Homotopies and polynomial system solving I:
Basic Principles

Ilias S. Kotsireas
University of Western Ontario
Computer Science Department
Ontario Research Centre for Computer Algebra
London, Ontario, N6A 5B7 Canada
http://www.orcca.on.ca/ilias
Ilias.Kotsireas@orcca.on.ca

Abstract

We present a survey of some basic ideas involved in the use of homotopies for solving systems of polynomial equations. These ideas are illustrated with many concrete examples. An introductory section on systems of polynomial equations and their solutions contains some necessary terminology that will be used in the sequel. We also describe a general algorithm to solve polynomial systems of \( n \) equations in \( n \) unknowns using homotopies. A Maple V implementation of the algorithm as well as a few accompanying Maple 6 worksheets are publicly available from the author’s web page.

1 Introduction

Homotopy methods have a very wide range of applicability. They can be used to compute approximately fixed points of smooth functions for example. We will show how to use homotopy methods to solve square\(^1\) systems of polynomial equations. In particular, homotopy methods can be used to locate all geometrically isolated solutions of a square polynomial system. The presentation of the material is self-contained and is based in part on the works [AG90] and [Mor87]. The aim of the paper is to introduce the basic ideas of homotopy methods as well as their application to polynomial system solving. We plan to treat more advanced aspects of the theory in forthcoming papers. A Maple V implementation of the algorithm is publicly available from the author’s web page. We plan to update the implementation for subsequent versions of Maple. A few accompanying Maple 6 worksheets are also available from the author’s web page.

2 Polynomials and Polynomial Systems

In this section we summarize some notions about polynomials and polynomial systems. We also introduce some terminology from Algebraic Geometry.

\(^1\)The number of equations is equal to the number of unknowns
2.1 Generalities

Let \( P(z) = z^n + a_{n-1}z^{n-1} + \ldots + a_1z + a_0 \) be a monic polynomial of degree \( n \) with complex coefficients \((a_i \in \mathbb{C} \text{ for } i = 0, 1, \ldots, n-1)\). The Fundamental Theorem of Algebra, states that \( P(z) \) has exactly \( n \) roots in \( \mathbb{C} \), counting multiplicities. Thus, \( P(z) \) can be factorized as follows:

\[
P(z) = (z - \alpha_1) \ldots (z - \alpha_n)
\]

where the \( n \) (not necessarily distinct) complex numbers \( \alpha_1, \ldots, \alpha_n \) are its \( n \) roots \((P(\alpha_i) = 0, i = 1, \ldots, n)\).

The multiplicity of a root \( \alpha_i \) of \( P(z) \) is defined as the number of times that the factor \((z - \alpha_i)\) appears in (1).

The (first) derivative of the polynomial \( P(z) \) is defined as:

\[
P'(z) = nz^{n-1} + (n-1)a_{n-1}z^{n-2} + \ldots + a_1.
\]

Derivatives of higher order are defined similarly. The notation \( P^{(i)}(z) \) is employed to denote the \( i \)-th derivative of \( P(z) \), for any non-negative integer \( i \). The convention \( P^{(0)}(z) = P(z) \) is also used. Derivatives give another way to define the multiplicity of a root. A root \( \alpha_i \) of \( P(z) \) is of multiplicity \( k \), when \( P^{(0)}(\alpha_i) = \ldots = P^{(k-1)}(\alpha_i) = 0 \) and \( P^{(k)}(\alpha_i) \neq 0 \).

When we pass from the case of one polynomial in one variable, to the case of many polynomials in many variables, the above notions do not generalize in any evident or easy way. In fact the study of these notions in the general case is the subject of a whole branch of Mathematics, called Algebraic Geometry.

We will present some basic notions from Algebraic Geometry, that we need for our purposes. For further reading on these topics as well as for thorough introductions to Algebraic Geometry, the reader is referred to [Ken38], [Ful89], and [Har97].

Let \( P_{i}(z_1, \ldots, z_n), i = 1, \ldots, n \) be a system of \( n \) polynomials with complex coefficients in \( n \) We are interested in the set of common zeros in \( \mathbb{C}^n \) of the polynomials \( P_i, i = 1, \ldots, n \), that is the set of solutions of the polynomial system:

\[
\begin{cases}
P_1(z_1, \ldots, z_n) = 0 \\
\vdots \\
P_n(z_1, \ldots, z_n) = 0
\end{cases}
\]

Definition 2.1. The degree of the monomial term \( a z_1^{r_1} \ldots z_n^{r_n} \) is the sum of the exponents: \( r_1 + \ldots + r_n \).

