Quelques applications des pseudozéros de polynômes

Stef Graillat
Université de Perpignan

graillat@univ-perp.fr
http://gala.univ-perp.fr/∼graillat

Séminaire du laboratoire MANO, 25 mars 2004

Introduction et motivations

But :

Travailler avec des polynômes ayant des données (coefficients ou racines) connues avec une incertitude : recherche de racines, primalité, calcul de PGCD, stabilité en automatique, etc.

Raisons :

- Résultats provenant d'expériences.
- Représentation des nombres en machine.

Applications :

- Traitement du signal et d'images.
- Robotique.
- Biologie moléculaire.
- Automatique.

Exemples du PGCD

Exemple 1 :

Soient p et q deux polynômes unitaires et $\deg p > 1$.
On suppose de plus que p divise $q \Longrightarrow \gcd(p, q) = p$.
Or pour toute constante $\varepsilon > 0$, on a $\gcd(p, q + \varepsilon) = 1$.

Exemple 2 :

$$\begin{aligned} p &= z^2 - 3.0001z + 1.9999 \approx (z-1)(z-2), \\ q &= z^2 - 1.9999z + 1.0001 \approx (z-1)^2. \end{aligned}$$

Pour ε une perturbation sur les coefficients, on aimerait dire :

- $\gcd(p, q) = z - 1$ pour $\varepsilon = 0.0001$.

Approche des problèmes

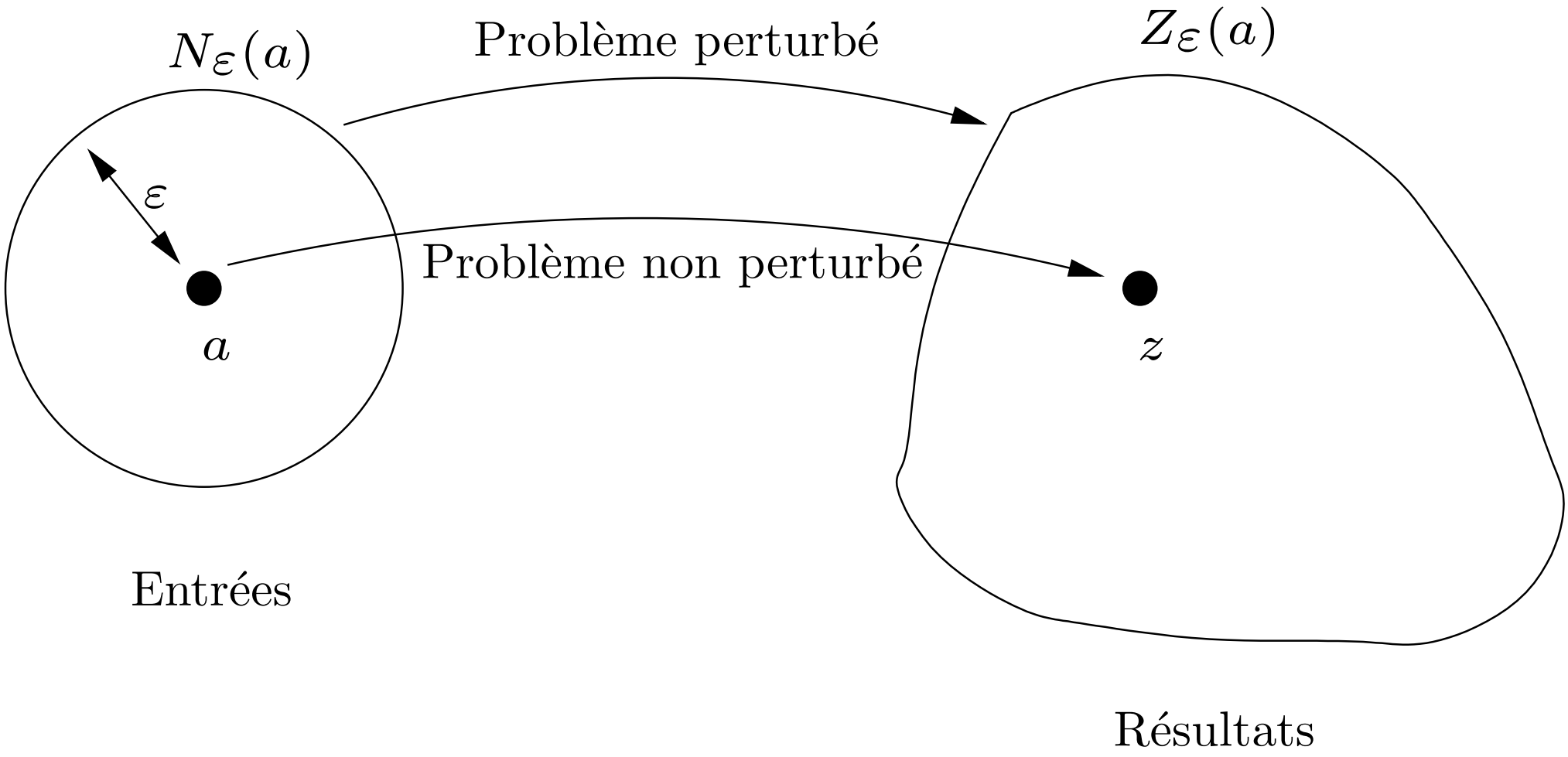
- Problèmes difficiles : on doit rechercher des « singularités ».
- La démarche générale :
  - Donner une définition précise de ce que l'on veut calculer.
  - Trouver des algorithmes pour ce calcul (souvent des heuristiques).
  - Certifier les résultats.
- appliquée à

  $\Rightarrow$ (ε-pseudo)zéros

  $\Rightarrow$ (ε-)primalité

  $\Rightarrow$ (ε-)stabilité en théorie du contrôle

Principe des calculs approchés



Plan de l'exposé

I – Les pseudozéros

- Définition et algorithme de calcul
- Problème du polynôme le plus proche avec une racine donnée

II – Application des pseudozéros à la primalité

- Définitions
- Présentation des algorithmes existants
- Apport des pseudozéros

III – Application des pseudozéros en théorie du contrôle

- Stabilité au sens de Hurwith et de Schur
- Calcul du rayon de stabilité

Les pseudozéros : définition, calcul et intérêt

Pseudozéros : définition

Perturbation :

Voisinage du polynôme $p \in \mathbf{C}_n[z]$

$$N_\varepsilon(p) = \{\widehat{p} \in \mathbf{C}_n[z] : \|p - \widehat{p}\| \leqslant \varepsilon\}\,.$$

Définition de l'ensemble des ε-pseudozéros :

$$Z_\varepsilon(p) = \{z \in \mathbb{C} : \widehat{p}(z) = 0 \text{ pour } \widehat{p} \in N_\varepsilon(p)\}\,.$$

$\|\cdot\|$, norme sur le vecteur des coefficients de p

Les pseudozéros

- Mosier (1986) : étude avec la norme $\|\cdot\|_\infty$.
- Trefethen et Toh (1994) : étude avec la norme $\|\cdot\|_2$.
  pseudozéros $\approx$ pseudospectres de la matrice compagnon.
- Chatelin et Frayssé (1996) : Synthèse des articles précédents dans *Lectures on Finite Precision Computations* (SIAM)
- Stetter (1999) : *algèbre numérique polynomiale*. Généralise les travaux précédents.
- Zhang (2001) : Étude en norme $\|\cdot\|_2$ de l'influence de la base (conditionnement de l'évaluation).

