Smale's Fundamental Theorem of Algebra reconsidered

Diego Armentano (joint work with Mike Shub)

Centro de Matemática, Universidad de la República.
E-mail: diego@cmat.edu.uy

May 8, 2012

Fields Institute, Toronto

Smale's Fundamental Theorem of Algebra

In 1981 Steve Smale initiated the complexity theory of finding a solution of polynomial equations of one complex variable.

Problem (*):

Given

$$f(z) = \sum_{i=0}^{d} a_i z^i, \quad a_i \in \mathbb{C}, \qquad \text{find} \quad \eta \in \mathbb{C} \quad \text{such that} \quad f(\eta) = 0$$

- η should be replaced by an approximate zero ("strong" Newton sink).
- Complexity = number of required steps.

In 1981 Steve Smale initiated the complexity theory of finding a solution of polynomial equations of one complex variable.

Problem (*):

Given

$$f(z) = \sum_{i=0}^{d} a_i z^i, \quad a_i \in \mathbb{C}, \qquad \text{find} \quad \eta \in \mathbb{C} \quad \text{such that} \quad f(\eta) = 0$$

- η should be replaced by an approximate zero ("strong" Newton sink).
- Complexity = number of required steps.

Smale's Fundamental Theorem of Algebra

In 1981 Steve Smale initiated the complexity theory of finding a solution of polynomial equations of one complex variable.

Problem (*):

Given

$$f(z) = \sum_{i=0}^{d} a_i z^i, \quad a_i \in \mathbb{C}, \qquad \text{find} \quad \eta \in \mathbb{C} \quad \text{such that} \quad f(\eta) = 0$$

- η should be replaced by an approximate zero ("strong" Newton sink).
- Complexity = number of required steps.

Smale introduced a STATISTICAL theory of cost:

Let $\mathcal{A}$ be an algorithm to solve (*), and consider a probability measure on the set of polynomials.

> *Given $\varepsilon > 0$, an allowable probability of failure, does the cost of $\mathcal{A}$ on a set of polynomials with probability $1 - \varepsilon$, grow at most polynomial in d?*

Smale gives a positive answer to this question, however this initial algorithm was not proven to be finite average cost.

Smale introduced a STATISTICAL theory of cost:

Let $\mathcal{A}$ be an algorithm to solve (*), and consider a probability measure on the set of polynomials.

> *Given $\varepsilon > 0$, an allowable probability of failure, does the cost of $\mathcal{A}$ on a set of polynomials with probability $1-\varepsilon$, grow at most polynomial in d?*

Smale gives a positive answer to this question, however this initial algorithm was not proven to be finite average cost.

Smale introduced a STATISTICAL theory of cost:

Let $\mathcal{A}$ be an algorithm to solve (*), and consider a probability measure on the set of polynomials.

> *Given* $\varepsilon > 0$*, an allowable probability of failure, does the cost of* $\mathcal{A}$ *on a set of polynomials with probability* $1 - \varepsilon$*, grow at most polynomial in* d*?*

Smale gives a positive answer to this question, however this initial algorithm was not proven to be finite average cost.

Smale's Algorithm:

Let $0 < h \leq 1$ and let $z_0 = 0$.

Inductively define

$$z_n = T_h(z_{n-1}),$$

where T_h is the modified Newton's method for f given by

$$T_h(z) = z - h\frac{f(z)}{f'(z)}.$$

(If h is small enough, $\{z_n\}$ approximate the trajectories of the Newton Flow $N(z) = -\frac{f(z)}{f'(z)}$.)

Smale's Algorithm:

Let $0 < h \leq 1$ and let $z_0 = 0$.

Inductively define

$$z_n = T_h(z_{n-1}),$$

where T_h is the modified Newton's method for f given by

$$T_h(z) = z - h\frac{f(z)}{f'(z)}.$$

(If h is small enough, $\{z_n\}$ approximate the trajectories of the Newton Flow $N(z) = -\frac{f(z)}{f'(z)}$.)

Smale's Algorithm:

Let $0 < h \leq 1$ and let $z_0 = 0$.

Inductively define

$$z_n = T_h(z_{n-1}),$$

where T_h is the modified Newton's method for f given by

$$T_h(z) = z - h\frac{f(z)}{f'(z)}.$$

(If h is small enough, $\{z_n\}$ approximate the trajectories of the Newton Flow $N(z) = -\frac{f(z)}{f'(z)}$.)

For $z_0 \in \mathbb{C}$, consider

$$f_t = f - (1-t)f(z_0), \qquad 0 \le t \le 1.$$

- f_t is a polynomial of the same degree as f;
- z_0 is a zero of f_0;
- $f_1 = f$.

We analytically continue z_0 to z_t a zero of f_t.

For $t = 1$ we arrive at a zero of f. Newton's method is then used to produce a discrete numerical approximation to the path (f_t, z_t).

For $z_0 \in \mathbb{C}$, consider

$$f_t = f - (1-t)f(z_0), \qquad 0 \leq t \leq 1.$$

- f_t is a polynomial of the same degree as f;
- z_0 is a zero of f_0;
- $f_1 = f$.

We analytically continue z_0 to z_t a zero of f_t.

For $t = 1$ we arrive at a zero of f. Newton's method is then used to produce a discrete numerical approximation to the path (f_t, z_t).

For $z_0 \in \mathbb{C}$, consider

$$f_t = f - (1-t)f(z_0), \qquad 0 \leq t \leq 1.$$

- f_t is a polynomial of the same degree as f;
- z_0 is a zero of f_0;
- $f_1 = f$.

We analytically continue z_0 to z_t a zero of f_t.

For $t = 1$ we arrive at a zero of f. Newton's method is then used to produce a discrete numerical approximation to the path (f_t, z_t).

- A tremendous amount of work has been done in the last 30 years following on Smale's initial contribution.
- In a series of papers (Bezout I-V) Shub-Smale made some further changes and achieved enough results for Smale 17th

Problem 17: Solving Polynomial Equations.

