Smale's Alpha Theory — Verifying Newton's Method

21 January 2016

Chris Swierczewski
cswiercz@uw.edu

Department of Applied Mathematics
University of Washington

Seattle, Washington

Newton's Method

Let $f : \mathbb{C} \to \mathbb{C}$ be a polynomial. Define,

$$N(f, x_0) = \begin{cases} x_0 - f(x_0)/f'(x_0) & \text{if } f'(x_0) \neq 0, \\ x_0 & \text{if } f'(x_0) = 0. \end{cases}$$

Newton's Method

Let $f : \mathbb{C} \to \mathbb{C}$ be a polynomial. Define,

$$N(f, x_0) = \begin{cases} x_0 - f(x_0)/f'(x_0) & \text{if } f'(x_0) \neq 0, \\ x_0 & \text{if } f'(x_0) = 0. \end{cases}$$

$$N^k(f, x_0) = \underbrace{(N \circ N \circ \cdots \circ N)}_{k \text{ times}}(f, x_0)$$

Newton's Method

Let $f : \mathbb{C} \to \mathbb{C}$ be a polynomial. Define,

$$N(f, x_0) = \begin{cases} x_0 - f(x_0)/f'(x_0) & \text{if } f'(x_0) \neq 0, \\ x_0 & \text{if } f'(x_0) = 0. \end{cases}$$

$$N^k(f, x_0) = \underbrace{(N \circ N \circ \cdots \circ N)}_{k \text{ times}}(f, x_0)$$

Convergence

$$\lim_{k \to \infty} N^k(f, x_0) = \xi \text{ a root of } f$$

Example

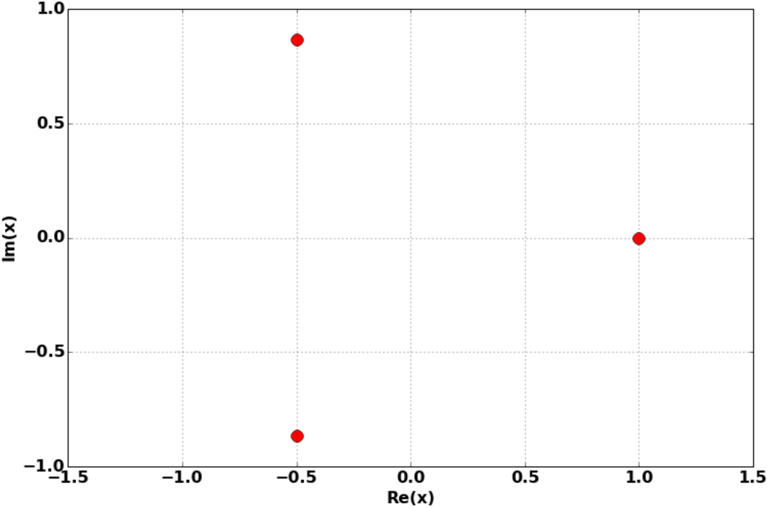
Roots of a Simple Cubic

$$f(x) = x^3 - 1$$

Actual roots:

$$\xi_k = e^{2\pi i k/3}, \text{for } k = 0, 1, 2.$$

Example: Actual Roots



Example: What is a “Good Guess”?

An initial guess “close” to the root should converge to that root:

Example: What is a “Good Guess”?

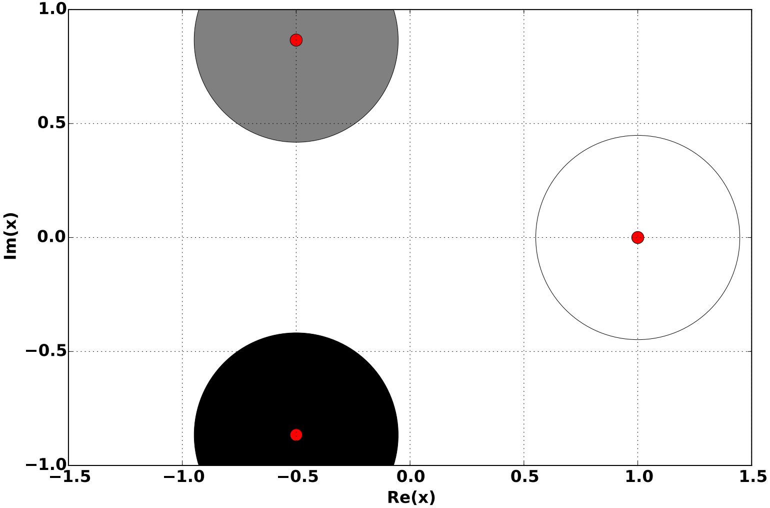
An initial guess “close” to the root should converge to that root:

Next Slide

- white region $\to$ guesses converging to ξ_0,
- grey region $\to$ guesses converging to ξ_1,
- black region $\to$ guesses converging to ξ_2,

(Apply Newton’s Method to each guess until we reach a root.)

