

## Finding All Solutions to a System of Polynomial Equations

By Alden H. Wright

**Abstract.** Given a polynomial equation of degree  $d$  over the complex domain, the Fundamental Theorem of Algebra tells us that there are  $d$  solutions, assuming that the solutions are counted by multiplicity. These solutions can be approximated by deforming a standard  $n$ th degree equation into the given equation, and following the solutions through the deformation. This is called the homotopy method. The Fundamental Theorem of Algebra can be proved by the same technique.

In this paper we extend these results and methods to a system of  $n$  polynomial equations in  $n$  complex variables. We show that the number of solutions to such a system is the product of the degrees of the equations (assuming that infinite solutions are included and solutions are counted by multiplicity)\*. The proof is based on a homotopy, or deformation, from a standard system of equations with the same degrees and known solutions. This homotopy provides a computational method of approximating all solutions. Computational results demonstrating the feasibility of this method are also presented.

**1. Introduction.** For  $k = 1, 2, \dots, n$ , let  $f_k(z_1, z_2, \dots, z_n)$  be a complex polynomial in the complex variables  $z_1, z_2, \dots, z_n$ . We are concerned with finding all solutions to the system  $\{f_k = 0\}$ . If a term of  $f_k$  is of the form:

$$a_1 z_1^{r_1} z_2^{r_2} \cdots z_n^{r_n},$$

we define the degree of this term to be  $r_1 + r_2 + \cdots + r_n$ , and we define the degree  $d_k$  of  $f_k$  to be the maximum of the degrees of its terms. In this paper we show how all solutions to the system  $f_k = 0$ ,  $k = 1, 2, \dots, n$ , can be computed numerically assuming that the solutions are discrete. We show that the system must have exactly  $\prod_{k=1}^n d_k$  solutions, if solutions are counted by multiplicity and infinite solutions are counted. (Infinite solutions are defined more precisely later in the paper.)

While in practice one usually wants all solutions to a system of equations that satisfy some conditions (such as finite real solutions), by finding all solutions with their multiplicities, one can be sure that all solutions with the desired properties have been found.

---

Received June 21, 1982; revised February 10, 1984.

1980 *Mathematics Subject Classification*. Primary 65H10, 14A04, 14A25; Secondary 12D10, 30G15.

*Key words and phrases*. Systems of nonlinear equations, homotopy methods, systems of polynomial equations.

\*This result can also be proved using the techniques of algebraic geometry. See, for instance, the sequence of papers of B. L. van der Waerden in *Mathematische Annalen*, Volumes 97 to 115 (1927-1938) or [10].

The method of solution used in this paper is the homotopy method. Another system of polynomial equations

$$Q_k(z_1, z_2, \dots, z_n) = 0, \quad k = 1, 2, \dots, n,$$

where  $Q_k$  has degree  $d_k$ , and the solutions to the system  $\{Q_k = 0\}$  are known, is chosen. Then a homotopy  $H: C^n \times [0, 1] \rightarrow C^n$  is defined by

$$H_k(z_1, \dots, z_n, t) = tf_k(z_1, \dots, z_n) + (1 - t)Q_k(z_1, \dots, z_n).$$

For each  $t \in [0, 1]$ ,  $H_k(z_1, \dots, z_n, t) = 0$ ,  $k = 1, \dots, n$ , is a system of polynomial equations. Thus, the homotopy continuously deforms the system  $\{Q_k = 0\}$  into the system  $\{f_k = 0\}$ . We show that if the  $Q_k$  are chosen correctly, then the  $\{H_k = 0\}$  system has  $\prod_{k=1}^n d_k$  distinct finite solutions for all  $t \in [0, 1]$ .

This means that the components of  $H^{-1}(0)$  can be characterized topologically as follows: First, a component may be a closed arc which intersects each slice of  $C^n \times \{t\}$  once. These components correspond to single roots of the system  $\{f_k = 0\}$ . Second, a component may consist of  $j$  arcs which meet in a single point of  $C^n \times \{1\}$ . This point of  $C^n \times \{1\}$  is a multiplicity  $j$  root of the system  $\{f_k = 0\}$ . Each slice of  $C^n \times \{t\}$ ,  $0 \leq t < 1$ , will intersect such a component in  $j$  points. Thirdly, a component may be a half-open arc which intersects each slice of  $C^n \times \{t\}$ ,  $0 \leq t < 1$ , in a single point which tends to infinity as  $t \rightarrow 1$ . Such a component corresponds to an infinite root. Infinite solutions can be described explicitly as points in complex projective space.