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

- Beltrán, Boito, Bürgisser, Cucker, Dedieu, Hirsch, Kim, Leykin, Li, Malajovich, Martens, Pardo, Renegar, Rojas, Sutherland.... and specially Mike Shub and Steve Smale.

• A tremendous amount of work has been done in the last 30 years following on Smale's initial contribution.

• In a series of papers (Bezout I-V) Shub-Smale made some further changes and achieved enough results for Smale 17th

Problem 17: Solving Polynomial Equations.

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

• Beltrán, Boito, Bürgisser, Cucker, Dedieu, Hirsch, Kim, Leykin, Li, Malajovich, Martens, Pardo, Renegar, Rojas, Sutherland.... and specially Mike Shub and Steve Smale.

- A tremendous amount of work has been done in the last 30 years following on Smale's initial contribution.
- In a series of papers (Bezout I-V) Shub-Smale made some further changes and achieved enough results for Smale 17th

Problem 17: Solving Polynomial Equations.

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

- Beltrán, Boito, Bürgisser, Cucker, Dedieu, Hirsch, Kim, Leykin, Li, Malajovich, Martens, Pardo, Renegar, Rojas, Sutherland.... and specially Mike Shub and Steve Smale.

• A tremendous amount of work has been done in the last 30 years following on Smale's initial contribution.

• In a series of papers (Bezout I-V) Shub-Smale made some further changes and achieved enough results for Smale 17th

Problem 17: Solving Polynomial Equations.

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

• Beltrán, Boito, Bürgisser, Cucker, Dedieu, Hirsch, Kim, Leykin, Li, Malajovich, Martens, Pardo, Renegar, Rojas, Sutherland.... and specially Mike Shub and Steve Smale.

- $\mathcal{H}_{(d)} := \mathcal{H}_{d_1} \times \cdots \times \mathcal{H}_{d_n}$ where $\mathcal{H}_{d_i}$ is the vector space of homogeneous polynomials of degree d_i in $n+1$ complex variables.
- For $f \in \mathcal{H}_{(d)}$ and $\lambda \in \mathbb{C}$,

$$f(\lambda\zeta) = \Delta\left(\lambda^{d_i}\right) f(\zeta),$$

where $\Delta(a_i)$ means the diagonal matrix whose i-th diagonal entry is a_i.
- Thus the zeros of $f \in \mathcal{H}_{(d)}$ are now complex lines so may be considered as points in projective space $\mathbb{P}(\mathbb{C}^{n+1})$.

- $\mathcal{H}_{(d)} := \mathcal{H}_{d_1} \times \cdots \times \mathcal{H}_{d_n}$ where $\mathcal{H}_{d_i}$ is the vector space of homogeneous polynomials of degree d_i in $n+1$ complex variables.
- For $f \in \mathcal{H}_{(d)}$ and $\lambda \in \mathbb{C}$,

$$f(\lambda\zeta) = \Delta\left(\lambda^{d_i}\right) f(\zeta),$$

where $\Delta(a_i)$ means the diagonal matrix whose i-th diagonal entry is a_i.

- Thus the zeros of $f \in \mathcal{H}_{(d)}$ are now complex lines so may be considered as points in projective space $\mathbb{P}(\mathbb{C}^{n+1})$.

- $\mathcal{H}_{(d)} := \mathcal{H}_{d_1} \times \cdots \times \mathcal{H}_{d_n}$ where $\mathcal{H}_{d_i}$ is the vector space of homogeneous polynomials of degree d_i in $n+1$ complex variables.
- For $f \in \mathcal{H}_{(d)}$ and $\lambda \in \mathbb{C}$,

$$f(\lambda\zeta) = \Delta\left(\lambda^{d_i}\right) f(\zeta),$$

where $\Delta(a_i)$ means the diagonal matrix whose i-th diagonal entry is a_i.

- Thus the zeros of $f \in \mathcal{H}_{(d)}$ are now complex lines so may be considered as points in projective space $\mathbb{P}(\mathbb{C}^{n+1})$.

- On $\mathcal{H}_{d_i}$ we put a unitarily invariant Hermitian structure:
  If $f(z) = \sum_{\|\alpha\|=d_i} a_\alpha z^\alpha$ and $g(z) = \sum_{\|\alpha\|=d_i} b_\alpha z^\alpha$ then the Weyl Hermitian structure is given by

$$\langle f, g \rangle = \sum_{\|\alpha\|=d_i} a_\alpha \overline{b_\alpha} \binom{d_i}{\alpha}^{-1}.$$

- On $\mathcal{H}_{(d)}$ we put the product structure

$$\langle f, g \rangle = \sum_{i=1}^{n} \langle f_i, g_i \rangle.$$

- On $\mathcal{H}_{d_i}$ we put a unitarily invariant Hermitian structure: If $f(z)=\sum_{\|\alpha\|=d_i} a_\alpha z^\alpha$ and $g(z)=\sum_{\|\alpha\|=d_i} b_\alpha z^\alpha$ then the Weyl Hermitian structure is given by

$$\langle f,g\rangle = \sum_{\|\alpha\|=d_i} a_\alpha \overline{b_\alpha} \binom{d_i}{\alpha}^{-1}.$$

- On $\mathcal{H}_{(d)}$ we put the product structure

$$\langle f,g\rangle = \sum_{i=1}^{n} \langle f_i, g_i\rangle.$$

- On $\mathbb{C}^{n+1}$ we put the usual Hermitian structure

$$\langle x, y \rangle = \sum_{k=0}^{n} x_k \, \overline{y_k}.$$

- $\mathbb{P}(\mathbb{C}^{n+1})$ inherits the Hermitian structure from $\mathbb{C}^{n+1}$ (Fubini-Study Herm. struct. $\langle w_1, w_2 \rangle_v = \frac{\langle w_1, w_2 \rangle}{\langle v, v \rangle}$, $w_i \in v^\perp$).
- $\mathcal{U}(n+1)$ (group of unitary transformations) acts on $\mathcal{H}_{(d)}$ and $\mathbb{C}^{n+1}$: $f \mapsto f \circ U^{-1}$, and $\zeta \mapsto U\zeta$, $U \in \mathcal{U}(n+1)$.
- This unitary action preserves the Hermitian structure on $\mathcal{H}_{(d)}$ and $\mathbb{C}^{n+1}$.

