A review on Homotopy Continuation Methods for Polynomial Systems

Weiran Wang

EECS, UC Merced

wwang5@ucmerced.edu

Problem Setting

- A polynomial $p \in \mathcal{C}[\mathbf{x}]$ is a finite sum of terms $c_{\mathbf{a}}\mathbf{x}^{\mathbf{a}}$. Each term is the product of a coefficient $c_{\mathbf{a}} \in \mathcal{C}$ and a monomial $\mathbf{x}^{\mathbf{a}} = x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}$. Exponent $\mathbf{a} = [a_1, a_2, \ldots, a_n]$ are n nonnegative integers. $a_1 + a_2 + \cdots + a_n$ is the degree of term $c_{\mathbf{a}}\mathbf{x}^{\mathbf{a}}$. The degree of a polynomial is the biggest degree of its terms. A homogeneous polynomial is a polynomial whose monomials with nonzero coefficients all have the same degree.
- We consider a polynomial system $f(\mathbf{x}) = \mathbf{0}$ of N equations $f = (f_1, f_2, \ldots, f_N)$ in n variables $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ with complex coefficients, $f_i(\mathbf{x}) \in \mathcal{C}[\mathbf{x}]$, for $i = 1, 2, \ldots, N$.
- The Jacobian matrix of the system $f(\mathbf{x}) = \mathbf{0}$ is the matrix of all first partial derivatives, denoted by $\frac{\partial f}{\partial \mathbf{x}} = \left[\frac{\partial f_i}{\partial x_j}\right]$, for $i = 1, 2, \ldots, \mathbf{N}$ and $j = 1, 2, \ldots, n$.
- Here we focus on square polynomial systems which contain as many equations as unknowns, i. e., $N = n$, and aim at obtaining all isolated zeros of the system.

Solutions

- Symbolic solutions: Groebner Basis, Resultant Method, Wu's Method
  - obtain accurate zeros
  - expensive in terms of both time and space
  - limited ability of expressing the zeros
- Numerical solutions: Newton-Raphson's Method
  - simple form

$$\frac{\partial f}{\partial \mathbf{x}}(\mathbf{x}^0)\Delta\mathbf{x} = -f(\mathbf{x}^0) \qquad \mathbf{x}^1 = \mathbf{x}^0 + \Delta\mathbf{x}$$

  - quadratic convergence
  - sensitive to initial value, locally convergent
- Homotopy Continuation Methods: globally convergent, exhaustive solvers

Linear Homotopy

To solve a target system $f(\mathbf{x}) = \mathbf{0}$, we construct a start system $g(\mathbf{x}) = \mathbf{0}$ and define the artificial-parameter homotopy

$$h(\mathbf{x}, t) = \gamma(1-t)g(\mathbf{x}) + tf(\mathbf{x}) = \mathbf{0}, \quad \gamma \in \mathcal{C}, \quad t \in [0, 1].$$

The constant γ is a random constant number, also called the accessibility constant. The t is the continuation parameter and varies from 0 to 1, deforming the start system g into the target system f, i. e., $h(\mathbf{x}, 0) = g(\mathbf{x})$ and $h(\mathbf{x}, 1) = f(\mathbf{x})$.

Start System

The start system g should be chosen correctly so that the following three properties hold.

- Property 0 (Triviality). The zeros of $g(\mathbf{x}) = \mathbf{0}$ are known.
- Property 1 (Smoothness). The zero set of $h(\mathbf{x}, t) = \mathbf{0}$ for $0 \le t \le 1$ consists of a finite number of smooth paths (called homotopy curves), each parameterized by t in $[0, 1)$.
- Property 2 (Accessibility). Every isolated zero of $h(\mathbf{x}, 1) = f(\mathbf{x}) = \mathbf{0}$ can be reached by some path originating at $t = 0$. It follows that this path starts at a zero of $h(\mathbf{x}, 0) = g(\mathbf{x}) = \mathbf{0}$.

