A Polyhedral Homotopy Continuation Method for Computing All Solutions of a Polynomial System of Equations in Complex Variables

Masakazu Kojima, **Yang** Dai,
Sunyoung Kim **and Akiko** Takeda

October 2001

“A polyhedral homotopy continuation method for computing all solutions of a polynomial system of equations in complex variables”

Masakazu Kojima, **Yang** Dai, **Sunyoung** Kim **and Akiko** Takeda

Contents

1. A polynomial equation system

$$f(x) = 0,$$

where

$$\begin{aligned} x &= (x_1, x_2, \ldots, x_n) \in \mathbb{C}^n, \\ f(x) &= (f_1(x), f_2(x), \ldots, f_n(x)), \\ f_j(x) &= \text{a polynomial in } n \text{ complex variables } x_1, x_2, \ldots, x_n. \end{aligned}$$

Example

$$\begin{aligned} & n = 3,\ x = (x_1, x_2, x_3),\ f(x) = (f_1(x), f_2(x), f_3(x)), \\ & f_1(x_1, x_2, x_3) = x_1^2 - (2.1 + i)x_1 x_2 x_3^2 + 8.5, \\ & f_2(x_1, x_2, x_3) = 1.5x_1^2 x_2 - x_1^2 x_2^2 x_3 - 1.6, \\ & f_3(x_1, x_2, x_3) = (3.6 + i)x_1 x_2^3 + 4.3x_1 x_2^2 x_3^2. \end{aligned}$$

Find all isolated solutions in $\mathbb{C}^n$.

Find all isolated solutions in $\mathbb{C}^n$.

- **A Fundamental problem in numerical mathematics.**
- **Various engineering applications.**
- **Global optimization:**
  If we compute all the Karush-Kuhn-Tucker stationary solutions, the we can pick up a global optimal solution among them. Here we assume that both objective and constraint functions are polynomials.

2. Typical benchmark test problem — 1 Economic-n polynomial:

$$
\begin{aligned}
&(x_1 + x_1x_2 + x_2x_3 + \cdots + x_{n-2}x_{n-1})x_n - 1 = 0\\
&(x_2 + x_1x_3 + \cdots + x_{n-3}x_{n-1})x_n - 2 = 0\\
&\cdots\\
&(x_{n-2} + x_1x_{n-1})x_n - (n-2) = 0\\
&x_{n-1}x_n - (n-1) = 0\\
&x_1 + x_2 + \cdots + x_{n-1} + 1 = 0.
\end{aligned}
$$

n	♯ of isolated solutions
10	256
11	512
12	1024
13	2048
14	4096
	$\cdots$
n	2^{n-2}

Typical benchmark test problem — 2: Cyclic-n polynomial

$$\begin{aligned}
f_1(x) &= x_1 + x_2 + \cdots + x_n, \\
f_2(x) &= x_1x_2 + x_2x_3 + \cdots + x_nx_1, \\
&\dots \\
f_{n-2}(x) &= x_1x_2\dots x_{n-2} + x_2x_3\dots x_{n-1} + \cdots + x_nx_1\cdots x_{n-2}, \\
f_{n-1}(x) &= x_1x_2\dots x_{n-1} + x_2x_3\dots x_n + \cdots + x_nx_1\cdots x_{n-1}, \\
f_n(x) &= x_1x_2\dots x_{n-1}x_n - 1.
\end{aligned}$$

n	♯ of nonsingular isolated solutions	♯/n	♯/($2n$)
10	34,940	3,494	1,747
11	184,756	16,796	8,398
12	367,488	30,624	15,312
13	?		
	…		

We can reduce the solutions to be computed to $1/n$ (or $1/(2n)$) using certain symmetries.

Enormous computational power for solving large scale problems

$\Rightarrow$ Parallel computation

3. Rough sketch of the polyhedral homotopy (continuation) method

$\Longrightarrow$

- Currently the most powerful and practical method for computing all solutions of a polynomial equation system.
- Suitable for parallel computation;
  all solutions can be computed independently in parallel.

Phase 1. Construct a family of homotopy functions.

- Branch-and-bound methods.
- Large scale linear programs.
- Nonlinear combinatorial optimization problems.

Phase 2. Trace homotopy paths by predictor-corrector methods.