- On $\mathbb{C}^{n+1}$ we put the usual Hermitian structure

$$\langle x, y \rangle = \sum_{k=0}^{n} x_k \, \overline{y_k}.$$

- $\mathbb{P}(\mathbb{C}^{n+1})$ inherits the Hermitian structure from $\mathbb{C}^{n+1}$ (Fubini-Study Herm. struct. $\langle w_1, w_2 \rangle_v = \frac{\langle w_1, w_2 \rangle}{\langle v, v \rangle}$, $w_i \in v^{\perp}$).
- $\mathcal{U}(n+1)$ (group of unitary transformations) acts on $\mathcal{H}_{(d)}$ and $\mathbb{C}^{n+1}$: $f \mapsto f \circ U^{-1}$, and $\zeta \mapsto U\zeta$, $U \in \mathcal{U}(n+1)$.
- This unitary action preserves the Hermitian structure on $\mathcal{H}_{(d)}$ and $\mathbb{C}^{n+1}$.

- On $\mathbb{C}^{n+1}$ we put the usual Hermitian structure

$$\langle x, y\rangle = \sum_{k=0}^{n} x_k\,\overline{y_k}.$$

- $\mathbb{P}(\mathbb{C}^{n+1})$ inherits the Hermitian structure from $\mathbb{C}^{n+1}$ (Fubini-Study Herm. struct. $\langle w_1, w_2\rangle_v = \frac{\langle w_1,w_2\rangle}{\langle v,v\rangle}$, $w_i \in v^\perp$).
- $\mathcal{U}(n+1)$ (group of unitary transformations) acts on $\mathcal{H}_{(d)}$ and $\mathbb{C}^{n+1}$: $f \mapsto f \circ U^{-1}$, and $\zeta \mapsto U\zeta$, $U \in \mathcal{U}(n+1)$.
- This unitary action preserves the Hermitian structure on $\mathcal{H}_{(d)}$ and $\mathbb{C}^{n+1}$.

- On $\mathbb{C}^{n+1}$ we put the usual Hermitian structure

$$\langle x, y \rangle = \sum_{k=0}^{n} x_k \, \overline{y_k}.$$

- $\mathbb{P}(\mathbb{C}^{n+1})$ inherits the Hermitian structure from $\mathbb{C}^{n+1}$ (Fubini-Study Herm. struct. $\langle w_1, w_2 \rangle_v = \frac{\langle w_1, w_2 \rangle}{\langle v, v \rangle}$, $w_i \in v^{\perp}$).
- $\mathcal{U}(n+1)$ (group of unitary transformations) acts on $\mathcal{H}_{(d)}$ and $\mathbb{C}^{n+1}$: $f \mapsto f \circ U^{-1}$, and $\zeta \mapsto U\zeta$, $U \in \mathcal{U}(n+1)$.
- This unitary action preserves the Hermitian structure on $\mathcal{H}_{(d)}$ and $\mathbb{C}^{n+1}$.

Extensions

Notations

The solution variety

$$\mathcal{V} = \{(f,\zeta) \in (\mathcal{H}_{(d)} - \{0\}) \times \mathbb{P}(\mathbb{C}^{n+1}) : \; f(\zeta) = 0\},$$

is a central object of study.

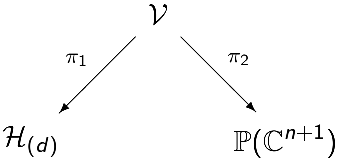
$\mathcal{V}$ is equipped with two projections

The solution variety

$$\mathcal{V} = \{(f,\zeta) \in (\mathcal{H}_{(d)} - \{0\}) \times \mathbb{P}(\mathbb{C}^{n+1}) : \ f(\zeta) = 0\},$$

is a central object of study.

$\mathcal{V}$ is equipped with two projections



- Choose $(g, \zeta) \in \mathcal{V}$ a known pair.
- Connect g to f by a C^1 curve f_t in $\mathcal{H}_{(d)}$, $0 \leq t \leq 1$, such that $f_0 = g$, $f_1 = f$, and continue $\zeta_0 = \zeta$ to ζ_t such that $f_t(\zeta_t) = 0$, so that $f_1(\zeta_1) = 0$.

Now homotopy methods numerically approximate the path (f_t, ζ_t).

One way to accomplish the approximation is via (projective) Newton's methods.

Given an approximation x_t to ζ_t define

$$x_{t+\Delta t} = N_{f_{t+\Delta t}}(x_t),$$

where

$$N_f(x) = x - (Df(x)|_{x^\perp})^{-1} f(x).$$

Homotopy Methods

- Choose $(g, \zeta) \in \mathcal{V}$ a known pair.
- Connect g to f by a C^1 curve f_t in $\mathcal{H}_{(d)}$, $0 \leq t \leq 1$, such that $f_0 = g$, $f_1 = f$, and continue $\zeta_0 = \zeta$ to ζ_t such that $f_t(\zeta_t) = 0$, so that $f_1(\zeta_1) = 0$.

Now homotopy methods numerically approximate the path (f_t, ζ_t).

One way to accomplish the approximation is via (projective) Newton's methods.

Given an approximation x_t to ζ_t define

$$x_{t+\Delta t} = N_{f_{t+\Delta t}}(x_t),$$

where

$$N_f(x) = x - (Df(x)|_{x^\perp})^{-1} f(x).$$

- Choose $(g, \zeta) \in \mathcal{V}$ a known pair.
- Connect g to f by a C^1 curve f_t in $\mathcal{H}_{(d)}$, $0 \leq t \leq 1$, such that $f_0 = g$, $f_1 = f$, and continue $\zeta_0 = \zeta$ to ζ_t such that $f_t(\zeta_t) = 0$, so that $f_1(\zeta_1) = 0$.

Now homotopy methods numerically approximate the path (f_t, ζ_t).