Each component of  $H^{-1}(0) \cap (C^n \times [0, 1])$  (which may be assumed to be a half-open arc), can be set up as the solution to a differential equation with a given initial value (the known solution of the system  $\{Q_k = 0\}$ ). These differential equations can be solved numerically to find the solutions to the system  $\{f_k = 0\}$ .

Previous papers using the homotopy method for system of polynomial equations include [2], [4], [5], [6], [7], [9], and [10]. This paper differs from earlier papers in its treatment of infinite roots. In [7], the equations  $Q_k = 0$  were taken to have degree one higher than the equations  $f_k = 0$  so as to prevent paths of  $H^{-1}(0)$  from going to infinity for finite roots. In [2], extra terms were added to the homotopy for the same reason. In this paper we view the problem over complex projective space, a compactification of  $C^n$  which allows an explicit representation of infinite roots. This paper shows that a relatively simple homotopy will give paths that do not go to infinity except for paths that go to infinite roots.

**2. Complex Projective Space.** Define an equivalence relation on  $C^{n+1} - \{0\}$  by letting  $x \sim y$  if  $x = \lambda y$  for some complex scalar  $\lambda \neq 0$ . Then complex projective space  $CP^n$  is  $C^{n+1} - \{0\}$  modulo this equivalence relation. We denote an element of  $CP^n$  by  $[z]$  where  $z = (z_0, z_1, \dots, z_n) \in C^{n+1} - \{0\}$ . Note that  $C^n$  is embedded in  $CP^n$  as the set of  $[z] = (z_0, z_1, \dots, z_n)$  such that  $z_0 \neq 0$ . Also, the set of  $[z]$  such that  $z_0 = 0$  corresponds to the set of "points at infinity".

Given  $f_k(z_1, z_2, \dots, z_n)$ , a complex polynomial of degree  $d_k$ , let  $P_k(z_0, z_1, z_2, \dots, z_n)$  be defined as follows: each term of  $P_k$  is obtained from the corresponding term of  $f_k$  by multiplying it by the power of  $z_0$  to bring the degree of the term up to  $d_k$ . Thus, a term of  $f_k$  of degree  $d$  is multiplied by  $z_0^{d_k-d}$ , and each term of  $P_k$  has degree  $d_k$ . Thus,  $P_k(\lambda z) = \lambda^{d_k} P_k(z)$ , and  $P_k$  takes all points of an

equivalence class of  $C^n - \{0\}$  to the same point. Then,  $P_k$  can be thought of as a map from  $CP^n$  to  $C$ , and  $\mathbf{P}$  (all  $n$  components  $P_k$  taken together) can be thought of as a map from  $CP^n$  to  $C^n$ . The map  $P_k$  is called the *homogenization* of the map  $f_k$ . The system

$$(1) \quad P_k(z_0, z_1, \dots, z_n) = 0, \quad k = 1, 2, \dots, n,$$

reduces to the system  $\{f_k = 0\}$  under the substitution  $z_0 = 1$ . Thus, the two systems can be considered to have the same set of roots in  $C^n$ .

**3. Defining the Homotopy.** In this section we define a system  $\{Q_k = 0\}$  of polynomial equations where each equation has the same degree as the corresponding equation of (1), and where the solutions are easily computable. Then, we define a homotopy which connects the new system to the given system  $\{P_k = 0\}$ .

For  $k = 1, 2, \dots, n$ , let

$$Q_k(\mathbf{z}, \mathbf{a}) = a_{k1}z_k^{d_k} - a_{k2}z_0^{d_k},$$

where  $\mathbf{z} = (z_0, z_1, \dots, z_n) \in C^{n+1}$ ,  $d_k$  is the degree of  $P_k$ , and  $\mathbf{a}$  is an  $n$  by 2 complex matrix with all  $a_{jk} \neq 0$ . Consider the system

$$(2) \quad Q_k(\mathbf{z}, \mathbf{a}) = 0 \quad \text{for } k = 1, 2, \dots, n.$$

For fixed  $\mathbf{a}$ , this system has  $d_1 d_2 \cdots d_n$  distinct solution classes in  $CP^n$  of the form  $[(1, z_1, z_2, \dots, z_n)]$ . These are obtained by letting each  $z_k$  be some  $d_k$ th root of  $a_{k2}/a_{k1}$ .

