Validated Numerics for Algebraic Path Tracking

Alexandre Guillemot, Pierre Lairez

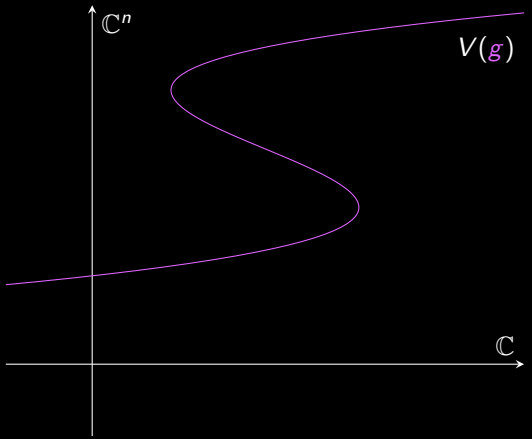
March 8, 2024

Introduction

Introduction

Setup

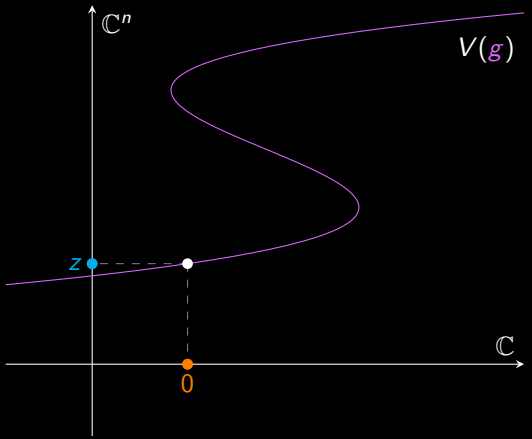
- Let $g : \mathbb{C} \times \mathbb{C}^n \to \mathbb{C}^n$ be a polynomial map.
- Notation : $g_t : \mathbb{C}^n \to \mathbb{C}^n$ defined by $g_t(z) = g(t, z)$.



Introduction

Setup

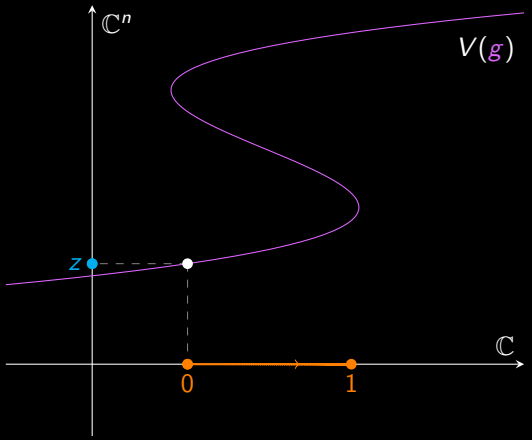
- Let $g : \mathbb{C} \times \mathbb{C}^n \to \mathbb{C}^n$ be a polynomial map.
- Notation : $g_t : \mathbb{C}^n \to \mathbb{C}^n$ defined by $g_t(z) = g(t, z)$.
- Let $z \in \mathbb{C}^n$ such that $g_0(z) = 0$.



Introduction

Setup

- Let $g : \mathbb{C} \times \mathbb{C}^n \to \mathbb{C}^n$ be a polynomial map.
- Notation : $g_t : \mathbb{C}^n \to \mathbb{C}^n$ defined by $g_t(z) = g(t, z)$.
- Let $z \in \mathbb{C}^n$ such that $g_0(z) = 0$.

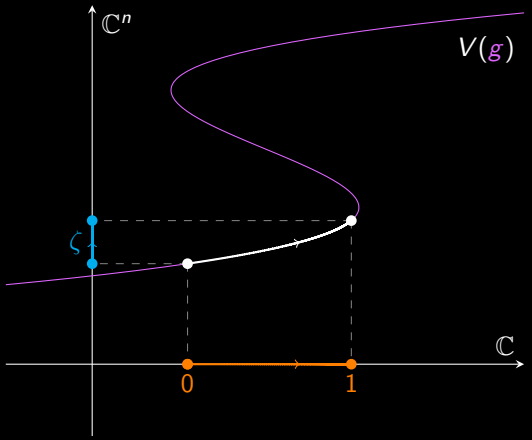


Introduction

Setup

- Let $g : \mathbb{C} \times \mathbb{C}^n \to \mathbb{C}^n$ be a polynomial map.
- Notation : $g_t : \mathbb{C}^n \to \mathbb{C}^n$ defined by $g_t(z) = g(t, z)$.
- Let $z \in \mathbb{C}^n$ such that $g_0(z) = 0$.

$\Rightarrow$ Moving the parameter from 0 to 1 induces $\zeta : [0, 1] \to \mathbb{C}^n$ s.t. $\zeta(0) = z$ and $g_t(\zeta(t)) = 0$.



