

Finding one root of a polynomial system

Smale's 17th problem

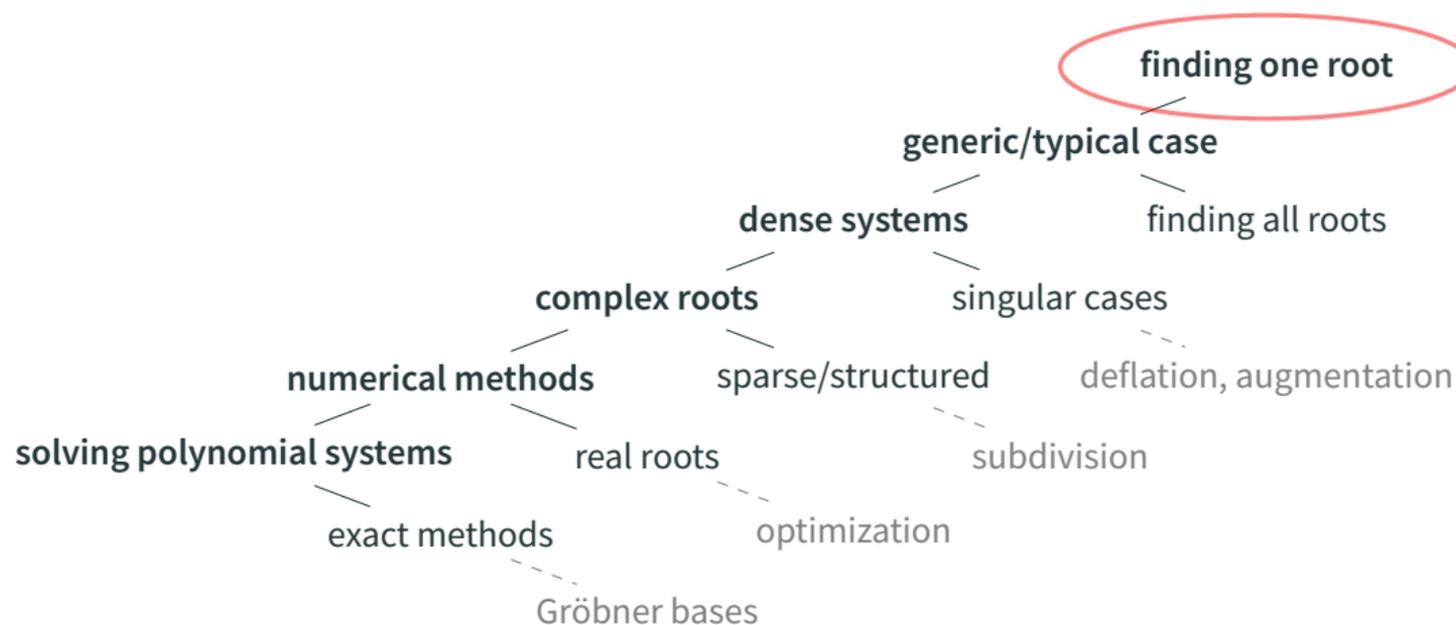
Pierre Lairez

Inria Saclay

FoCM 2017

Foundations of computational mathematics

15 July 2017, Barcelona



Finding one root: a purely numerical question

Bézout bound vs. input size

n polynomial equations

n variables, degree D

degree	input size	#roots
D	$n \binom{D+n}{n}$	D^n
2	$\sim \frac{1}{2} n^3$	2^n
n	$\sim \frac{1}{\sqrt{\pi}} n^{\frac{1}{2}} 4^n$	n^n
$D \gg n$	$\sim \frac{1}{(n-1)!} D^n$	D^n

#roots \gg input size

To compute a single root, do we have to pay for #roots?

Exact computation

Having one root is having them all (generically).

Numerical computation

One may approximate one root disregarding the others.

Polynomial complexity?

Maybe, but only with numerical methods.

Smale 17th problem

“Can a zero of n complex polynomial equations in n unknowns be found approximately, on the average, in polynomial time with a uniform algorithm?”

— *S. Smale, 1998*

approximate root A point from which Newton’s iteration converges quadratically.

polynomial time with respect to the input size.

on the average with respect to some input distribution.

uniform algorithm A Blum–Shub–Smale machine (a.k.a. real random access machine):

- registers store exact real numbers,
- unit cost arithmetic operations,
- branching on positivity testing.