Definition 2.2. The degree \( d_i \) of the polynomial \( P_i \) is the maximum of the degrees of its monomial terms.

Definition 2.3. The homogeneous part of the polynomial \( P_i \) is the polynomial \( \tilde{P}_i \) obtained from \( P \) by deleting all terms of degree less than \( d_i \).

Definition 2.4. The homogenization of the polynomial \( P_i(z_1, \ldots, z_n) \), of degree \( d_i \), is a polynomial \( \tilde{P}_i(z_1, \ldots, z_n, z_{n+1}) \) in \( n+1 \) variables defined as:

\[
\tilde{P}_i(z_1, \ldots, z_n, z_{n+1}) = z_{n+1}^{d_i} P_i \left( \frac{z_1}{z_{n+1}}, \ldots, \frac{z_n}{z_{n+1}} \right).
\]

The polynomial \( \tilde{P}_i \) is obtained from \( P_i \) by multiplying each term of \( P_i \) with an appropriate power of \( z_{n+1} \) such that the degree of the term becomes \( d_i \).
The notions of homogeneous part and homogenization are defined for systems of the form (2), just by applying definitions 2.3 and 2.4 to each polynomial separately.

**Example 2.1.** Consider the polynomial system:

\[
\begin{align*}
f &= \begin{cases} \frac{z_1^3}{z_1^2} - z_1 = 0 \\
\frac{z_1^2}{z_1^2} + 1 = 0 \end{cases}
\end{align*}
\]  

(3)

The system (3) has the homogeneous part \( \hat{f} \), and the homogenization \( \tilde{f} \):

\[
\begin{align*}
\hat{f} &= \begin{cases} z_1^3 = 0 \\
\frac{z_1^2}{z_1^2} = 0 \end{cases} \\
\tilde{f} &= \begin{cases} \frac{z_1^3}{z_3^2} - \frac{z_3^2}{z_3^2}z_1 = 0 \\
\frac{z_1^2}{z_1^2} + \frac{z_3^4}{z_3^4} = 0 \end{cases}
\end{align*}
\]  

(4)

The homogenization of system (3) has one more variable, \( z_3 \) than the original system.

### 2.1.1 Solutions of Polynomial Systems

We define here the important concepts of geometrically isolated and singular solutions as well as the notion of solutions at infinity. We also mention a multivariate analog of the Fundamental Theorem of Algebra, Bezout’s theorem.

**Definition 2.5.** The Jacobian matrix of the system (2) is defined as:

\[
J = \begin{pmatrix}
\frac{\partial P_1}{\partial z_1} & \cdots & \frac{\partial P_1}{\partial z_n} \\
\vdots & \ddots & \vdots \\
\frac{\partial P_n}{\partial z_1} & \cdots & \frac{\partial P_n}{\partial z_n}
\end{pmatrix}
\]

**Definition 2.6.** A solution \((z_1, \ldots, z_n)\) of the system (2) is called singular, when the determinant of the Jacobian matrix evaluated at the solution vanishes.

**Definition 2.7.** A solution \((z_1, \ldots, z_n)\) of the system (2) is called geometrically isolated when there is a ball centered at the solution which contains no other solution.

When \( n = 1 \) a singular solution is just a solution where the derivative vanishes. Geometrically, a singular solution means that the tangents to the curves intersecting at the solution are equal. Solutions that is not geometrically isolated, are singular. Geometrically isolated solutions can be either singular or non-singular. The following theorem clarifies further the relationship between singular solutions and multiplicities.

**Theorem 2.1.** A solution of the system (2) has multiplicity greater than one if and only if it is singular.

The theorem implies that nonsingular solutions have multiplicity one and that solutions that have multiplicity one are nonsingular.
Example 2.2. Consider the polynomial system:

\[
\begin{align*}
& x^2 + y^2 - 25 = 0 \\
& x^2 - y - 5 = 0
\end{align*}
\] (5)

The system has three solutions, \((x, y) = (3, 4), (-3, 4), (0, -5)\). The determinant of the Jacobian is \(-2x(2y + 1)\). The first two solutions are nonsingular, because the determinant of the Jacobian does not vanish. The third solution is singular, because the determinant of the Jacobian vanishes at this solution. All three solutions are geometrically isolated.