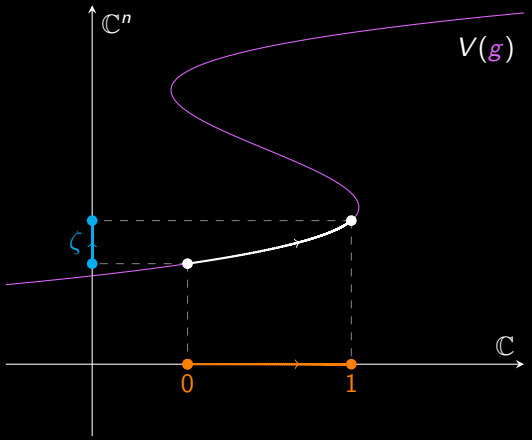
Introduction

Setup

- Let $g : \mathbb{C} \times \mathbb{C}^n \to \mathbb{C}^n$ be a polynomial map.
- Notation : $g_t : \mathbb{C}^n \to \mathbb{C}^n$ defined by $g_t(z) = g(t, z)$.
- Let $z \in \mathbb{C}^n$ such that $g_0(z) = 0$.

$\Rightarrow$ Moving the parameter from 0 to 1 induces $\zeta : [0, 1] \to \mathbb{C}^n$ s.t. $\zeta(0) = z$ and $g_t(\zeta(t)) = 0$.

- Goal : "Track" ζ, with some topological guarantees.



Motivations

To solve multivariate square polynomial systems

- State of the art method,
- **Only need guarantees on the endpoints** when tracking homotopies,
- Possible to use a non certified tracker and to certify *a posteriori* the endpoints,
- Efficient implementations of non-certified trackers : **HomotopyContinuation.jl**, **PHCpack**, **Bertini**, ...

Motivations

To solve multivariate square polynomial systems

- State of the art method,
- **Only need guarantees on the endpoints** when tracking homotopies,
- Possible to use a non certified tracker and to certify *a posteriori* the endpoints,
- Efficient implementations of non-certified trackers : **HomotopyContinuation.jl**, **PHCpack**, **Bertini**, ...

Effective algebraic topology

- **The certification of the whole path** is required,
- Many theoritical works, but only one implementation in **Macaulay2**,
- There is a huge gap between non-certified and certified methods in terms of the number of steps required to track a root,
- Goal : narrow this gap. Tools : **Interval arithmetic, Krawczyk's method, predictor**.

Path certification

Root isolation

Root isolation criterion (Krawczyk, Moore, Rump)

- $f : \mathbb{C}^n \to \mathbb{C}^n$ polynomial,
- $z \in \mathbb{C}^n$, $r \in \mathbb{R}_{>0}$, $A \in \mathbb{C}^{n \times n}$,
- $\rho \in (0, 1)$,

such that for all $u, v \in B_r$,

$$-Af(z) + [\mathrm{I}_n - A \cdot Jf(z + u)]v \in B_{\rho r}.$$

Then there exists a unique $\tilde{z} \in z + B_{\rho r}$ s.t. $f(\tilde{z}) = 0$.

Root isolation

Root isolation criterion (Krawczyk, Moore, Rump)

- $f : \mathbb{C}^n \to \mathbb{C}^n$ polynomial,
- $z \in \mathbb{C}^n$, $r \in \mathbb{R}_{>0}$, $A \in \mathbb{C}^{n \times n}$,
- $\rho \in (0, 1)$,

such that for all $u, v \in B_r$,

$$-Af(z) + [\mathrm{I}_n - A \cdot Jf(z + u)]v \in B_{\rho r}.$$

Then there exists a unique $\tilde{z} \in z + B_{\rho r}$ s.t. $f(\tilde{z}) = 0$.

Proof sketch

We show that $\varphi : z + B_{\rho r} \to \mathbb{C}^n$ defined by $\varphi(z) = z - Af(z)$ is a ρ-contraction map with values in $z + B_{\rho r}$.

Moore boxes

Let $f : \mathbb{C}^n \to \mathbb{C}^n$ be polynomial map and $\rho \in (0, 1)$.

Moore boxes

A ρ-*Moore box* for f is a triple (z, r, A) satisfying

$$-Af(z) + [\mathrm{I}_n - A \cdot Jf(z + B_r)]B_r \subseteq B_{\rho r}.$$

Moore boxes

Let $f : \mathbb{C}^n \to \mathbb{C}^n$ be polynomial map and $\rho \in (0, 1)$.

Moore boxes

A ρ-*Moore box* for f is a triple (z, r, A) satisfying

$$-Af(z) + [\mathrm{I}_n - A \cdot Jf(z + B_r)]B_r \subseteq B_{\rho r}.$$

Moore boxes are our data structure to represent zeros of f.

Moore boxes

Let $f : \mathbb{C}^n \to \mathbb{C}^n$ be polynomial map and $\rho \in (0, 1)$.

