs intermédiaires prend la forme $\operatorname{Ind}_a^b(\frac{1}{P}) = V_a^b(P)$.

La formule d'inversion

Proposition (Cauchy 1837)

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

$$\operatorname{Ind}_a^b\Big(\frac{Q}{P}\Big) + \operatorname{Ind}_a^b\Big(\frac{P}{Q}\Big) = V_a^b(P,Q).$$

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

Si $[a,b]$ ne contient aucun pôle, alors le coté gauche est nul ;
le coté droite est nul par le théorème des valeurs intermédiaires.

La formule est additive par rapport au découpage de l'intervalle $[a,b]$.
On peut donc supposer que a est l'unique pôle, puis $P(a) = 0$ et $Q(a) \neq 0$.

Ainsi Q est de signe constant sur $[a,b]$, et P est de signe constant sur $]a,b]$.

À gauche on trouve $\operatorname{Ind}_a^b(\frac{P}{Q}) = 0$ et $\operatorname{Ind}_a^b(\frac{Q}{P}) = \frac{1}{2}\operatorname{Ind}_a^+(\frac{Q}{P})$.

À droite on trouve $V_a(P,Q) = \frac{1}{2}$. Pour V_b on distingue deux cas :

- Si $V_b(P,Q) = 0$, alors $PQ > 0$ sur $]a,b]$, donc $\lim_a^+(\frac{Q}{P}) = +\infty$.
- Si $V_b(P,Q) = 1$, alors $PQ < 0$ sur $]a,b]$, donc $\lim_a^+(\frac{Q}{P}) = -\infty$.

Dans les deux cas on trouve $\frac{1}{2}\operatorname{Ind}_a^+(\frac{Q}{P}) = V_a^b(P,Q)$. □

Suites de Sturm

Définition

Une suite $(S_0, \dots, S_n)$ dans $\mathbf{R}[X]$ est dite *de Sturm* par rapport à un intervalle $[a,b] \subset \mathbf{R}$ si elle satisfait à la condition suivante :

Si $S_k(x) = 0$ où $0 < k < n$ et $x \in [a,b]$, alors $S_{k-1}(x)S_{k+1}(x) < 0$.

Corollaire (de la formule d'inversion)

Si $(S_0, S_1, \dots, S_{n-1}, S_n)$ est une suite de Sturm dans $\mathbf{R}[X]$, alors

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

Démonstration. Pour $n = 2$ la formule d'inversion nous donne

$$\operatorname{Ind}_a^b\Big(\frac{S_1}{S_0}\Big) + \operatorname{Ind}_a^b\Big(\frac{S_0}{S_1}\Big) + \operatorname{Ind}_a^b\Big(\frac{S_2}{S_1}\Big) + \operatorname{Ind}_a^b\Big(\frac{S_1}{S_2}\Big) = V_a^b\big(S_0, S_1, S_2\big).$$

Les contributions au milieu s'annulent. On conclut par récurrence sur n. □

Suites de Sturm par l'algorithme d'Euclide

Pour $P_0, P_1 \in \mathbf{R}[X]^*$ on itère la division euclidienne : $P_{k+1} = Q_k P_k - P_{k-1}$.
Ce processus s'arrête avec $P_{n+1} = 0$ et $P_n \sim \gcd(P_0, P_1)$.

Définition & proposition

La suite de Sturm euclidienne $(S_0, S_1, \dots, S_n)$ associée à (P_0, P_1) est définie par $S_k := P_k/P_n$. Il s'agit effectivement d'une suite de Sturm.