Example: What is a “Good Guess”?



Example: What is a "Bad Guess"?

What if the initial Newton guess is further away?

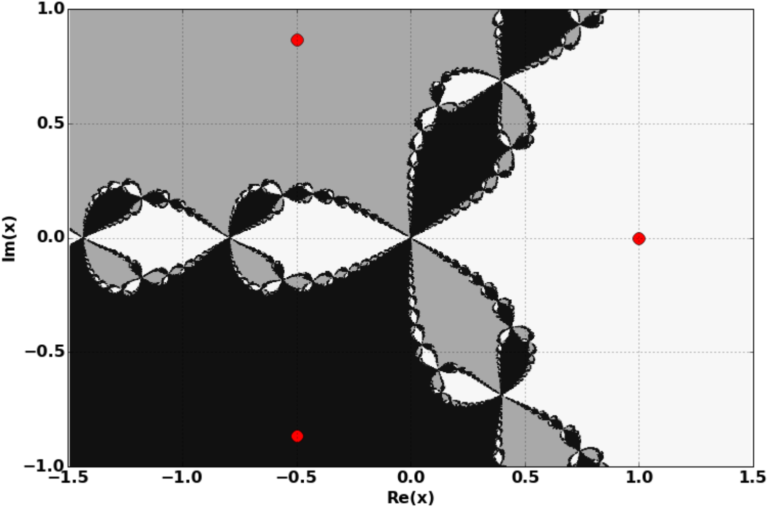
Example: What is a “Bad Guess”?

What if the initial Newton guess is further away?

Next Slide

- white region → guesses converging to ξ_0,
- grey region → guesses converging to ξ_1,
- black region → guesses converging to ξ_2,

Example: What is a “Bad Guess”?

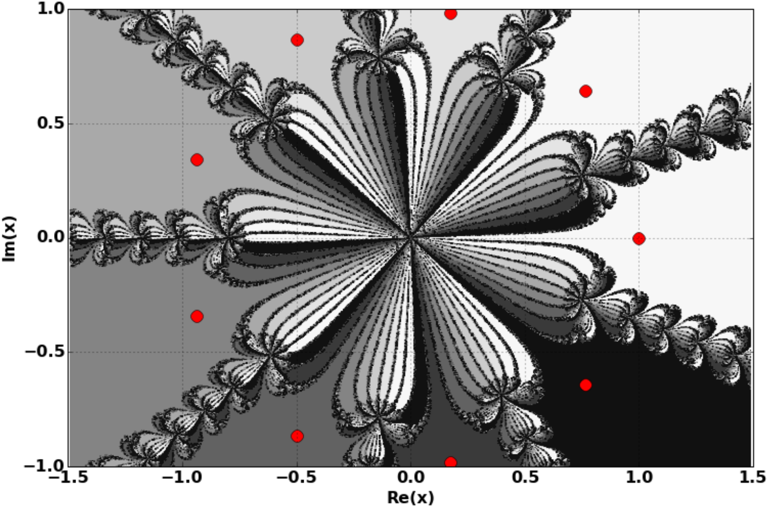


There are many terrible guesses.

There are many terrible guesses.

(Even guesses closer to some roots converge to other roots.)

Example: Roots of $f(x) = x^9 - 1$



Two Questions

Question #1

Can we ensure our guesses are far away from nasty fractal areas?

Two Questions

Question #1

Can we ensure our guesses are far away from nasty fractal areas?

Question #2

Given two guesses can we determine if they will converge to different roots? (Or the same root?)

Two Questions

Question #1

Can we ensure our guesses are far away from nasty fractal areas?

Question #2

Given two guesses can we determine if they will converge to different roots? (Or the same root?)

But...

...can we do these a priori? *(w/o knowing location of roots)*

Terminology

Let $f : \mathbb{C} \to \mathbb{C}$ be a polynomial.

- Define: $x \in \mathbb{C}$ is an **approximate solution** to f with **associated solution** $\xi \in \mathbb{C}$ if

$$\left| N^{(k)}(f, x) - \xi \right| \leq \left(\tfrac{1}{2}\right)^{2^k - 1} |x - \xi|$$

Terminology

Let $f : \mathbb{C} \to \mathbb{C}$ be a polynomial.

- Define: $x \in \mathbb{C}$ is an **approximate solution** to f with **associated solution** $\xi \in \mathbb{C}$ if

$$\left| N^{(k)}(f, x) - \xi \right| \leq \left(\tfrac{1}{2}\right)^{2^k - 1} |x - \xi|$$

- approximate solutions converge quadratically to their associated soltuions

Terminology

Let $f : \mathbb{C} \to \mathbb{C}$ be a polynomial.