Les pseudozéros sont facilement calculables

Théorème :
L'ensemble des ε-pseudozéros vérifie

$$Z_\varepsilon(p) = \left\{ z \in \mathbb{C} : |g(z)| := \frac{|p(z)|}{\|\underline{z}\|_*} \leqslant \varepsilon \right\},$$

où $\underline{z} = (1, z, \ldots, z^n)$ et $\|\cdot\|_*$ est la norme duale de $\|\cdot\|$.

La démonstration nécessite de connaître « le » polynôme le plus proche de p ayant une racine donnée.

Le polynôme le plus proche ayant une racine donnée p_u

Soient p dans $\mathbf{C}_n[z]$ et $u \in \mathbf{C}$.

Énoncé du problème :

Trouver un polynôme $p_u \in \mathbf{C}_n[z]$ vérifiant $p_u(u) = 0$ et tel que s'il existe un polynôme $q \in \mathbf{C}_n[z]$ avec $q(u) = 0$ alors on ait $\|p - p_u\| \leqslant \|p - q\|$.

On cherche :

- une expression explicite de p_u ;
- un résultat d'unicité.

Calcul de p_u

Notons $\underline{u} := (1, u, u^2, \ldots, u^n) \in \mathbf{C}^{n+1}$.
Il existe $d \in \mathbf{C}^{n+1}$ vérifiant $d^* \underline{u} = \|\underline{u}\|_*$ et $\|d\| = 1$.
Définissons les polynômes r et p_u par

$$r(z) = \sum_{k=0}^{n} r_k z^k \quad \text{avec} \quad r_k = \overline{d}_k,$$

$$p_u(z) = p(z) - \frac{p(u)}{r(u)} r(z).$$

p_u est le polynôme le plus proche de p
ayant u comme racine.

Unicité de p_u

Une condition suffisante d'unicité :

Théorème. *Si la norme $\|\cdot\|$ est strictement convexe, alors p_u est unique.*

C'est le cas, par exemple, pour les normes $\|\cdot\|_p$ pour $1 < p < \infty$.

On n'a pas unicité pour $\|\cdot\|_1$ et $\|\cdot\|_\infty$. Pour $p(z) = 1 + z$

	$\|\cdot\|_1, \quad u = 1$		$\|\cdot\|_\infty, \quad u = 0$	
p_u	$p_1^{(1)}(z) = 0$	$p_1^{(2)}(z) = \frac{1}{3}(1 - z)$	$p_0^{(1)}(z) = z$	$p_0^{(2)}(z) = \frac{1}{2}z$
$p - p_i$	$z - 1$	$\frac{4}{3}z - \frac{2}{3}$	1	$\frac{1}{2}z + 1$
$\|p - p_i\|$	2	2	1	1

Cas polynôme réel, racine réelle

Soient p dans $\mathbf{R}_n[x]$ et $u \in \mathbf{R}$: identique au cas complexe. $\underline{u} := (1, u, u^2, \ldots, u^n) \in \mathbf{R}^{n+1}$.
Il existe $d \in \mathbf{R}^{n+1}$ vérifiant $d^* \underline{u} = \|\underline{u}\|_*$ et $\|d\| = 1$.
Définissons les polynômes $r \in \mathbf{R}[x]$ et $p_u \in \mathbf{R}[x]$ par

$$r(z) = \sum_{k=0}^{n} r_k x^k \quad \text{avec} \quad r_k = d_k,$$

$$p_u(x) = p(x) - \frac{p(u)}{r(u)} r(x).$$

p_u est le polynôme le plus proche de p
ayant u comme racine.

Problème ouvert

Soient p dans $\mathbf{R}_n[x]$ et $u \in \mathbf{C}\backslash\mathbf{R}$.

Énoncé du problème :

Trouver un polynôme $p_u \in \mathbf{R}_n[x]$ annulant u et tel que s'il existe un polynôme $q \in \mathbf{R}_n[x]$ avec $q(u) = 0$ alors on ait $\|p - p_u\| \leqslant \|p - q\|$.

- Pas encore de formule explicite comme dans le cas complexe.
- Solution par des méthodes d'optimisation.

Algorithme de calcul des pseudozéros

Tracé de ε-pseudozéros :

1. On maille un carré contenant toutes les racines de p (commande MATLAB : `meshgrid`).
2. On calcule $g(z) := \frac{|p(z)|}{\|\underline{z}\|}$ pour tous les points z de la grille.
3. On affiche la ligne de niveau $|g(z)| = \varepsilon$ (commande MATLAB : `contour`).

Problèmes :

- Localisation : trouver un carré contenant toutes les racines de p et tous les pseudozéros.
- Séparation : trouver un pas de grille qui sépare toutes les racines.

Choix de la grille

Soit p un polynôme unitaire de degré n et $\{z_i\}$ l'ensemble de ses n racines. Notons $r = \max_{i=1;\ldots;n} |z_i|$. On a [Mignotte, 1989]

$$\boxed{r \leqslant 1 + \|p\|_\infty.}$$

Notons $R := 1 + \|p\| + \varepsilon$. On montre que

$Z_\varepsilon(p) \subset B(0, R)$ boule fermée de centre 0 et de rayon R.

Complexité du tracé

Notons L la longueur du carré et h le pas de discrétisation. L'évaluation de $g(u)$ nécessite

- l'évaluation d'un polynôme, ce qui se fait en $\mathcal{O}(n)$,
- le calcul de la norme d'un vecteur qui se fait aussi en général en $\mathcal{O}(n)$.

La complexité de l'algorithme précédent est donc en

$$\boxed{\mathcal{O}((L/h)^2 n)}.$$

L et h dépendent de n mais aussi des coefficients du polynôme.

