Roots of a Polynomial System

(Com S 477/577 Notes)

Yan-Bin Jia

Sep 27, 2022

Systems of polynomial equations show up in many applications areas such as robotics (kinematics, motion planning, collision detection, etc.), computer vision (object modeling, surface fitting, recognition, etc.), graphics, geometric modeling (curve and surface intersections), computer-aided design, mechanical design, and chemical equilibrium systems. In this lecture, we introduce a powerful technique that finds the roots of a polynomial system by tracking the solutions of "nearby" systems.

Many problems in geometry and robot kinematics can be formulated in terms of trigonometric functions, which often can be converted to polynomials. Figure 1(a) shows a manipulator with two

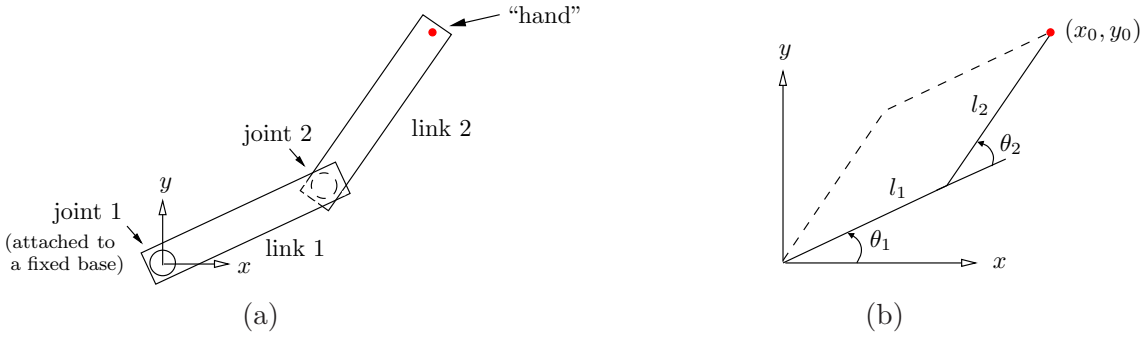

Figure 1: A 2-linkage manipulator.

links of lengths l_1 and l_2, respectively, and the first joint attached to a based on the table. The "hand", or the end-effector, is located at the tip of the second link. A coordinate system is set up at the first joint with the x-axis parallel the table plane and the y-axis vertical to it. Part (b) of the figure abstracts the manipulator configuration, which is completely determined by the angles θ_1 and θ_2 of the two joints. The end-effector is located at

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = l_1 \begin{pmatrix} \cos\theta_1 \\ \sin\theta_1 \end{pmatrix} + l_2 \begin{pmatrix} \cos(\theta_1 + \theta_2) \\ \sin(\theta_1 + \theta_2) \end{pmatrix}. \tag{1}$$

Determining the end-effector position of the manipulator from its joint angles is called the *forward kinematics* problem.

Very often we want to move the end-effector to a specified location $\binom{x_0}{y_0}$ and thus need to know the joint angle values. In other words, we would like to solve the system (1) for θ_1 and θ_2. This is referred to as the *inverse kinematics* problem. One solution can be found from the triangle defined

by the origin (joint 1), joint 2, and $\binom{x_0}{y_0}$. Apply the law of cosine to obtain the angle $\pi - \theta_2$ between the two links, and the angle α between the first link and the edge joining the origin and $\binom{x_0}{y_0}$. Then, θ_1 is the difference between the polar angle of $\binom{x_0}{y_0}$ and α. The other solution, describing the symmetric pose of the arm drawn in dashed lines in Figure 1, easily follows from the first one.

Alternatively, we can solve (1) by turning it into a polynomial system. Here is how. For $k = 1, 2$ let us write c_k for $\cos\theta_k$ and s_k for $\sin\theta_k$, and add the equation $c_k^2 + s_k^2 = 1$. The result is a quadratic system in four variables c_1, s_1, c_2, and s_2:

$$\begin{aligned} l_1 c_1 + l_2(c_1 c_2 - s_1 s_2) &= x_0, \\ l_1 s_1 + l_2(s_1 c_2 + s_2 c_1) &= y_0, \\ c_1^2 + s_1^2 &= 1, \\ c_2^2 + s_2^2 &= 1. \end{aligned}$$

From the values of c_k and s_k in a solution of the above system, we can uniquely determine the angle as $\theta_k = \text{atan2}(c_k, s_k)$. There are two solutions of the system which correspond to the two different arm poses in Figure 1. This polynomial approach is more general than the trignometric one earlier, and applicable to the inverse kinematics of a robotic arm with more than two links.

1 Multivariate Polynomial Systems

We have already studied polynomials in one variable, and learned some global root finding techniques such as Müller's method. Now we look at how to solve a number of polynomial equations in the same number of variables. First, we need to generalize the definition of a polynomial in one variable to that of one in multiple variables.