- Define: $x \in \mathbb{C}$ is an **approximate solution** to f with **associated solution** $\xi \in \mathbb{C}$ if

$$\left| N^{(k)}(f, x) - \xi \right| \leq \left(\tfrac{1}{2}\right)^{2^k - 1} |x - \xi|$$

- approximate solutions converge quadratically to their associated soltuions
- *"x lies inside the quadratic convergence region of ξ"*

Quadratic Convergence Region

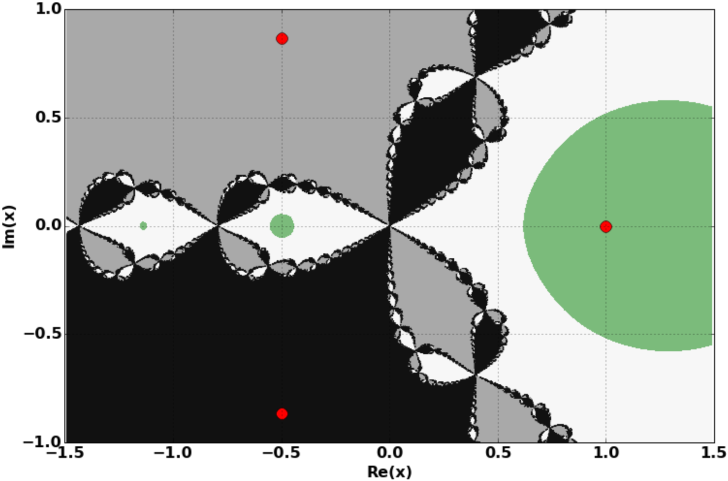


Figure: Quadratic convergence region of $\xi = 1$ for $f(x) = x^3 - 1$.

Question #1: Ensuring Quadratic Convergence

Determine if x is an **approximate solution**.

Question #1: Ensuring Quadratic Convergence

Determine if x is an **approximate solution**.

- **Problem**: the condition

$$\left| N^{(k)}(f, x) - \xi \right| \leq \left(\tfrac{1}{2}\right)^{2^k - 1} |x - \xi|$$

  requires knowing ξ!

Question #1: Ensuring Quadratic Convergence

Determine if x is an **approximate solution**.

- **Problem**: the condition

  $$\left|N^{(k)}(f,x) - \xi\right| \leq \left(\tfrac{1}{2}\right)^{2^k - 1} |x - \xi|$$

  requires knowing ξ!
- **Smale's Alpha Theory**: sufficient conditions for x to be in *some* quadratic convergence region

Question #1: Ensuring Quadratic Convergence

Smale's alpha theory: let

Question #1: Ensuring Quadratic Convergence

Smale's alpha theory: let

$$\alpha(f, x) := \beta(f, x)\gamma(f, x)$$

Question #1: Ensuring Quadratic Convergence

Smale's alpha theory: let

$$\alpha(f, x) := \beta(f, x)\gamma(f, x)$$
$$\beta(f, x) := |x - N(f, x)| = \left|f(x)/f'(x)\right|$$

Question #1: Ensuring Quadratic Convergence

Smale's alpha theory: let

$$\alpha(f,x) := \beta(f,x)\gamma(f,x)$$
$$\beta(f,x) := |x - N(f,x)| = \left|f(x)/f'(x)\right|$$
$$\gamma(f,x) := \max_{k\geq 2}\left|\frac{f^{(k)}(x)/f'(x)}{k!}\right|^{\frac{1}{k-1}}$$

Question #1: Converging to a Given Root

Smale Theorem #1

If $f : \mathbb{C} \to \mathbb{C}$ is a polynomial and $x \in \mathbb{C}$ such that

$$\alpha(f, x) \leq \frac{13 - 3\sqrt{17}}{4} \approx 0.157671$$

then x is an **approximate solution** to f.

Question #1: Converging to a Given Root

Smale Theorem #1

If $f : \mathbb{C} \to \mathbb{C}$ is a polynomial and $x \in \mathbb{C}$ such that

$$\alpha(f, x) \leq \frac{13 - 3\sqrt{17}}{4} \approx 0.157671$$

then x is an **approximate solution** to f.
Additionally,

$$|x - \xi| \leq 2\beta(f, x)$$

where ξ is the **associated solution** to x.