Moore boxes

A ρ-*Moore box* for f is a triple (z, r, A) satisfying

$$-Af(z) + [\mathrm{I}_n - A \cdot Jf(z + B_r)]B_r \subseteq B_{\rho r}.$$

Moore boxes are our data structure to represent zeros of f.

Problem

How do we check that a triple (z, r, A) is a Moore box on a computer ?

Interval arithmetic

We choose a set $\mathbb{F} \subset \mathbb{R}$ of representable numbers.

Interval space

- $\square\mathbb{R} = \{[a, b] \subseteq \mathbb{R} \mid a \leq b,\ a, b \in \mathbb{F}\}$,
- $\square\mathbb{C}$: pairs of elements of $\square\mathbb{R}$,
- $\square\mathbb{R}^n$ resp. $\square\mathbb{C}^n$: vectors of boxes in $\mathbb{R}$ resp. $\mathbb{C}$ of size n.

Interval arithmetic

We choose a set $\mathbb{F} \subset \mathbb{R}$ of representable numbers.

Interval space

- $\square\mathbb{R} = \{[a, b] \subseteq \mathbb{R} \mid a \leq b,\ a, b \in \mathbb{F}\}$,
- $\square\mathbb{C}$: pairs of elements of $\square\mathbb{R}$,
- $\square\mathbb{R}^n$ resp. $\square\mathbb{C}^n$: vectors of boxes in $\mathbb{R}$ resp. $\mathbb{C}$ of size n.

Interval extension

Let $f : \mathbb{C}^n \to \mathbb{C}^m$. An *interval extension* of f is a map $\square f : \square\mathbb{C}^n \to \square\mathbb{C}^m$ such that

$$\forall Z \in \square\mathbb{C}^n,\ f(Z) \subseteq \square f(Z).$$

Interval arithmetic

We choose a set $\mathbb{F} \subset \mathbb{R}$ of representable numbers.

Interval space

- $\square\mathbb{R} = \{[a, b] \subseteq \mathbb{R} \mid a \leq b,\ a, b \in \mathbb{F}\}$,
- $\square\mathbb{C}$: pairs of elements of $\square\mathbb{R}$,
- $\square\mathbb{R}^n$ resp. $\square\mathbb{C}^n$: vectors of boxes in $\mathbb{R}$ resp. $\mathbb{C}$ of size n.

Interval extension

Let $f : \mathbb{C}^n \to \mathbb{C}^m$. An *interval extension* of f is a map $\square f : \square\mathbb{C}^n \to \square\mathbb{C}^m$ such that

$$\forall Z \in \square\mathbb{C}^n,\ f(Z) \subseteq \square f(Z).$$

Interval arithmetic gives an effective way to build an extension $\square f$ given a (polynomial) map f.

Interval certification of a Moore box

Algorithm

1 **def** $M(\square f, \square Jf, z, r, A, \rho)$**:**

2 **if** $-A \cdot \square f(z) + [\mathrm{I}_n - A \cdot \square Jf(z + B_r)]B_r \subseteq B_{\rho r}$**:**

3 **return** `True`

4 **else:**

5 **return** `False`

Interval certification of a Moore box

Algorithm

```
def M(□f, □Jf, z, r, A, ρ):
    if −A · □f(z) + [I_n − A · □Jf(z + B_r)]B_r ⊆ B_ρr:
        return True
    else:
        return False
```

Proposition

If $\square f$ is an extension of f, $\square Jf$ is an extension of Jf and $M(\square f, \square Jf, z, r, A, \rho)$ returns `True`, then (z, r, A) is a ρ-Moore box for f.

Interval certification of a Moore box

Algorithm

```
def M(□f, □Jf, z, r, A, ρ):
    if −A · □f(z) + [I_n − A · □Jf(z + B_r)]B_r ⊆ B_ρr:
        return True
    else:
        return False
```

Proposition

If $\square f$ is an extension of f, $\square Jf$ is an extension of Jf and $M(\square f, \square Jf, z, r, A, \rho)$ returns `True`, then (z, r, A) is a ρ-Moore box for f.

Proof

Fundamental property of interval arithmetic !

Algorithm

General idea

Input

- A polynomial map $g : \mathbb{C} \times \mathbb{C}^n \to \mathbb{C}^n$,
- A zero $\tilde{z}$ of g_0 (represented by a $\frac{7}{8}$-Moore box (z, r, A)).

General idea

Input

- A polynomial map $g : \mathbb{C} \times \mathbb{C}^n \to \mathbb{C}^n$,
- A zero $\tilde{z}$ of g_0 (represented by a $\frac{7}{8}$-Moore box (z, r, A)).

Let $\zeta : [0, 1] \to \mathbb{C}^n$ s.t. $\zeta(0) = \tilde{z}$ and $g_t(\zeta(t)) = 0$ for all $t \in [0, 1]$.