Démonstration. Pour $0 < k < n$ nous avons

$$S_{k+1} = Q_k S_k - S_{k-1}.$$

Si $S_k(x) = 0$, alors $S_{k+1}(x) = -S_{k-1}(x)$.
Si l'on avait $S_{k-1}(x) = S_{k+1}(x) = 0$, alors on aurait $S_j(x) = 0$ pour tout j.
Or, $S_n = 1$ par construction, donc cette dégénérescence est impossible. □

Définition

Pour $\frac{P_1}{P_0} \in \mathbf{R}(X)$ et $a, b \in \mathbf{R}$ on définit *l'indice de Sturm* par

$$V_a^b\Big(\frac{P_1}{P_0}\Big) := V_a^b\big(S_0, S_1, \dots, S_n\big).$$

Cas exceptionnels : si ou $P_0 = 0$ ou $P_1 = 0$, on pose $V_a^b(\frac{P_1}{P_0}) = 0$.
Ceci définit l'indice de Sturm sur toute la droite projective $\mathbf{PR}(X)$.

Complément : fractions continues

La suite $S_{k+1} = Q_k S_k - S_{k-1}$ est le développement en fraction continue :

$$\frac{P_1}{P_0} = \frac{S_1}{S_0} = \frac{S_1}{Q_1 S_1 - S_2} = \frac{1}{Q_1 - \dfrac{S_2}{S_1}} = \dots = \cfrac{1}{Q_1 - \cfrac{1}{Q_2 - \cfrac{\dots}{Q_{n-1} - \cfrac{1}{Q_n}}}}.$$

Ainsi les quotients $(Q_1, \dots, Q_n)$ suffisent pour reconstruire $\frac{P_1}{P_0} = \frac{S_1}{S_0}$.
Ils sont souvent plus efficaces à stocker que la suite $(S_0, S_1, \dots, S_n)$.
Génériquement on s'attend à $\deg(Q_k) = 1$, donc $Q_k = a_k X + b_k$.

Le théorème de Sturm

Théorème (Sturm 1829/35, Cauchy 1831/37)

Pour tout $R, S \in \mathbf{R}[X]$ nous avons l'égalité

$$\operatorname{Ind}_a^b\Big(\frac{R}{S}\Big) = V_a^b\Big(\frac{R}{S}\Big).$$

Démonstration. Soit $(S_0, S_1, \dots, S_n)$ la suite de Sturm euclidienne pour $\frac{R}{S}$:

$$\operatorname{Ind}_a^b\Big(\frac{R}{S}\Big) = \operatorname{Ind}_a^b\Big(\frac{S_1}{S_0}\Big) + \operatorname{Ind}_a^b\Big(\frac{S_{n-1}}{S_n}\Big) = V_a^b\big(S_0, S_1, \dots, S_n\big) = V_a^b\Big(\frac{R}{S}\Big).$$

Corollaire (Sturm 1829/35)

Pour tout polynôme $P \in \mathbf{R}[X]^*$ nous avons

$$\#\big\{x \in [a,b] \;\big|\; P(x) = 0\big\} = \operatorname{Ind}_a^b\Big(\frac{P'}{P}\Big) = V_a^b\Big(\frac{P'}{P}\Big).$$

D'éventuelles racines sur le bord comptent pour un demi.

Commentaires

Remarque

Le théorème de Sturm réduit un problème 1-dimensionnel sur l'intervalle $[a, b]$ à un problème 0-dimensionnel au bord $\{a, b\}$. L'approche se généralise au cas complexe, où l'on réduit un problème 2-dimensionnel sur un rectangle Γ à un problème 1-dimensionnel sur le bord $\partial\Gamma$. (Voir plus bas.)

Remarque

Le théorème de Sturm est traditionnellement formulé sous deux hypothèses supplémentaires, à savoir $\operatorname{pgcd}(R, S) = 1$ et $S(a)S(b) \neq 0$. Ici l'hypothèse $\operatorname{pgcd}(R, S) = 1$ est rendue superflue en définissant Ind_a^b et V_a^b correctement sur la droite projective $\mathbf{PR}(X)$. Le cas $S(a)S(b) = 0$ est résolu en comptant correctement les points au bord.