Alpha Region

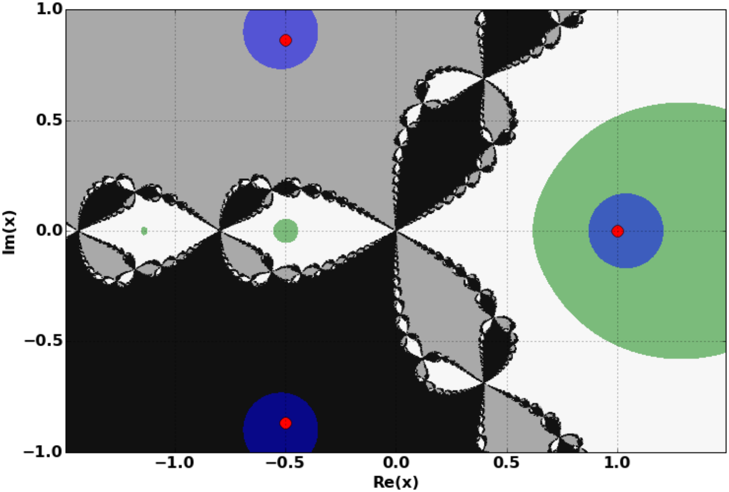


Figure: Region where $\alpha(f, x) < 0.157\ldots$ for $f(x) = x^3 - 1$.

Alpha Region: Discussion

- Pros

Alpha Region: Discussion

- Pros
  - quadratic convergence condition *without* knowing roots,

Alpha Region: Discussion

- Pros
  - quadratic convergence condition *without* knowing roots,
  - approximates how far away you are

$$|x - \xi| < 2\beta(f, x),$$

Alpha Region: Discussion

- Pros
  - quadratic convergence condition *without* knowing roots,
  - approximates how far away you are

$$|x - \xi| < 2\beta(f, x),$$

- Cons

Alpha Region: Discussion

- Pros
  - quadratic convergence condition *without* knowing roots,
  - approximates how far away you are

$$|x - \xi| < 2\beta(f, x),$$

- Cons
  - doesn't say which root (but β gives us an idea)

Alpha Region: Discussion

- Pros
  - quadratic convergence condition *without* knowing roots,
  - approximates how far away you are

$$|x - \xi| < 2\beta(f, x),$$

- Cons
  - doesn't say which root (but β gives us an idea)
  - alpha region much smaller than quad. conv. region

Question #2: Converging to Distinct Roots

Ensure two **approximate solutions** x_1, x_2 have distinct **associated solutions** ξ_1, ξ_2.

Question #2: Converging to Distinct Roots

Ensure two **approximate solutions** x_1, x_2 have distinct **associated solutions** ξ_1, ξ_2.

Smale Theorem #2

If

$$|x_1 - x_2| > 2\Big(\beta(f, x_1) + \beta(f, x_2)\Big)$$

then

$$\xi_1 \neq \xi_2$$

Question #2: Converging to Distinct Roots

Ensure two **approximate solutions** x_1, x_2 have distinct **associated solutions** ξ_1, ξ_2.

Smale Theorem #2

If

$$|x_1 - x_2| > 2\Big(\beta(f, x_1) + \beta(f, x_2)\Big)$$

then

$$\xi_1 \neq \xi_2$$

- Follows from Smale Theorem #1:

$$|x - \xi| \leq 2\beta(f, x)$$

Question #2: Converging to Distinct Roots

Ensure two **approximate solutions** x_1, x_2 have distinct **associated solutions** ξ_1, ξ_2.

Smale Theorem #2

If

$$|x_1 - x_2| > 2\Big(\beta(f, x_1) + \beta(f, x_2)\Big)$$

then

$$\xi_1 \neq \xi_2$$

- Follows from Smale Theorem #1:

$$|x - \xi| \leq 2\beta(f, x)$$

- **Homework:** prove this

Application: Analytic Continuation

Let $f(x, y) = y^3 - x$.

Application: Analytic Continuation

Let $f(x, y) = y^3 - x$.

- function of y with x as a parameter,

Application: Analytic Continuation

Let $f(x, y) = y^3 - x$.

- function of y with x as a parameter,
- given an x we can find roots y_1, y_2, y_3 to $f(x, y) = 0$,

Application: Analytic Continuation

Let $f(x, y) = y^3 - x$.

- function of y with x as a parameter,
- given an x we can find roots y_1, y_2, y_3 to $f(x, y) = 0$,
- **fact:** polynomial roots vary continuously as function of coefficients

Application: Analytic Continuation

Let $f(x, y) = y^3 - x$.

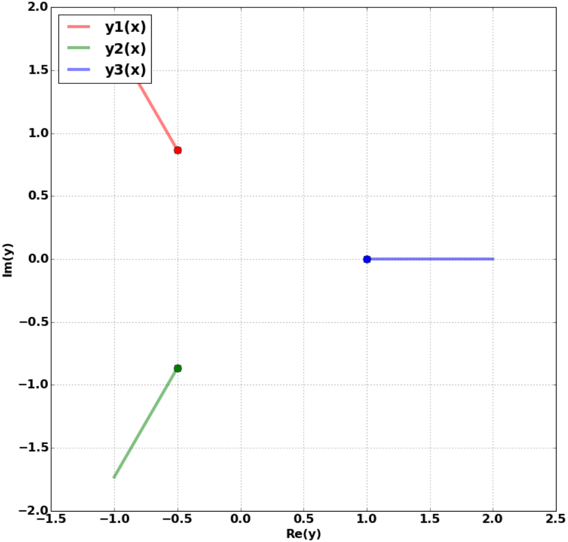
- function of y with x as a parameter,
- given an x we can find roots y_1, y_2, y_3 to $f(x, y) = 0$,
- **fact:** polynomial roots vary continuously as function of coefficients