General idea

Input

- A polynomial map $g : \mathbb{C} \times \mathbb{C}^n \to \mathbb{C}^n$,
- A zero $\tilde{z}$ of g_0 (represented by a $\frac{7}{8}$-Moore box (z, r, A)).

Let $\zeta : [0,1] \to \mathbb{C}^n$ s.t. $\zeta(0) = \tilde{z}$ and $g_t(\zeta(t)) = 0$ for all $t \in [0,1]$.

Output

Moore boxes (z_t, r_t, A_t) for g_t isolating $\zeta(t)$, this for all $t \in [0,1]$.

General idea

Input

- A polynomial map $g : \mathbb{C} \times \mathbb{C}^n \to \mathbb{C}^n$,
- A zero $\tilde{z}$ of g_0 (represented by a $\frac{7}{8}$-Moore box (z, r, A)).

Let $\zeta : [0,1] \to \mathbb{C}^n$ s.t. $\zeta(0) = \tilde{z}$ and $g_t(\zeta(t)) = 0$ for all $t \in [0,1]$.

Output

Moore boxes (z_t, r_t, A_t) for g_t isolating $\zeta(t)$, this for all $t \in [0,1]$.

Output in practice

- Subintervals $I_1, \cdots, I_k$ covering $[0,1]$,
- triples $(z_1, r_1, A_1), \cdots, (z_k, r_k, A_k)$

such that (z_i, r_i, A_i) is a $\frac{7}{8}$-Moore box for g on I_i.

Subroutines

Refine

input : $f : \mathbb{C}^n \to \mathbb{C}^n$; (z, r, A) a $\frac{7}{8}$-Moore box

output : a $\frac{1}{8}$-Moore box with same associated zero as (z, r, A)

algorithm : A balance between pseudo-Newton iterations and reducing r

Subroutines

Refine

input : $f : \mathbb{C}^n \to \mathbb{C}^n$; (z, r, A) a $\frac{7}{8}$-Moore box

output : a $\frac{1}{8}$-Moore box with same associated zero as (z, r, A)

algorithm : A balance between pseudo-Newton iterations and reducing r

Thicken

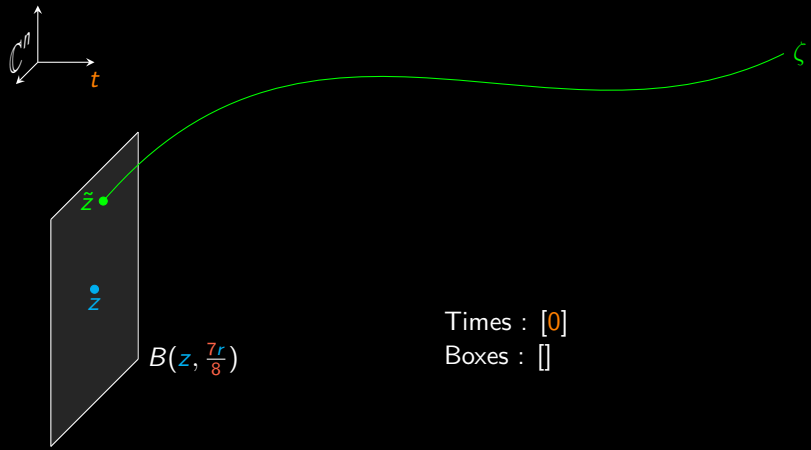
input : $g : \mathbb{C} \times \mathbb{C}^n \to \mathbb{C}^n$; $t \in [0, 1]$; (z, r, A) a $\frac{7}{8}$-Moore box for g_t

output : $t' \in [0, 1]$ s.t. for all $s \in [t, t']$, (z, r, A) is a $\frac{7}{8}$-Moore box for g_s

algorithm : Decrease t' until $M(\square g_T, \square J g_T, z, r, A, \rho)$ returns `True`, where

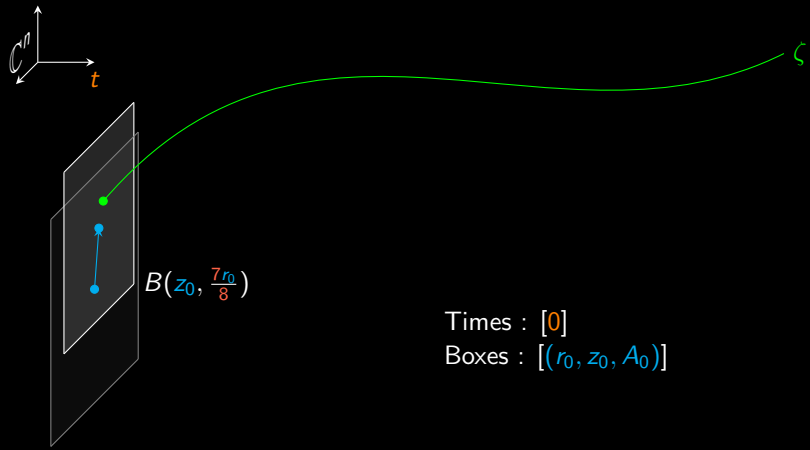
- $T = [t, t']$,
- $\square g_T : \square\mathbb{C}^n \to \square\mathbb{C}^n$ is an extension of g_s for all $s \in T$,
- $\square J g_T : \square\mathbb{C}^n \to \square\mathbb{C}^{n \times n}$ is an extension of $J g_s$ for all $s \in T$.