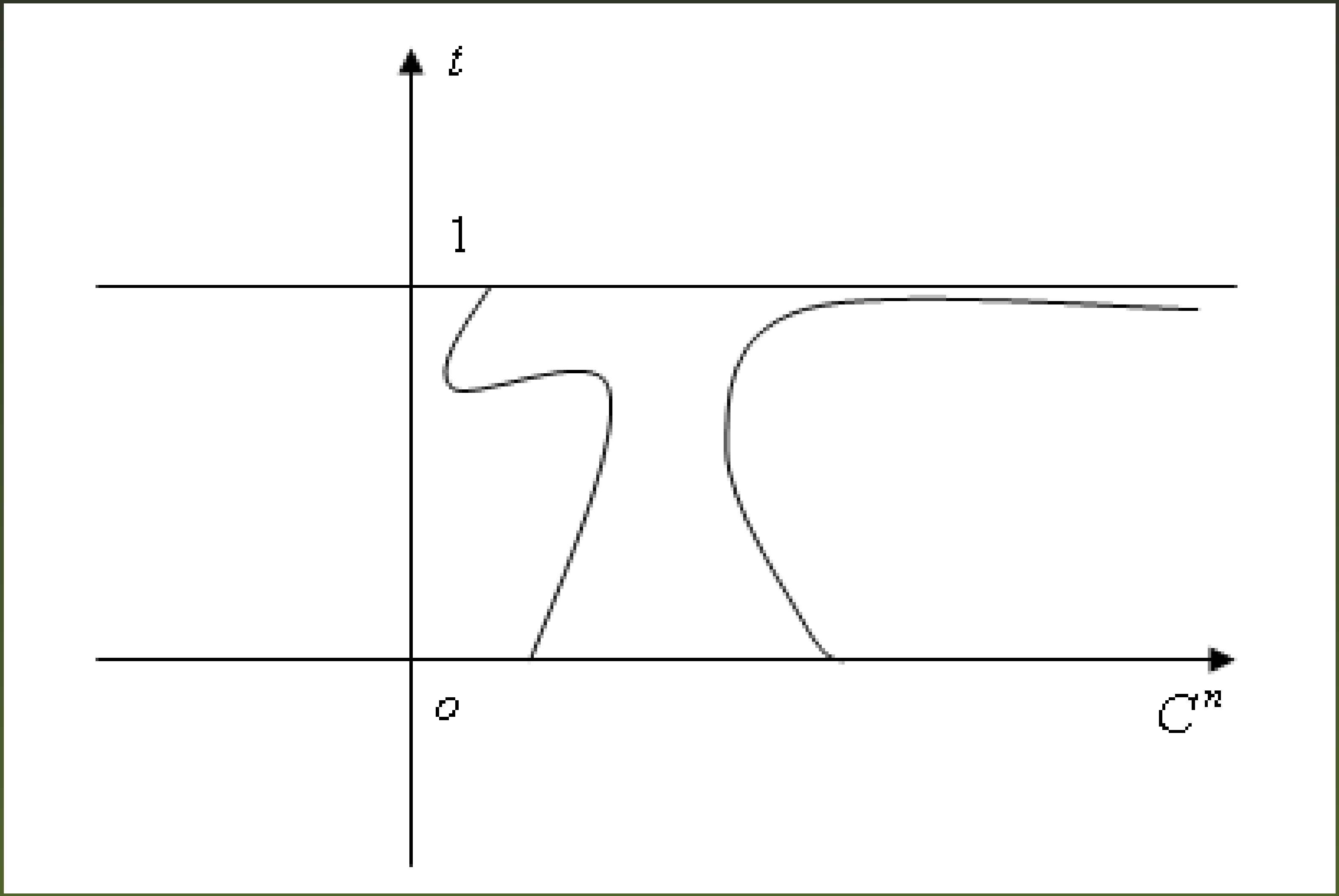
Property 1 & 2 are guaranteed *generalized Sard's Theorem*.

Homotopy Curves

- (the gamma trick) For almost all choices of complex constant γ, all solution paths defined by the homotopy $h(\mathbf{x}, t) = \gamma(1-t)g(\mathbf{x}) + tf(\mathbf{x}) = \mathbf{0}$ are regular, i. e.: for all $t \in [0, 1)$, the Jacobian matrix of $h(\mathbf{x}, t)$ is regular and no path diverges.
- We only need to consider two types of homotopy curves
  - starts from a zero of $g(\mathbf{x}) = \mathbf{0}$ and ends at a zero of $f(\mathbf{x}) = \mathbf{0}$
  - starts from a zero of $g(\mathbf{x}) = \mathbf{0}$ but diverge to infinity when $t \longrightarrow 1$
- In other words, it is quite possible for $g(\mathbf{x}) = \mathbf{0}$ to have more solutions than $f(\mathbf{x}) = \mathbf{0}$.