$$\text{roots “above” } x : \qquad y_1(x),\ y_2(x),\ y_3(x)$$

Example: $f(x, y) = y^3 - x$

Let x_i range from $x_0 = 1$ to $x_N = 8$:

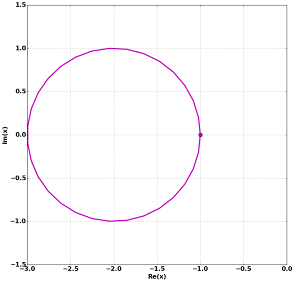
$$y_1(1) = 1$$
$$y_2(1) = e^{2\pi i/3}$$
$$y_3(1) = e^{4\pi i/3}$$
$$\vdots$$
$$y_1(8) = 2$$
$$y_2(8) = 2e^{2\pi i/3}$$
$$y_3(8) = 2e^{4\pi i/3}$$

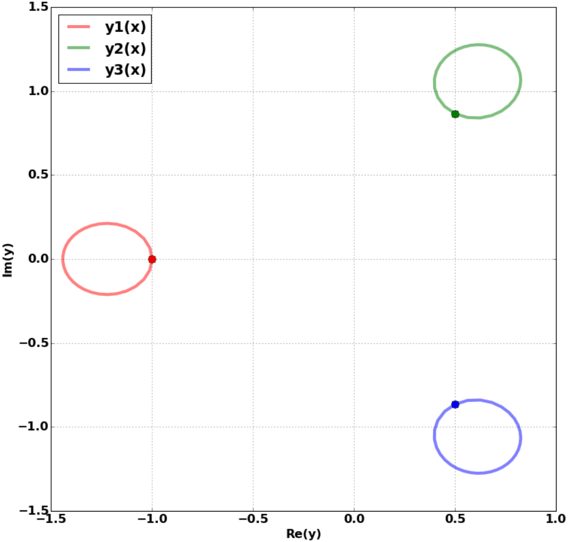


Example: $f(x, y) = y^3 - x$

Let x_i range along the complex circle

$$x(t) = e^{2\pi i t} - 2$$
$$t \in [0, 1]$$





Computing These y-Paths

Let $y_1^{(i)}, y_2^{(i)}, y_3^{(i)}$ be the y-roots computed above x_i:

$$f(x_i, y) = 0.$$

Computing These y-Paths

Let $y_1^{(i)}, y_2^{(i)}, y_3^{(i)}$ be the y-roots computed above x_i:

$$f(x_i, y) = 0.$$

Goal: compute corresponding y-roots

$$y_1^{(i+1)}, y_2^{(i+1)}, y_3^{(i+1)}$$

above x_{i+1}: *solution to* $f(x_{i+1}, y) = 0$.

Computing These y-Paths

Let $y_1^{(i)}, y_2^{(i)}, y_3^{(i)}$ be the y-roots computed above x_i:

$$f(x_i, y) = 0.$$

Goal: compute corresponding y-roots

$$y_1^{(i+1)}, y_2^{(i+1)}, y_3^{(i+1)}$$

above x_{i+1}: *solution to* $f(x_{i+1}, y) = 0$.

- **Idea:** use $y_1^{(i)}$ as Newton iteration guess in

  $$g(y) := f(x_{i+1}, y) = 0$$

  to get $y_1^{(i+1)}$

Computing These y-Paths

Let $y_1^{(i)}, y_2^{(i)}, y_3^{(i)}$ be the y-roots computed above x_i:

$$f(x_i, y) = 0.$$

Goal: compute corresponding y-roots

$$y_1^{(i+1)}, y_2^{(i+1)}, y_3^{(i+1)}$$

above x_{i+1}: *solution to* $f(x_{i+1}, y) = 0$.