Simulation numérique

Ensemble des pseudozéros du polynôme « de Wilkinson »

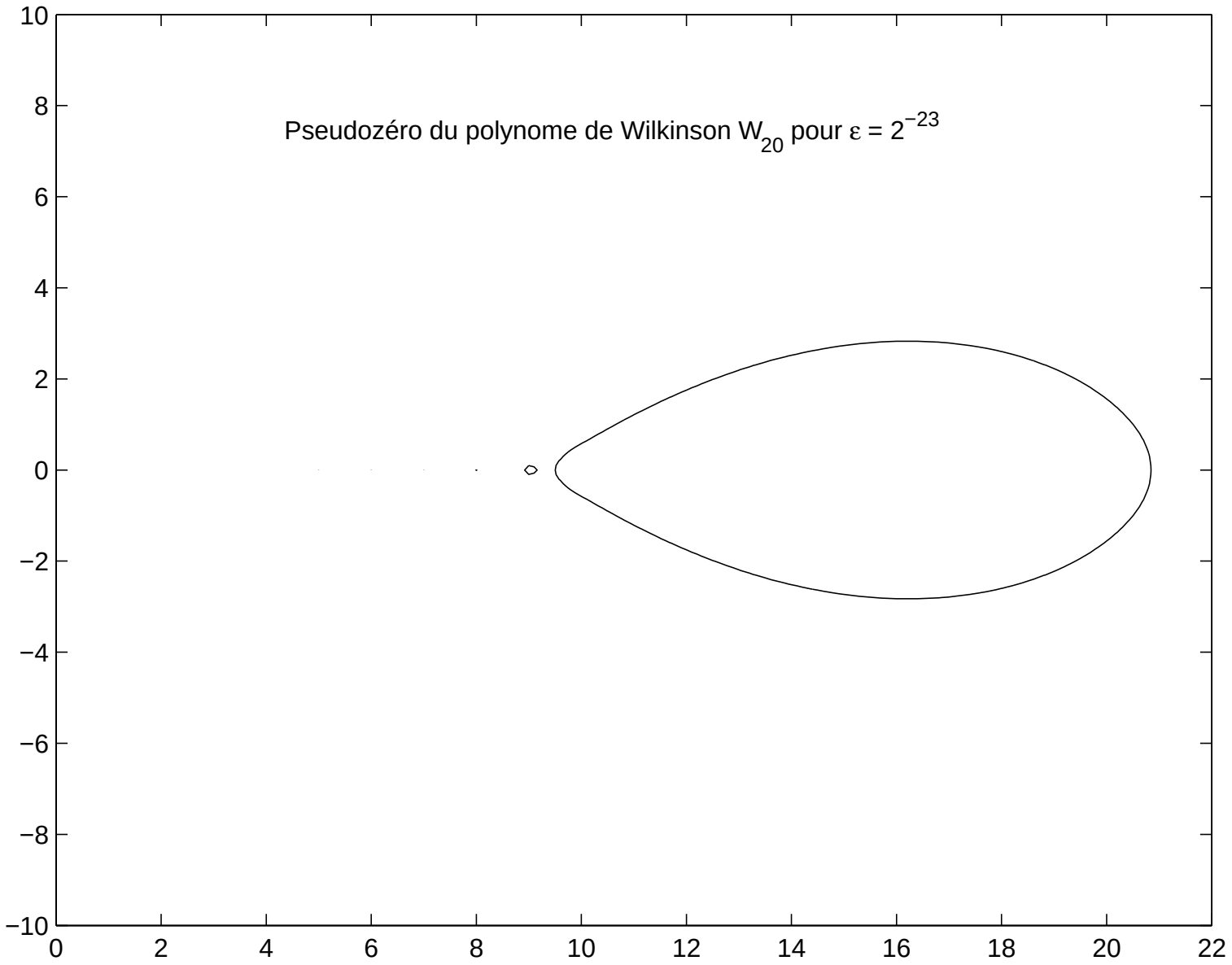
$$\begin{aligned} W_{20} &= (z-1)(z-2)\cdots(z-20), \\ &= z^{20} - 210z^{19} + \cdots + 20!, \end{aligned}$$

en ne perturbant que le coefficient de z^{19} avec une perturbation inférieure à $\varepsilon = 2^{-23}$.

On utilise une norme $\|\cdot\|_\infty$ pondérée :

$$\|p\|_\infty = \max_i \frac{|p_i|}{m_i} \text{ avec } m_i \text{ réels positifs}$$

et la convention $m/0 = \infty$ si $m > 0$ et $0/0 = 0$.


Pseudozéro du polynome de Wilkinson W_{20} pour $\varepsilon = 2^{-23}$
10
8
6
4
2
0
−2
−4
−6
−8
−10
0
2
4
6
8
10
12
14
16
18
20
22

Intérêts des pseudozéros

Intérêts

- une étude qualitative du polynôme
- comprendre les résultats des algorithmes
- utiliser des polynômes avec une incertitude sur leurs coefficients (mesure physique / précision finie)

Inconvénients

- le coût

Deux théorèmes relatifs aux pseudozéros

Théorème. [Mosier] *Si on suppose que l'ensemble des pseudozéros est borné et si $q \in N_\varepsilon(p)$ alors q et p ont le même nombre de racines (en comptant les multiplicités) dans chaque composante connexe de $Z_\varepsilon(p)$. De plus, il y a au moins une racine du polynôme p dans chaque composante connexe de $Z_\varepsilon(p)$.*

Théorème. [Mosier] *Un voisinage d'une racine de p contient deux racines de p si et seulement si il contient un $u \in \mathbf{C}$ comme racine double de p_u .*

Pseudozéros de polynômes réels

Si $p \in \mathbf{R}_n[x]$, on définit

$$N_\varepsilon(p) := \{q \in \mathbf{R}_n[x] : \|p - q\| \leqslant \varepsilon\}.$$

Deux cas :

- on cherche les pseudozéros réels : identique au cas complexe ;
- on cherche tous les pseudozéros complexes non réels.

On définit alors l'ensemble des pseudozéros par

$$Z_\varepsilon(p) := \{z \in \mathbf{C} : \widehat{p}(z) = 0 \text{ pour } \widehat{p} \in N_\varepsilon(p)\}.$$

$Z_\varepsilon(p)$ est symétrique par rapport à l'axe des réels.

Deux théorèmes relatifs au pseudozéros de polynômes réels

Théorème. [Stetter] *Soit* Z_μ *une composante connexe de* $Z_\varepsilon(p)$ *de multiplicité* 1. *Alors ou bien* $Z_\mu \subset \mathbf{R}$ *ou bien* $Z_\mu \cap \mathbf{R} = \varnothing$.