A multivariate polynomial $f(x_1, x_2, \ldots, x_n)$ is a finite sum of terms $c_\alpha x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$, where $\alpha_1, \ldots, \alpha_n$ are non-negative integers, and c_α is a coefficient with its index α denoting the vector $(\alpha_1, \ldots, \alpha_n)$. The term $c_\alpha x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$ is called a *monomial*.

The degree of $c_\alpha x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$ is $\alpha_1 + \alpha_2 + \cdots + \alpha_n$. The polynomial f has degree equal to the maximum degree of all of its monomials. A *homogeneous* polynomial is a sum of monomials of the same degree.

Example 1. The real zeros (x, y) of the system

$$\begin{aligned} x^2 + y^2 &= 1, \\ xy &= c, \end{aligned} \tag{2}$$

where $c > 0$, if exist, locate the intersections of the unit circle with a hyperbola. As shown in Figure 2, the situation depends on the value of c.

To solve the system, we multiply the second equation in (2) with 2 and add the result to the first equation, obtaining $x + y = \pm\sqrt{1 + 2c}$. Then, substitute $y = \pm\sqrt{1 + 2c} - x$ into the second equation to obtain x. This yields four complex common roots:

$$\begin{aligned} &\left(\tfrac{\sqrt{1+2c}+\sqrt{1-2c}}{2}, \tfrac{\sqrt{1+2c}-\sqrt{1-2c}}{2}\right), && \left(\tfrac{\sqrt{1+2c}-\sqrt{1-2c}}{2}, \tfrac{\sqrt{1+2c}+\sqrt{1-2c}}{2}\right), \\ &-\left(\tfrac{\sqrt{1+2c}+\sqrt{1-2c}}{2}, \tfrac{\sqrt{1+2c}-\sqrt{1-2c}}{2}\right), && -\left(\tfrac{\sqrt{1+2c}-\sqrt{1-2c}}{2}, \tfrac{\sqrt{1+2c}+\sqrt{1-2c}}{2}\right). \end{aligned} \tag{3}$$

We infer that the two curves do not intersect when $c > \frac{1}{2}$, intersect at two points $(\sqrt{2}/2, \sqrt{2}/2)$ and $(-\sqrt{2}/2, -\sqrt{2}/2)$ when $c = \frac{1}{2}$, or at the four points given in the equation when $0 < c < \frac{1}{2}$.

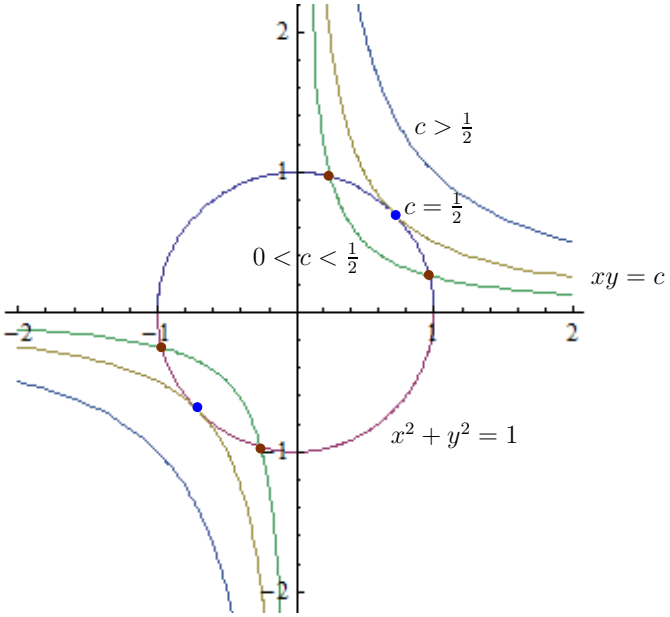

Figure 2: The unit circle $x^2 + y^2 = 1$ intersects a hyperbola $xy = c$, $c > 0$ at no point if $c > \frac{1}{2}$, two points if $c = \frac{1}{2}$, and four points if $0 < c < \frac{1}{2}$.

The above example tells us that real zeros may or may not exist for a polynomial system, and in case they do, their vary in number. In the below, we will focus on finding all complex zeros, for which bounds and techniques are more established. Real zeros are easily selected out once all complex ones are found.

2 Root Counting

Consider a general polynomial system $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$, where $\boldsymbol{f} = (f_1, f_2, \ldots, f_n)^\top$ and $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$, fully expanded as

$$\begin{aligned} f_1(x_1, x_2, \ldots, x_n) &= 0, \\ f_2(x_1, x_2, \ldots, x_n) &= 0, \\ &\vdots \\ f_n(x_1, x_2, \ldots, x_n) &= 0. \end{aligned} \tag{4}$$

The system, with as many equations as unknowns, is called a *square polynomial system*. Let $d_1, \ldots, d_n$ be the degrees of $f_1, \ldots, f_n$, respectively.