- **Idea:** use $y_1^{(i)}$ as Newton iteration guess in

  $$g(y) := f(x_{i+1}, y) = 0$$

  to get $y_1^{(i+1)}$
- **Important:** must satisfy

  $$y_1(x_i) = y_1^{(i)} \quad \text{and} \quad y_1(x_{i+1}) = y_1^{(i+1)}$$

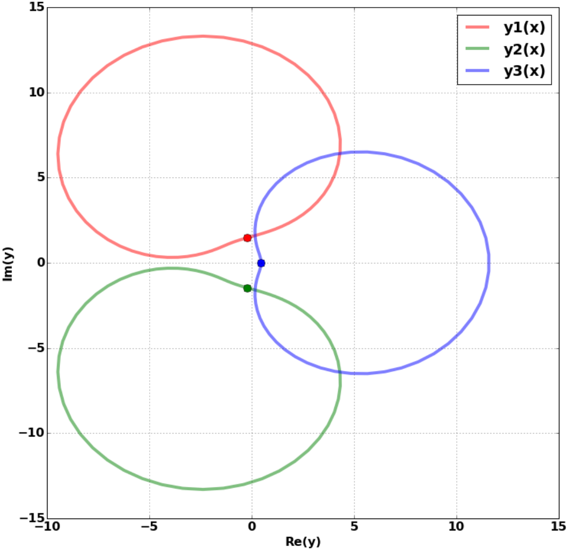
Example: $f(x, y) = y^3 - 2x^3y + x^7$

Let x_i range along the complex circle

$$x(t) = e^{2\pi it} - 2$$
$$t \in [0, 1]$$

64 different x-values

small Δx means $y^{(i)}$ are good guesses for $y^{(i+1)}$



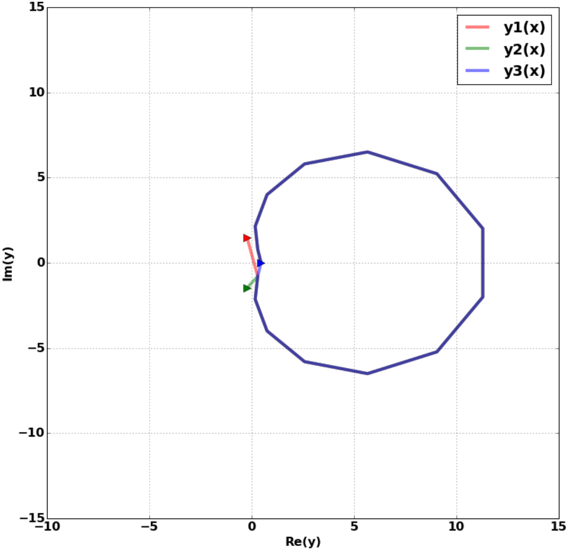
Example: $f(x, y) = y^3 - 2x^3y + x^7$

Let x_i range along the complex circle

$$x(t) = e^{2\pi it} - 2$$
$$t \in [0, 1]$$

16 different x-values

Something wrong happened. (Too large Δx.)



Problem

"Just take the Δx *steps to be really small."*

Problem

"Just take the Δx steps to be really small."

- What does "small" mean?

Problem

"Just take the Δx steps to be really small."

- What does "small" mean?
- Heuristics in programming should be avoided. (Understatement of the year.)

Problem

"Just take the Δx steps to be really small."

- What does "small" mean?
- Heuristics in programming should be avoided. (Understatement of the year.)
- Too small $\rightarrow$ computationally inefficient.

Problem

"Just take the Δx steps to be really small."

- What does "small" mean?
- Heuristics in programming should be avoided. (Understatement of the year.)
- Too small $\rightarrow$ computationally inefficient.
- **Smale:** determine when Δx is small enough such that

Problem

"Just take the Δx steps to be really small."

- What does "small" mean?
- Heuristics in programming should be avoided. (Understatement of the year.)
- Too small $\rightarrow$ computationally inefficient.
- **Smale:** determine when Δx is small enough such that
  - each $y_j^{(i)}$ will converge under Newton **(Use Smale Theorem #1)**

Problem

"Just take the Δx steps to be really small."

- What does "small" mean?
- Heuristics in programming should be avoided. (Understatement of the year.)
- Too small $\rightarrow$ computationally inefficient.
- **Smale:** determine when Δx is small enough such that
  - each $y_j^{(i)}$ will converge under Newton **(Use Smale Theorem #1)**
  - each $y_j^{(i)}$ will converge to *distinct* $y_j^{(i+1)}$ **(Use Smale Theorem #2)**

Algorithm: Analytic Continutation

Algorithm: `analytic`$(f, x_i, x_{i+1}, y^{(i)})$

Algorithm: Analytic Continutation

Algorithm: `analytic`$(f, x_i, x_{i+1}, y^{(i)})$
Input:

- polynomial $f = f(x, y)$,
- x-points x_i and x_{i+1},
- ordered y-roots $y^{(i)} = \left(y_1^{(i)}, \ldots, y_d^{(i)}\right)$ above x_i.

Algorithm: $\texttt{analytic}(f, x_i, x_{i+1}, y^{(i)})$
Input:

- polynomial $f = f(x, y)$,
- x-points x_i and x_{i+1},
- ordered y-roots $y^{(i)} = \left(y_1^{(i)}, \ldots, y_d^{(i)}\right)$ above x_i.