One way to accomplish the approximation is via (projective) Newton's methods.

Given an approximation x_t to ζ_t define

$$x_{t+\Delta t} = N_{f_{t+\Delta t}}(x_t),$$

where

$$N_f(x) = x - (Df(x)|_{x^\perp})^{-1} f(x).$$

Mike Shub prove that Δt may be chosen so that:

- $t_0 = 0,\ t_k = t_{k-1} + \Delta t_k$;
- x_{t_k} is an approx. zero of f_{t_k} with associated zero ζ_{t_k} and
- $t_K = 1$ for

$$K = K(f, g, \zeta) \leq C\, D^{3/2} \int_0^1 \mu(f_t, \zeta_t)\, \|(\dot{f}_t, \dot{\zeta}_t)\|_{(f_t, \zeta_t)}\, dt = (I).$$

(C universal constant, $D = \max d_i$),

$$\mu(f, \zeta) = \|f\| \cdot \|(Df(\zeta)|_{\zeta^\perp})^{-1} \Delta(\|\zeta\|^{d_i - 1} \sqrt{d_i})\|$$

is the condition number of f at ζ, and

$\|(\dot{f}_t, \dot{\zeta}_t)\|_{(f_t, \zeta_t)}$ is the norm of the tangent vector to the projected curve in (f_t, ζ_t) in $\mathcal{V}_{\mathbb{P}} \subset \mathbb{P}\left(\mathcal{H}_{(d)}\right) \times \mathbb{P}(\mathbb{C}^{n+1})$.($\Delta t_k$ is made explicit in Dedieu-Malajovich-Shub).

Homotopy Methods

Mike Shub prove that Δt may be chosen so that:

- $t_0 = 0,\ t_k = t_{k-1} + \Delta t_k;$
- x_{t_k} is an approx. zero of f_{t_k} with associated zero ζ_{t_k} and
- $t_K = 1$ for

$$K = K(f,g,\zeta) \leq C\, D^{3/2} \int_0^1 \mu(f_t,\zeta_t)\, \|(\dot{f}_t,\dot{\zeta}_t)\|_{(f_t,\zeta_t)}\, dt = (I).$$

(C universal constant, $D = \max d_i$),

$$\mu(f,\zeta) = \|f\| \cdot \|(Df(\zeta)|_{\zeta^\perp})^{-1}\Delta(\|\zeta\|^{d_i-1}\sqrt{d_i})\|$$

is the condition number of f at ζ, and

$\|(\dot{f}_t,\dot{\zeta}_t)\|_{(f_t,\zeta_t)}$ is the norm of the tangent vector to the projected curve in (f_t,ζ_t) in $\mathcal{V}_{\mathbb{P}} \subset \mathbb{P}\left(\mathcal{H}_{(d)}\right) \times \mathbb{P}(\mathbb{C}^{n+1})$.($\Delta t_k$ is made explicit in Dedieu-Malajovich-Shub).

Homotopy Methods

Mike Shub prove that Δt may be chosen so that:

- $t_0 = 0$, $t_k = t_{k-1} + \Delta t_k$;
- x_{t_k} is an approx. zero of f_{t_k} with associated zero ζ_{t_k} and
- $t_K = 1$ for

$$K = K(f, g, \zeta) \leq C\, D^{3/2} \int_0^1 \mu(f_t, \zeta_t)\, \|(\dot{f}_t, \dot{\zeta}_t)\|_{(f_t, \zeta_t)}\, dt = (I).$$

(C universal constant, $D = \max d_i$),

$$\mu(f, \zeta) = \|f\| \cdot \|(Df(\zeta)|_{\zeta^\perp})^{-1} \Delta(\|\zeta\|^{d_i - 1} \sqrt{d_i})\|$$

is the condition number of f at ζ, and

$\|(\dot{f}_t, \dot{\zeta}_t)\|_{(f_t, \zeta_t)}$ is the norm of the tangent vector to the projected curve in (f_t, ζ_t) in $\mathcal{V}_{\mathbb{P}} \subset \mathbb{P}\left(\mathcal{H}_{(d)}\right) \times \mathbb{P}(\mathbb{C}^{n+1})$.($\Delta t_k$ is made explicit in Dedieu-Malajovich-Shub).

An affirmative probabilistic solution to Smale's 17th problem is proven by Beltrán and Pardo (2009). They prove that a random point (g, ζ) is good in the sense that with random fixed starting point $(g, \zeta) = (f_0, \zeta_0)$ the average value of K is bounded by $O(nN)$.

Bürgisser and Cucker (2011) produce a deterministic starting point with polynomial average cost, except for a narrow range of dimensions. Precisely, $D \leq n^{\frac{1}{1+\varepsilon}}$ (lin. h.m) or $D \geq n^{1+\varepsilon}$ (variant Renegar).

So Smale's 17th problem in its deterministic form remains open for a narrow range of degrees and variables.

An affirmative probabilistic solution to Smale's 17th problem is proven by Beltrán and Pardo (2009). They prove that a random point (g,ζ) is good in the sense that with random fixed starting point $(g,\zeta) = (f_0,\zeta_0)$ the average value of K is bounded by $O(nN)$.

Bürgisser and Cucker (2011) produce a deterministic starting point with polynomial average cost, except for a narrow range of dimensions. Precisely, $D \le n^{\frac{1}{1+\varepsilon}}$ (lin. h.m) or $D \ge n^{1+\varepsilon}$ (variant Renegar).

So Smale's 17th problem in its deterministic form remains open for a narrow range of degrees and variables.

An affirmative probabilistic solution to Smale's 17th problem is proven by Beltrán and Pardo (2009). They prove that a random point (g, ζ) is good in the sense that with random fixed starting point $(g, \zeta) = (f_0, \zeta_0)$ the average value of K is bounded by $O(nN)$.

Bürgisser and Cucker (2011) produce a deterministic starting point with polynomial average cost, except for a narrow range of dimensions. Precisely, $D \leq n^{\frac{1}{1+\varepsilon}}$ (lin. h.m) or $D \geq n^{1+\varepsilon}$ (variant Renegar).