Flow of the algorithm



Times : [0]
Boxes : []

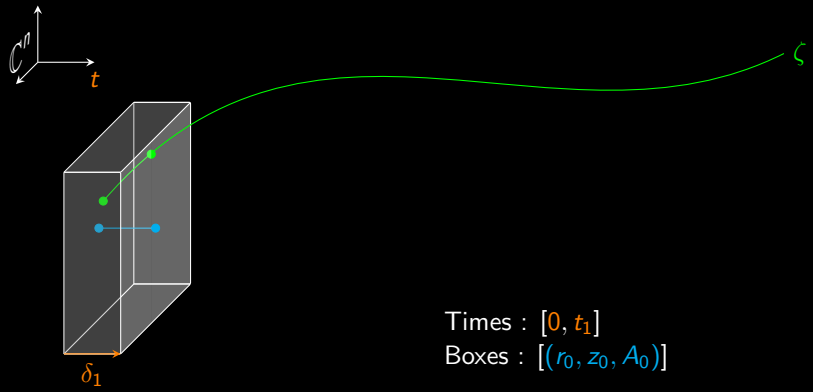
Flow of the algorithm



Times : $[0]$

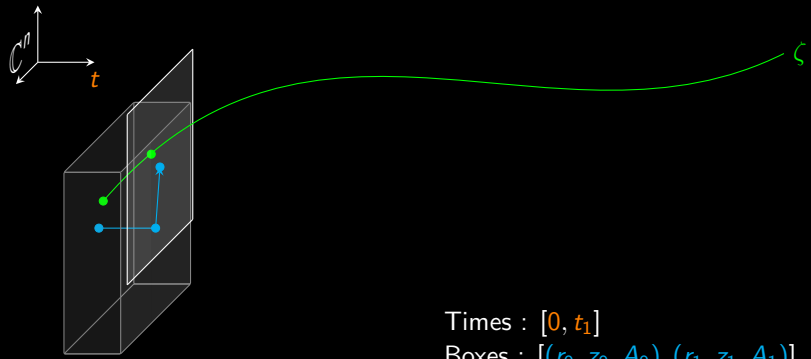
Boxes : $[(r_0, z_0, A_0)]$

Flow of the algorithm



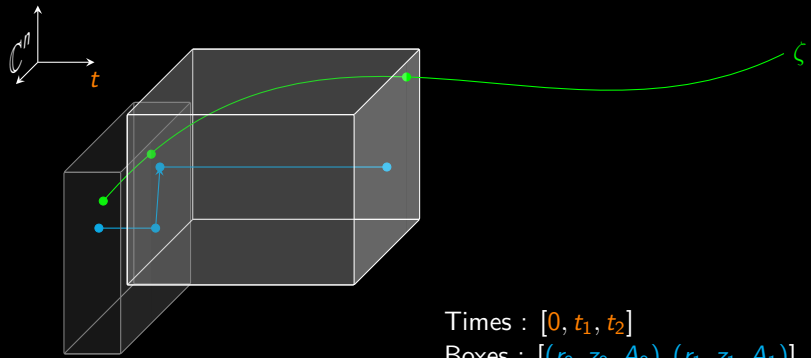
Times : $[0, t_1]$
Boxes : $[(r_0, z_0, A_0)]$

Flow of the algorithm



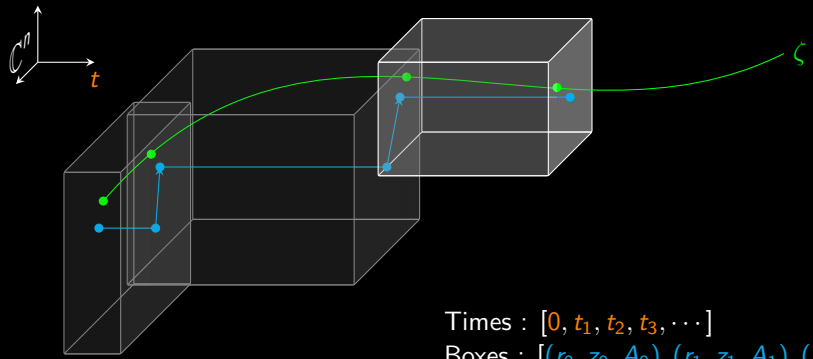
Times : $[0, t_1]$
Boxes : $[(r_0, z_0, A_0), (r_1, z_1, A_1)]$