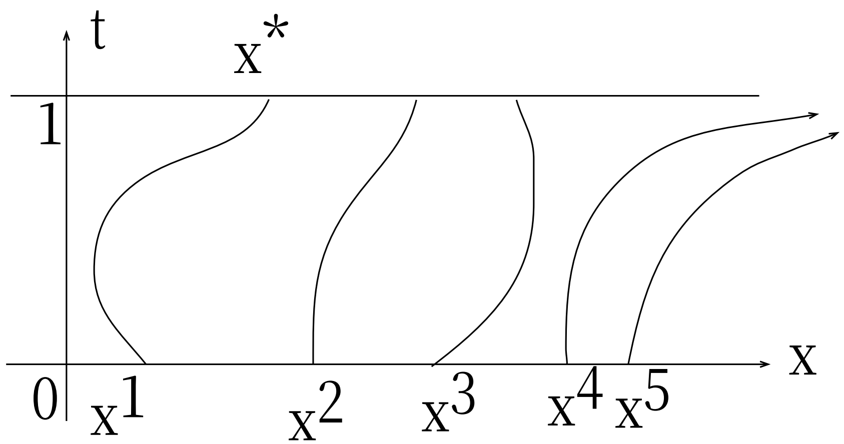
- Highly nonlinear homotopy paths that require sophisticated step length control techniques.

Phase 3. Verify that all isolated solutions are computed.

- **The number of solutions is unknown in general.**
- **Approximate solutions are computed but exact solutions are never computed.**

4. Basic ideas of Phases 1 and 2.

Phase 1. Let x^* be a solution of $f(x) = 0$. We construct a homotopy equation system $h(x,t) = 0$ such that (i) all solutions of the initial system $h(x,0) = 0$ are known, (ii) $h(x,1) = f(x)$ for every $x \in \mathbb{C}^n$; hence if $h(x,1) = 0$, x is a solution of $f(x) = 0$, and (iii) x^* is connected to a solution x^1 of $h(x,0) = 0$ through the solution path of $h(x,t) = 0$.



Phase 2. Starting from each known solution of the initial system $h(x,0) = 0$, we trace the solution path of $h(x,t) = 0$ till t attains 1 by a predictor-corrector method to obtain a solution of $f(x) = 0$.

- **This idea is common for the traditional linear homotopy method and the polyhedral homotopy method.**
- **Some solution paths diverge as $t \to \infty$; tracing such paths are useless.**
- **The number of useless divergent paths is much less in the polyhedral homotopy method than in the traditional homotopy method.**
- **Multiple homotopy functions are employed in polyhedral homotopy methods while a common single h is employed for all solutions of $f(x) = 0$ in the traditional linear homotopy method.**

Notation

For $\forall a \in \mathrm{Z}_+^n \equiv \{(a_1, , \ldots, a_n) \geq 0 : a_j$ **is integer**$\}$ **and** $\forall x \in \mathrm{C}^n$, **define**

$$x^{a} = x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}.$$

Then we can write $\forall\ f_j(x)$ **of a poly. system** $f(x) = (f_1(x), \ldots, f_n(x))$ **as**

$$f_j(x) = \sum_{a \in \mathcal{A}_j} c_j(a) x^{a},$$

where $c_j(a) \in \mathrm{C}$ $(a \in \mathcal{A}_j)$ **and** $\mathcal{A}_j$ **a finite subset of** Z_+^n $(j = 1, \ldots, n)$.
We call $\mathcal{A}_j$ **the support of** $f_j(x)$.

For example, $n = 3$,

$$\begin{aligned} f_3(x_1, x_2, x_3) &= (3.6 + i)x_1 x_2^3 + 4.3 x_1 x_2^2 x_3^2 \\ &= c_3((1,3,0))x^{(1,3,0)} + c_3((1,2,2))x^{(1,2,2)} \\ &= \textstyle\sum_{a \in \mathcal{A}_3} c_3(a) x^{a} \end{aligned}$$

where

$$\begin{aligned} \mathcal{A}_3 &= \{(1,3,0),\ (1,2,2)\}, \\ c_3((1,3,0)) &= 3.6 + i,\ \ c_3((1,2,2)) = 4.3. \end{aligned}$$

5. Phase 1. Construction of a family of homotopy functions

$$h^k(x,t) \in \mathrm{C}^n,\ (x,t) \in \mathrm{C}^n \times [0,1]\ (k = 1,2,\ldots,\ell).$$

(a) For $\forall k = 1,2,\ldots,\ell$,

$$h_j^k(x,t) = \sum_{a \in \mathcal{A}_j} c_j(x) x^{a} t^{\rho_j^k(a)}\ (j = 1,2,\ldots,n),$$

where exactly two of $\{\rho_j^k(a) : a \in \mathcal{A}_j\}$ are zero and all others are positive $(j = 1,2,\ldots,n)$.