Homotopy Curves (cont.)

Bounded and unbounded homotopy curves



A typical choice of start system

A typical choice which satisfies Properties 0–2 is,

$$\begin{aligned} g_1(x_1, \ldots, x_n) &= a_1 x_1^{d_1} - b_1, \\ &\vdots \\ g_n(x_1, \ldots, x_n) &= a_n x_n^{d_n} - b_n, \end{aligned}$$

where $d_1, \ldots, d_n$ are the degrees of $f_1(\mathbf{x}), \ldots, f_n(\mathbf{x})$ respectively, and a_j, b_j are random complex numbers (and therefore nonzero with probability one).

This system produces $d = d_1 \ldots d_n$ paths since $g(\mathbf{x}) = \mathbf{0}$ has d isolated nonsingular solutions. System $f(\mathbf{x}) = \mathbf{0}$ is called deficient if it has less than d solutions.

Bezout's Theorem

(Bezout's Theorem) The number of isolated zeros, counting multiplicities, of $f(\mathbf{x})$ in $\mathcal{C}^n$, is bounded above by the *Bezout number* $d = d_1 \cdots d_n$.

In practice, this bound could be very loose for deficient systems. For example, the following Cassou-Nogues system

$$\begin{aligned}
p_1 =\ & 15b^4cd^2 + 6b^4c^3 + 21b^4c^2d - 144b^2c - 8b^2c^2e \\
& -28b^2cde - 648b^2d + 36b^2d^2e + 9b^4d^3 - 120, \\
p_2 =\ & 30b^4c^3d - 32cde^2 - 720b^2cd - 24b^2c^3e - 432b^2c^2 + 576ce - 576de \\
& +16b^2cd^2e + 16d^2e^2 + 16c^2e^2 + 9b^4c^4 + 39b^4c^2d^2 + 18b^4cd^3 \\
& -432b^2d^2 + 24b^2d^3e - 16b^2c^2de - 240c + 5184, \\
p_3 =\ & 216b^2cd - 162b^2d^2 - 81b^2c^2 + 1008ce - 1008de + 15b^2c^2de \\
& -15b^2c^3e - 80cde^2 + 40d^2e^2 + 40c^2e^2 + 5184, \\
p_4 =\ & 4b^2cd - 3b^2d^2 - 4b^2c^2 + 22ce - 22de + 261.
\end{aligned}$$

has only 16 isolated zeros, but the above start system will produce $d_1 \times d_2 \times d_3 \times d_4 = 7 \times 8 \times 6 \times 4 = 1344$ paths. (mixed volume: 24)

Mixed Volume

- Consider the polynomial system $f(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_n(\mathbf{x})) = \mathbf{0}$. For $i = 1, 2, \ldots, n$, the support A_i of f_i collects the exponents of $\mathbf{x}$ of those monomials which occur in f_i with nonzero coefficient, so we can write $f_i(\mathbf{x}) = \sum_{\mathbf{a} \in A_i} c_{\mathbf{a}} \mathbf{x}^{\mathbf{a}}$. The Newton polytope P_i of f_i is the convex hull of the support A_i of f_i, i. e., $P_i = \operatorname{conv}(A_i)$.
- (Minkowski's theorem) Given an n-tuple of polytopes $(P_1, P_2, \ldots, P_n)$, the volume of a general linear combination of these polytopes, denoted by $V(\lambda_1 P_i + \lambda_2 P_2 + \cdots + \lambda_n P_n)$, is a homogeneous polynomial of degree n in $\lambda_1, \lambda_2, \ldots, \lambda_n$. The coefficient of $\lambda_1 \lambda_2 \cdots \lambda_n$ in this polynomial is the mixed volume $V(P_1, P_2, \ldots, P_n)$ of the n-tuple.