Flow of the algorithm



Times : $[0, t_1, t_2]$
Boxes : $[(r_0, z_0, A_0), (r_1, z_1, A_1)]$

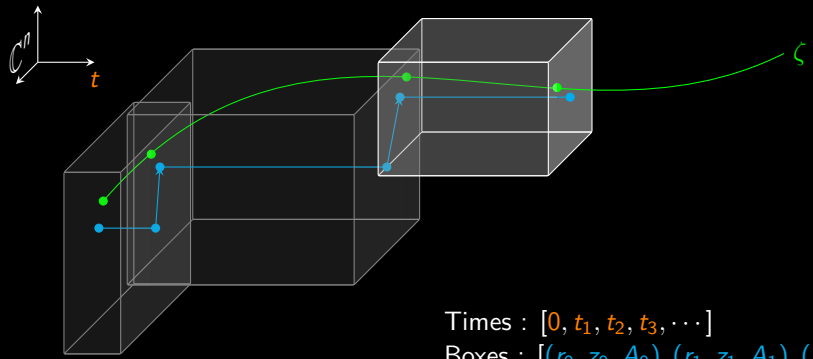
Flow of the algorithm



Times : $[0, t_1, t_2, t_3, \cdots]$

Boxes : $[(r_0, z_0, A_0), (r_1, z_1, A_1), (r_2, z_2, A_2), \cdots]$

Flow of the algorithm

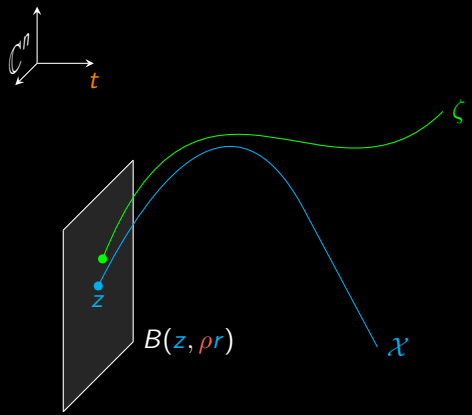


Times : $[0, t_1, t_2, t_3, \cdots]$
Boxes : $[(r_0, z_0, A_0), (r_1, z_1, A_1), (r_2, z_2, A_2), \cdots]$

Theorem (Guillemot, Lairez)

In a computation model with dynamic floating-point precision, the algorithm is correct and terminates.

Thickening with a predictor

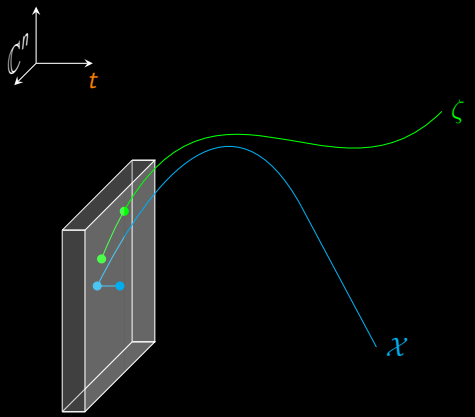


Predictor

A map $\mathcal{X} : \mathbb{R} \to \mathbb{C}^n$ such that $\mathcal{X}(t) = z$.

$\mathcal{X}$ should stay close to ζ around t.

Thickening with a predictor



Predictor

A map $\mathcal{X} : \mathbb{R} \to \mathbb{C}^n$ such that $\mathcal{X}(t) = z$.

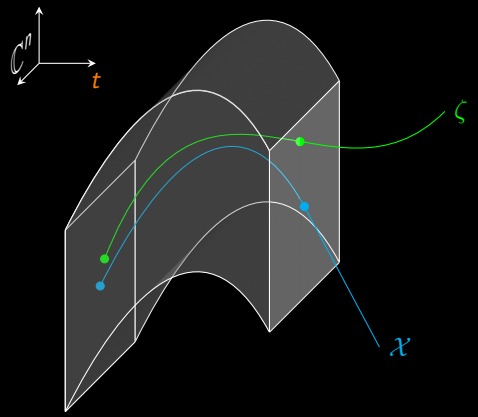
$\mathcal{X}$ should stay close to ζ around t.

Certifying the prediction

Problem : check that for all $s \in [t, t']$, (z, r, A) is a ρ-Moore box for g_s.

Solution : try $M(\square g_T, \square J g_T, z, r, A, \rho)$, where $T = [t, t']$.

Thickening with a predictor



Predictor

A map $\mathcal{X} : \mathbb{R} \to \mathbb{C}^n$ such that $\mathcal{X}(t) = z$.