Theorem 1 (Bézout) *[10] Unless $\boldsymbol{f}(\boldsymbol{x})$ has an inifite number of zeros, the number of its isolated zeros in $\mathbb{C}^n$, counting multiplicities, does not exceed the number $d = d_1 d_2 \cdots d_n$.*

The number d is called the *Bézout number*. It often turns out to be a rather loose bound. For example, the Cassou-Nogues system in four variables b, c, d, e below:

$$0 = 15b^4cd^2 + 6b^4c^3 + 21b^4c^2d - 144b^2c - 8b^2c^2e - 28b^2cde - 648b^2d + 36b^2d^2e + 9b^4d^3 - 120,$$

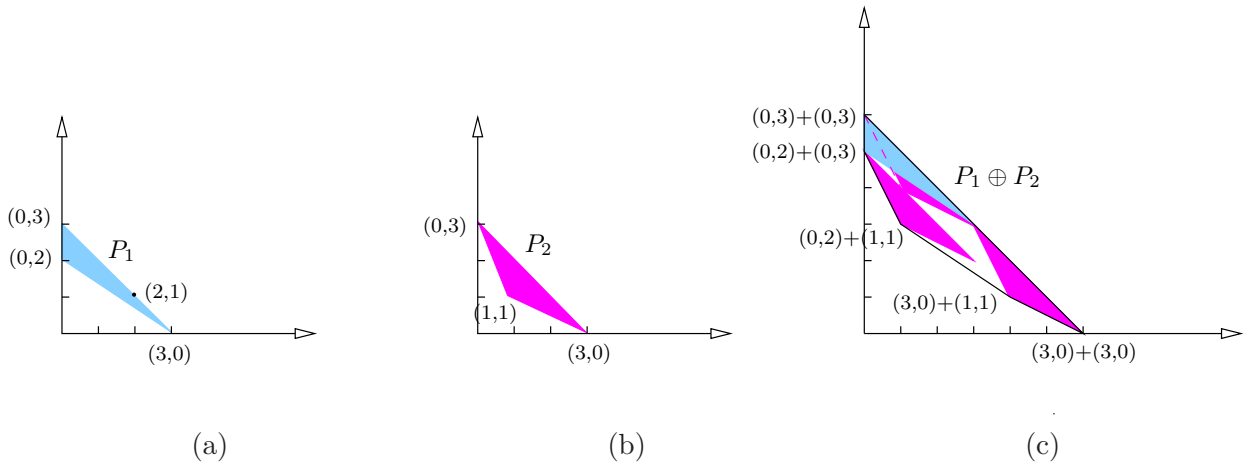

Figure 3: Newton polytopes for the two polynomials (6): (a) P_1 and (b) P_2. (c) Their Minkowski sum $P_1 \oplus P_2$.

$$\begin{aligned}
0 &= 30b^4c^3d - 32cde^2 - 720b^2dc - 24b^2c^3e - 432b^2c^2 + 576ce - 576de + 16b^2cd^2e + 16d^2e^2 \\
&\quad + 16c^2e^2 + 9b^4c^4 + 39b^4c^2d^2 + 18b^4cd^3 - 432b^2d^2 + 24b^2d^3e - 16b^2c^2de - 240c + 5184, \\
0 &= 216b^2cd - 162b^2d^2 - 81b^2c^2 + 1008ce - 1008de + 15b^2c^2de - 15b^2c^3e - 80cde^2 \\
&\quad + 40d^2e^2 + 40c^2e^2 + 5184, \\
0 &= 4b^2cd - 3b^2d^2 - 4b^2c^2 + 22ce - 22de + 261,
\end{aligned} \tag{5}$$

is known to have 16 distinct zeros in the four-dimensional complex domain $\mathbb{C}^4$, and only 4 distinct zeros in the four-dimensional real domain $\mathbb{R}^4$. However, its Bézout number is $7 \times 8 \times 6 \times 4 = 1344$.

A tighter bound was provided by Bernstein [3] based on the notion of mixed volume. Introduce the *support* S_k of a polynomial f_k in the system (4) as the set of the exponents of its monomials. To illustrate, we use an example from [2] that deals with the following two-variable polynomial system:

$$\begin{aligned}
f_1(x,y) &= 3x^3 - 4x^2y + y^2 + 2y^3, \\
f_2(x,y) &= -6x^3 + 2xy - 5y^3.
\end{aligned} \tag{6}$$

The support S_1 of f_1 contains $(3,0), (2,1), (0,2), (0,3)$, while the support S_2 of f_2 contains $(3,0)$, $(1,1)$, $(0,3)$. The *Newton polytope* P_k of f_k is the convex hull of the support S_k. Figure 3(a) and (b) draws the Newton polytopes P_1 and P_2 of f_1 and f_2, respectively.

The *Minkowski sum* of the Newton polytopes $P_1, \ldots, P_n$ for the polynomials $f_1, \ldots, f_n$ is the set

$$P_1 \oplus \cdots \oplus P_n = \{p_1 + \cdots p_n \mid p_k \in P_k, 1 \le k \le n\}.$$

For system (6), the Minskowski sum of its Newton polytopes in Figure 3 (a) and (b) is shown in part (c). It is swept out by P_2 with its upper vertex sliding along the boundary of P_1 followed by an upward translation of the whole swept area by 3.