Infinite precision?! Yes, but we still have to deal with stability issues.

The model is very relevant for this problem.

Problem solved!

Shub, Smale (1990s) Quantitative theory of Newton's iteration
Complexity of numerical continuation

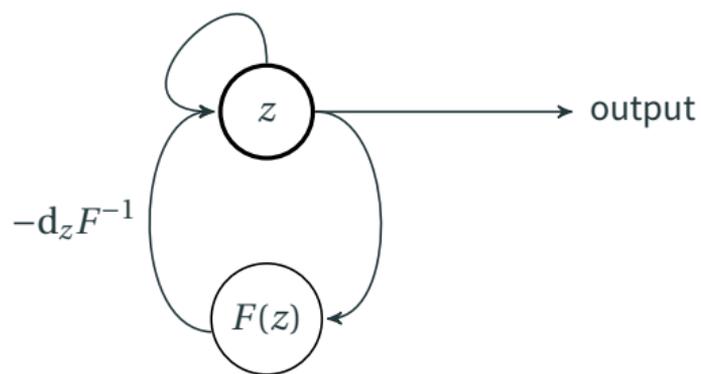
Beltrán, Pardo (2009) Randomization

Bürgisser, Cucker (2011) Deterministic polynomial average time when $D \ll n$ or $D \gg n$
Smoothed analysis

Lairez (2017) Derandomization

Numerical continuation

Newton's iteration

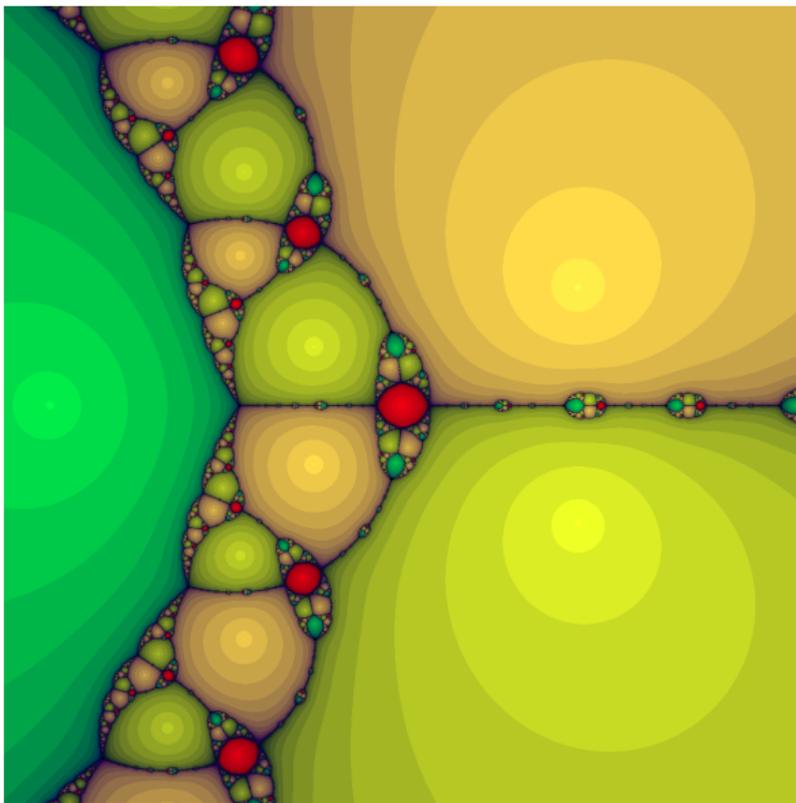


$F: \mathbb{C}^n \rightarrow \mathbb{C}^n$ a polynomial map,

$$z_{k+1} = z_k - d_{z_k} F^{-1} \cdot F(z_k).$$

- Converges quadratically fast close to a regular root.
- May diverge on a open set of initial point.

The geometry of the basins of attraction is complex...



Convergence of Newton's iteration for the polynomial $z^3 - 2z + 2$.
In red, the points from which Newton's iteration do not converge.

(Picture by Henning Makholm.)