Théorème. [Stetter] *Une composante connexe* Z_μ *de* $Z_\varepsilon(p)$ *vérifiant* $\varnothing \neq Z_\mu \cap \mathbf{R} \neq Z_\mu$ *est de multiplicité au moins* 2.

Application des pseudozéros à la primalité

Définition d'un PGCD approché

Définition classique :

Soient p et q des polynômes de degrés respectifs n et m et soit ε un nombre positif. On appelle :

- **ε-diviseur** (ou diviseur approché) : tout diviseur des polynômes perturbés $\widehat{p}$ et $\widehat{q}$ vérifiant
$$\deg \widehat{p} \leqslant n, \deg \widehat{q} \leqslant m \text{ et } \max(\|p - \widehat{p}\|, \|q - \widehat{q}\|) \leqslant \varepsilon.$$
- **ε-PGCD** (PGCD approché) : un ε-diviseur de degré maximum.

Remarques :

- Tolérance sur les coefficients (nombres flottants / mesures).
- Unicité du degré mais non du ε-PGCD.
- Dépendance par rapport au corps de base.

Définition de la ε-primalité

Définition :

Deux polynômes p et q sont **ε-premiers entre eux**
si leur **ε-PGCD** est 1.

Calcul :

- Optimisation : algorithme de Karmarkar et Lakshman (1995).
- Borne de Sylvester : algorithme COPRIME [Beckermann et Labahn 1998].
- Graphique : les pseudozéros.

Algorithme de Karmarkar et Lakshman

Il s'agit d'un algorithme d'optimisation (moindre carré).

- **Entrée :** p et q deux polynômes unitaires de même degré n.
- **Sortie :** $\widehat{p}$, $\widehat{q}$ deux polynômes unitaires de degré n et α tels que α soit racine commune de $\widehat{p}$ et $\widehat{q}$ avec $\|p - \widehat{p}\|_2^2 + \|q - \widehat{q}\|_2^2$ minimum.
- **Complexité :** polynomiale en n.

[Karmarkar & Lakshman - 98]

Étape de l'algorithme

1. Le minimum de $\|p - \widehat{p}\|_2^2 + \|q - \widehat{q}\|_2^2$ comme fonction de α sous les contraintes $\widehat{p}(\alpha) = \widehat{q}(\alpha) = 0$ est

$$\mathcal{N}_M(\alpha) := \frac{p(\alpha)\overline{p(\alpha)} + q(\alpha)\overline{q(\alpha)}}{\sum_{j=0}^{n-1}(\alpha\overline{\alpha})^j}, \quad \alpha = a + ib \in \mathbf{C}.$$

2. Soit α tel que $\frac{\partial \mathcal{N}_M}{\partial a}(\alpha) = 0$, $\frac{\partial \mathcal{N}_M}{\partial b}(\alpha) = 0$ et $\mathcal{N}_M(\alpha)$ minimum.
3. On a

$$\widehat{p}_i = p_i - \frac{\overline{\alpha}^i p(\alpha)}{\sum_{k=0}^{n-1}(\overline{\alpha}\alpha)^k}, \quad \widehat{q}_i = q_i - \frac{\overline{\alpha}^i q(\alpha)}{\sum_{k=0}^{n-1}(\overline{\alpha}\alpha)^k}.$$

Algorithme COPRIME

$\|p\| = \sum |p_i|$, $\|(p,q)\| = \max\{\|p\|, \|q\|\} = \max\{\sum |p_i|, \sum |q_i|\}$.

Algorithme de Beckermann et Labahn (1998).

- **Entrée :** p et q deux polynômes.
- **Sortie :** borne inférieure de $\epsilon(p,q)$ défini par

$$\epsilon(p,q) = \inf\{\|(p-\widehat{p}, q-\widehat{q})\| : (\widehat{p},\widehat{q}) \text{ ont une racine commune et}$$

$$\deg\widehat{p} \leqslant n, \deg\widehat{q} \leqslant m\}.$$

- **Complexité :** en $\mathcal{O}((n+m)^2)$.

Matrice de Sylvester

$$S(p,q) = \begin{bmatrix} p_0 & 0 & \cdots & 0 & q_0 & 0 & \cdots & 0 \\ p_1 & p_0 & \ddots & \vdots & q_1 & q_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & \ddots & \ddots & 0 \\ p_n & & \ddots & p_0 & q_m & & \ddots & q_0 \\ 0 & p_n & & p_1 & 0 & q_m & & q_1 \\ \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & p_n & 0 & \dots & 0 & q_m \end{bmatrix} \in \mathbf{C}^{(n+m)\times(n+m)}.$$

Critère de Sylvester : p et q sont premier entre eux $\Longleftrightarrow$ la matrice $S(p,q)$ est régulière.