So Smale's 17th problem in its deterministic form remains open for a narrow range of degrees and variables.

Given $\zeta \in \mathbb{P}(\mathbb{C}^{n+1})$ we define for $f \in \mathcal{H}_{(d)}$ the straight line segment $f_t \in \mathcal{H}_{(d)}$, $0 \leq t \leq 1$, by

$$(f_t)_i = f_i - (1-t)\frac{\langle \cdot, \zeta\rangle^{d_i}}{\langle \zeta, \zeta\rangle^{d_i}} f_i(\zeta), \qquad (i = 1, \ldots, n).$$

So $f_0(\zeta) = 0$ and $f_1 = f$. Therefore we may apply homotopy methods to this line segment.

Note that if we restrict f to the affine chart $\zeta + \zeta^{\perp}$ then

$$f_t(z) = f(z) - (1-t)f(\zeta),$$

and if we take $\zeta = (1, 0, \ldots, 0)$ and $n = 1$ we recover Smale's algorithm.

Given $\zeta \in \mathbb{P}(\mathbb{C}^{n+1})$ we define for $f \in \mathcal{H}_{(d)}$ the straight line segment $f_t \in \mathcal{H}_{(d)}$, $0 \leq t \leq 1$, by

$$(f_t)_i = f_i - (1-t)\frac{\langle \cdot, \zeta \rangle^{d_i}}{\langle \zeta, \zeta \rangle^{d_i}} f_i(\zeta), \qquad (i = 1, \ldots, n).$$

So $f_0(\zeta) = 0$ and $f_1 = f$. Therefore we may apply homotopy methods to this line segment.

Note that if we restrict f to the affine chart $\zeta + \zeta^{\perp}$ then

$$f_t(z) = f(z) - (1-t)f(\zeta),$$

and if we take $\zeta = (1, 0, \ldots, 0)$ and $n = 1$ we recover Smale's algorithm.

Given $\zeta \in \mathbb{P}(\mathbb{C}^{n+1})$ we define for $f \in \mathcal{H}_{(d)}$ the straight line segment $f_t \in \mathcal{H}_{(d)}$, $0 \leq t \leq 1$, by

$$(f_t)_i = f_i - (1-t)\frac{\langle \cdot, \zeta\rangle^{d_i}}{\langle \zeta, \zeta\rangle^{d_i}} f_i(\zeta), \qquad (i = 1, \ldots, n).$$

So $f_0(\zeta) = 0$ and $f_1 = f$. Therefore we may apply homotopy methods to this line segment.

Note that if we restrict f to the affine chart $\zeta + \zeta^{\perp}$ then

$$f_t(z) = f(z) - (1-t)f(\zeta),$$

and if we take $\zeta = (1, 0, \ldots, 0)$ and $n = 1$ we recover Smale's algorithm.

Given $\zeta \in \mathbb{P}(\mathbb{C}^{n+1})$ we define for $f \in \mathcal{H}_{(d)}$ the straight line segment $f_t \in \mathcal{H}_{(d)}$, $0 \leq t \leq 1$, by

$$(f_t)_i = f_i - (1-t)\frac{\langle \cdot, \zeta \rangle^{d_i}}{\langle \zeta, \zeta \rangle^{d_i}} f_i(\zeta), \qquad (i = 1, \ldots, n).$$

So $f_0(\zeta) = 0$ and $f_1 = f$. Therefore we may apply homotopy methods to this line segment.

Note that if we restrict f to the affine chart $\zeta + \zeta^{\perp}$ then

$$f_t(z) = f(z) - (1-t)f(\zeta),$$

and if we take $\zeta = (1, 0, \ldots, 0)$ and $n = 1$ we recover Smale's algorithm.

Let $\mathcal{V}_\zeta = \pi_1(\pi_2^{-1}(\zeta))$ be the subspace of $\mathcal{H}_{(d)}$ given by

$$\mathcal{V}_\zeta = \{f \in \mathcal{H}_{(d)} : \; f(\zeta) = 0\},$$

then

$$f_0 = f - \Delta\left(\frac{\langle \cdot, \zeta \rangle^{d_i}}{\langle \zeta, \zeta \rangle^{d_i}}\right) f(\zeta),$$

is the orthogonal projection $\Pi_\zeta(f)$ of f on $\mathcal{V}_\zeta$.

We have

$$\|f - \Pi_\zeta(f)\| = \|\Delta(\|\zeta\|^{-d_i}) f(\zeta)\|,$$

Let $\mathcal{V}_\zeta = \pi_1(\pi_2^{-1}(\zeta))$ be the subspace of $\mathcal{H}_{(d)}$ given by

$$\mathcal{V}_\zeta = \{f \in \mathcal{H}_{(d)} : \ f(\zeta) = 0\},$$

then

$$f_0 = f - \Delta\left(\frac{\langle \cdot, \zeta \rangle^{d_i}}{\langle \zeta, \zeta \rangle^{d_i}}\right) f(\zeta),$$

is the orthogonal projection $\Pi_\zeta(f)$ of f on $\mathcal{V}_\zeta$.

We have

$$\|f - \Pi_\zeta(f)\| = \|\Delta(\|\zeta\|^{-d_i}) f(\zeta)\|,$$

Let $\mathcal{V}_\zeta = \pi_1(\pi_2^{-1}(\zeta))$ be the subspace of $\mathcal{H}_{(d)}$ given by

$$\mathcal{V}_\zeta = \{f \in \mathcal{H}_{(d)} : \ f(\zeta) = 0\},$$

then

$$f_0 = f - \Delta\left(\frac{\langle \cdot, \zeta \rangle^{d_i}}{\langle \zeta, \zeta \rangle^{d_i}}\right) f(\zeta),$$

is the orthogonal projection $\Pi_\zeta(f)$ of f on $\mathcal{V}_\zeta$.

We have

$$\|f - \Pi_\zeta(f)\| = \|\Delta(\|\zeta\|^{-d_i}) f(\zeta)\|,$$

Then we can write $f_t = (1-t)\Pi_\zeta(f) + tf$.