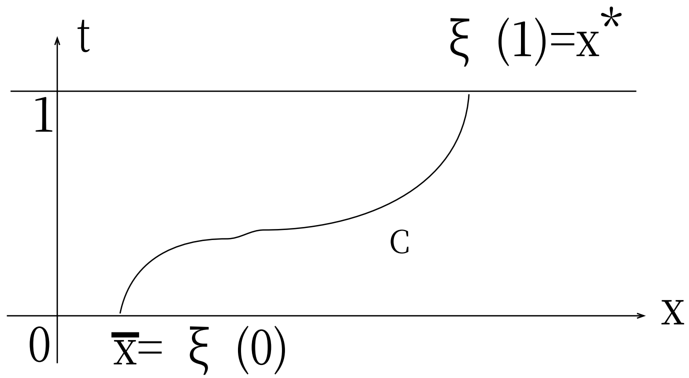
$\Downarrow$

Each component $h_j^k(x,0)$ of $h^k(x,0)$ consists of two terms; hence the starting equation system turn out to be a binomial equation system

$$h_j^k(x,0) \equiv c_j(a^j)x^{a^j} + c_j(\tilde{a}^j)x^{\tilde{a}^j} = 0 \ (j = 1, 2, \ldots, n).$$

$\Rightarrow$ We can easily compute all solutions by linear algebra (or elimination technique).

(b) $\forall$ **sol.** $\boldsymbol{x^*}$ **of** $\boldsymbol{f(x) = 0}$, $\exists \boldsymbol{k}$, $\exists$ **sol.** $\tilde{\boldsymbol{x}}$ **of** $\boldsymbol{h^k(x, 0) = 0}$; $\tilde{\boldsymbol{x}}$ **is connected to** $\boldsymbol{x^*}$ **through a sol. path** $\boldsymbol{C = \{(\xi(t), t) : t \in \times[0, 1]\}}$ **of** $\boldsymbol{h^k(x, t) = 0}$.



How do we construct such a family of homotopy functions?

Choose $\omega_j(a) \in \mathrm{R}$ (randomly) ($a \in \mathcal{A}_j,\ j = 1, 2, \ldots, n$).

Let $(\alpha, \beta) = (\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n) \in \mathrm{R}^{2n}$ whose value we will determine later. Define

$$\begin{aligned} h^{\alpha,\beta}(x,t) &= (h_1^{\alpha,\beta}(x,t), h_2^{\alpha,\beta}(x,t), \ldots, h_n^{\alpha,\beta}(x,t)), \\ h_j^{\alpha,\beta}(x,t) &= \textstyle\sum_{a \in \mathcal{A}_j} c_j(a) x^{a} t^{\rho_j(a,\alpha,\beta)} \ (j = 1, 2, \ldots, n), \end{aligned}$$

where

$$\rho_j(a, \alpha, \beta) = \langle a, \alpha \rangle + \omega_j(a) - \beta_j \geq 0 \ (a \in \mathcal{A}_j,\ j = 1, \ldots, n), \quad (1)$$

$$\text{for } \forall j, \quad \text{exactly } 2 \text{ of } \{\langle a, \alpha \rangle + \omega_j(a) - \beta_j : a \in \mathcal{A}_j\} \text{ are } 0. \quad (2)$$

Nondegeneracy assumpt.: $\forall$ sol. $(\alpha, \beta) \in \mathrm{R}^{2n}$ of (1), at most $2n$ equalities.

In the polyhedral homotopy theory, it is known that

- $\Gamma \equiv \{(\alpha, \beta) :$ **solutions of (1) and (2)**$\}$ **is finite.**
- **The family** $h^{\alpha,\beta}(x, t)$ $((\alpha, \beta) \in \Gamma)$ **satisfy the desired properties we have mentioned;**

(a) **The starting system** $h^{\alpha,\beta}(x, 0) = 0$ **is binomial** $((\alpha, \beta) \in \Gamma)$.