Output: ordered y-roots $y^{i+1} = \left(y_1^{(i+1)}, \ldots, y_d^{(i+1)}\right)$ above x_{i+1}.

- such that $y_j^{(i)} \to y_j^{(i+1)}$ (same position j)

Algorithm: Analytic Continutation

Algorithm: `analytic`$\big(f,\, x_i,\, x_{i+1},\, y^{(i)}\big)$

1. Check that each $y_j^{(i)}$ is an **approximate solution** to

$$g(y) := f(x_{i+1}, y) = 0$$

using $\alpha(g, y_j^{(i)}) < 0.157\ldots$ If any are not, **refine step**:

Algorithm: Analytic Contintutation

Algorithm: `analytic`$\left(f,\, x_i,\, x_{i+1},\, y^{(i)}\right)$

1. Check that each $y_j^{(i)}$ is an **approximate solution** to

$$g(y) := f(x_{i+1}, y) = 0$$

   using $\alpha(g, y_j^{(i)}) < 0.157\ldots$ If any are not, **refine step**:
   - $x_{i+1/2} \leftarrow (x_i + x_{i+1})/2$

Algorithm: Analytic Continutation

Algorithm: $\texttt{analytic}\big(f,\, x_i,\, x_{i+1},\, y^{(i)}\big)$

1. Check that each $y_j^{(i)}$ is an **approximate solution** to

$$g(y) := f(x_{i+1}, y) = 0$$

   using $\alpha(g, y_j^{(i)}) < 0.157\ldots$ If any are not, **refine step**:
   - $x_{i+1/2} \leftarrow (x_i + x_{i+1})/2$
   - $y^{(i+1/2)} \leftarrow \texttt{analytic}\big(f,\, x_i,\, x_{i+1/2},\, y^{(i)}\big)$

Algorithm: Analytic Continutation

Algorithm: $\texttt{analytic}\big(f,\, x_i,\, x_{i+1},\, y^{(i)}\big)$

1. Check that each $y_j^{(i)}$ is an **approximate solution** to

$$g(y) := f(x_{i+1}, y) = 0$$

   using $\alpha(g, y_j^{(i)}) < 0.157\ldots$ If any are not, **refine step**:

   - $x_{i+1/2} \leftarrow (x_i + x_{i+1})/2$
   - $y^{(i+1/2)} \leftarrow \texttt{analytic}\big(f,\, x_i,\, x_{i+1/2},\, y^{(i)}\big)$
   - $y^{(i+1)} \leftarrow \texttt{analytic}\big(f,\, x_{i+1/2},\, x_{i+1},\, y^{(i+1/2)}\big)$

Algorithm: Analytic Continutation

Algorithm: `analytic`$\left(f,\, x_i,\, x_{i+1},\, y^{(i)}\right)$

1. Check that each $y_j^{(i)}$ is an **approximate solution** to

$$g(y) := f(x_{i+1}, y) = 0$$

   using $\alpha(g, y_j^{(i)}) < 0.157\ldots$ If any are not, **refine step**:
   - $x_{i+1/2} \leftarrow (x_i + x_{i+1})/2$
   - $y^{(i+1/2)} \leftarrow$ `analytic`$\left(f,\, x_i,\, x_{i+1/2},\, y^{(i)}\right)$
   - $y^{(i+1)} \leftarrow$ `analytic`$\left(f,\, x_{i+1/2},\, x_{i+1},\, y^{(i+1/2)}\right)$

2. Determine if all **approximate solutions** $y_j^{(i)}$ will converge to distinct **associated solutions** $y_j^{(i+1)}$:

$$|y_j^{(i)} - y_k^{(i)}| > 2\left(\beta(f, y_j^{(i)}) + \beta(f, y_k^{(i)})\right), \quad \forall j, k = 1, \ldots, d.$$

   If any are not, **refine step**.

Algorithm: Analytic Continutation

Algorithm: $\texttt{analytic}\big(f,\, x_i,\, x_{i+1},\, y^{(i)}\big)$

1. Check that each $y_j^{(i)}$ is an **approximate solution** to

$$g(y) := f(x_{i+1}, y) = 0$$

   using $\alpha(g, y_j^{(i)}) < 0.157\ldots$ If any are not, **refine step**:
   - $x_{i+1/2} \leftarrow (x_i + x_{i+1})/2$
   - $y^{(i+1/2)} \leftarrow \texttt{analytic}\big(f,\, x_i,\, x_{i+1/2},\, y^{(i)}\big)$
   - $y^{(i+1)} \leftarrow \texttt{analytic}\big(f,\, x_{i+1/2},\, x_{i+1},\, y^{(i+1/2)}\big)$

2. Determine if all **approximate solutions** $y_j^{(i)}$ will converge to distinct **associated solutions** $y_j^{(i+1)}$:

$$|y_j^{(i)} - y_k^{(i)}| > 2\big(\beta(f, y_j^{(i)}) + \beta(f, y_k^{(i)})\big), \quad \forall j, k = 1, \ldots, d.$$

   If any are not, **refine step**.