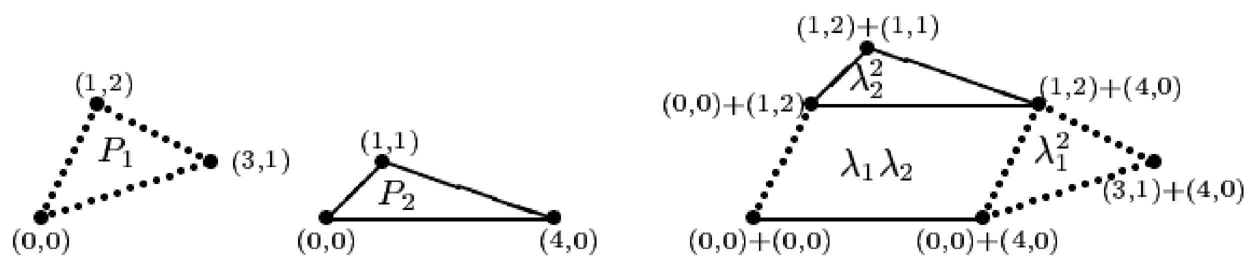


Figure 2: A subdivision of the sum of two polygons P_1 and P_2. The sum is the convex hull of all sums of the vertices of the polygons. The cells in the subdivision are labeled by the multipliers for the area of $\lambda_1 P_1 + \lambda_2 P_2$.

Root Count

(Bernshtein's theorem, 1965) Let $f(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_n(\mathbf{x})) = \mathbf{0}$, and A_i is the support of f_i, $i = 1, \dots, n$. The number of isolated zeros, counting multiplicities, of $f(\mathbf{x})$ in $\mathcal{C}^n$, is bounded above by the mixed volume $\mathcal{M} = V(\mathrm{conv}\,(A_1 \cup \{\mathbf{0}\}), \dots, \mathrm{conv}\,(A_n \cup \{\mathbf{0}\}))$.

Consider the following start system

$$
\begin{aligned}
g_1(x_1, \dots, x_n) &= \sum_{\mathbf{a} \in A_1'} c_{1,\mathbf{a}} \mathbf{x}^{\mathbf{a}}, \\
&\vdots \\
g_n(x_1, \dots, x_n) &= \sum_{\mathbf{a} \in A_n'} c_{n,\mathbf{a}} \mathbf{x}^{\mathbf{a}},
\end{aligned}
$$

where $A_i' = A_i \cup \{\mathbf{0}\}$, and $c_{i,\mathbf{a}}$ are generated randomly, $i = 1, \dots, n$. Then $g(\mathbf{x}) = \mathbf{0}$ has exactly $\mathcal{M}$ isolated zeros in $(\mathcal{C} \setminus \{0\})^n$ with probability one.

Polyhedral Homotopy

Let t denote a new complex variable and consider the polynomial system $q(\mathbf{x},t) = (q_1(\mathbf{x},t),\ldots,q_n(\mathbf{x},t)) = \mathbf{0}$ in the $n+1$ variables given by

$$
\begin{aligned}
q_1(\mathbf{x},t) &= \sum_{\mathbf{a}\in A_1'} c_{1,\mathbf{a}}\mathbf{x}^{\mathbf{a}} t^{w_1(a)}, \\
&\vdots \\
q_n(\mathbf{x},t) &= \sum_{\mathbf{a}\in A_n'} c_{n,\mathbf{a}}\mathbf{x}^{\mathbf{a}} t^{w_n(a)},
\end{aligned}
$$

where each $w_i : A_i' \to \mathcal{R}$ for $i = 1,\ldots,n$ is chosen generically. It can be regarded as a homotopy defined on $(\mathcal{C}\setminus\{0\})^n \times [0,1]$, known as polyhedral homotopy. It also satisfies Properties 0–2 and has $\mathcal{M}$ homotopy curves.

Polyhedral Homotopy (cont.)

By some substitution of variables $\mathbf{x} = \mathbf{y}t^{\alpha}$, we could obtain system $q_\alpha(\mathbf{y}, t) = \mathbf{0}$. Each homotopy curve of $q_\alpha(\mathbf{y}, t) = \mathbf{0}$ ends at an isolated zero of $g(\mathbf{y}) = q_\alpha(\mathbf{y}, 1) = \mathbf{0}$, and originates from a zero of the following Binomial System

