A deterministic solution to Smale’s 17th problem

Algorithms and complexity in algebraic geometry
Simons Institute, Berkeley, December 16, 2015

Pierre Lairez
TU Berlin
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.

Average polynomial time
Polynomial w. r. t. input size, on average w. r. t. a reasonable input distribution, typically Gaussian.

Uniform algorithm
A BSS machine: unit cost arithmetic operations on exact real numbers.
Symbolic vs. numeric

Symbolic
Knowing one root is knowing them all; the number of root is overpolynomial.

Numeric
Homotopy methods allow to approximate one root, disregarding the others.
\( \sim \) a polynomial complexity is not ruled out.

Typically
- \( n \) equations of degree 2 with \( n \) unknowns.
- Input size: \( N = n\binom{n+2}{2} \sim \frac{1}{2}n^3 \).
- Number of roots: \( \mathcal{D} = 2^n \), this is overpolynomial in \( N \).
Symbolic vs. numeric

Symbolic

Knowing one root is knowing them all; the number of root is overpolynomial.

Numeric

Homotopy methods allow to approximate one root, disregarding the others.
\[ \sim \text{a polynomial complexity is not ruled out.} \]

Typically

- \( n \) equations of degree \( n \) with \( n \) unknowns.
- Input size: \( N = n \left( \frac{2^n}{n} \right) \sim Cn^{1/2}4^n \).
- Number of roots: \( D = n^n \), this is overpolynomial in \( N \).
Notations

- $n$ and $D$, positive integers.
- $\mathcal{H}$, the linear space of all systems of $n$ equations of degree $D$ with $n$ unknowns; also functions $\mathbb{C}^n \to \mathbb{C}^n$.
- $N$, the complex dimension of $\mathcal{H}$.
- $\mathcal{H}$ is endowed with a hermitian inner product.
- $\mathbb{S}(\mathcal{H})$, the systems with unit norm.
The homotopy method

Input

\[ f \in \mathcal{H}, \text{a system to solve.} \]

Starting point

Choose another \( g \in \mathcal{H} \) of which we know a root \( \zeta \in \mathbb{C}^n \).

Homotopy

\[ h_0 = g \quad h_{k+1} = h_k + \delta_k \cdot (f - g) \]
\[ z_0 = \zeta \quad z_{k+1} = z_k - (d_z h_{k+1})^{-1}(h_{k+1}(z_k)). \]

End point

If \( h_K = f \), then \( z_K \) is an approximate root of \( f \).

- How to choose the step size \( \delta_k \)?
- How to choose the starting pair \((g, \zeta)\)?
The complexity of the homotopy method
Shub, Smale, 90’s

Shub and Smale:

- Gave a method to choose the $\delta_k$ in terms of a condition number $\mu(f, z)$;
- For each $n$ and $D$, proved the existence of a starting point $(g, \zeta)$ from which the homotopy method is efficient on the average.
- Gave a bound on the number of iteration in the homotopy method:

$$\text{number of iterations} \leq cD^{3/2} \int_g^f \mu(h, \eta)^2 dh.$$
Random starting points
Beltrán, Pardo, 2009

Beltrán and Pardo:
▷ Proved that a random starting point \((g, \zeta)\) is efficient on the (twofold) average.
▷ Discovered how to pick a random pair \((g, \zeta)\).

For us, Beltrán-Pardo algorithm is a function
\[
\text{BP} : \mathbb{S}(\mathcal{H}) \times [0, 1]^\mathbb{N} \rightarrow \mathbb{C}^n
\]
such that
▷ \(\text{BP}(f, a)\) is an approximate root of \(f\), for almost all \(f\) and \(a\);
▷ if \(f\) and \(a\) are uniformly distributed, then \(\mathbb{E}(\text{cost}_{\text{BP}}(f, a)) = O(nD^{3/2}N^2)\).
Smoothed analysis

*Bürgisser, Cucker, 2011*

Bürgisser and Cucker:

- Proved that the smoothed complexity of Beltrán-Pardo algorithm is polynomial:
  \[
  \sup_{f \in \mathcal{H}} \mathbb{E} (\text{cost}_{\text{BP}}(f)) = \infty
  \]

  \[
  \text{but } \sup_{f \in \mathcal{H}} \mathbb{E} (\text{cost}_{\text{BP}}(f + \varepsilon)) = O \left( \frac{1}{\sigma} n D^{3/2} N^2 \right),
  \]

  where \( \varepsilon \in \mathcal{H} \) is a random non centered Gaussian variable with variance \( \sigma^2 \).