Example 2.3. Consider the polynomial system:

\[
\begin{align*}
& x^2 - 25 = 0 \\
& xy - x - 5y + 5 = 0
\end{align*}
\] (6)

The system has the solution \((x, y) = (-5, 1)\) and the infinity of solutions \((x, y) = (5, y)\), which represents a straight line \(^2\). The determinant of the Jacobian is \(2x(x - 5)\). The first solution is nonsingular and geometrically isolated. The infinity of solutions \((5, y)\) are all singular and non geometrically isolated.

Let \(P\) denote the polynomial system (2). If \([w_0, w_1, \ldots, w_n]\) is a solution of the homogenization \(\bar{P}\), then for every \(\zeta \in \mathcal{C}, [\zeta w_0, \ldots, \zeta w_n]\) is also a solution of \(\bar{P}\). From the point of view of Algebraic Geometry the point \([w_0, w_1, \ldots, w_n]\) and the point \([\zeta w_0, w_1, \ldots, \zeta w_n]\) (for arbitrary \(\zeta \in \mathcal{C}\)) are identified and are considered as a point in the complex projective space \(\mathbb{P}\).

The solutions of a polynomial system and its homogenization are closely related. In fact, the solutions of the homogenized system either correspond to solutions of the initial system, or give rise to solutions at infinity of the initial system. More precisely, there are two cases to consider:

1. The solution \([w_0, w_1, \ldots, w_n]\) of \(\bar{P}\) intersects the hyperplane \(z_0 = 0\) transversely (that means that without loss of generality we can take \(w_0 = 1\)). This corresponds to a solution \([w_0, w_1, \ldots, w_n]\) of \(P\). Conversely, each solution \([w_1, \ldots, w_n]\) of \(P\) corresponds to a solution \([1, w_1, \ldots, w_n]\) of \(\bar{P}\).

\(^2\)This is clear by noticing that the first equation factorizes as \((x - 5)(x + 5)\) and that the second equations factorizes as \((x - 5)(y - 1)\)
2. The solution \([w_0, w_1, \ldots, w_n]\) of \(\tilde{P}\) lies in the hyperplane \(z_0 = 0\) (that means that \(w_0 = 0\)). This corresponds to a nontrivial solution \([w_1, \ldots, w_n]\) of the homogeneous part \(\tilde{P}\). Such solutions are called solutions of \(P\) at infinity.

In practice, the solutions at infinity of the original system \(f\), are the solutions of the corresponding homogeneous part \(\hat{f}\) whose first nonzero entry is equal to one. We make the above discussion more concrete with the following generic example of two equations in two variables.

**Example 2.4.** Let \(P(x, y) = ax^2 + by^2 + cxy + dx + ey + f\) and \(Q(x, y) = a'x^2 + b'y^2 + c'xy + d'x + e'y + f'\) be two generic polynomials of degree 2 in two variables \(x, y\) with real or complex coefficients \(a, b, c, d, e, f, a', b', c', d', e', f'\). For simplicity we suppose that at least the six coefficients \(a, b, c, d, e, f\) are non-zero. The homogeneous part of the system \(P(x, y) = 0, Q(x, y) = 0\) is:

\[
\hat{P}(x, y) = ax^2 + by^2 + cxy, \quad \hat{Q}(x, y) = a'x^2 + b'y^2 + c'xy. \quad (7)
\]

The nonzero solutions of (7), are called solutions at infinity of the original system. It is clear that if \((x_0, y_0)\) is a solution of (7), then \((rx_0, ry_0)\) is also a solution, for every \(r \in \mathbb{C}\). We are interested in solutions of (7), such that no two are multiples of each other. A way to achieve this is to look for solutions only in the form \((1, y)\) or \((0, 1)\), because any other solution of (7), will be a scalar multiple of a solution in one of these two forms.

**Example 2.5.** The systems (5) and (6) do not have solutions at infinity.

### 2.1.2 Bezout’s Theorem

In the case of a general system of \(n\) polynomial equations in \(n\) variables \(x_1, \ldots, x_n\),

\[
\begin{align*}
f_1(x_1, \ldots, x_n) &= 0 \\
f_2(x_1, \ldots, x_n) &= 0 \\
& \vdots \\
f_n(x_1, \ldots, x_n) &= 0
\end{align*}
\quad (8)
\]
a generalization of the Fundamental Theorem of Algebra is Bezout’s Theorem.