$F: \mathbb{C}^n \rightarrow \mathbb{C}^n$, a polynomial map.

$$\gamma(F, x) \triangleq \sup_{k>1} \left\| \frac{1}{k!} d_x F^{-1} \cdot d_x^k F \right\|^{\frac{1}{k-1}}.$$

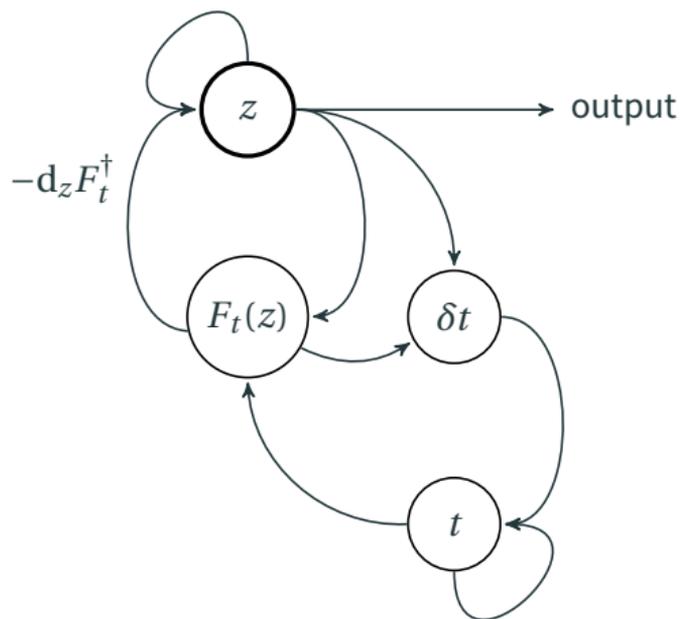
γ -Theorem

If $F(\zeta) = 0$ and $\|z - \zeta\| \gamma(F, \zeta) \leq \frac{3-\sqrt{7}}{2}$ then

$$\left\| \text{Newton}^{(k)}(z) - \zeta \right\| \leq 2^{1-2^k}.$$

Numerical continuation

$F_t : \mathbb{C}^n \rightarrow \mathbb{C}^n$ a polynomial system depending continuously on $t \in [0, 1]$; z_0 a root of F_0 .



$$z_{k+1} = z_k - \mathbf{d}_{z_k} F_{t_k}^\dagger \cdot F_{t_k}(z_k)$$

$$t_{k+1} = t_k + \delta t_k$$

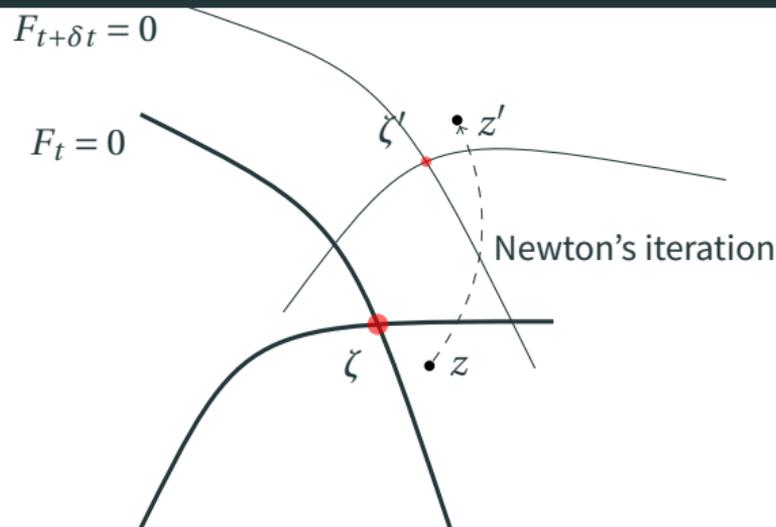
- Solves any generic system
- How to set the step size δt ?
- How to choose the start system F_0 ?
- How to choose a path?
- How many steps do we need to go from F_0 to F_1 ?

\mathcal{H} the space of homogeneous polynomial systems of n equations of degree D in $n + 1$ variables, embedded with some Hermitian norm that is invariant under unitary change of variables.