Remarque

La propriété des valeurs intermédiaires est essentielle : sur $\mathbb{Q}$ l'énoncé serait faux : $f(x) = 2x/(x^2 - 2)$ n'a pas de pôles, d'où $\operatorname{Ind}_1^2(f) = 0$. La suite de Sturm euclidienne est $S_0 = X^2 - 2$ and $S_1 = 2X$ and $S_2 = 2$, d'où $V_1^2(S_0, S_1, S_2) = 1$. On voit que l'indice de Sturm ne compte pas les pôles dans $\mathbb{Q}$ mais dans la clôture réelle $\mathbb{Q}^c$.

Objectif

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

Nous allons construire une fonction

$$\operatorname{ind}\colon \left\{ \begin{array}{c} \text{lacets } \gamma\colon [0,1] \to \mathbf{C}^* \\ \text{polynomiaux par morceaux} \end{array} \right\} \to \mathbb{Z}$$

qui compte le nombre de tours autour de 0.

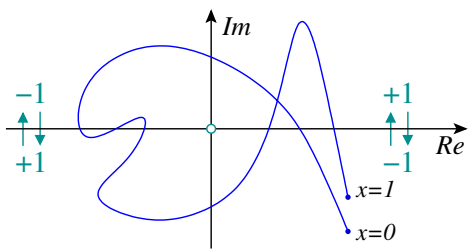
Plus précisément :

1. Pour tout rectangle $\Gamma \subset \mathbf{C}$ on a $\operatorname{ind}(\partial\Gamma) = \begin{cases} 1 & \text{si } 0 \in \operatorname{Int}\Gamma, \\ 0 & \text{si } 0 \in \mathbf{C} \smallsetminus \Gamma. \end{cases}$
2. $\operatorname{ind}(\gamma_1 \cdot \gamma_2) = \operatorname{ind}(\gamma_1) + \operatorname{ind}(\gamma_2)$.
3. $\operatorname{ind}(\gamma_0) = \operatorname{ind}(\gamma_1)$ si $\gamma_0 \sim \gamma_1$ sont homotopes dans $\mathbf{C}^*$.

Bénéfice algorithmique : l'indice se calcule par les suites de Sturm !

L'indice d'un chemin dans le plan complexe

Pour $F \in \mathbf{C}[X]$ la fonction $\gamma\colon [0,1] \to \mathbf{C}$, $x \mapsto F(x)$ décrit un chemin dans $\mathbf{C}$.



Observation

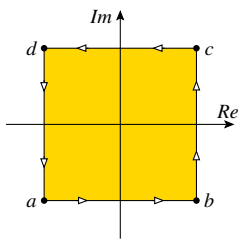
L'indice $\operatorname{ind}_0^1(F) := \frac{1}{2}\operatorname{Ind}_0^1\left(\frac{\operatorname{re} F}{\operatorname{im} F}\right)$ compte les tours autour de 0.

Pour $a, b \in \mathbf{C}$ on considère le chemin $\gamma\colon [0,1] \to \mathbf{C}$,
$\gamma(x) = (b-a)x + a$ allant de $\gamma(0) = a$ vers $\gamma(1) = b$.
Il paramétrise le segment $[a, b] = \{(1-x)a + xb \mid 0 \leq x \leq 1\}$.

Définition

Pour $F \in \mathbf{C}[Z]$ et $a, b \in \mathbf{C}$ on pose $\operatorname{ind}_a^b(F) = \operatorname{ind}_0^1 F\big((b-a)X + a\big)$.

L'indice par rapport à un rectangle



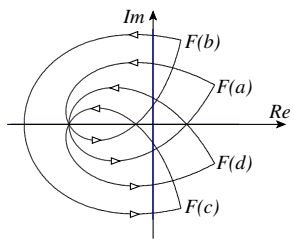


Illustration : $F = Z^5 - 5Z^4 - 2Z^3 - 2Z^2 - 3Z - 12$ sur $\Gamma = [-1, +1] \times [-1, +1]$.