Let $d_i = \deg f_i$ for $i = 1, \ldots, n$. The total degree $d$ of the system (8) is defined as the product of the degrees:

$$d = d_1 d_2 \ldots d_n.$$  

**Theorem 2.2 (Bezout).** For the system of polynomial equations (8) we have:

1. The total number of geometrically isolated solutions and solutions at infinity is no more than $d$.

2. If the system has neither an infinite number of solutions nor an infinite number of solutions at infinity, then it has exactly $d$ solutions and solutions at infinity, counting multiplicities.

**Example 2.6.** The system (5) is of total degree $d = 4$ and the second case of Bezout’s theorem applies. In this case we infer from Bezout’s theorem that the singular solution is of multiplicity 2. The system (6) is of total degree $d = 4$ and the first case of Bezout’s theorem applies. Indeed, since the system does not have solutions at infinity, the total number of geometrically isolated solutions and solutions at infinity is 1.

## 3 Homotopies

The word homotopy comes from the greek word ομοτοπία and expresses etymologically the notion of continuous deformation of one object to another. We illustrate the use of homotopies as a means of continuously deforming a function into a target function with two examples. Plotting the graphs of the functions as well as the intermediate functions is very instructive and helps gaining a better understanding of the situation.

**Example 3.1.** Consider two univariate polynomials $p_1$ and $q_1$ of the same degree:

$$p_1(x) = x^2 - 3, \quad q_1(x) = -x^2 - x + 1.$$  

Define the convex homotopy $H : \mathbb{C} \times [0, 1] \mapsto \mathbb{C}$ with:

$$H(x, \lambda) = (1 - \lambda)p_1(x) + \lambda q_1(x).$$

In geometric terms, the homotopy $H$ provides us with a continuous deformation from $p_1$ (which is obtained for $\lambda = 0$ by $H(x, 0)$) to $q_1$ (which is obtained for $\lambda = 1$ by $H(x, 1)$).

The graph in Figure 3 shows $H(x, \lambda)$ for several values of $\lambda$ in the $x$-interval $[-2, 1.5]$. It should be noted that for $\lambda = 1/2$, we have that $H(x, \lambda) = -\frac{1}{2}x - 1$ is the equation of a straight line passing through $(0, -1)$.

**Example 3.2.** Consider two univariate polynomials $p_2$ and $q_2$ of different degrees:

$$p_2(x) = x^2 - 1, \quad q_2(x) = (x^2 - \frac{1}{4})(x^2 - 4).$$

Define the convex homotopy $H : \mathbb{C} \times [0, 1] \mapsto \mathbb{C}$ with:

$$H(x, \lambda) = (1 - \lambda)p_2(x) + \lambda q_2(x).$$

The graph in Figure 4 shows $H(x, \lambda)$ for several values of $\lambda$ in the $x$-interval $[-2.5, 2.5]$.

\footnote{Another way to see that the multiplicity of the singular solution is two, is shown in the example following the description of the general homotopy algorithm.}
Figure 3: The homotopy $H$ for the polynomials $p_1$ and $q_1$

Figure 4: The homotopy $H$ for the polynomials $p_2$ and $q_2$
We will now look at how we can use the notion of homotopy to solve a simple polynomial equation of degree 2 in one variable.

Consider the polynomial equation of degree 2:

$$q(x) = x^2 + 8x - 9 = 0.$$  \hfill (9)

We remark that we can immediately solve by inspection the simpler equation

$$p(x) = x^2 - 9 = 0.$$  \hfill (10)

which is obtained by (9) by deleting the middle term. As we have seen before, homotopies provide us with a way of deforming \( p(x) \) into \( q(x) \) by means of the homotopy equation \( H(x, \lambda) = (1 - \lambda) p(x) + \lambda q(x) \). In our case, \( H(x, \lambda) = x^2 + 8\lambda x - 9 \). Can we also deform the solutions of \( p(x) = 0 \) into the solutions of \( q(x) = 0 \)? A natural way to do that would be to solve the equation \( H(x, \lambda) = 0 \) for many values of the parameter \( \lambda \in [0, 1] \) and look at the sequences of values generated by the two initial solutions of (10).