- Described a deterministic algorithm with average complexity \( N^{O(\log \log N)} \).
Deterministic algorithm with complexity $O(nD^{3/2}N^2)$. 
**Duplication of the uniform dist. on $[0, 1]$**

- $q > 0$ an integer.
- $x \in [0, 1]$ a uniformly distributed random variable.

- $\lfloor x \rfloor_q \overset{\text{def}}{=} 2^{-q}[2^q x] \in [0, 1]$, the truncature of $x$ to precision $q$.
- $\{x\}_q \overset{\text{def}}{=} 2^q x - \lfloor 2^q x \rfloor \in [0, 1]$, the fractionary part.

**Proposition**

- The probability distribution of $\lfloor x \rfloor_q$ converges to the uniform distribution $[0, 1]$ when $q \to \infty$.
- $\{x\}_q$ is uniformly distributed on $[0, 1]$.
- $\lfloor x \rfloor_q$ and $\{x\}_q$ are independent.
**Duplication of the uniform dist. on $\mathbb{S}(\mathcal{H})$**

- $q > 0$ an integer.
- $x \in \mathbb{S}(\mathcal{H})$ a uniformly distributed random variable.

- $\lfloor x \rfloor_q \overset{\text{def}}{=} [\ldots] \in \mathbb{S}(\mathcal{H})$, the truncature of $x$ to precision $q$.
- $\{x\}_q \overset{\text{def}}{=} [\ldots] \in \mathbb{S}(\mathcal{H})$, the fractionary part.

**Proposition**

- The probability distribution of $\lfloor x \rfloor_q$ converges to the uniform distribution $\mathbb{S}(\mathcal{H})$ when $q \to \infty$.
- $\{x\}_q$ is *almost* uniformly distributed on $\mathbb{S}(\mathcal{H})$.
- $\lfloor x \rfloor_q$ and $\{x\}_q$ are *almost* independent.
A deterministic algorithm
Derandomization of Beltrán-Pardo algorithm

Beltrán-Pardo algorithm
\[ \text{BP} : \mathcal{S}(\mathcal{H}) \times [0, 1]^\mathbb{N} \rightarrow \mathbb{C}^n. \]
A deterministic algorithm

Derandomization of Beltrán-Pardo algorithm

Modified Beltrán-Pardo algorithm

\[ \text{BP} : \mathbb{S}(\mathcal{H}) \times \mathbb{S}(\mathcal{H}) \rightarrow \mathbb{C}^n. \]
A deterministic algorithm
Derandomization of Beltrán-Pardo algorithm

Modified Beltrán-Pardo algorithm

$$\text{BP} : \mathbb{S}(\mathcal{H}) \times \mathbb{S}(\mathcal{H}) \rightarrow \mathbb{C}^n.$$
A deterministic algorithm
Derandomization of Beltrán-Pardo algorithm

Modified Beltrán-Pardo algorithm

\[ \text{BP} : \mathcal{S}(\mathcal{H}) \times \mathcal{S}(\mathcal{H}) \rightarrow \mathbb{C}^n. \]

The algorithm, 1st try

**procedure** DBP(f)

\[ q \leftarrow \text{a large enough integer} \]

**return** \( \text{BP} \left( \lfloor f \rfloor_q, \{f\}_q \right) \)

**end procedure**
A deterministic algorithm
Derandomization of Beltrán-Pardo algorithm

Modified Beltrán-Pardo algorithm
BP : \( S(H) \times S(H) \rightarrow \mathbb{C}^n \).

The algorithm, 2nd try

```
procedure DBP(f)
    q ← \lfloor \log_2 N \rfloor
    repeat
        q ← 2q
        z ← BP(\lfloor f \rfloor_q, \{f\}_q)
    until z is an approximate root of f
    return z
end procedure
```
A deterministic algorithm
Derandomization of Beltrán-Pardo algorithm

Modified Beltrán-Pardo algorithm
\[ \text{BP} : \mathbb{S}(\mathcal{H}) \times \mathbb{S}(\mathcal{H}) \rightarrow \mathbb{C}^n. \]

The algorithm, final version

\begin{verbatim}
procedure DBP(f)
  q ← \lceil \log_2 N \rceil
  repeat
    q ← 2q
    z ← BP(\lfloor f \rfloor_q, \{f\}_q) with early abort
  until z is an approximate root of f
  return z
end procedure
\end{verbatim}
Homotopy continuation with early abort

**procedure** \( \text{HC}'(f, g, z, q) \)

\[
t \leftarrow 1/ \left( 101 D^{3/2} \mu(g, z)^2 d_S(f, g) \right)
\]