Définition

Étant donné un polynôme $F \in \mathbf{C}[Z]$ et un rectangle $\Gamma \subset \mathbf{C}$, on pose

$$\operatorname{ind}_{\partial\Gamma}(F) := \operatorname{ind}_a^b(F) + \operatorname{ind}_b^c(F) + \operatorname{ind}_c^d(F) + \operatorname{ind}_d^a(F).$$

Exemple

$$\operatorname{ind}_{\partial\Gamma}(Z - z_0) = \begin{cases} 1 & \text{si } z_0 \text{ est dans l'intérieur de } \Gamma, \\ \frac{1}{2} & \text{si } z_0 \text{ est sur une arête de } \Gamma, \\ \frac{1}{4} & \text{si } z_0 \text{ est un sommet de } \Gamma, \\ 0 & \text{si } z_0 \text{ est à l'extérieur de } \Gamma. \end{cases}$$

La formule du produit

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

Lemme

Si aucune des paires (P, Q) et (R, S) n'a de zéro commun en a ni en b, alors

$$\mathrm{Ind}_a^b\Big(\frac{PR - QS}{PS + QR}\Big) = \mathrm{Ind}_a^b\Big(\frac{P}{Q}\Big) + \mathrm{Ind}_a^b\Big(\frac{R}{S}\Big) - V_a^b\big(1, QS(PS + QR)\big).$$

Pour $P = S$ et $Q = R$ on retrouve la formule d'inversion.
La preuve du cas général suit les mêmes principes.

Théorème

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

$$\mathrm{ind}_{\partial\Gamma}(F \cdot G) = \mathrm{ind}_{\partial\Gamma}(F) + \mathrm{ind}_{\partial\Gamma}(G).$$

Corollaire

Supposons que $F \in \mathbf{C}[Z]$ soit scindé, $F = c(Z - z_1)\cdots(Z - z_n)$, et qu'aucune des racines $z_1, \ldots, z_n$ ne tombe sur un sommet de Γ. Alors $\mathrm{ind}_{\partial\Gamma}(F)$ compte le nombre des racines dans Γ.

Il faut encore s'affranchir de l'hypothèse que F soit scindé !

Localisation grossière des racines

Définition (rayon de Cauchy)

Soit $F = a_n Z^n + a_{n-1}Z^{n-1} + \cdots + a_1 Z + a_0$ dans $\mathbf{C}[Z]$ où $a_n \neq 0$.
On pose $M := \max\{0, |a_0|, \ldots, |a_{n-1}|\}$ et on définit $\rho_F := 1 + M/|a_n|$.

Théorème

Si $z \in \mathbf{C}$ vérifie $|z| \geq \rho_F$, alors $|F(z)| \geq |a_n| > 0$.

Ainsi toutes les racines de F dans $\mathbf{C}$ sont dans $B(\rho_F) = \{z \in \mathbf{C} \mid |z| < \rho_F\}$.

Démonstration. L'affirmation est vraie pour $F = a_n Z^n$ où $M = 0$ et $\rho_F = 1$.
Dans la suite nous pouvons donc supposer que $M > 0$ et $\rho_F > 1$.
Pour tout $z \in \mathbf{C}$ vérifiant $|z| \geq \rho_F$ nous trouvons

$$|F(z) - a_n z^n| = |a_0 + a_1 z + \cdots + a_{n-1}z^{n-1}| \leq |a_0| + |a_1||z| + \cdots + |a_{n-1}||z^{n-1}|$$
$$\leq M + M|z| + \cdots + M|z|^{n-1} = M\tfrac{|z|^n - 1}{|z| - 1} \leq |a_n|(|z|^n - 1).$$

Pour la dernière inégalité on utilise que $|z| \geq \rho_F$ implique
$|z| - 1 \geq \rho_F - 1 = M/|a_n|$. On a

$$|a_n z^n| = |a_n z^n - F(z) + F(z)| \leq |a_n z^n - F(z)| + |F(z)|, \quad \text{d'où}$$
$$|F(z)| \geq |a_n z^n| - |F(z) - a_n z^n| \geq |a_n||z|^n - |a_n|(|z|^n - 1) = |a_n| > 0.$$