F a polynomial system in \mathcal{H}

z a root of F in \mathbb{P}^n

$$\begin{aligned} \mu(F, z) &= \sup \frac{d_{\mathbb{P}}(z, z')}{\|F' - F\|} \quad \text{with } F' \sim F \text{ and } F'(z') = 0 \\ &= \left\| (d_z F)^\dagger \right\| = \frac{1}{\text{least singular value of } d_z F} \\ &\simeq \sup \frac{1}{\|F - F'\|} \quad \text{where } z \text{ is a singular root of } F' \\ &\geq 2D^{-\frac{3}{2}} \gamma_{\text{proj}}(F, z). \end{aligned}$$



We have $d(\zeta, \zeta') \lesssim \mu(F_t, \zeta) \|\dot{F}_t\| \delta t$.

We need $d(z, \zeta') \lesssim \frac{1}{D^{\frac{3}{2}} \mu(F_{t+\delta t}, \zeta')}$.

It suffices that $\delta t \lesssim \frac{1}{D^{\frac{3}{2}} \mu(F_t, \zeta)^2}$.

Theorem (Shub 2009)

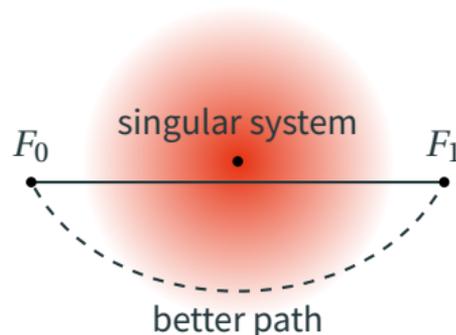
One can compute an approximate root of F_1 given an approximate root of F_0 with

$$\# \text{steps} \leq 136 D^{\frac{3}{2}} \int_0^1 \mu(F_t, \zeta_t)^2 \|\dot{F}_t\| dt.$$

How to choose the path?

linear interpolation $F_t = tF_1 + (1 - t)F_0$

a better path? We can imagine the notion of adaptive path, but it is difficult to make it works.



difficulty It is not enough to control the conditioning of the start system, we need a grasp on what happens along the continuation path.

A $\{x_i = 0, \quad 1 \leq i \leq n \quad (\text{homogeneized in degree } D)$

with its root $(0, \dots, 0)$.

Works well when $D \gg n$ (Armentano, Beltrán, Bürgisser, Cucker, Shub 2016).

B $\{x_i^D = 1, \quad 1 \leq i \leq n$

with its D^n roots.

Works well when $D \ll n$ (Bürgisser, Cucker 2011).

C $F(x_1, \dots, x_n) - F(0, \dots, 0) = 0$

with its root $(0, \dots, 0)$.

Randomization of the start system

Conditioning of a random system

- $F \in \mathcal{H}$, random polynomial system, uniformly distributed in $\mathbb{S}(\mathcal{H})$.
- ζ a random root of $F = 0$, uniformly chosen among the D^n roots.

Theorem (Beltrán, Pardo 2011; Bürgisser, Cucker 2011)

$$\mathbb{E}(\mu(F, \zeta)^2) \leq n \cdot \underbrace{\dim \mathcal{H}}_{\text{the input size}}$$

- How to sample (F, ζ) ? Chicken-and-egg problem?

Complexity of numerical continuation with random endpoints

F_0, F_1 random polynomial systems of norm 1, uniformly distributed.

ζ_0 a random root of F_0 , uniformly distributed.

F_t linear interpolation (normalized to have norm 1).

ζ_t continuation of ζ_0 .

lemma $\forall t, F_t$ is uniformly distributed and ζ_t is uniformly distributed among its roots.

$$\#\text{steps} \leq 136 D^{\frac{3}{2}} d_{\mathbb{S}}(F_0, F_1) \int_0^1 \mu(F_t, \zeta_t)^2 dt \quad (\text{Shub 2009})$$

$$\mathbb{E}[\#\text{steps}] \leq 136\pi D^{\frac{3}{2}} \mathbb{E} \left[\int_0^1 \mu(F_t, \zeta_t)^2 dt \right]$$

$$\leq 136\pi D^{\frac{3}{2}} \int_0^1 \mathbb{E} [\mu(F_t, \zeta_t)^2] dt \quad (\text{Tonelli's theorem})$$

$$= \mathcal{O} \left(n D^{\frac{3}{2}} (\text{input size}) \right) \quad (\text{Beltrán, Pardo 2011; Bürgisser, Cucker 2011})$$

first try Sample $\zeta \in \mathbb{P}^n$ uniformly,
sample F uniformly in $\{F \text{ s.t. } F(\zeta) = 0 \text{ and } \|F\| = 1\}$.