Let  $\mathbf{P}: C^{n+1} \rightarrow C^n$  be the function whose  $k$ th component is  $P_k(\mathbf{z})$ , and let  $\mathbf{Q}: C^{n+1} \times C^{2n} \rightarrow C^n$  be defined similarly. Define  $\mathbf{H}: C^{n+1} \times R \times C^{2n} \rightarrow C^n$  by

$$\mathbf{H}(\mathbf{z}, t, \mathbf{a}) = t\mathbf{P}(\mathbf{z}) + (1 - t)\mathbf{Q}(\mathbf{z}, \mathbf{a}).$$

When we consider  $\mathbf{a}$  as being fixed, we will let  $\mathbf{H}_\mathbf{a}(\mathbf{z}, t) = \mathbf{H}(\mathbf{z}, t, \mathbf{a})$ .

For  $j = 0, 1, \dots, n$ , define  $\varphi^j: C^n \rightarrow CP^n$  by

$$\varphi^j(z_0, \dots, \hat{z}_j, \dots, z_n) = [(z_0, \dots, z_{j-1}, 1, z_{j+1}, \dots, z_n)],$$

where  $(z_0, \dots, \hat{z}_j, \dots, z_n) = (z_0, \dots, z_{j-1}, z_{j+1}, \dots, z_n)$ . Then  $\varphi^j$  is a natural chart homeomorphism for the complex manifold  $CP^n$ . Similarly, define  $H^j: C^n \times R \times C^{2n} \rightarrow C^n$  by

$$\mathbf{H}^j(z_0, \dots, \hat{z}_j, \dots, z_n, t, \mathbf{a}) = \mathbf{H}(z_0, \dots, z_{j-1}, 1, z_{j+1}, \dots, z_n, t, \mathbf{a}).$$

Let

$$Y_\mathbf{a} = \{([z], t) \in CP^n \times [0, 1) : \mathbf{H}(\mathbf{z}, t, \mathbf{a}) = \mathbf{0}\},$$

and, for  $j = 0, 1, \dots, n$ , let

$$Y_\mathbf{a}^j = \{(z_0, \dots, \hat{z}_j, \dots, z_n, t) : \mathbf{H}^j(z_0, \dots, \hat{z}_j, \dots, z_n, t, \mathbf{a}) = \mathbf{0}\}.$$

Clearly,  $\varphi^j(Y_\mathbf{a}^j) = Y_\mathbf{a} \cap \{([z], t) \in CP^n \times [0, 1) : z_j \neq 0\}$ .

The idea of the proof is to show that if  $\mathbf{a}$  is chosen randomly, then  $Y_\mathbf{a}$  consists of  $d_1 d_2 \cdots d_n$  disjoint half-open arcs in  $CP^n \times [0, 1)$ , where the endpoint of each arc is a known root of (2) in  $CP^n \times \{0\}$ , and where the limit of the other end of the arc is a root of (1). These arcs can be followed numerically to approximate the roots. The major theoretical tool is the transversality theorem which is stated in the next section.

**4. Transversality.** In this section we state the special case of the topological transversality theorem that we will use. In applying transversality, we consider complex vector spaces as real vector spaces of twice the complex dimension.

Let  $U \subseteq R^m$  and let  $f: U \rightarrow R^p$  be a  $C^1$  map. ( $C^1$  means once continuously differentiable.) We say that  $y \in R^p$  is a *regular value* for  $f$  if

$$\text{Range } Df(\mathbf{x}) = R^p$$

for all  $\mathbf{x} \in f^{-1}(y)$ , where  $Df(\mathbf{x})$  denotes the  $p$  by  $m$  Jacobian matrix of partial derivatives of  $f$  evaluated at  $\mathbf{x}$ . Note that if  $m < p$ , and if  $y$  is a regular value of  $f$ , then  $y \notin f(U)$ .

**TRANSVERSALITY THEOREM.** Let  $V \subseteq R^q$ ,  $U \subseteq R^m$  be open, and let  $F: V \times U \rightarrow R^p$  be  $C^r$  where  $r > \max\{0, m - p\}$ . If  $\mathbf{0} \in R^p$  is a regular value of  $F$ , then, for almost all (in the sense of Lebesgue measure)  $\mathbf{a} \in V$ ,  $\mathbf{0}$  is a regular value of  $F_{\mathbf{a}}: U \rightarrow R^p$ , where  $F_{\mathbf{a}}(\mathbf{x}) = F(\mathbf{x}, \mathbf{a})$ .

An outline of a proof of the above theorem from the implicit function theorem and Sard's theorem is given in [3]. For a more complete treatment of transversality theorems see [1] or [8].

**COROLLARY.** If  $m = p + 1$ , then, for almost all  $\mathbf{a} \in V$ ,  $F_{\mathbf{a}}^{-1}(\mathbf{0})$  is a collection of  $C^1$  curves (one-dimensional manifolds). Furthermore,  $F_{\mathbf{a}}^{-1}(\mathbf{0})$  is a closed subset of  $U$ .