Let ζ_t be the homotopy continuation of ζ along the path f_t (in case it is defined). Then $\{(f_t, \zeta_t)\}_{\in[0,1]} \subset \mathcal{V}$, and ζ_1 is a root of f.

- For a.e. $f \in \mathcal{H}_{(d)}$ the set of $\zeta \in \mathbb{P}(\mathbb{C}^{n+1})$ such that ζ_t is defined for all $t \in [0,1]$ has full measure. Moreover, the boundary of this full measure set is a stratified set.
- Suppose η is a non-degenerate zero of $h \in \mathcal{H}_{(d)}$. Let $B(h,\eta)$ be the basin of η, i.e. the set of those $\zeta \in \mathbb{P}(\mathbb{C}^{n+1})$ such that the zero ζ of $\Pi_\zeta(h)$ continues to η for the homotopy h_t. ($B(h,\eta)$ is an open set.)

Then we can write $f_t = (1-t)\Pi_\zeta(f) + tf$.

Let ζ_t be the homotopy continuation of ζ along the path f_t (in case it is defined). Then $\{(f_t, \zeta_t)\}_{\in[0,1]} \subset \mathcal{V}$, and ζ_1 is a root of f.

- For a.e. $f \in \mathcal{H}_{(d)}$ the set of $\zeta \in \mathbb{P}(\mathbb{C}^{n+1})$ such that ζ_t is defined for all $t \in [0,1]$ has full measure. Moreover, the boundary of this full measure set is a stratified set.
- Suppose η is a non-degenerate zero of $h \in \mathcal{H}_{(d)}$. Let $B(h,\eta)$ be the basin of η, i.e. the set of those $\zeta \in \mathbb{P}(\mathbb{C}^{n+1})$ such that the zero ζ of $\Pi_\zeta(h)$ continues to η for the homotopy h_t. ($B(h,\eta)$ is an open set.)

Then we can write $f_t = (1-t)\Pi_\zeta(f) + tf$.

Let ζ_t be the homotopy continuation of ζ along the path f_t (in case it is defined). Then $\{(f_t, \zeta_t)\}_{\in[0,1]} \subset \mathcal{V}$, and ζ_1 is a root of f.

- For a.e. $f \in \mathcal{H}_{(d)}$ the set of $\zeta \in \mathbb{P}(\mathbb{C}^{n+1})$ such that ζ_t is defined for all $t \in [0,1]$ has full measure. Moreover, the boundary of this full measure set is a stratified set.
- Suppose η is a non-degenerate zero of $h \in \mathcal{H}_{(d)}$. Let $B(h,\eta)$ be the basin of η, i.e. the set of those $\zeta \in \mathbb{P}(\mathbb{C}^{n+1})$ such that the zero ζ of $\Pi_\zeta(h)$ continues to η for the homotopy h_t. ($B(h,\eta)$ is an open set.)

Then we can write $f_t = (1-t)\Pi_\zeta(f) + tf$.

Let ζ_t be the homotopy continuation of ζ along the path f_t (in case it is defined). Then $\{(f_t, \zeta_t)\}_{\in[0,1]} \subset \mathcal{V}$, and ζ_1 is a root of f.

- For a.e. $f \in \mathcal{H}_{(d)}$ the set of $\zeta \in \mathbb{P}(\mathbb{C}^{n+1})$ such that ζ_t is defined for all $t \in [0,1]$ has full measure. Moreover, the boundary of this full measure set is a stratified set.
- Suppose η is a non-degenerate zero of $h \in \mathcal{H}_{(d)}$. Let $B(h,\eta)$ be the basin of η, i.e. the set of those $\zeta \in \mathbb{P}(\mathbb{C}^{n+1})$ such that the zero ζ of $\Pi_\zeta(h)$ continues to η for the homotopy h_t. ($B(h,\eta)$ is an open set.)

Then we can write $f_t = (1-t)\Pi_\zeta(f) + tf$.

Let ζ_t be the homotopy continuation of ζ along the path f_t (in case it is defined). Then $\{(f_t, \zeta_t)\}_{\in[0,1]} \subset \mathcal{V}$, and ζ_1 is a root of f.

- For a.e. $f \in \mathcal{H}_{(d)}$ the set of $\zeta \in \mathbb{P}(\mathbb{C}^{n+1})$ such that ζ_t is defined for all $t \in [0,1]$ has full measure. Moreover, the boundary of this full measure set is a stratified set.
- Suppose η is a non-degenerate zero of $h \in \mathcal{H}_{(d)}$. Let $B(h,\eta)$ be the basin of η, i.e. the set of those $\zeta \in \mathbb{P}(\mathbb{C}^{n+1})$ such that the zero ζ of $\Pi_\zeta(h)$ continues to η for the homotopy h_t. ($B(h,\eta)$ is an open set.)

The main result is

THEOREM

$$\mathbb{E}((I)) = \frac{\text{“}C\text{”}\, D^{3/2}}{(2\pi)^N} \int_{h\in\mathcal{H}_{(d)}} \Big[\sum_{\eta/\ h(\eta)=0} \frac{\mu^2(h,\eta)}{\|h\|^2} \Theta(h,\eta) \Big] e^{-\|h\|^2/2}\, dh,$$

where

$$\Theta(h,\eta) = \int_{\zeta\in B(h,\eta)} \frac{\|\Pi_\zeta(h)\|}{\|\Delta(\|\zeta\|^{-d_i})h(\zeta)\|^{2n-1}} e^{\|\Delta(\|\zeta\|^{-d_i})h(\zeta)\|^2/2}\, d\zeta.$$

Essentially nothing is known about the integrals.