The volume of a general linear combination $\lambda_1 P_1 \oplus \lambda_2 P_2 \oplus \cdots \oplus \lambda_n P_n$, where $\lambda_1, \lambda_2, \ldots, \lambda_n > 0$, is a homogeneous polynomial in $\lambda_1, \lambda_2, \ldots, \lambda_n$. The coefficient of $\lambda_1 \lambda_2 \cdots \lambda_n$ in this polynomial is called the *mixed volume* M of $P_1, P_2, \ldots, P_n$, or of the supports $S_1, S_2, \ldots, S_n$. Let $\mathbb{C}^* = \mathbb{C} \setminus 0$ be the set of all non-zero complex numbers.

Theorem 2 (Bernstein, 1975) *Suppose the polynomial system (4) has finitely many roots in* $(\mathbb{C}^*)^n$. *Then the number of these roots is bounded from above by the mixed volume of its Newton polytopes* P_k, $1 \le k \le n$.

In fact, the root count equals the mixed volume provided that the polynomials are sufficiently "generic". Here we are not going to make the meaning of "generic" exact. In practice, it is often assumed that the polynomials in a given system are "generic". So the Bernstein bound turns out to be the exact root count.

We denote by vol_n the volume operator in $\mathbb{R}^n$. The mixed volume can be calculated as below [8, p. 140]:

$$\begin{aligned}
M(P_1, P_2) &= \mathrm{vol}_2(P_1 \oplus P_2) - \mathrm{vol}_2(P_1) - \mathrm{vol}_2(P_2), \\
M(P_1, P_2, P_3) &= \mathrm{vol}_3(P_1 \oplus P_2 \oplus P_3) - \mathrm{vol}_3(P_1 \oplus P_2) - \mathrm{vol}_3(P_2 \oplus P_3) - \mathrm{vol}_3(P_3 \oplus P_1) \\
&\quad + \mathrm{vol}_3(P_1) + \mathrm{vol}_3(P_2) + \mathrm{vol}_3(P_3), \\
M(P_1, \ldots P_n) &= \sum_{i=1}^{n} (-1)^{n-i} \sum_{\substack{\{j_1, \ldots, j_k\} \text{ a combination} \\ \text{of } k \text{ indices from } 1, \ldots, n}} \mathrm{vol}_i(P_{j_1} \oplus \cdots \oplus P_{j_i}).
\end{aligned}$$

For system (6) illustrated in Figure 3, P_1 and P_2 each has volume (area) $\frac{3}{2}$, and $P_1 \oplus P_2$ has volume 7. Thus, the system has exactly $7 - \frac{3}{2} - \frac{3}{2} = 4$ non-zero roots. Computation of the mixed volume is a combinatorial problem for which efficient methods have been developed. These methods are highly technical and beyond the scope of this course.

Example 2. Let us apply Bernstein's theorem to the circle-hyperbola system (2) from Example 1. We let

$$\begin{aligned}
f_1(x, y) &= x^2 + y^2 - 1, \\
f_2(x, y) &= xy - c.
\end{aligned}$$

The supports are

$$\begin{aligned}
S_1 &= \{(2, 0), (0, 2), (0, 0)\}, \\
S_2 &= \{(1, 1), (0, 0)\}.
\end{aligned}$$

The Newton polytope P_1 of S_1 is a triangle with area 2, while P_2 of S_2 is a line segment with length $\sqrt{2}$ and zero area.

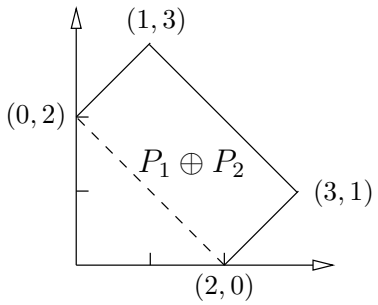

Figure 4: Minkowsky sum $P_1 \oplus P_2$.

It is easy to see that $P_1 \oplus P_2$ is a pentagon that is the union of P_1 with the rectangle swept out by P_2 in translation as its lower vertex $(0,0)$ moves along the tilted edge of P_1 connecting $(2,0)$ and $(0,2)$. The rectangle has area $\sqrt{2} \cdot 2\sqrt{2} = 4$, which is the Bernstein count. The system has exactly four complex roots given in (3).

3 Homotopy Continuation

Solution of polynomial systems has been active in computer algebra and computational algebraic geometry. Well-known symbolic techniques include Gröbner basis [5], which can be viewed as a multivariate, non-linear generalization of both Euclid's algorithm and Gaussian elimination for linear systems; the resultant method [11], which performs variable elimination via construction of matrices and determinants; and Wu's method [14], which is based on polynomial division and the concept of characteristic set. Though capable of finding accurate zeros, these methods are computationally expensive in both time and space.