✘ F is not uniformly distributed.

BP method Sample a *linear* system L uniformly,
compute its unique root $\zeta \in \mathbb{P}^n$,
sample F uniformly in $\{F \text{ s.t. } F(\zeta) = 0, d_\zeta F = L \text{ and } \|F\| = 1\}$.

✔ F and ζ are uniformly distributed.

Solves Smale's problem *with randomization*.

Total average complexity $\mathcal{O}\left(nD^{\frac{3}{2}}(\text{input size})^2\right)$.

average analysis gives little information on the complexity of solving *one* given system.

worst-case analysis is irrelevant here (unbounded close to a system with a singular root).

smoothed analysis bridges the gap and gives information on a single system F perturbed by a Gaussian noise ε of variance σ^2 . This models an input data that is only approximate.

$$\sup_{\text{system } F} \mathbb{E} [\text{cost of computing one root of } F + \varepsilon] = \mathcal{O}(\sigma^{-1} n D^{\frac{3}{2}} N^2).$$

average-case w.r.t. the noise

worst-case

Derandomization

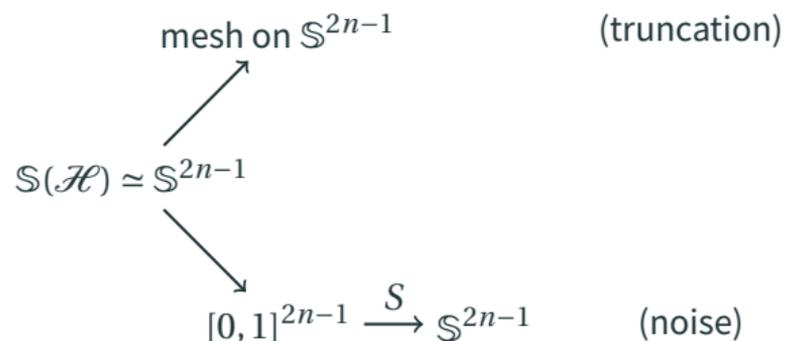
x , a random uniformly distributed variable in $[0, 1]$.

$$x = \underbrace{0.6044025624180895161178081249104686}_{\text{truncation} \downarrow} \overbrace{50529019746531591013322667888500001621027}^{\text{noise extraction} \uparrow}$$

0.6044025624180895161178081249104686

- The truncation is a random variable that is close to x .
- The noise is an independent from x and uniformly distributed in $[0, 1]$.

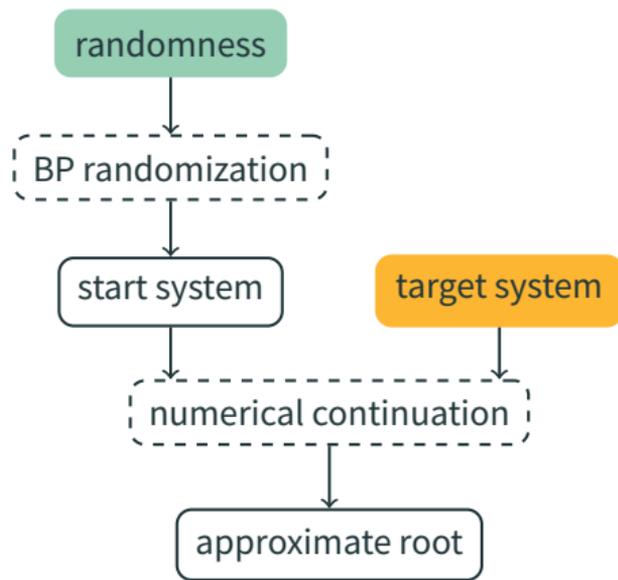
Truncation and noise extraction on an odd-dimensional sphere