Comptage des racines

Lemme (version locale)

Si $F \in \mathbf{C}[X, Y]$ vérifie $F(x, y) \neq 0$ dans un point $(x, y) \in \mathbf{R}^2$, alors il existe $\delta > 0$ tel que $\mathrm{ind}_{\partial\Gamma}(F) = 0$ pour tout $\Gamma \subset [x - \delta, x + \delta] \times [y - \delta, y + \delta]$.

Théorème (version globale)

Soit $\Gamma = [x_0, x_1] \times [y_0, y_1]$ un rectangle dans $\mathbf{C}$. Si $F \in \mathbf{C}[X, Y]$ vérifie $F(x, y) \neq 0$ pour tout $(x, y) \in \Gamma$, alors $\mathrm{ind}_{\partial\Gamma}(F) = 0$.

Corollaire (comptage des racines)

Supposons que $F \in \mathbf{C}[Z]^*$ ne s'annule pas dans les sommets de $\Gamma \subset \mathbf{C}$. Alors $\mathrm{ind}_{\partial\Gamma}(F)$ compte les racines de F dans Γ.

Démonstration. On factorise $F = (Z - z_1)\cdots(Z - z_m)G$
tel que le facteur restant $G \in \mathbf{C}[Z]^*$ n'ait pas de racines dans $\mathbf{C}$.
L'affirmation découle de la formule du produit et du théorème ci-dessus. □

Invariance par homotopie

Théorème

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

Démonstration. L'absence des zéros sur $[0, 1] \times [a, b]$ nous assure que

$$\mathrm{ind}_a^b(F \mid T = 0) - \mathrm{ind}_a^b(F \mid T = 1) = \mathrm{ind}_0^1(F \mid Z = a) - \mathrm{ind}_0^1(F \mid Z = b).$$

La somme sur les quatre cotés de Γ donne $\mathrm{ind}_{\partial\Gamma}(F_0) - \mathrm{ind}_{\partial\Gamma}(F_1) = 0$. □

Corollaire

Pour tout polynôme $F \in \mathbf{C}[Z]^*$ et tout rectangle $\Gamma \subset \mathbf{C}$ contenant $B(\rho_F)$ nous avons $\mathrm{ind}_{\partial\Gamma}(F) = \deg F$.

Démonstration. Soit $F = a_n Z^n + a_{n-1}Z^{n-1} + \cdots + a_0$ où $a_n \neq 0$.
On déforme $F_1 = F$ en $F_0 = a_n Z^n$ par $F_t = a_n Z^n + t(a_{n-1}Z^{n-1} + \cdots + a_0)$.
Le rayon de Cauchy $r_t = 1 + tM/|a_n|$ décroît de $r_1 = \rho_F$ à $r_0 = 1$.
Ainsi F_t n'a pas de racines sur $\partial\Gamma$, donc $\mathrm{ind}_{\partial\Gamma}(F_1) = \mathrm{ind}_{\partial\Gamma}(F_0) = n$. □

Ceci achève notre démonstration du théorème fondamental de l'algèbre :
Le rectangle $\Gamma \subset \mathbf{C}$ contient n racines de F.