$$\begin{aligned} q_1^\alpha(\mathbf{y}, t) &= c_{11}\mathbf{y}^{\mathbf{a}_1^{(1)}} + c_{12}\mathbf{y}^{\mathbf{a}_2^{(1)}} = 0, \\ &\vdots \\ q_n^\alpha(\mathbf{y}, t) &= c_{n1}\mathbf{y}^{\mathbf{a}_1^{(n)}} + c_{n2}\mathbf{y}^{\mathbf{a}_2^{(n)}} = 0, \end{aligned}$$

which has exactly

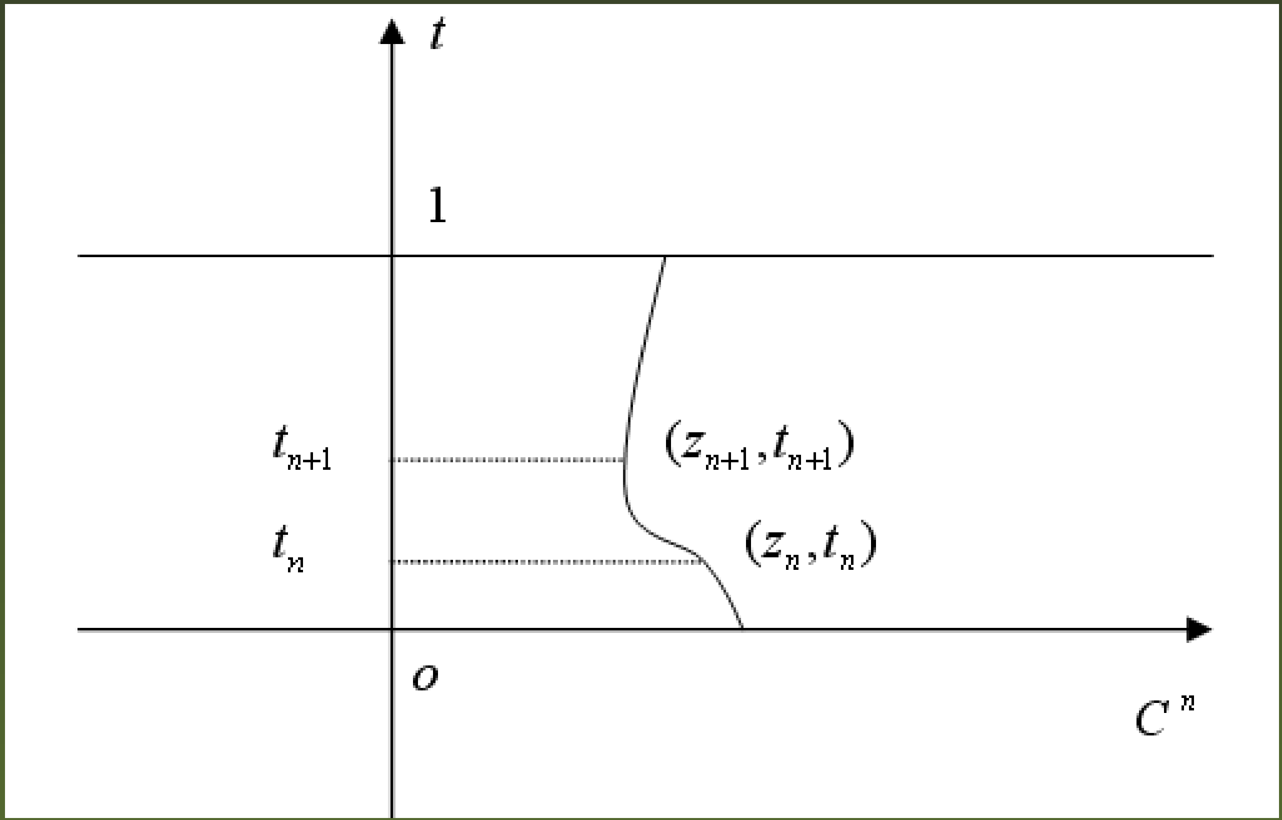
$$k_\alpha \equiv \left| \det \begin{pmatrix} \mathbf{a}_1^{(1)} & - & \mathbf{a}_2^{(1)} \\ & \vdots & \\ \mathbf{a}_1^{(n)} & - & \mathbf{a}_2^{(n)} \end{pmatrix} \right|$$

nonsingular solutions in $(\mathcal{C} \setminus \{0\})^n$.