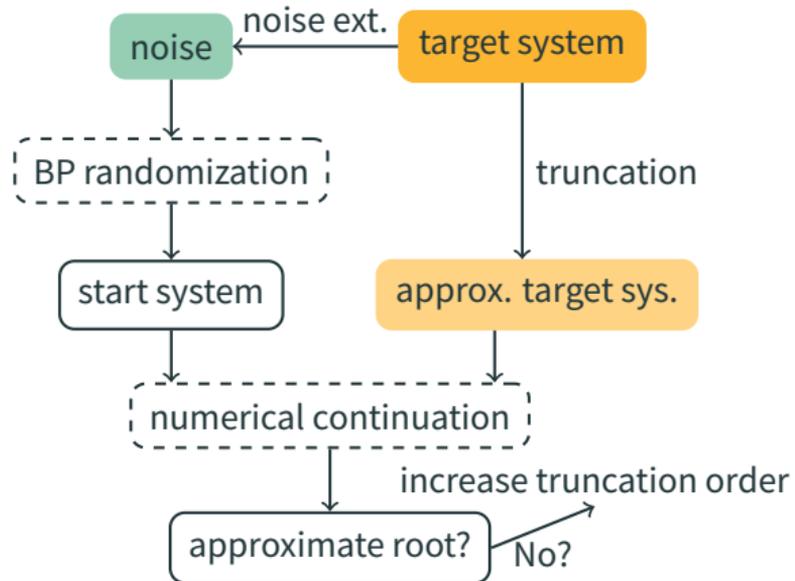
- S is a measure preserving map due to Sibuya (1962).
- The noise is *nearly* uniformly distributed and *nearly* independent from the truncation.

Derandomization

Beltràn and Pardo's randomization



Lairez's derandomization



Solves Smale's problem with a *deterministic algorithm*.

Randomness is in Smale's question from its very formulation asking for an average analysis.

Beyond Smale's problem

Complexity of numerical algorithms

theory  applications and observations
gap

structured system Can we have interesting complexity bounds, supported by probabilistic analysis, for structured systems, especially sparse systems and low evaluation complexity systems?

singular roots Can we design algorithms that find singular roots within nice complexity bounds?

better complexity In the setting of Smale's question, can we reach a quasi-optimal $(\text{input size})^{1+o(1)}$ average complexity?

Complexity exponent in Smale's problem

$$\text{total cost} = \mathcal{O}\left(\underbrace{(\text{input size})}_{\text{cost of Newton's iteration}} \cdot \#\text{steps}\right).$$

Beltrán, Pardo (2009) $\mathbb{E}(\#\text{steps}) = (\text{input size})^{1+o(1)}$

Armentano, Beltrán, Bürgisser, Cucker, Shub (2016)

$$\mathbb{E}(\#\text{steps}) = (\text{input size})^{\frac{1}{2}+o(1)}$$

work in progress $\mathbb{E}(\#\text{steps}) = \text{poly}(n, D) = (\text{input size})^{o(1)}$

question Given polynomial systems F_0 and F_1 and a root ζ of F_0 , how large is

$$\inf_{\text{path } F_0 \rightarrow F_1} \int_0^1 \mu(F_t, \zeta_t) \sqrt{\|\dot{F}_t\|^2 + \|\dot{\zeta}_t\|^2} dt?$$

(This upper bounds the minimal number of continuation steps required to go from F_0 to F_1 .)

answer Not much!

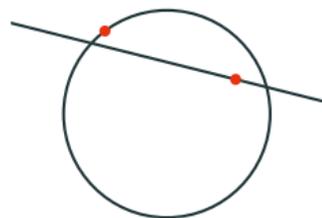
$$\begin{aligned} \# \text{steps} &= \mathcal{O} \left(nD^{\frac{3}{2}} + n^{\frac{1}{2}} \log(\mu(F_0, \zeta_0) \mu(F_1, \zeta_1)) \right) \\ \rightsquigarrow \mathbb{E}(\# \text{steps}) &= \mathcal{O} \left(nD^3 \log(\text{input size}) \right) \text{ with } F_0 \text{ and } F_1 \text{ random} \end{aligned}$$

but... The construction is not algorithmically useful (one need to know a root of the target system to construct the path).

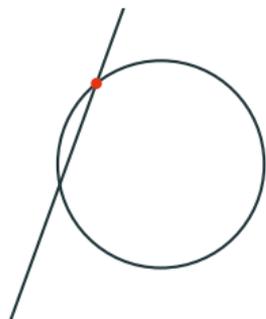
Bigger steps with unitary paths

observation In \mathcal{H} , relatively small perturbation of a typical system F changes everything.
Makes it difficult to make bigger steps.