![Figure 5: The two paths converging to the solutions of equation (9)](image)

To implement this program we choose a small \( \varepsilon > 0 \) (we take \( \varepsilon = 0.001 \) here) and solve the equation \( H(x, \varepsilon) = 0 \) using Newton’s method with initial conditions the two initial solutions of (10). We use the resulting two values as initial conditions to solve \( H(x, 2\varepsilon) = 0 \) using Newton’s method and we repeat this process until we reach \( H(x, 1) = 0 \) which is the target equation (9). Since the parameter \( \lambda \) takes its values in the interval \([0, 1]\), and we have taken \( \varepsilon = 0.001 \), we will need 100 iterations to reach \( H(x, 1) \). Figure 5 shows the two paths generated by this process. The upper path emanates from the initial solution \( x = 3 \) of (10) and terminates to the solution \( x = 1 \) of (9) and the lower path emanates from the initial solution \( x = -3 \) of (10) and terminates to the solution \( x = -9 \) of (9).

In spite of the success of the homotopy approach in the previous example, one should be careful when using homotopy methods to solve polynomial systems. One of the first things that can go wrong is path crossing. Two paths may converge to the same solution of the target equation. This can happen when we
encounter singular paths for example. These are paths with the property that for specific values $\lambda_s$ and $x_s$ we have

$$H(x_s, \lambda_s) = \frac{dH}{dx}(x_s, \lambda_s).$$  \hspace{1cm} (11)

Geometrically, the two paths cross and then they merge into one path converging into one solution. Thus, the other solution is lost. One way to remedy this problem and recover the lost solution is to use another homotopy equation. Suppose that we want to solve the equation $p_s(x) = x^2 + ax + b = 0$. Define the homotopy

$$H_s(x, \lambda) = (1 - \lambda)(x^2 - q^2) + \lambda p_s(x).$$  \hspace{1cm} (12)

We see that $H_s(x, \lambda) = x^2 + a\lambda x + (\lambda b - (1 - \lambda)q^2)$ and that we still have the desired property $H_s(x, 1) = p_s(x)$. Moreover, the initial system is now the easy-to-solve system $H_s(x, 0) = x^2 - q^2$. Taking $q$ to be a random complex number we are sure to circumvent the difficulty that a path becomes singular, because the resultant

$$\mathcal{R} \left( H_s(x, \lambda), \frac{dH_s}{dx}(x, \lambda) \right) = -a^2\lambda^2 + (4q^2 + 4b)\lambda - 4q^2$$

will generically be different than zero and thus (11) will never occur.

### 3.1 Generalities

Homotopy (or continuation) methods are a powerful method for solving polynomial systems of dimension zero, that is polynomial systems with a finite number of solutions. More precisely, they are used to compute approximations to all geometrically isolated solutions of such systems. In fact homotopy methods are used for solving more general systems and also for locating fixed points of smooth functions in general.

The basic ideas employed by homotopy methods to solve a system $f(x) = 0$, where $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a smooth map, can be described succinctly in the following steps:

1. Define a polynomial system $g(x) = 0$, where $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a smooth map, with known and easily computable solutions.

2. Define a smooth homotopy map $H : \mathbb{R}^n \times [0, 1] \rightarrow \mathbb{R}^n$ with $H(x, 0) = g(x)$ and $H(x, 1) = f(x)$.

3. Trace an implicitly defined curve $c(s) \in H^{-1}(0)$, from a starting point $(x_0, 0)$ such that $H(x_0, 0) = g(x_0) = 0$, to an end point $(x_1, 1)$ such that $H(x_1, 1) = f(x_1) = 0$.

There are several questions arising from the above formulation of the general homotopy method for solving polynomial systems:

1. How to define the homotopy map $H(x, \lambda)$?

2. When does the curve $c(s)$ exists, is smooth and the point $(x_1, 1)$ belongs to its range?

3. If the curve $c(s)$ exists, when it is assured that it will intersect the target homotopy level $\lambda = 1$ in a finite length?
4. What are the available methods to numerically trace such a curve?