Path Tracking

predictor-corrector methods:

- from an approximate solution point to the path, a predictor gives a new approximate point a given step size along the path (for example, an Euler predictor steps ahead along the tangent to the path)
- then a corrector brings this new point closer to the path (usually Newton's method)



Step length control and failure

- **Initialize** Select: an initial step size, s; the number of corrector iterations allowed per step, $N \leq 1$; the step adjustment factor, $a \in (0, 1)$; the step expansion integer, $M \leq 1$; and a minimum step size s_{min}.
- **Predict** Estimate a new point near the path whose distance from the current point is the step size s.
- **Correct** Iteratively improve the new path point, constraining its distance from the prior path point. Allow at most N iterations to reach the specified tolerance.
- **On success** If the tolerance is achieved:
  - Update the current path point to be the newly found path point.
  - If we have reached the final value of t, exit with success.
  - If there have been M successes in a row, expand the step size to $s = s/a$.
- **On failure** If the tolerance is not achieved:
  - Cut the step size to $s = as$.
  - If $s < s_{min}$, exit with failure.
- **Loop** Go back to **Predict**.

N should be set to be small (2 or 3). Otherwise a bad prediction might ultimately converge, but it may wander first and become attracted to a completely different path in the homotopy.

Isolated Singularities

The convergence of Newton's method slows down to a halt when approaching a singular solution.

Solution: *Deflation* — effective reconditioner, restore quadratic convergence

Consider polynomial system $f(\mathbf{x}) = \mathbf{0}$ of N equations in n unknowns. Let R be the rank of the Jacobian at an isolated singular root $\mathbf{x}^*$. We define a deflation of f at $\mathbf{x}^*$ as the system

$$F(\mathbf{x}, \lambda) = \begin{cases} f(\mathbf{x}) & = & \mathbf{0} \\ \frac{\partial f}{\partial \mathbf{x}} B\lambda & = & \mathbf{0} \\ \mathbf{h}\lambda & = & 1 \end{cases}$$

using $R + 1$ multipliers $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_{R+1})$ as extra variables, together with an N-by-$(R+1)$ matrix B of random complex numbers and a random vector of $R + 1$ complex coefficients h.

Isolated Singularities (cont.)

Theorem: The number of deflations needed to restore the quadratic convergence of Newton's method converging to an isolated solution is strictly less than the multiplicity of the isolated solution.

Let $f_1, \ldots, \mathbf{f}_s \in \mathcal{C}[x_1, \ldots, x_n]$, $I =< f_1, \ldots, f_s >$ is the ideal generated by the system, $\mathbf{p} = (p_1, \ldots, p_n)$ is an isolated zero of the system, $\mathcal{M} =< x_1 - p_1, \ldots, \mathbf{x}_n - p_n >$, then the multiplicity of $\mathbf{p}$ is defined as

$$m(\mathbf{p}) = \dim_{\mathcal{C}} \mathcal{C}[x_1, \ldots, \mathbf{x}_n]_{\mathcal{M}} / IC[x_1, \ldots, \mathbf{x}_n]_{\mathcal{M}},$$

and it is finite.

Example with PHC package

PHCpack: a general-purpose solver for polynomial systems
by homotopy continuation
Algorithm 795 in ACM Trans. Math. Softw.
Jan Verschelde

Available at `http://www.math.uic.edu/~jan/`

References

[1] Jan Verschelde. Homotopy Methods for Solving Polynomial Systems, tutorial at ISSAC 2005, July 24-27, 2005. Key Laboratory of Mathematics Mechanization, Beijing, China

[2] Tien-Yien Li. Solving Polynomial Systems By Polyhedral Homotopies. Taiwanese Journal Of Mathematics, Vol. 3, No. 3, pp. 251-279, September 1999.

[3] Daniel J. Bates et al. Adaptive Multiprecision Path Tracking. SIAM Journal on Numerical Analysis, Vol. 46, pp. 722-746, 2008.

[4] Anton Leykin, Jan Verschelde, Ailing Zhao. Newton's method with deflation for isolated singularities of polynomial systems. Theor. Comput. Sci. 359(1-3): 111-122 (2006).

[5] David A. Cox, John B. Little, Don O'Shea. Using Algebraic Geometry (Second Edition), Graduate Texts in Mathematics, Volume 185, Springer, 2005.