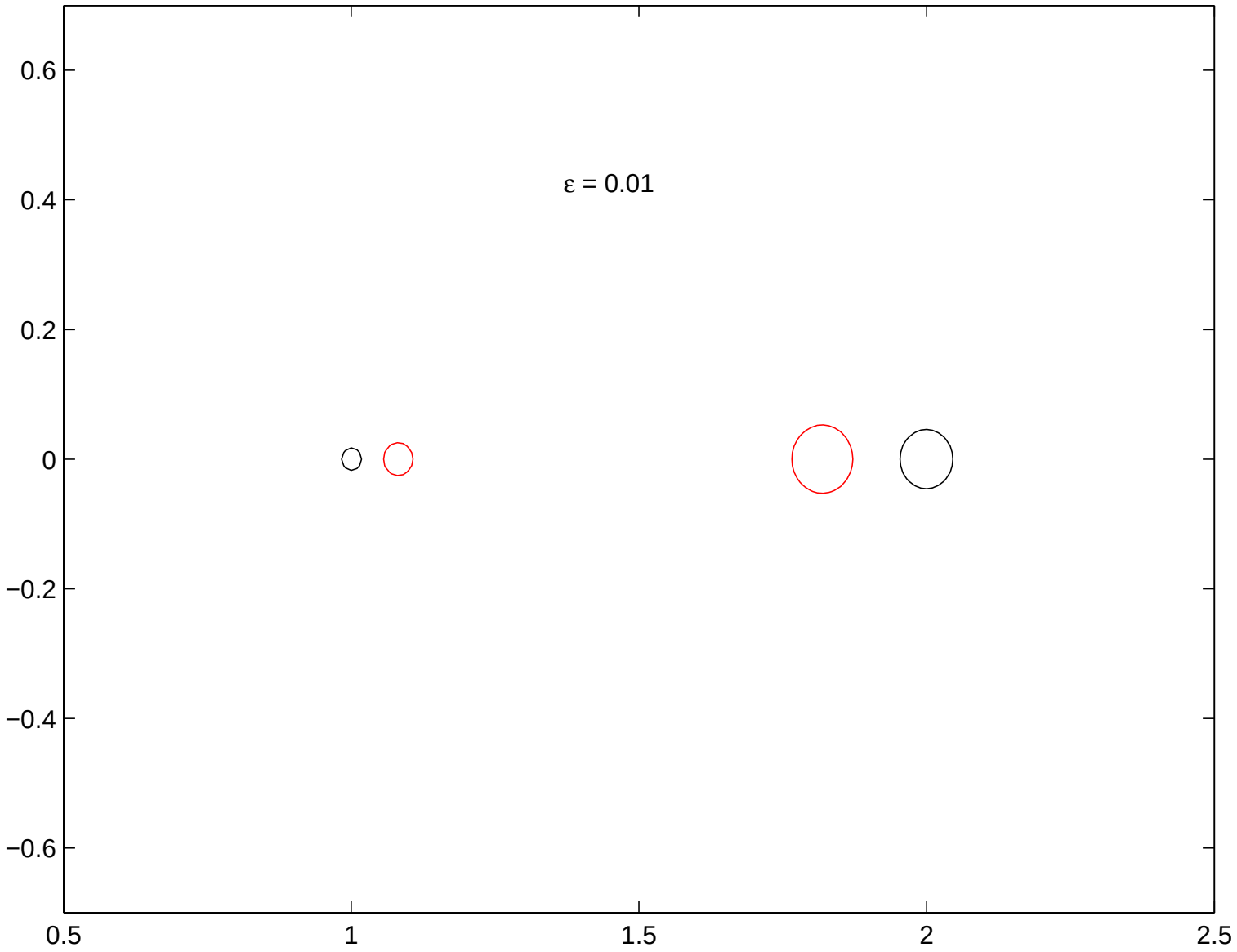
Présentation de la méthode

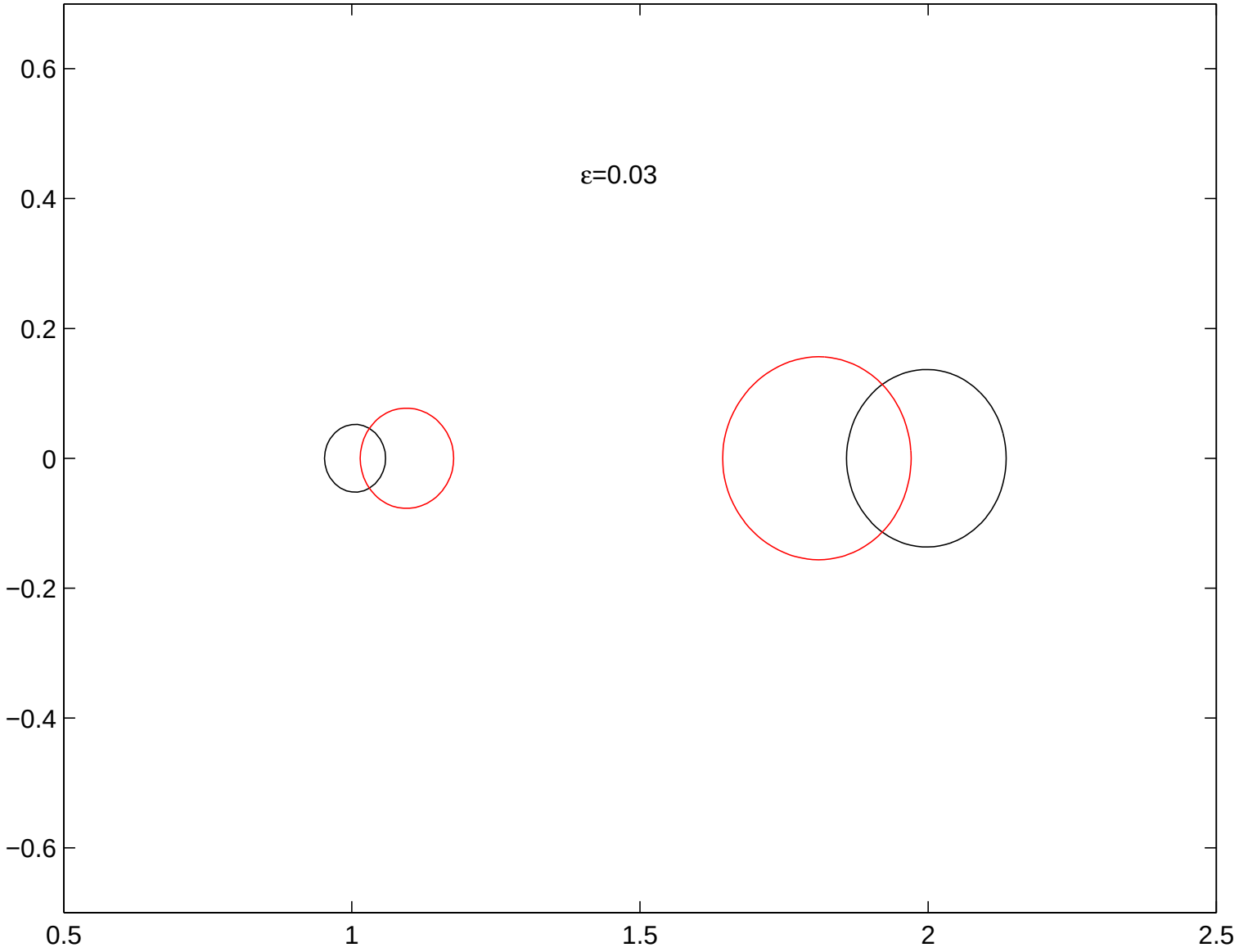
$$\epsilon(p,q) \geqslant \frac{1}{\|S(p,q)^{-1}\|}$$

- Estimation de $\|S(p,q)^{-1}\|$ basée sur une SVD trop coûteux.
- On cherche une borne supérieure de $\|S(p,q)^{-1}\|$.