The first question is usually answered by using the convex homotopy

\[ H(x, \lambda) = (1 - \lambda)g(x) + \lambda f(x), \] (13)

which satisfies the requirements \( H(x, 0) = g(x) \) and \( H(x, 1) = f(x) \). Other homotopy maps have also been proposed to deal with particular types of problems.

The second question is answered by the Implicit Function Theorem. If \((x_0, 0)\) is a regular zero point of \( H \) then a curve \( c(s) \in H^{-1}(0) \) with \( c(0) = (x_0, 0) \) exists on some open interval around zero. Moreover, if zero is a regular value of \( H \), then \( c(s) \) is diffeomorphic to a circle or the real line.

The third question is answered by imposing some boundary condition which guarantees that the curve will reach the target homotopy level \( \lambda = 1 \) in a finite length. Such conditions are usually studied with respect to some particular problem.

The fourth question gives rise to more complicated considerations. When the curve \( c(s) \) can be parameterized with respect to the homotopy parameter \( \lambda \), then we choose a sufficiently small increment \( \Delta \lambda \) and solve successively the systems \( H(x, 0), H(x, \Delta \lambda), \ldots, H(x, 1) \) by using the solution to each system as a starting point for the next. It is clear that if we use a Newton iteration to solve the systems and since the starting points will already be good approximations to the solutions, this iteration will generally converge.

4 A General Homotopy Algorithm

Now we will present a general homotopy algorithm, along the lines of the answer to the fourth question as it is depicted at the end of paragraph 3.1. Then we will describe our implementation of the algorithm in Maple and finally we will give a practical application of the algorithm in determining the multiplicity of a singular solution of a polynomial system.

4.1 Description of the algorithm

Let \( f(x) = 0 \), denote a system of \( n \) polynomial equations with complex coefficients, in \( n \) unknowns. Thus in expanded form we have: \( f_1(x_1, \ldots, x_n) = 0, \ldots, f_n(x_1, \ldots, x_n) = 0 \). Let \( d_i \) denote the total degree of equation \( f_i(x_1, \ldots, x_n) = 0 \) for \( i = 1, \ldots, n \). Let \( d = d_1 \ldots d_n \) denote the total degree of the system \( f(x) = 0 \). Finally let us denote by \( \varepsilon \) the accuracy.

With these notations we are ready to state a general homotopy algorithm.

INPUT: The system \( f(x) = 0 \), the accuracy \( \varepsilon \) and \( 2n \) random complex numbers.

OUTPUT: The \( d \) paths that converge to geometrically isolated solutions of the system \( f(x) = 0 \).
1.- Choose randomly the complex numbers \( p_1, \ldots, p_n \) and \( q_1, \ldots, q_n \)
2.- Form the auxiliary initial system \( g(x) = 0 \) with \( g_i(x) = p_i^{d_i} x^{d_i} - q_i^{d_i} \) for \( i = 1, \ldots, n \)
3.- Compute the \( d \) roots \( x^0 \) of \( g(x) = 0 \), by the formula:
   \[
   x^0 = \left( r_1(d_1, k_1) \frac{q_1}{p_1}, \ldots, r_n(d_n, k_n) \frac{q_n}{p_n} \right)
   \]
   where \( k_i = 1, \ldots, d_i \) for \( i = 1, \ldots, n \) and
   \[
   r_i(d_i, k_i) \text{ are the } d_i\text{-th roots of unity } r_i(d_i, k_i) = \cos \frac{2\pi k_i}{d_i} + i \sin \frac{2\pi k_i}{d_i}, k_i = 1, \ldots, d_i
   \]
4.- Define the homotopy map \( H: \mathbb{A}^n \times [0, 1] \rightarrow \mathbb{A}^n \) by
   \[
   H(x, \lambda) = (1 - \lambda)g(x) + \lambda f(x).
   \]
5.- Set \( s = \left\lfloor \frac{1}{\varepsilon} \right\rfloor \) the number of steps taken on each path.
6.- For each of the \( d \) roots \( x^0 \), generate a path towards a solution of \( f(x) = 0 \) as follows:
   - Use the solution computed at 3.- as an initial condition to
     solve the system \( H(x, \varepsilon) = 0 \) by a Newton iteration.
   - Use the computed solution of \( H(x, \varepsilon) = 0 \) as an initial condition
     to solve the system \( H(x, 2\varepsilon) = 0 \) by a Newton iteration.
   - 
   - Use the computed solution of \( H(x, (s-1)\varepsilon) = 0 \) as an initial condition
     to solve the system \( H(x, s\varepsilon) (== H(x, 1)) = 0 \) by a Newton iteration.
   - Return the path (set of all computed solutions)