**while** \( 1 > t \) **do**

\[
h \leftarrow \Gamma(g, f, t) \quad \triangleright \text{“} tf + (1 - t)g \text{” on the sphere}
\]

\[
z \leftarrow \text{Newton}(h, z)
\]

\[
t \leftarrow t + 1/ \left( 101 D^{3/2} \mu(h, z)^2 d_S(f, g) \right)
\]

**abort if** \( 151 D^{3/2} \mu(h, z)^2 > 2^q \)

**end while**

**return** \( z \)

**end procedure**

- If \( \|f - \tilde{f}\| \leq 2^{-q} \), then \( \text{HC}'(f, g, z, q) \) fails or returns an approximate root of \( \tilde{f} \).

- In any case, it performs at most \( cD^{3/2} \int_g^{\tilde{f}} \mu(h, z)^2 dh \) steps.
Complexity analysis

- Let $f \in \mathcal{S}(\mathcal{H})$ be a uniformly distributed random variable.
- Let $\Omega$ be the number of iterations in $\text{DBP}(f)$.

**Proposition** — $\mathbb{E}(\Omega) \leq 7$. (And the distribution is very light-tailed.)

$\leadsto$ The precision $q$ is typically no more than $128 \log N$. 
Complexity analysis

- Let \( f \in \mathbb{S}(\mathcal{H}) \) be a uniformly distributed random variable.
- Let \( \Omega \) be the number of iterations in DBP\((f)\).

**Proposition** — \( \mathbb{E}(\Omega) \leq 7 \). (And the distribution is very light-tailed.)

\( \leadsto \) The precision \( q \) is typically no more than 128 \( \log N \).

**Complexity analysis**

- \( \text{cost}_{\text{DBP}}(f) = \sum_{k=1}^{\Omega} \left( \mathcal{O}(Nq_k) + \text{cost}_{\text{BP}'}(\lfloor f \rfloor q_k, \{f\}_q) \right) \)
Complexity analysis

- Let $f \in \mathbb{S}(\mathcal{H})$ be a uniformly distributed random variable.
- Let $\Omega$ be the number of iterations in DBP($f$).

**Proposition** — $\mathbb{E}(\Omega) \leq 7$. (And the distribution is very light-tailed.)

$\leadsto$ The precision $q$ is typically no more than $128 \log N$.

**Complexity analysis**

- $\text{cost}_{\text{DBP}}(f) = \sum_1^\Omega \left( \mathcal{O}(Nq_k) + \text{cost}_{\text{BP}'}(\lfloor f \rfloor_{q_k}, \{f\}_{q_k}) \right)$
- $\text{cost}_{\text{BP}'}(\lfloor f \rfloor_{q_k}, \{f\}_{q_k}) \sim \text{cost}_{\text{BP}}(\lfloor f \rfloor_{q_k}, g) \sim \text{cost}_{\text{BP}}(f, g)$
**Complexity analysis**

- Let $f \in \mathbb{S}(\mathcal{H})$ be a uniformly distributed random variable.
- Let $\Omega$ be the number of iterations in DBP($f$).

**Proposition** — $\mathbb{E}(\Omega) \leq 7$. (And the distribution is very light-tailed.)

$\leadsto$ The precision $q$ is typically no more than $128 \log N$.

**Complexity analysis**

- $\text{cost}_{\text{DBP}}(f) = \sum_{k=1}^{\Omega} \left( O(Nq_k) + \text{cost}_{\text{BP'}} \left( \lfloor f \rfloor_{q_k}, \{f\}_{q_k} \right) \right)$
- $\text{cost}_{\text{BP'}} \left( \lfloor f \rfloor_{q_k}, \{f\}_{q_k} \right) \sim \text{cost}_{\text{BP}}(\lfloor f \rfloor_{q_k}, g) \sim \text{cost}_{\text{BP}}(f, g)$
- $\mathbb{E}(\text{cout}_{\text{BPD}}(f)) = O(nD^{3/2}N^2)$
Conclusion

Randomness is part of Smale’s 17th problem from its very formulation asking for an average analysis.

Problème no. 17bis — Can a zero of $n$ complex polynomial equations in $n$ unknowns be found approximately in polynomial time with respect to the evaluation complexity of the input and the logarithm of its conditionning?
Conclusion

Randomness is part of Smale’s 17th problem from its very formulation asking for an average analysis.

Problème no. 17bis — Can a zero of $n$ complex polynomial equations in $n$ unknowns be found approximately in polynomial time with respect to the evaluation complexity of the input and the logarithm of its conditionning?

Thank you!