Les pseudozéros

- **Entrée :** p et q deux polynômes.
- **Sortie :** un graphique.
- **Inconvénient :** outil qualitatif.
- **Exemple en $\|\cdot\|_2$:**
  $p = (z-1)(z-2) = z^2 - 3z + 2$
  $q = (z-1.08)(z-1.82) = z^2 - 2.9z + 1.9656$


ε = 0.01
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
0.5
1
1.5
2
2.5

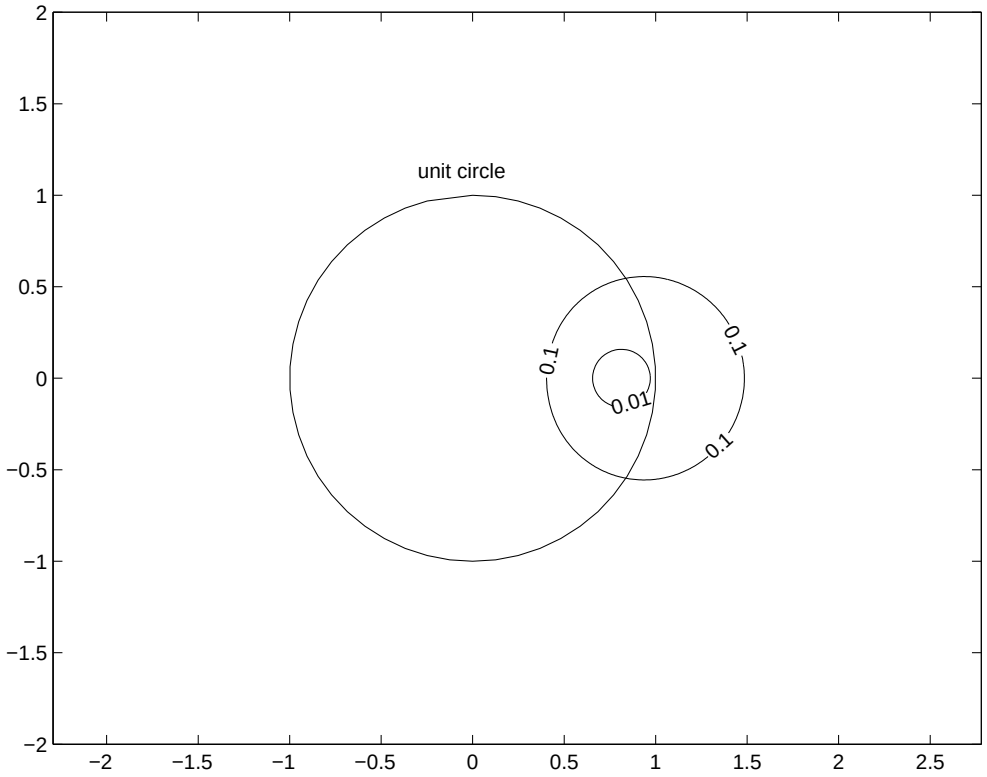

ε=0.03
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
0.5
1
1.5
2
2.5

Application des pseudozéros en théorie du contrôle

La stabilité robuste de Schur en théorie du contrôle

Stabilité de Schur : |racines de p| < 1.

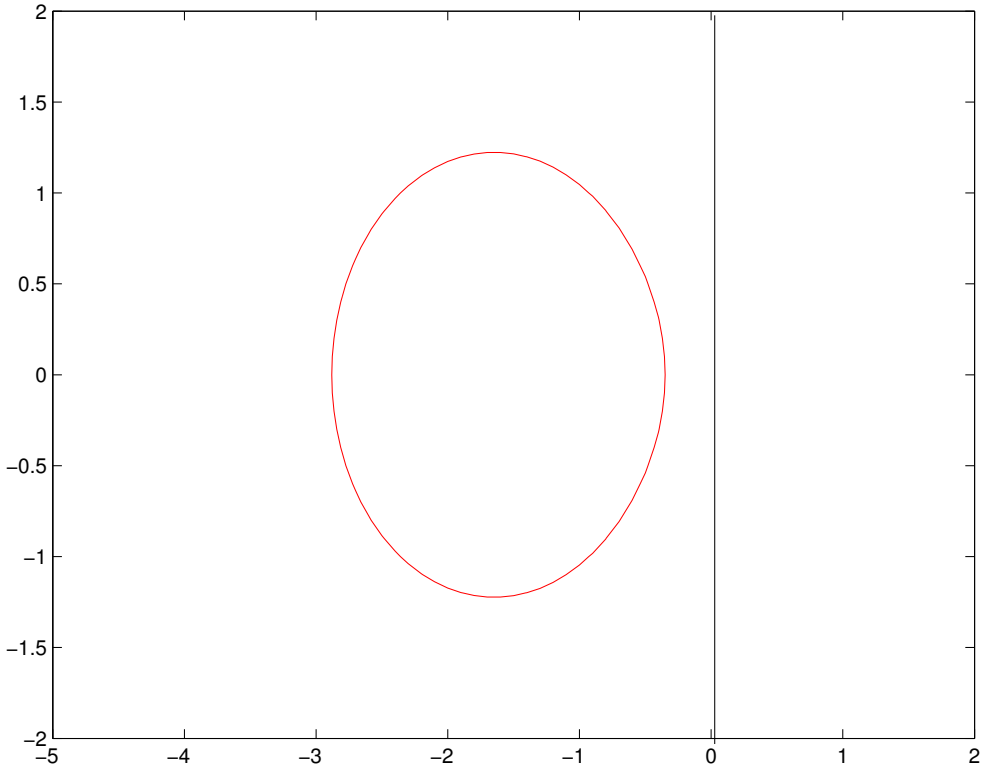
ε-pseudozero de $p(z) = (z - 0.8)^2$ pour $\varepsilon = 0.1$ and $\varepsilon = 0.01$.



La stabilité robuste de Hurwitz en théorie du contrôle

La stabilité de Hurwitz : la partie réelle des racines de $p < 0$.

ε-pseudozero de $p(z) = (z+1)^2$ pour $\varepsilon = 0.4$.



Stabilité d'un polynôme

$\mathcal{P}$: $\mathbf{C}[X]$ muni de la norme 2, $\|\cdot\|$
$\mathcal{P}_n$: éléments de $\mathcal{P}$ de degré inférieur ou égal à n
$\mathcal{M}_n$: éléments de $\mathcal{P}_n$ unitaires

Définition. *Un polynôme est dit* stable *si toutes ses racines ont une partie réelle strictement négative et* instable *sinon.*

La fonction *abscisse* $a : \mathcal{P} \to \mathbf{R}$ est définie par

$$a(p) = \max\{\mathrm{Re}(z) : p(z) = 0\}.$$

Un polynôme p est stable $\iff a(p) < 0$

Motivation

En théorie du contrôle, une fonction de transfert s'écrit souvent sous la forme $H(p) = \frac{N(p)}{D(p)}$ ou N et D sont des polynômes.

Le système est stable si D est un polynôme stable.

Question : si D est stable, est-on loin d'un système instable ?

Problème : Trouver la distance au système instable le plus proche.