Numerical root finding methods such as Newton's give up exactness for efficiency. With only local convergence, they are quite sensitive to initial values, which sometimes have to be picked at random in repeated trials due to lack of prior information.

Here we will describe the homotopy continuation method credited to Garcia and Zangwill [4]. This method solves a polynomial system $\boldsymbol{f}$ by tracking the solution of the "nearby" systems. The idea is to construct a parametrized family of polynomial systems such that at one endpoint of the parametrization interval is $\boldsymbol{f}$ to be solved while at the other endpoint is a system $\boldsymbol{g}$ with known zeros. As the parameter value varies across the domain, the system $\boldsymbol{g}$ morphs into the system $\boldsymbol{f}$. During the period, the solution is tracked from a known one to $\boldsymbol{g}$ across all systems in the family to reach a solution to $\boldsymbol{f}$.

To make the above idea more concrete, let us assume that we can solve

$$\boldsymbol{g}(\boldsymbol{x}) = \begin{pmatrix} g_1(\boldsymbol{x}) \\ \vdots \\ g_n(\boldsymbol{x}) \end{pmatrix} = \boldsymbol{0}. \tag{7}$$

We form the homotopy[1]

$$\boldsymbol{h}(\boldsymbol{x}, t) = t\boldsymbol{f}(\boldsymbol{x}) + (1-t)\boldsymbol{g}(\boldsymbol{x}), \tag{8}$$

where $t \in [0,1]$. Note that $\boldsymbol{h}(\boldsymbol{x},0) = \boldsymbol{g}(\boldsymbol{x})$ and $\boldsymbol{h}(\boldsymbol{x},1) = \boldsymbol{f}(\boldsymbol{x})$. For every root $\boldsymbol{x}_0$ of (7), the equation

$$\boldsymbol{h}(\boldsymbol{x}, t) = \boldsymbol{0}$$

implicitly defines a curve $\boldsymbol{x} = \boldsymbol{x}(t)$ over $[0,1]$ that is the homotopy evolution of $\boldsymbol{x}_0$ toward some root $\boldsymbol{x}^* = \boldsymbol{x}(1)$ of $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$. See Figure 5 for a plot. This is one solution curve to the family of equations $\boldsymbol{h}(\boldsymbol{x}, t) = \boldsymbol{0}$ parametrized with t.

Starting at $t = 0$ and $\boldsymbol{x}(0) = \boldsymbol{x}_0$, we would like to stay on the solution curve $\boldsymbol{x}(t)$ such that $\boldsymbol{h}(\boldsymbol{x}, t) = \boldsymbol{0}$ holds for any t value we step over, until $t = 1$ is reached. This is conducted via numerically solving the following ordinary differential equation (ODE):

$$\frac{d\boldsymbol{h}(\boldsymbol{x}, t)}{dt} = \boldsymbol{0},$$

[1]There are many other homotopies. For example, the mathematics software Maple uses $\boldsymbol{h}(\boldsymbol{x}, t) = (1-(1-t)^2)\boldsymbol{f}(\boldsymbol{x}) + (1-t)^2\boldsymbol{g}(\boldsymbol{x})$.

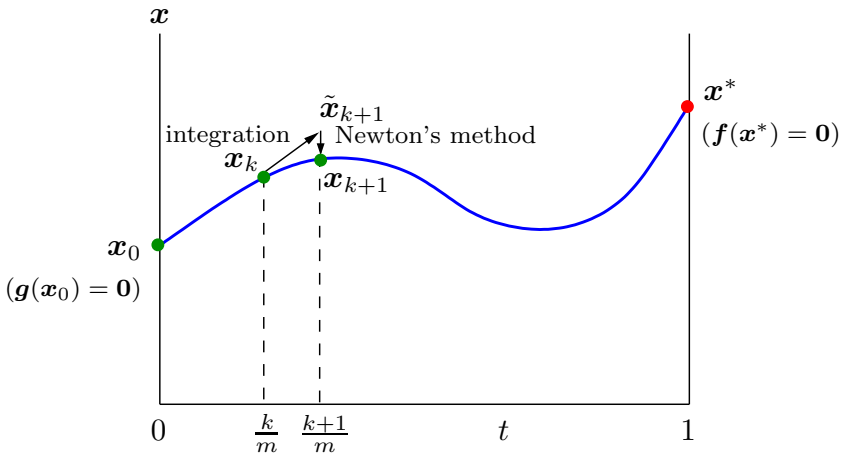

Figure 5: Homotopy continuation method.