The main result is

THEOREM

$$\mathbb{E}((I)) = \frac{\text{“}C\text{”}\, D^{3/2}}{(2\pi)^N} \int_{h\in\mathcal{H}_{(d)}} \Big[\sum_{\eta/\ h(\eta)=0} \frac{\mu^2(h,\eta)}{\|h\|^2} \Theta(h,\eta) \Big] e^{-\|h\|^2/2}\, dh,$$

where

$$\Theta(h,\eta) = \int_{\zeta\in B(h,\eta)} \frac{\|\Pi_\zeta(h)\|}{\|\Delta(\|\zeta\|^{-d_i})h(\zeta)\|^{2n-1}} e^{\|\Delta(\|\zeta\|^{-d_i})h(\zeta)\|^2/2}\, d\zeta.$$

Essentially nothing is known about the integrals.

The main result is

Theorem

$$\mathbb{E}((I)) = \frac{\text{“}C\text{''}\, D^{3/2}}{(2\pi)^N} \int_{h\in\mathcal{H}_{(d)}} \Big[\sum_{\eta/\ h(\eta)=0} \frac{\mu^2(h,\eta)}{\|h\|^2} \Theta(h,\eta) \Big] e^{-\|h\|^2/2}\, dh,$$

where

$$\Theta(h,\eta) = \int_{\zeta\in B(h,\eta)} \frac{\|\Pi_\zeta(h)\|}{\|\Delta(\|\zeta\|^{-d_i})h(\zeta)\|^{2n-1}} e^{\|\Delta(\|\zeta\|^{-d_i})h(\zeta)\|^2/2}\, d\zeta.$$

Essentially nothing is known about the integrals.

(A) Is $\mathbb{E}(\mathrm{I})$ finite for all or some n?

(B) Might $\mathbb{E}(\mathrm{I})$ even be polynomial in N for some range of dimensions and degrees?

(C) What are the basins like?

The integral

$$\frac{1}{(2\pi)^N}\int_{h\in\mathcal{H}_{(d)}}\sum_{\eta/\ h(\eta)=0}\frac{\mu^2(h,\eta)}{\|h\|^2}\cdot e^{-\|h\|^2/2}\,dh \leq \frac{e(n+1)}{2}\mathcal{D},$$

where $\mathcal{D}=d_1\cdots d_n$ is the Bézout number (Shub-Smale, Bürgisser-Cucker). So the question is how does the factor $\Theta(h,\eta)$ affect the integral.

(A) Is $\mathbb{E}(\mathrm{I})$ finite for all or some n?

(B) Might $\mathbb{E}(\mathrm{I})$ even be polynomial in N for some range of dimensions and degrees?

(C) What are the basins like?

The integral

$$\frac{1}{(2\pi)^N}\int_{h\in\mathcal{H}_{(d)}}\sum_{\eta/\ h(\eta)=0}\frac{\mu^2(h,\eta)}{\|h\|^2}\cdot e^{-\|h\|^2/2}\,dh \leq \frac{e(n+1)}{2}\mathcal{D},$$

where $\mathcal{D}=d_1\cdots d_n$ is the Bézout number (Shub-Smale, Bürgisser-Cucker). So the question is how does the factor $\Theta(h,\eta)$ affect the integral.

(A) Is $\mathbb{E}(I)$ finite for all or some n?

(B) Might $\mathbb{E}(I)$ even be polynomial in N for some range of dimensions and degrees?

(C) What are the basins like?

The integral

$$\frac{1}{(2\pi)^N}\int_{h\in\mathcal{H}_{(d)}}\sum_{\eta/\ h(\eta)=0}\frac{\mu^2(h,\eta)}{\|h\|^2}\cdot e^{-\|h\|^2/2}\,dh\leq\frac{e(n+1)}{2}\mathcal{D},$$

where $\mathcal{D}=d_1\cdots d_n$ is the Bézout number (Shub-Smale, Bürgisser-Cucker). So the question is how does the factor $\Theta(h,\eta)$ affect the integral.

(A) Is $\mathbb{E}(\mathsf{I})$ finite for all or some n?

(B) Might $\mathbb{E}(\mathsf{I})$ even be polynomial in N for some range of dimensions and degrees?

(C) What are the basins like?

The integral

$$\frac{1}{(2\pi)^N}\int_{h\in\mathcal{H}_{(d)}}\sum_{\eta/\ h(\eta)=0}\frac{\mu^2(h,\eta)}{\|h\|^2}\cdot e^{-\|h\|^2/2}\,dh \le \frac{e(n+1)}{2}\mathcal{D},$$

where $\mathcal{D}=d_1\cdots d_n$ is the Bézout number (Shub-Smale, Bürgisser-Cucker). So the question is how does the factor $\Theta(h,\eta)$ affect the integral.

(A) Is $\mathbb{E}(\mathsf{I})$ finite for all or some n?

(B) Might $\mathbb{E}(\mathsf{I})$ even be polynomial in N for some range of dimensions and degrees?

(C) What are the basins like?

The integral

$$\frac{1}{(2\pi)^N}\int_{h\in\mathcal{H}_{(d)}}\sum_{\eta/\ h(\eta)=0}\frac{\mu^2(h,\eta)}{\|h\|^2}\cdot e^{-\|h\|^2/2}\,dh \le \frac{e(n+1)}{2}\mathcal{D},$$

where $\mathcal{D} = d_1\cdots d_n$ is the Bézout number (Shub-Smale, Bürgisser-Cucker). So the question is how does the factor $\Theta(h,\eta)$ affect the integral.

Evaluate or estimate

$$\int_{\zeta\in\mathbb{P}(\mathbb{C}^{n+1})} \frac{1}{\|\Delta(\|\zeta\|^{-d_i})h(\zeta)\|^{2n-1}} \cdot e^{\frac{1}{2}\|\Delta(\|\zeta\|^{-d_i})h(\zeta)\|^2}\, d\zeta.$$

If this integral can be controlled, if the integral on the $\mathcal{D}$ basins are reasonably balanced, the factor of $\mathcal{D}$ in

$$\frac{1}{(2\pi)^N}\int_{h\in\mathcal{H}_{(d)}} \sum_{\eta/\ h(\eta)=0} \frac{\mu^2(h,\eta)}{\|h\|^2} \cdot e^{-\|h\|^2/2}\, dh \leq \frac{e(n+1)}{2}\mathcal{D},$$

may cancel!.