Localisation des racines complexes

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

On peut construire l'indice algébrique, ayant les bonnes propriétés,

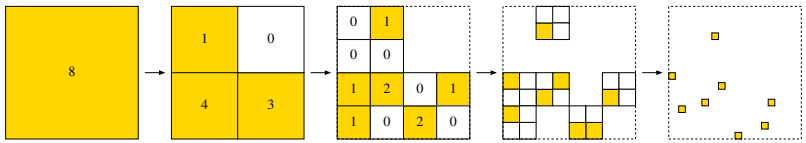
$$\mathrm{ind}\colon \left\{ \begin{array}{c} \text{lacets } \gamma\colon [0,1] \to \mathbf{C}^* \\ \text{polynomiaux par morceaux} \end{array} \right\} \to \mathbb{Z}.$$

La construction est réelle algébrique.
$\Longrightarrow$ arithmétique des polynômes au lieu de l'analyse
$\Longrightarrow$ calcul formel au lieu de calcul numérique

Cet outil permet de prouver le théorème fondamental de l'algèbre :

Le corps $\mathbf{C}$ est algébriquement clos.

Mieux, l'indice permet de localiser les racines de tout polynôme $F \in \mathbb{C}[Z]$:



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

Complément : cross-over vers la méthode de Newton

Newton itère l'application $\Phi\colon \mathbb{C} \smallsetminus \mathscr{Z}(F') \to \mathbb{C}$, $\Phi(z) = z - F(z)/F'(z)$:

Théorème

Les points fixes de Φ sont les zéros simples de F, c'est-à-dire, les points $z_0 \in \mathbb{C}$ tels que $F(z_0) = 0$ et $F'(z_0) \neq 0$. Ce sont des points fixes super-attractifs : il existe $\delta > 0$ tel que toute valeur initiale $u_0 \in B(z_0, \delta)$ satisfasse $|\Phi^n(u_0) - z_0| \leq 2^{1-2^n} \cdot |u_0 - z_0|$ for all $n \in \mathbb{N}$.

La convergence est donc extrêmement rapide. La difficulté est de trouver des bonnes valeurs initiales $u_0 \approx z_0$. Le critère suivant en donne une solution :

Proposition (cross-over du global au local)

Soit $F \in \mathbb{C}[Z]$ de degré n et séparable. Supposons que nous avons séparé les racines en n disques disjoints $B(u_k, \delta_k)$ où $k = 1, \ldots, n$ tels que

$$3n\delta_k \leq |u_k - u_j| \qquad \text{pour tout } j \neq k.$$

Alors la méthode de Newton converge pour toute valeur initiale u_k vers la racine correspondante $z_k \in B(u_k, \delta_k)$. Dès le début la convergence de Newton est au moins aussi rapide que la dichotomie :

$$|\Phi^n(u_k) - z_k| \leq 2^{-n}|u_k - z_k| \qquad \text{for all } n \in \mathbb{N}.$$

Complément : complexité algorithmique

Pour mettre cette méthode en œuvre sur ordinateur, il vaut mieux calculer avec des coefficients dans $\mathbb{Q}[i]$: calculs exacts sans erreurs d'arrondi !

Il est souvent plus efficace de calculer dans $\mathbb{Z}[i]$. On peut toujours se ramener à ce cas en multipliant par le ppcm des dénominateurs.

Soit $F = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$ un polynôme de degré n dans $\mathbb{Z}[i][Z]$ tel que $|\operatorname{re} a_k| \leq 2^a$ et $|\operatorname{im} a_k| \leq 2^a$ pour tout k.

Supposons que toutes les racines sont dans le disk $B(r)$.

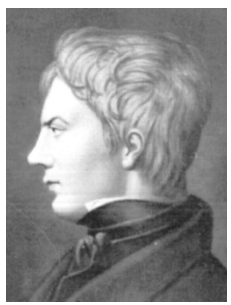
Théorème

Pour approcher les racines de F à une précision de $3r/2^b$ près l'algorithme ci-dessus nécessite $\tilde{O}(n^3 b(a + nb))$ opérations binaires.

Voir l'article pour des explications et des références.

Pour $a \approx nb$ la méthode algébrique est donc de complexité $\tilde{O}(n^4 b^2)$.

Le record mondial (Schönhage 1982) pour localiser les racines est $\tilde{O}(n^2(n + b))$, donc juste un ordre de grandeur meilleur.



Je vous remercie de votre attention !