with the initial value $\boldsymbol{x}(0) = \boldsymbol{x}_0$. Expand the left hand side of the above equation:

$$\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}} \cdot \frac{d\boldsymbol{x}}{dt} + \frac{\partial \boldsymbol{h}}{\partial t} = \boldsymbol{0},$$

where

$$\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}} = \begin{pmatrix} \partial h_1/\partial \boldsymbol{x} \\ \partial h_2/\partial \boldsymbol{x} \\ \vdots \\ \partial h_n/\partial \boldsymbol{x} \end{pmatrix} = \begin{pmatrix} \frac{\partial h_1}{\partial x_1} & \frac{\partial h_1}{\partial x_2} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \frac{\partial h_2}{\partial x_1} & \frac{\partial h_2}{\partial x_2} & \cdots & \frac{\partial h_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_n}{\partial x_1} & \frac{\partial h_n}{\partial x_2} & \cdots & \frac{\partial h_n}{\partial x_n} \end{pmatrix} \quad \text{and} \quad \frac{\partial \boldsymbol{h}}{\partial t} = \boldsymbol{f} - \boldsymbol{g}.$$

Assuming the matrix to be invertible at the current t value, we obtain

$$\frac{d\boldsymbol{x}}{dt} = -\left(\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}}\right)^{-1} \frac{\partial \boldsymbol{h}}{\partial t}. \tag{9}$$

Integrate $d\boldsymbol{x}/dt$ over t on $[0, 1]$ starting with $\boldsymbol{x}_0$. We choose a step size $\frac{1}{m}$ for some large integer $m > 0$. Denote by $\boldsymbol{x}_k$ the estimate of $\boldsymbol{x}(\frac{k}{m})$. One step of Euler's integration method yields

$$\begin{aligned} \tilde{\boldsymbol{x}}_{k+1} &= \boldsymbol{x}_k + \frac{1}{m}\frac{d\boldsymbol{x}}{dt}\left(\frac{k}{m}\right) \\ &= \boldsymbol{x}_k - \frac{1}{m}\left(\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}}\right)^{-1}\Big|_{(x_k, k/m)} \frac{\partial \boldsymbol{h}}{\partial t}\Big|_{(x_k, k/m)}. \end{aligned} \tag{10}$$

The obtained $\tilde{\boldsymbol{x}}_{k+1}$ is an estimate for the ordinate of the point on the solution curve $\boldsymbol{x}(t)$ at $t = \frac{k+1}{m}$ in Figure 5, in other words, for the root of the system $\boldsymbol{h}(\boldsymbol{x}, (k+1)/m) = \boldsymbol{0}$. It is then polished using Newton's method over this system as follows. Set $\boldsymbol{y}_0 = \tilde{\boldsymbol{x}}_{k+1}$. The iteration formula is

$$\boldsymbol{y}_{l+1} = \boldsymbol{y}_l - \left(\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{x}}\right)^{-1}\Big|_{(y_l, (k+1)/m)} \boldsymbol{h}|_{(y_l, (k+1)/m)}. \tag{11}$$

Assign the polished value to $\boldsymbol{x}_{k+1}$.

Figure 5 illustrates the working of the algorithm. Starting at a zero $\boldsymbol{x}_0$ of $\boldsymbol{g}(x)$, the drawn solution path is part of the zero set $\boldsymbol{h}^{-1}(\boldsymbol{0})$ of $\boldsymbol{h}(\boldsymbol{x}, t)$. It is referred to as a *homotopy curve*. The algorithm is essentially a predictor-corrector method. Euler's integration (10) serves as the predictor to yield a new approximate point along the homotopy curve. Newton's method (11) serves as the corrector to bring this new point very close to the curve through an iterative polishing process.

EXAMPLE 3. Let us illustrate the algorithm over the system (6). By Bernstein's theorem, it has exactly four non-zero complex roots. Using the following start system:

$$\begin{aligned} g_1(x, y) &= x^3 - 1, \\ g_2(x, y) &= y^3 - 1, \end{aligned}$$

we define

$$\begin{aligned} h_1(x, y, t) &= t f_1(x, y) + (1 - t) g_1(x, y), \\ h_2(x, y, t) &= t f_2(x, y) + (1 - t) g_2(x, y), \end{aligned}$$

with $t \in [0, 1]$.

The system $\boldsymbol{g}(x, y) = 0$ has a total of nine complex zeros:

$$(x, y) = \left(\cos \frac{2j\pi}{3} + i \sin \frac{2j\pi}{3}, \cos \frac{2k\pi}{3} + i \sin \frac{2k\pi}{3} \right), \qquad j, k = 0, 1, 2.$$

Starting with these nine zeros, the PHCpack software [12] implementing the homotopy continuation method finds all four roots:

$$\begin{array}{ll} (-0.0316950027102798 + 0.181213765826737i, & -0.110462903286809 - 0.204565439823804i), \\ (-0.0316950027102798 - 0.181213765826737i, & -0.1.10462903286809 + 0.204565439823804i), \\ (0.411875744374350 - 0.120485325502200i & 0.167490014536420 - 0.332145613080015i), \\ (0.411875744374350 + 0.120485325502200i & 0.167490014536420 + 0.332145613080015i). \end{array}$$

3.1 Completeness and Start System

The fundamental result on homotopy continuation is from [4, 6]. Choose the start system $\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{0}$ which consists of equations

$$g_k(x_1, \ldots, x_n) = x_k^{d_k+1} - c_k, \qquad 1 \leq k \leq n,$$

where $c_1, \ldots, c_n$ are complex numbers. By separating the real part of a complex number from its imaginary part, we can represent $\nabla \boldsymbol{h} = (\partial \boldsymbol{h}/\partial \boldsymbol{x}, \partial \boldsymbol{h}/\partial t)$ as a $2n \times (2n + 1)$ matrix.