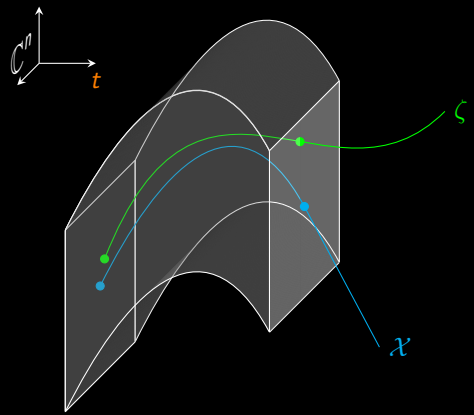
$\mathcal{X}$ should stay close to ζ around t.

Certifying the prediction

Problem : check that for all $s \in [t, t']$, $(\mathcal{X}(s), r, A)$ is a ρ-Moore box for g_s.

Solution : try $M(\square g_T, \square J g_T, \mathcal{X}(T), r, A, \rho)$, where $T = [t, t']$.

Thickening with a predictor



Predictor

A map $\mathcal{X} : \mathbb{R} \to \mathbb{C}^n$ such that $\mathcal{X}(t) = z$.

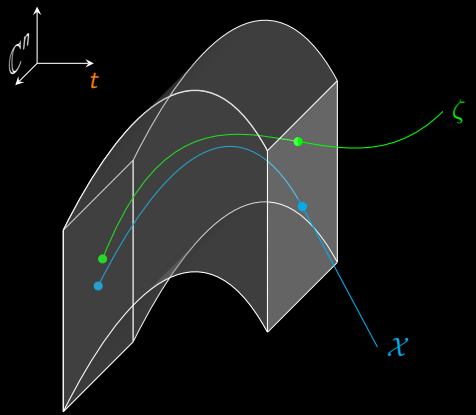
$\mathcal{X}$ should stay close to ζ around t.

Certifying the prediction

Problem : check that for all $s \in [t, t']$, $(\mathcal{X}(s), r, A)$ is a ρ-Moore box for g_s.
Solution : try $M(\square g_T, \square J g_T, \mathcal{X}(T), r, A, \rho)$, where $T = [t, t']$.

This is too strong !

Thickening with a predictor



Predictor

A map $\mathcal{X} : \mathbb{R} \to \mathbb{C}^n$ such that $\mathcal{X}(t) = z$.

$\mathcal{X}$ should stay close to ζ around t.

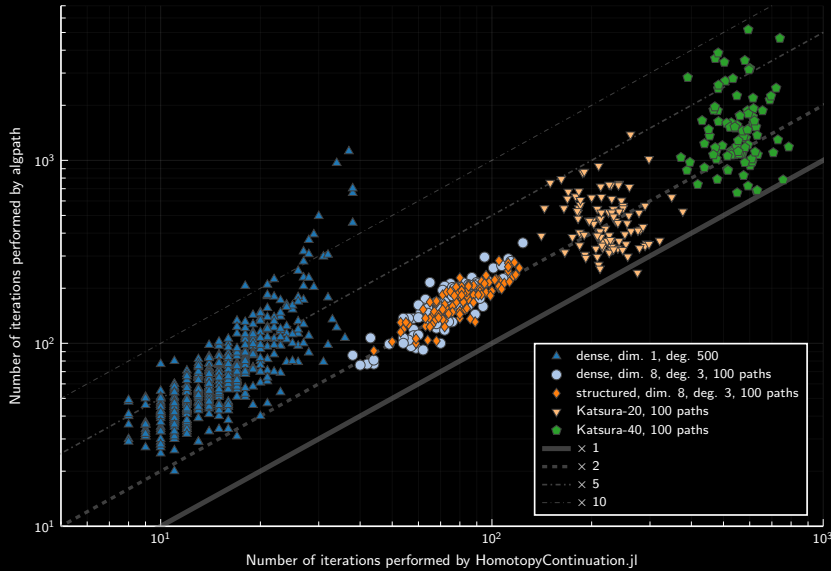
Certifying the prediction

Problem : check that for all $s \in [t, t']$, $(\mathcal{X}(s), r, A)$ is a ρ-Moore box for g_s.
Solution : compute

$$-Ag_s(\mathcal{X}(s)) + [\mathrm{I}_n - A \cdot Jg_s(\mathcal{X}(s)) + B_r)]B_r$$

using Taylor models on T.