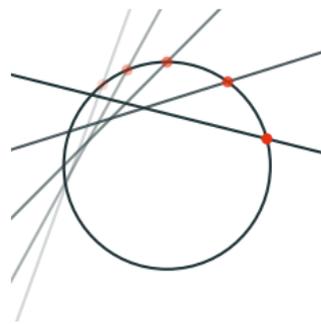
idea Perform the continuation in a lower dimensional parameter space:
We allow only rigid motions of the equations rather than arbitrary deformations.



compute one solution
of each equation



move the hypersurfaces
to make the solution match



continuously return
to the original position

More formally...

parameter space $U(n+1) \times \cdots \times U(n+1)$, that is n copy of the unitary group.

This has dimension $\sim n^3$, compare with $n \cdot \binom{D+n}{n}$.

paths Geodesics in the parameter space.

randomization Same principle as Beltràn and Pardo's randomization.

complexity $\mathbb{E}(\#\text{steps}) = \text{poly}(n, D)$.

Gràcies!

Merci !

¡Gracias!

Thank you!

Danke!

Present slides are online at *pierre.lairez.fr* with bibliographic references.

References I

-  Armentano, D., C. Beltrán, P. Bürgisser, F. Cucker, M. Shub (2016). “Condition Length and Complexity for the Solution of Polynomial Systems”. In: *Found. Comput. Math.*
-  Beltrán, C., L. M. Pardo (2009). “Smale’s 17th Problem: Average Polynomial Time to Compute Affine and Projective Solutions”. In: *J. Amer. Math. Soc.* 22.2, pp. 363–385.
-  – (2011). “Fast Linear Homotopy to Find Approximate Zeros of Polynomial Systems”. In: *Found. Comput. Math.* 11.1, pp. 95–129.
-  Beltrán, C., M. Shub (2009). “Complexity of Bezout’s Theorem. VII. Distance Estimates in the Condition Metric”. In: *Found. Comput. Math.* 9.2, pp. 179–195.
-  Bürgisser, P., F. Cucker (2011). “On a Problem Posed by Steve Smale”. In: *Ann. of Math. (2)* 174.3, pp. 1785–1836.
-  Lairez, P. (2017). “A Deterministic Algorithm to Compute Approximate Roots of Polynomial Systems in Polynomial Average Time”. In: *Found. Comput. Math.*

References II

-  Shub, M. (1993). “Some Remarks on Bezout’s Theorem and Complexity Theory”. In: *From Topology to Computation: Proceedings of the Smalefest*. Springer, New York, pp. 443–455.
-  – (2009). “Complexity of Bezout’s Theorem. VI. Geodesics in the Condition (Number) Metric”. In: *Found. Comput. Math.* 9.2, pp. 171–178.
-  Shub, M., S. Smale (1993a). “Complexity of Bézout’s Theorem. I. Geometric Aspects”. In: *J. Amer. Math. Soc.* 6.2, pp. 459–501.
-  – (1993b). “Complexity of Bezout’s Theorem. II. Volumes and Probabilities”. In: *Computational Algebraic Geometry (Nice, 1992)*. Vol. 109. Progr. Math. Birkhäuser Boston, Boston, MA, pp. 267–285.
-  – (1993c). “Complexity of Bezout’s Theorem. III. Condition Number and Packing”. In: *J. Complexity* 9.1, pp. 4–14.
-  – (1994). “Complexity of Bezout’s Theorem. V. Polynomial Time”. In: *Theoret. Comput. Sci.* 133.1. Selected papers of the Workshop on Continuous Algorithms and Complexity (Barcelona, 1993), pp. 141–164.

References III

-  Shub, M., S. Smale (1996). “Complexity of Bezout’s Theorem. IV. Probability of Success; Extensions”. In: *SIAM J. Numer. Anal.* 33.1, pp. 128–148.
-  Sibuya, M. (1962). “A Method for Generating Uniformly Distributed Points on N -Dimensional Spheres”. In: *Ann. Inst. Statist. Math.* 14, pp. 81–85.
-  Smale, S. (1986). “Newton’s Method Estimates from Data at One Point”. In: *The Merging of Disciplines: New Directions in Pure, Applied, and Computational Mathematics (Laramie, Wyo., 1985)*. Springer, New York, pp. 185–196.
-  – (1998). “Mathematical Problems for the next Century”. In: *The Mathematical Intelligencer* 20.2, pp. 7–15.