(b) $\forall$ **sol.** x^* **of** $f(x) = 0$, $\exists(\alpha, \beta) \in \Gamma$, $\exists$ **sol.** $\tilde{x}$ **of** $h^{\alpha,\beta}(x, 0) = 0$; $\tilde{x}$ **is connected to** x^* **through a sol. path** $C = \{(\xi(t), t) : t \in \times[0, 1]\}$ **of** $h^{\alpha,\beta}(x, t) = 0$.

Therefore "computing all solutions Γ of the linear ineq. system (1) with the comb. cond. (2)" forms an important subprob. in Phase 1.

$$\Uparrow$$

- **Implicit enum. tech. (or b-and-b. methods) used in optimization.**
- **The simplex method for linear programs.**
- **Parallel computation.**

6. Computational results on the solution of (1) & (2) — 1

DEC Alpha 21164 (600MHz) with 1GB memory

Parallel Comp. on the sol. of (1) & (2) — Eco-n problems

Intel Pentium III (824MHz) with 640MB memory

Parallel Comp. on the sol. of (1) & (2) — Cyc-n problems

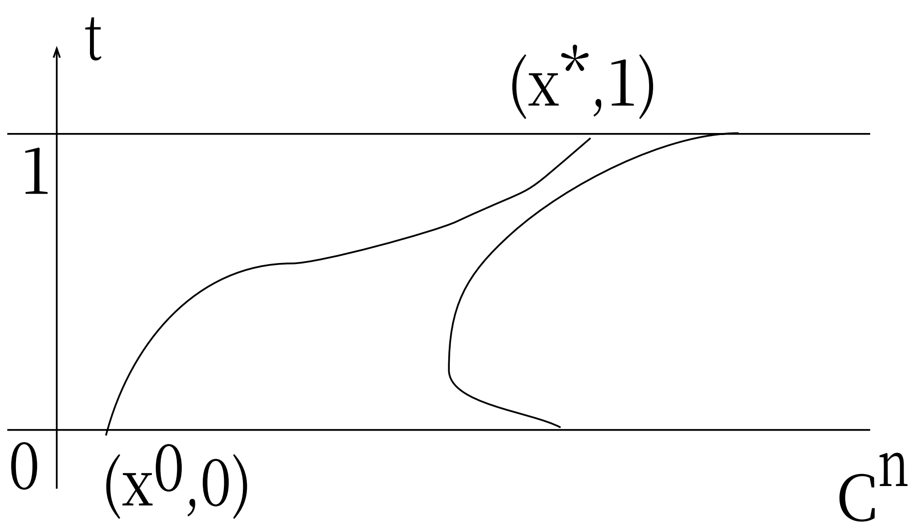
Intel Pentium III (824MHz) with 640MB memory

7. Phase 2 - Tracing homotopy paths by predictor-corrector methods

Homotopy equation system

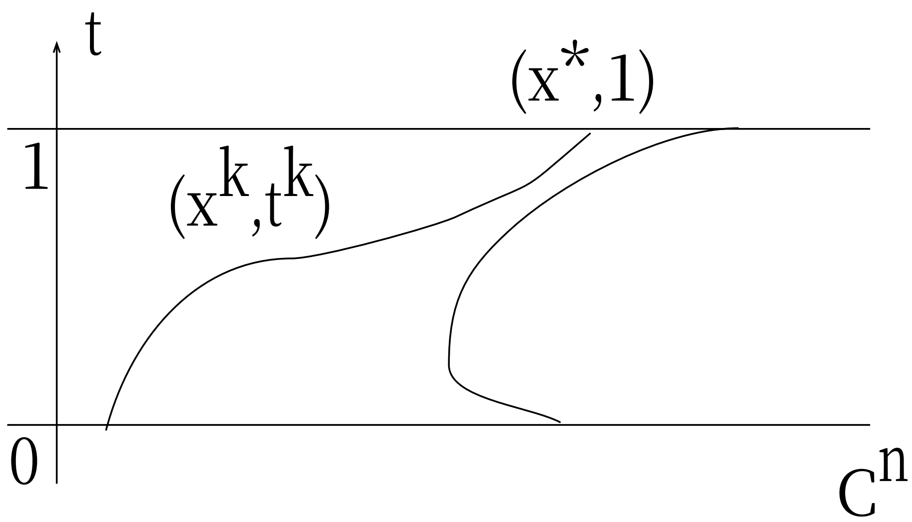
$$h_j(x,t) \equiv \sum_{a \in \mathcal{A}_j} c_j(a) x^{a} t^{\rho_j(a)} = 0, \ (x,t) \in \mathrm{C}^n \times [0,1] \quad (j = 1,2,\ldots,n) \tag{3}$$