Implementation and benchmarks



References

- Beltrán, Carlos and Anton Leykin. "Certified Numerical Homotopy Tracking". In: *Experimental Mathematics* 21.1 (Mar. 2012). Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/10586458.2011.606184, pp. 69–83.
- Breiding, Paul and Sascha Timme. "HomotopyContinuation.jl: A Package for Homotopy Continuation in Julia". en. In: *Mathematical Software – ICMS 2018*. Ed. by James H. Davenport et al. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2018, pp. 458–465.
- Duff, Timothy and Kisun Lee. *Certified homotopy tracking using the Krawczyk method*. arXiv:2402.07053 [cs, math]. Feb. 2024.
- Hauenstein, Jonathan D. and Alan C. Liddell. "Certified predictor–corrector tracking for Newton homotopies". In: *Journal of Symbolic Computation* 74 (May 2016), pp. 239–254.
- Kranich, Stefan. *An epsilon-delta bound for plane algebraic curves and its use for certified homotopy continuation of systems of plane algebraic curves*. arXiv:1505.03432 [math]. May 2016.
- Marco-Buzunariz, Miguel Ángel and Marcos Rodríguez. "SIROCCO: A Library for Certified Polynomial Root Continuation". en. In: *Mathematical Software – ICMS 2016*. Ed. by Gert-Martin Greuel et al. Vol. 9725. Series Title: Lecture Notes in Computer Science. Cham: Springer International Publishing, 2016, pp. 191–197.
- Moore, R. E. "A Test for Existence of Solutions to Nonlinear Systems". In: *SIAM Journal on Numerical Analysis* 14.4 (1977). Publisher: Society for Industrial and Applied Mathematics, pp. 611–615.
- van der Hoeven, Joris. *Reliable homotopy continuation*. Research Report. LIX, Ecole polytechnique, Nov. 2011.
- Verschelde, Jan. "Algorithm 795: PHCpack: a general-purpose solver for polynomial systems by homotopy continuation". In: *ACM Transactions on Mathematical Software* 25.2 (June 1999), pp. 251–276.

Test data

We tested systems of the form $g_t(z) = tf^{\odot}(z) + (1-t)f^{\triangleright}(z)$ ($f^{\triangleright}$ is the start system, $f^{\odot}$ is the target system).

Test data

We tested systems of the form $g_t(z) = tf^{\odot}(z) + (1-t)f^{\triangleright}(z)$ ($f^{\triangleright}$ is the start system, $f^{\odot}$ is the target system).

Target systems

- Dense : $f_i^{\odot}$'s of given degree with random coefficients
- Structured : $f_i^{\odot}$'s of the form $\pm 1 + \sum_{i=1}^{5} \left(\sum_{j=1}^{n} a_{i,j} z_j\right)^d$, $a_{i,j} \in_R \{-1, 0, 1\}$
- Katsura family (sparse - high dimension - low degree)

Test data

We tested systems of the form $g_t(z) = tf^{\odot}(z) + (1-t)f^{\triangleright}(z)$ ($f^{\triangleright}$ is the start system, $f^{\odot}$ is the target system).

Target systems

- Dense : $f_i^{\odot}$'s of given degree with random coefficients
- Structured : $f_i^{\odot}$'s of the form $\pm 1 + \sum_{i=1}^{5} \left(\sum_{j=1}^{n} a_{i,j} z_j \right)^d$, $a_{i,j} \in_R \{-1, 0, 1\}$
- Katsura family (sparse - high dimension - low degree)

Start systems

- Total degree homotoies : $f_i^{\triangleright}$'s of the form $\gamma_i(z_i^{d_i} - 1)$, $\gamma_i \in_R \mathbb{C}$, $d_i = \deg f_i^{\odot}$
- Newton homotopies : $f^{\triangleright}(z) = f^{\odot}(z) - f^{\odot}(z_0)$

Benchmarks table

			HomotopyContinuation.jl			algpath			Macaulay2		
name	dim.	max deg	med.	ksteps/s	time (s)	med.	ksteps/s	time (s)	med.	ksteps/s	time (s)
dense	1	10	6	29	1.8	11	55	< 0.1	629	49	0.2
dense	1	30	10	43	1.9	23	25	< 0.1	830 k	28	19 min
dense	1	50	12	38	1.8	30	13	0.7		> 1 h	
dense	2	10	22	33	2.4	53	9.2	0.7	33 k	27	165
dense	2	30	24	5.8	6.4	85	1.4	72		> 1 h	
dense	2	50	27	2.3	32	117	0.53	12 min		> 1 h	
katsura *	11	2	112	38	4.4	199	6.0	3.5	20 k	11	203
katsura *	21	2	222	13	6.0	461	1.3	37		> 1 h	
katsura *	41	2	554	2.8	24	1371	0.19	13 min		> 1 h	
dense *	8	3	73	2.1	6.2	157	0.86	19	21 k	8.3	281
structured *	8	3	81	40	3.9	182	7.9	2.3	36 k	9.9	371
structured ^{N}	10	10	53	0.19	3.0	123	4.9	< 0.1		> 1 h	
structured ^{N}	20	20		> 8 GB		1591	1.2	1.5		> 8 GB	
structured ^{N}	30	30		> 8 GB		1989	0.43	5.2		> 8 GB	