(on se restreindra au cas où D est unitaire)

Énoncé du problème

Rayon de stabilité $\beta(p)$: distance du polynôme $p \in \mathcal{M}_n$ à l'ensemble des polynômes unitaires instables.

$$\beta(p) = \min\{\|p - q\| : q \in \mathcal{M}_n \text{ et } a(q) \geqslant 0\}.$$

Énoncé du problème :

Étant donné un polynôme $p \in \mathcal{M}_n$, calculer $\beta(p)$.

Solution proposée

Outils utilisés

- la formule explicite donnant les pseudozéros
- la dépendance continue des racines par rapport aux coefficients des polynômes
- l'algorithme de Sturm (compte les racines réelles)

Les résultats

- un algorithme calculant $\beta(p)$ avec une tolérance arbitraire τ
- un graphique montrant les pseudozéros à la distance $\beta(p)$
  $\longrightarrow$ permet une analyse qualitative du résultat
  $\longrightarrow$ visualisation du résultat

Une autre caractérisation de $Z_\varepsilon(p)$

Notons $h_{p,\varepsilon} : \mathbf{R}^2 \to \mathbf{R}$ la fonction définie par

$$h_{p,\varepsilon}(x,y) = |p(x+iy)|^2 - \varepsilon^2 \sum_{j=0}^{n-1} (x^2+y^2)^j.$$

On a alors

$$\boxed{Z_\varepsilon(p) = \{(x,y) \in \mathbf{R}^2 : h_{p,\varepsilon}(x,y) \leqslant 0\}}$$

$\Longrightarrow$ $h_\varepsilon(\cdot, y)$ et $h_\varepsilon(x, \cdot)$ sont des polynômes de degré $2n$.

Résultats théoriques utilisés

Proposition. *La fonction* abscisse

$$a : \mathcal{P}_n \to \mathbf{R}$$

définie par $a(p) = \max\{\mathrm{Re}(z) : p(z) = 0\}$ *est continue sur* $\mathcal{M}_n$.

Proposition. *On a la relation suivante*

$$\beta(p) = \min\{\|p - q\| : q \in \mathcal{M}_n \text{ et } a(q) = 0\}.$$

Théorème. *L'équation* $h_{p,\varepsilon}(0, y) = 0$ *a une solution* y *réelle si et seulement si* $\beta(p) \leqslant \varepsilon$.

Algorithme (dichotomie)

Entrée : un polynôme stable p et une tolérance τ sur la précision de $\beta(p)$ calculé

Sortie : un nombre α tel que $|\alpha - \beta(p)| \leqslant \tau$

1: $\gamma := 0, \quad \delta := \|p - z^n\|$
2: **tant que** $|\gamma - \delta| > \tau$ **faire**
3: $\quad \varepsilon := \frac{\gamma+\delta}{2}$
4: $\quad$ **si** l'équation $h_{p,\varepsilon}(0, y) = 0$ a une solution y réelle **alors**
5: $\quad\quad \delta := \varepsilon$
6: $\quad$ **autrement**
7: $\quad\quad \gamma := \varepsilon$
8: $\quad$ **fin si**
9: **fin tant que**
10: **retourne** $\alpha = \frac{\gamma+\delta}{2}$

Complexité de l'algorithme

Nécessite l'application de l'algorithme de Sturm qui se fait au moins en $\mathcal{O}(n^3)$.

Le nombre d'itérations de la dichotomie pour obtenir le résultat avec une tolérance τ est supérieur à

$$\left\lceil \frac{\ln(\|p - z^n\|/\tau)}{\ln 2} \right\rceil$$

Simulations numériques

Pour le polynôme $p(z) = z + 1$, l'algorithme donne $\beta(p) \approx 0.999996$

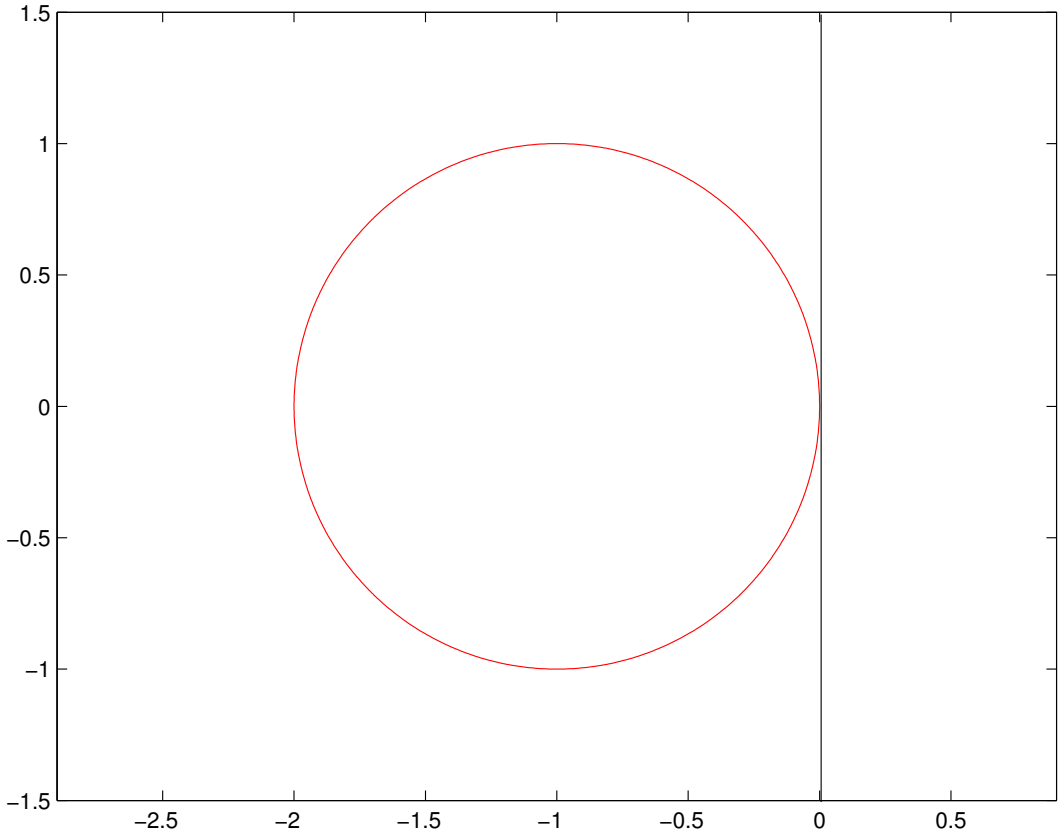


FIG. 1: $\beta(p)$-pseudozéros de $p(z) = z + 1$

Simulations numériques (suite)