A general homotopy algorithm

The above algorithm will always produce \( d \) paths that will converge to geometrically isolated solutions of the system \( f(x) = 0 \). Each path can be viewed as a set containing \( s \) elements of the form \( (\lambda, [x_1(\lambda), \ldots, x_n(\lambda)]) \), where the values of \( \lambda \) range from 0 to 1 and the \( x_i \)'s are the solution of the corresponding system \( H(x, \lambda) = 0 \).

In the case \( n = 1 \), that is when we are solving a univariate equation, and if the solutions to the intermediate systems are real, then we can actually visualize the paths by plotting them, using \( \lambda \) as the independent variable. For higher values of \( n \) this nice geometric interpretation is no longer available.

The algorithm finds geometrically isolated solutions of the system \( f(x) = 0 \) regardless of their multiplicity. Actually the algorithm provides us with a way of resolving the multiplicity of a geometrically isolated solution \( x \) of the system \( f(x) = 0 \). Perturb \( f \) by adding arbitrarily small complex numbers to each coefficient of \( f \) (including the zero coefficients) in such a way that the perturbed system has only nonsingular solutions. The perturbed system has \( m \) solutions arbitrarily close to \( x_0 \). The number \( m \) is defined as the multiplicity of \( x_0 \).

4.2 Implementation of the algorithm

The algorithm described in the previous paragraph has been implemented in the Computer Algebra System Maple ([CGG+91]). The program handles a system of \( n \) polynomial equations in \( n \) unknowns. If the total degree of the system is \( d \), then the program generates \( d \) paths that will converge to geometrically isolated solutions of the system. Due to the fact that we use the convex homotopy map (13), it is guaranteed that the paths will not cross, or exhibit any other erratic behaviors, like for example diverging to infinity, bifurcating, backtracking, exploding to a hypersurface, or spiraling.
The initial starting values $p_1, \ldots, p_n$ and $q_1, \ldots, q_n$ might be complex or real and they are chosen randomly. If all these values are real, then the paths might be real or complex. If at least one of the values is complex, then the paths will be complex and the real roots are identified as complex numbers with very small imaginary parts. In the case or real paths the program contains routines for plotting them. In the case of complex paths the program also contains some plotting facilities. The Newton iteration algorithm used is a classical one. Since the initial value for the Newton iteration is already a good approximation of the solution, convergence is achieved usually in a few steps.

### 4.3 Example: Resolution of multiplicities

Let us show how to use our implementation of the homotopy algorithm to resolve the multiplicity of the singular solution $(x, y) = (0, -5)$ of the system (5).

We choose 12 random integers $a_{11}, b_{11}, c_{11}, d_{11}, e_{11}, f_{11}, a_{22}, b_{22}, c_{22}, d_{22}, e_{22}, f_{22}$ from the interval $[0, 1]$ and for $k$ taking the values $4, 8, 16, 32, 64$ we define the following perturbed systems of the system (5):

$$P_k = \begin{cases} 
(1 + a_{11} 10^{-k})x^2 + (1 + b_{11} 10^{-k})y^2 + c_{11} 10^{-k}xy + d_{11} 10^{-k}x + e_{11} 10^{-k}y + (f_{11} 10^{-k} - 25) = 0 \\
(1 + a_{22} 10^{-k})x^2 + b_{22} 10^{-k}y^2 + c_{22} 10^{-k}xy + d_{22} 10^{-k}x + (e_{22} 10^{-k} - 1)y + (f_{22} 10^{-k} - 5) = 0
\end{cases}$$

The perturbed systems $P_4, P_8, P_{16}, P_{32}, P_{64}$ all have 4 nonsingular solutions each. Moreover, each of these five perturbed systems has two solutions close to the singular solution $(x, y) = (0, -5)$ of the system (5). That means that the solution $(x, y) = (0, -5)$ of the system (5), is of multiplicity 2. This result is in agreement with the Bezout count of the roots for system (5).

### 5 Improvements of the algorithm

In this section we mention some ways to improve the general algorithm presented in the previous section. The main part of this section focuses on more efficient ways to follow the path that leads to a solution.