Theorem 3 (Garcia and Zangwill, 1975) *If, for all $\boldsymbol{x} \in \mathbb{C}^n$, $t \in [0, 1]$,*

1. *$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$ implies that $\nabla \boldsymbol{f}(\boldsymbol{x})$ is non-singular, and*
2. *$\boldsymbol{h}(\boldsymbol{x}, t) = \boldsymbol{0}$ implies that $\nabla \boldsymbol{h}(\boldsymbol{x}, t)$ has rank $2n$,*

then all solutions to $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$ can be found by tracking the homotopy curves.

The choice of $\boldsymbol{g}$ leads to generation of some spurious homotopy paths whose total number $(d_1+1)(d_2+1)\cdots(d_n+1)$ exceeds the maximum number $d_1 d_2 \cdots d_n$ of zeros by Bézout's theorem.

Garcia also made the observation that with the alternate choice

$$g_k = x_k^{d_k} - c_k, \qquad 1 \le k \le n, \tag{12}$$

the algorithm will work in many cases.[2] The resulting system $\boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{0}$ produces $d = d_1 \cdots d_n$ paths, the same number as Bézout's upper bound on the number of isolated zeros of $\boldsymbol{f}(\boldsymbol{x})$. We hope that the homotopy paths $\boldsymbol{x}(t)$ starting at these d zeros of $\boldsymbol{g}$ will lead to all zeros of $\boldsymbol{f}$, as t increases from 0 to 1.

Generally, the start system $\boldsymbol{g}$ should be chosen to satisfy the following three properties [13].

- *Triviality*. The zeros of $\boldsymbol{g}$ are known.
- *Smoothness*. The zero set $\boldsymbol{h}^{-1}(\boldsymbol{0})$ of $\boldsymbol{h}(\boldsymbol{x}, t)$ consists of a finite number of smooth homotopy curves over $[0, 1]$.
- *Accessibility*. Every isolated zero of $\boldsymbol{h}(\boldsymbol{x}, 1) = \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{0}$ can be reached by some path originating at a zero of $\boldsymbol{h}(\boldsymbol{x}, 0) = \boldsymbol{g}(\boldsymbol{x}) = \boldsymbol{0}$.

The choice of $\boldsymbol{g}$ according to (12) satisfies all three properties. Under the choice, the system produces d paths. Assuming no path diverges, multiple of these paths may converge to one zero of $\boldsymbol{f}$ if $\boldsymbol{f}$ has less than d zeros.

3.2 Pitfalls

We use an example from [6] to show that singularities in homotopy paths can break the root finding method. The example is concerned with a univariate polynomial:

$$f(x) = -x^2 + 1.$$

Our first choice for g is

$$g(x) = x - c,$$

where c is a non-zero complex number. The choice yields

$$h(x, t) = (1-t)(x-c) + t(-x^2+1). \tag{13}$$

We solve this quadratic system algebraically and find that vanishing of $h(x, t)$ happens when either $t = 0$ and $x = c$, or

$$x = \frac{1}{2t}\left((1-t) \pm \sqrt{(1-t)^2 - 4ct(1-t) + 4t^2}.\right)$$

Take the limit of the above expressions:

$$\lim_{t\to 0} x = \begin{cases} \infty, & \text{if } + \text{ is selected;} \\ c, & \text{if } - \text{ is selected.} \end{cases}$$

The $+$ and $-$ signs respectively specify two branches of $h^{-1}(0)$ that are drawn in Figure 6(a). It is

[2] It was not a coincidence that in Example 3 we chose g_1 and g_2 to be polynomials of the same degrees as f_1 and f_2.