Pour le polynôme $p(z) = z^2+z+1/2$, l'algorithme donne $\beta(p) \approx 0.485868$

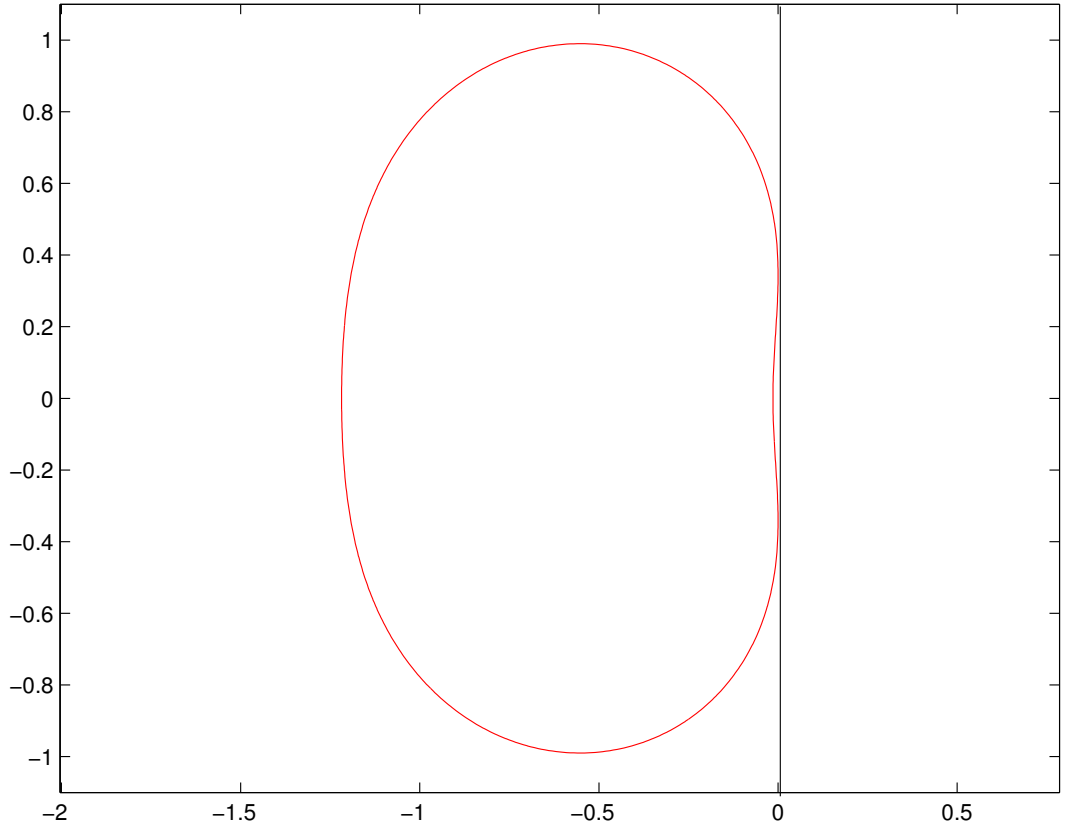


FIG. 2: $\beta(p)$-pseudozéros de $p(z) = z^2 + z + 1/2$

Simulations numériques (suite)

Pour le polynôme $p(z) = z^3 + 4z^2 + 6z + 4$, l'algorithme donne $\beta(p) \approx 2.610226$

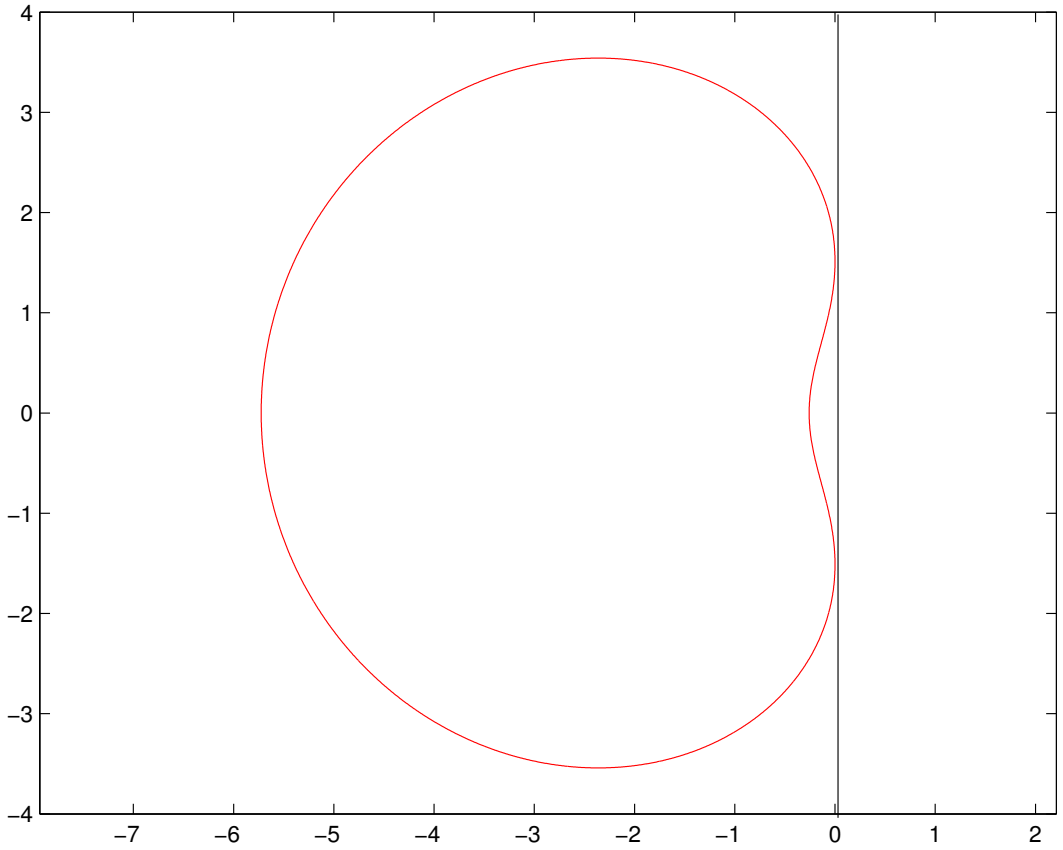


FIG. 3: $\beta(p)$-pseudozéros de $p(z) = z^3 + 4z^2 + 6z + 4$

Conclusion

Nous avons un outil qui a

- des applications en calcul formel
- des applications pour le calcul en précision finie
- des applications en théorie du contrôle