Starting from a known init. sol. $(x^0, 0)$,trace the sol. path $\ni (x^0, 0)$.



Pred. with a step len. $dt > 0$: $Dh_x(x^0,0)dx + Dh_t(x^0,0)dt = 0$

Corr.: Newton meth. to $h(x, 0+dt) = 0$ **with the init. pt.** $\tilde{x}^0 = x^0 + dx$.



Predictor with $dt > 0$ at (x^k, t^k): $Dh_x(x^k, t^k)dx + Dh_t(x^k, t^k)dt = 0$

Too large step length $dt \Longrightarrow$ Jump into a different solution path.

Too small step length $dt \Longrightarrow$ more pred. iter. and more cpu time.

Step length control is essential!

Difficulty in Phase 2 — High nonlinearity in $h(x,t)$. Some $\rho_j(a)$'s are huge, for example

$$h_j(x,t) = \cdots + c_j(a)x^{a}t^{10} + c_j(a')x^{a'}t^{1,000} + c_j(a")x^{a"}t^{100,000} + \cdots$$

- **Sophisticated step length control.**
- **Construct homotopies with less power $\Longrightarrow$ Optimization problems.**

Change of t^p as $t \to 1$, $p = 10,\ 1,000,\ 10,000$

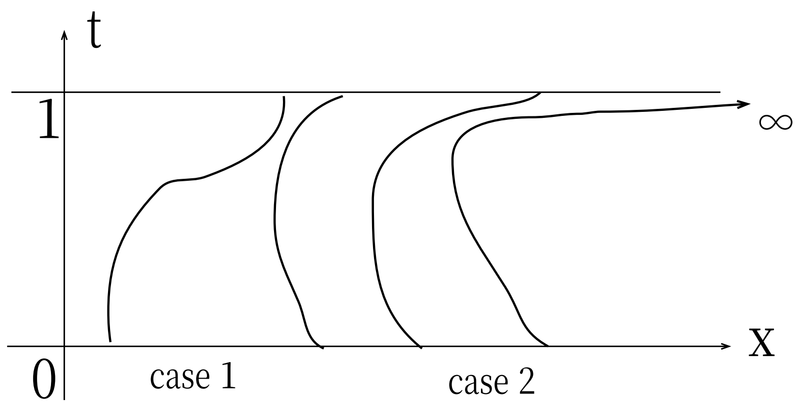
8. Numerical results — Economic-n problems

Numerical results — Cyclic-n problems

9. Concluding Remarks

(a) **While we trace a homotopy path numerically, a jump into another path sometime occurs $\Longrightarrow$ Not 100% reliable.**
But the reliability is very high; for example, less than 0.1% solutions are missing in our numerical experiments. There are two ways to compensate such a fault.

(a-1) **Suppose that numerical tracing of two paths led to a common solution as in case 1 below. Then we know there is an illegal jump while tracing one of them. Hence, recompute those two paths using smaller predictor step lengths.**



(a-2) Construct multiple sets of homotopy functions each of which theoretically covers all solutions. Then apply the polyhedral homotopy method to each set of homotopy functions to generate multiple sets of solutions. Even if a solution is missing in a set, the same solution is unlikely to be missed in all other sets. Therefore "merging all the sets of solutions" increases the reliability much.

Concluding Remarks — 2

(b) The polyhedral homotopy continuation method involves various optimization techniques such as branch-and-bound methods, linear programs, and predictor-corrector methods.

(c) Reducing the powers of the continuation parameter t is crucial to achieve the numerical stability and efficiency in tracing homotopy paths. This problem can be formulated as a nonlinear combinatorial optimization problem.

(d) An important feature of the homotopy continuation method is that all homotopy paths can be computed independently and simultaneously in parallel.