3. Finally, Newton iterate each $y_j^{(i)}$ to $y_j^{(i+1)}$ and return.

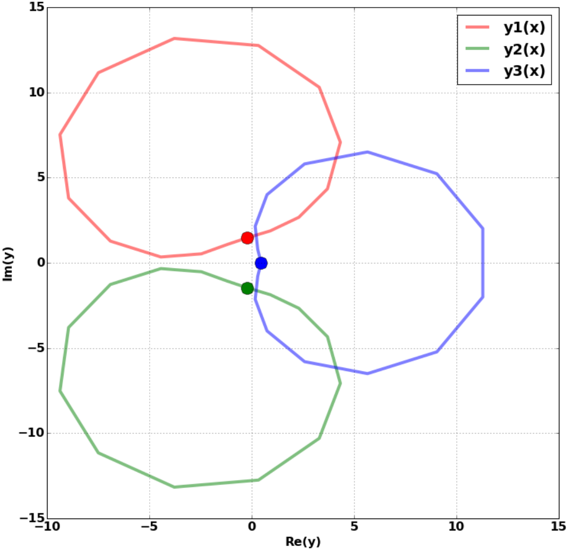
Example: $f(x, y) = y^3 - 2x^3y + x^7$

Let x_i range along the complex circle

$$x(t) = e^{2\pi it} - 2$$
$$t \in [0, 1]$$

16 different x-values

Smale guarantees we converge to the correct roots.



Example: $f(x, y) = y^3 - 2x^3y + x^7$

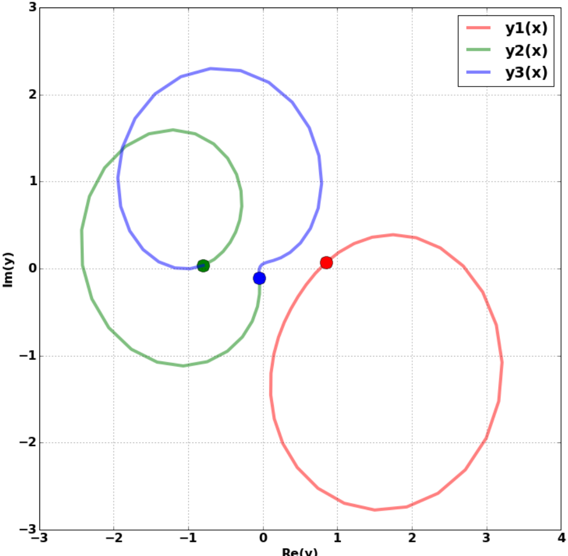
Let x_i range along the complex circle

$$x(t) = \frac{1}{2}e^{2\pi it} + \beta$$
$$t \in [0, 1]$$

where

$$\beta \approx -0.8369 - 0.6081j.$$

(Branch point of curve.)



Final Remarks

- Works for square systems of polynomials $f : \mathbb{C}^n \to \mathbb{C}^n$.

Final Remarks

- Works for square systems of polynomials $f : \mathbb{C}^n \to \mathbb{C}^n$.
  - Derivative $f' \to$ Jacobian Df.

Final Remarks

- Works for square systems of polynomials $f : \mathbb{C}^n \to \mathbb{C}^n$.
  - Derivative $f' \to$ Jacobian Df.
- Even works for smooth functions $f : \mathbb{C}^n \to \mathbb{C}^n$.

Final Remarks

- Works for square systems of polynomials $f : \mathbb{C}^n \to \mathbb{C}^n$.
  - Derivative $f' \to$ Jacobian Df.
- Even works for smooth functions $f : \mathbb{C}^n \to \mathbb{C}^n$.
  - Definition of $\gamma(f, x)$: "max $\to$ sup".

Final Remarks

- Works for square systems of polynomials $f : \mathbb{C}^n \to \mathbb{C}^n$.
  - Derivative $f' \to$ Jacobian Df.
- Even works for smooth functions $f : \mathbb{C}^n \to \mathbb{C}^n$.
  - Definition of $\gamma(f, x)$: "max $\to$ sup".
  - Some simpler bounds on γ: results in much smaller α-region.

Thank you

Talk and code available at `www.cswiercz.info`. GitHub repo at `github.com/cswiercz/smale`.

References

- S. Smale, "Newton's method estimates from data at one point", Springer New York, 1986.
- J. D. Hauenstein, F. Sottile, "AlphaCertified: certifying solutions to polynomial systems", ACM Trans. Math. Softw., vol. 38, no. 4, pp. 1-20, 2012.