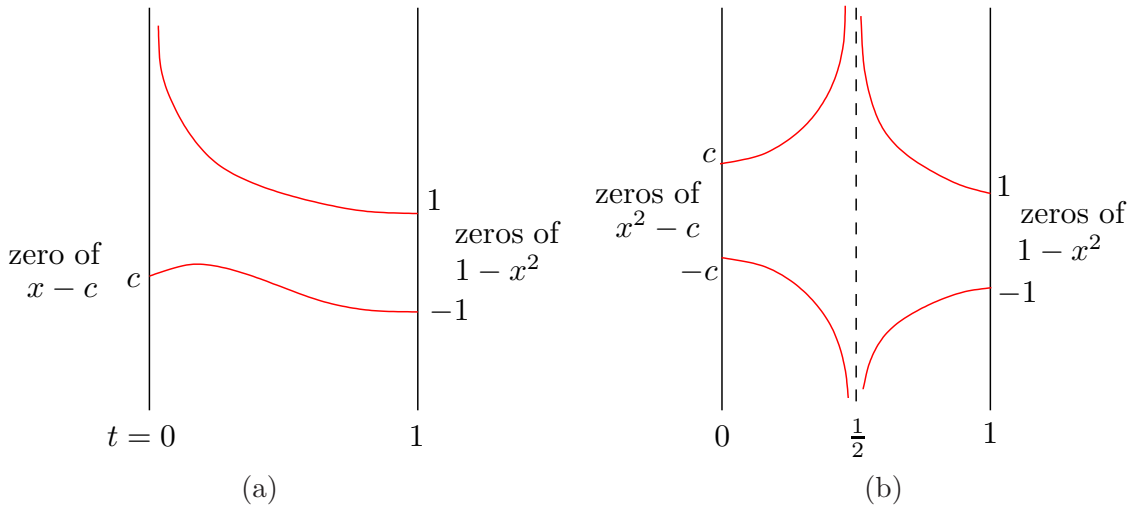

Figure 6: Graphs of the solution set $h^{-1}(0)$ with singularities: (a) system (13) and (b) system (14).

clear that the upper homotopy curve of the zero set will never be stepped on with the choice of g.

Now, let us consider a second choice:

$$g(x) = x^2 - c,$$

with $c \neq 1$. Here,

$$h(x,t) = (1-t)(x^2 - c) + t(-x^2 + 1). \tag{14}$$

It can be shown that four branches of $h^{-1}(0)$ exist as illustrated in Figure 6(b). No solution to $g(x) = 0$ will connect to a solution to $f(x) = 0$. All four branches become unbounded as $t \to \frac{1}{2}$. The partial derivative $\frac{\partial h}{\partial x} = 2x(1-t) - 2xt$ is zero at the singular point $t = \frac{1}{2}$. It cannot be inverted to compute $\frac{dx}{dt}$ according to (9). As a result, the two homotopy paths starting at $t = 0$ will not cross the line $t = \frac{1}{2}$. The homotopy continuation method will not find a single root of $\boldsymbol{f}$ with the choice of $\boldsymbol{g}$.

Overall, the homotopy continuation method is a powerful exhaustive solver. It is numerically stable, has global convergence, and can find multiple solutions.

References

[1] Eugene L. Allgower and Kurt Georg. *Numerical Continuation Methods: an Introduction*. Springer-Verlag, 1990.

[2] Max Buot and Donald Richards. Homotopy Continuation. Online PowerPoint notes, 2005.

[3] D. Bernstein. The number of roots of a system of equations. *Functional Analysis and Its Applications*, 9:183–185, 1975.

[4] C. B. Garcia and W. I. Zangwill. Finding all solutions to polynomial systems and other systems of equations. *Mathematical Programming*, 16:159-176, 1979.

[5] D. Lazard. GrÖbner bases, Gaussian elimination and resolution of systems of algebraic equations. *Proceedings of the European Computer Algebra Conference on Computer Algebra*, pp. 146–156, 1983.

[6] Alexander P. Morgan. A method for computing all solutions to systems of polynomial equations. *ACM Transactions on Mathematical Software*, 9(1):1–17, 1983.

[7] Alexander Morgan. Solving Polynomial Systems Using Continuation for Engineering and Scientific Problems. Prentice-Hall, Inc., 1987.

[8] Andrew J. Sommese and Charles W. Wampler. *The Numerical Solution of Systems of Polynomials Arising in Engineering and Science*. World Scientific Publishing Co. Pte. Ltd., 2005.

[9] S. Petitjean. Algebraic geometry and computer vision: polynomial systems, real and complex roots. *Journal of Mathematical Imaging and Vision*, 10:1–32, 1999.

[10] B. Sturmfels. Polynomial equations and convex polytopes. *The American Mathematical Monthly*, 105(10):907–922, 1998.

[11] B. L. van der Waerden. *Algebra*, vol. 1. Springer-Verlag, 1991.

[12] Jan Verschelde. Algorithm 795: PHCpack: a general-purpose solver for polynomial systems by homotopy continuation. *ACM Transactions on Mathematical Software*, 25(2):251–276, 1999. `http://homepages.math.uic.edu/~jan/PHCpack/phcpack.html`.

[13] Weiran Wang. A Review on Homotopy Continuation Methods for Polynomial Systems. `https://eng.ucmerced.edu/people/wwang5/misc/phc.pdf`.

[14] W.-T. Wu. On the decision problem and the mechanization of theorem-proving in elementary geometry. In W. Bledsoe and D. Loveland, ed., *Automated Theorem Proving: After 25 Years*, pp. 213-234. American Mathematical Society, 1983.