Evaluate or estimate

$$\int_{\zeta \in \mathbb{P}(\mathbb{C}^{n+1})} \frac{1}{\|\Delta(\|\zeta\|^{-d_i})h(\zeta)\|^{2n-1}} \cdot e^{\frac{1}{2}\|\Delta(\|\zeta\|^{-d_i})h(\zeta)\|^2}\, d\zeta.$$

If this integral can be controlled, if the integral on the $\mathcal{D}$ basins are reasonably balanced, the factor of $\mathcal{D}$ in

$$\frac{1}{(2\pi)^N} \int_{h \in \mathcal{H}_{(d)}} \sum_{\eta /\ h(\eta)=0} \frac{\mu^2(h,\eta)}{\|h\|^2} \cdot e^{-\|h\|^2/2}\, dh \leq \frac{e(n+1)}{2}\mathcal{D},$$

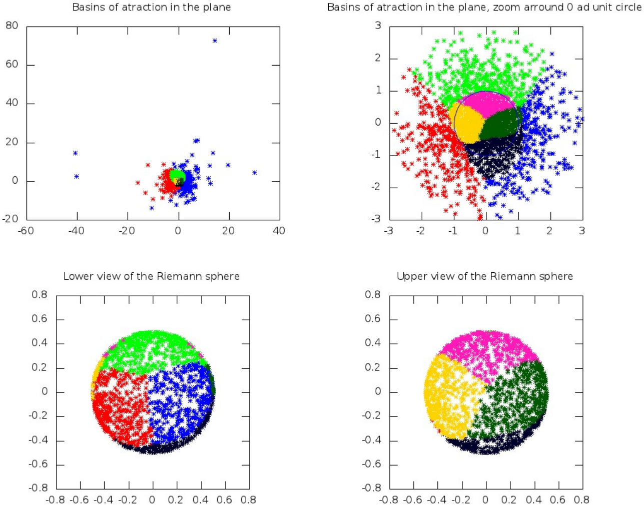
may cancel!.

Numerical experiments performed by Carlos Beltrán ($n = 1$ and $d = 7$) in the Altamira super-computer.

Roots in $\mathbb{C}$	$\mu(h, \cdot)$	$\Theta(h, \cdot)$	$\text{vol}(B(h, \cdot))$
$3.260883 - i1.658800$	1.712852	1.487095	0.140509π
$-2.357860 - i1.329208$	1.738380	1.728768	0.138576π
$-0.210068 + i1.868947$	1.608231	1.586398	0.144054π
$0.227994 - i0.782004$	1.909433	1.544021	0.125685π
$-0.044701 + i0.384342$	3.231554	3.152883	0.147277π
$-0.308283 + i0.049618$	3.183603	2.793696	0.152433π
$0.213950 - i0.068700$	2.948318	2.647258	0.151466π

TABLE: Degree 7 random polynomial.

Smale's algorithm reconsidered: Experiments



Comparison with roots of unity case: The errors for the root of unity case does not seem enough to explain the variation of $\Theta(h, \cdot)$. So it is likely that they are not all equal.

On the other hand, the ratios of the volumes of the basins of the random and roots of unity examples do seem to be of the same order of magnitude. So perhaps they are all equal?

- There appear to be 7 connected regions with a root in each. So there is some hope that this is true in general. That is there may generically be a root in each connected component of the basins and all these basins may have equal volume. This would be very interesting and would be very good start on understanding the integrals.

Comparison with roots of unity case: The errors for the root of unity case does not seem enough to explain the variation of $\Theta(h, \cdot)$. So it is likely that they are not all equal.

On the other hand, the ratios of the volumes of the basins of the random and roots of unity examples do seem to be of the same order of magnitude. So perhaps they are all equal?

- There appear to be 7 connected regions with a root in each. So there is some hope that this is true in general. That is there may generically be a root in each connected component of the basins and all these basins may have equal volume. This would be very interesting and would be very good start on understanding the integrals.

Comparison with roots of unity case: The errors for the root of unity case does not seem enough to explain the variation of $\Theta(h, \cdot)$. So it is likely that they are not all equal.

On the other hand, the ratios of the volumes of the basins of the random and roots of unity examples do seem to be of the same order of magnitude. So perhaps they are all equal?

- There appear to be 7 connected regions with a root in each. So there is some hope that this is true in general. That is there may generically be a root in each connected component of the basins and all these basins may have equal volume. This would be very interesting and would be very good start on understanding the integrals.

Comparison with roots of unity case: The errors for the root of unity case does not seem enough to explain the variation of $\Theta(h, \cdot)$. So it is likely that they are not all equal.
On the other hand, the ratios of the volumes of the basins of the random and roots of unity examples do seem to be of the same order of magnitude. So perhaps they are all equal?

- There appear to be 7 connected regions with a root in each. So there is some hope that this is true in general. That is there may generically be a root in each connected component of the basins and all these basins may have equal volume. This would be very interesting and would be very good start on understanding the integrals.

Also Santiago Laplagne is doing more experimental examples and it seems again that the volumes of the basins are equal.

More questions:

- The boundary of the basins are contained in a stratified set, the structure of which should be persistent by the isotopy theorem on the connected components of the complement of the critical values of the projection $(f, \zeta) \mapsto f$.

  Is there more than one component?

Also Santiago Laplagne is doing more experimental examples and it seems again that the volumes of the basins are equal.

More questions:

- The boundary of the basins are contained in a stratified set, the structure of which should be persistent by the isotopy theorem on the connected components of the complement of the critical values of the projection $(f, \zeta) \mapsto f$.

  Is there more than one component?

Also Santiago Laplagne is doing more experimental examples and it seems again that the volumes of the basins are equal.

More questions:

- The boundary of the basins are contained in a stratified set, the structure of which should be persistent by the isotopy theorem on the connected components of the complement of the critical values of the projection $(f,\zeta) \mapsto f$.

  Is there more than one component?

FIGURE: Mike and Jean-Pierre in FOCM Semester, Fields Institute 2009

.

GRACIAS MIKE!!