In the general algorithm we used a natural but rather naive approach to follow the path. Indeed the approach consisted in fixing a certain increment $\Delta t$, setting $x_0$ to be a solution of the initial auxiliary system $g(x) = 0$ (corresponding to $H(x, 0) = 0$), and solving the system $H(x, \Delta t) = 0$ with Newton’s method using $x_0$ as a starting point. Then $x_0$ is updated to the solution $x^*$ of this new system and we solve $H(x, 2\Delta t) = 0$ with Newton’s method using $x_0$ as a starting point. We continue to increment $\lambda$ by $\Delta t$ and update $x_0$ until we reach $\lambda = 1$ and then $x_0$ is a solution of the system $H(x, 1) = 0$ (corresponding to $f(x) = 0$).

This particular method of following the path is a special case of a general class of methods called ”predictor-corrector” methods. From the point of view of this new terminology, what we do is that if $(x_0, \lambda_0)$ is the current point on the path, then we predict that another nearby point on the path will be $(x_1, \lambda_0 + \Delta t)$ and using Newton’s method we correct it to $(x^*, \lambda_0 + \Delta t)$. This approach is very inefficient with respect to implementation issues, mainly because the actual geometry of the path is not taken into account. This is due to the fact that the step-size $\Delta t$ is fixed. A way to improve the efficiency of this approach, is to perform adjustments to the step-size $\Delta t$ according to the quality of the approximation in the Newton corrector step. If the norm of the residual of Newton’s method is sufficiently small, within a prescribed number of iterations, then the step-size may be increased, otherwise the step-size is decreased and the corrector step is repeated. Since increasing the step-size without precautions can be dangerous, a variation on the same theme would be to increase the step size only after a fixed number of Newton corrector steps with sufficiently small norm residuals have been recorded.
A more sophisticated method to follow the path is given by using the tangent vectors to the path. Suppose that \((x_0, \lambda_0)\) is the current point on the path. Then we solve \(H(x, \lambda_1) = 0\) with starting point \(x_1\), where

\[
(x_1, \lambda_1) = (x_0, \lambda_0) + \left( \frac{dx}{dt}(\lambda_0), 1 \right)
\]

The quantity \(\frac{dx}{dt}\) represents a tangent vector to the curve, pointing in the direction of increasing \(t\). We will now show how to actually compute \(\frac{dx}{dt}\).

Since we generate the paths by incrementing \(t\), it is natural to suppose that we have a parameterized path \(x(t)\) as \(t\) goes from 0 to 1. Then we will have:

\[
H(x(t), t) = 0, \quad \text{for } t \in [0, 1],
\]

since by definition the paths are solutions of \(H(x, t) = 0\). By differentiating (14) we have

\[
\frac{dH(x(t), t)}{dt} = 0,
\]

Applying the chain rule to equation (15) and using the notation \(H = (h_1, \ldots, h_n)\) we get

\[
\begin{bmatrix}
\frac{\partial h_1}{\partial x_1} & \ldots & \frac{\partial h_1}{\partial x_n} & \frac{\partial h_1}{\partial t} \\
\vdots & \ddots & \vdots & \vdots \\
\frac{\partial h_n}{\partial x_1} & \ldots & \frac{\partial h_n}{\partial x_n} & \frac{\partial h_n}{\partial t}
\end{bmatrix}
\begin{bmatrix}
\frac{dx_1}{dt} \\
\vdots \\
\frac{dx_n}{dt} \\
1
\end{bmatrix}
= \begin{bmatrix}
0 \\
\vdots \\
0 \\
0
\end{bmatrix}
\]

Writing equation (16) in block-matrix form, we have:

\[
\begin{bmatrix}
dh_x & dh_t
\end{bmatrix}
\begin{bmatrix}
\frac{dx}{dt}
\end{bmatrix}
= 0
\]

which gives the following linear system for \(\frac{dx}{dt}\):

\[
dh_x \frac{dx}{dt} + dh_t = 0
\]

The linear system (17) gives rise to the following system of differential equations of the first order:

\[
\frac{dx}{dt} = -dh_x^{-1}dh_t
\]

where the matrix \(dh_x\) is invertible because of the fact that we choose the arbitrary constants \(p_1, \ldots, p_n, q_1, \ldots, q_n\) randomly.
References