*Proof.* This follows easily from the implicit function theorem. (See pages 9 and 10 of [8].)

**5. The Relation Between a Complex Linear Transformation and Its Real Counterpart.** Recall that  $C^n = \{(z_1, z_2, \dots, z_n): z \in C\}$ . If  $z_i = x_i + iy_i$ , then there is a bijection  $\alpha$  from  $C^n$  into  $R^{2n}$  given by  $(z_1, z_2, \dots, z_n) \rightarrow (x_1, y_1, x_2, y_2, \dots, x_n, y_n)$ . If  $A$  is an  $m$  by  $n$  complex matrix representing a linear transformation from  $C^n$  into  $C^m$ , there is a  $2m$  by  $2n$  real matrix which we will denote by  $\alpha(A)$  such that  $\alpha(Az) = \alpha(A)\alpha(z)$ . The matrix  $\alpha(A)$  is obtained by replacing the  $jk$ -entry  $a_{jk} = \xi_{jk} + i\eta_{jk}$  of  $A$  by the 2 by 2 block

$$\begin{bmatrix} \xi_{jk} & -\eta_{jk} \\ \eta_{jk} & \xi_{jk} \end{bmatrix}.$$

**LEMMA 5.1.** Let  $A$  be an  $m$  by  $n$  complex matrix. Then the (real) rank of  $\alpha(A)$  is twice the (complex) rank of  $A$ .

*Proof.* Let  $E$  be a nonsingular  $m$  by  $m$  complex matrix such that  $EA$  is in row-reduced echelon form. Then, the rank of  $A$  is the number of nonzero rows in  $EA$ , and each nonzero row of  $EA$  has a leading entry of 1. Then  $\alpha(E)\alpha(A) = \alpha(EA)$  is also in row-reduced echelon form, and has two nonzero rows for each nonzero row of  $A$ .

## 6. The Main Theorem and Its Proof.

**MAIN THEOREM.** For almost all  $\mathbf{a} \in C^{2n}$  (in the sense of Lebesgue measure),

(1)  $\mathbf{0} \in C^n$  is a regular value of  $H_{\mathbf{a}}^j: C^n \times [0, 1) \rightarrow C^n$  for  $j = 0, 1, \dots, n$ ;

(2) (a)  $Y_{\mathbf{a}}^j \subseteq \{(z, t) \in C^n \times [0, 1): z_0 \neq 0\}$  for  $j = 1, 2, \dots, n$ ,

(b)  $Y_{\mathbf{a}} \subseteq \varphi^0(C^n) \times [0, 1) = \{([z], t) \in CP^n \times [0, 1): z_0 \neq 0\}$ ;

- (3) (a) Each component of  $Y_{\mathbf{a}}$  is a half-open arc that can be parametrized by a function  $r: [0, 1) \rightarrow CP^n \times [0, 1)$  such that  $r(t) \in CP^n \times \{t\}$ ,
- (b)  $Y_{\mathbf{a}}$  has  $\prod_{k=1}^n d_k$  components;
- (4) (a) If  $([z^*], 1)$  is in the closure of  $Y_{\mathbf{a}} \subseteq CP^n \times [0, 1]$ , then  $\mathbf{P}(z^*) = 0$ ,
- (b) If  $[z^*]$  is an isolated solution class of  $\mathbf{P}$ , then there is a component of  $Y_{\mathbf{a}}$  with a parametrization  $r(t)$  as in (3a) such that  $\lim_{t \rightarrow 1} r(t) = [z^*]$ .

*Proof.* The transversality theorem can be applied to  $\mathbf{H}^0 | C^n \times (-\infty, 1) \times C^{2n} \rightarrow C^n$  to show that  $\mathbf{0}$  is a regular value of  $\mathbf{H}_{\mathbf{a}}^0$  for almost all  $\mathbf{a} \in C^{2n}$ . Now let  $j \in \{1, 2, \dots, n\}$  be fixed. Let  $N = \{(z_0, \dots, \hat{z}_j, \dots, z_n) \in C^n: z_0 \neq 0\}$ . The transversality theorem can be applied to  $\mathbf{H}^j | N \times (-\infty, 1) \times C^{2n} \rightarrow C^n$  to show that  $\mathbf{0}$  is a regular value of  $\mathbf{H}_{\mathbf{a}}^j | N \times (-\infty, 1)$  for almost all  $\mathbf{a} \in C^{2n}$ . Next, we will show that  $Y_{\mathbf{a}}^j \subseteq N \times (-\infty, 1)$  for almost all  $\mathbf{a} \in C^{2n}$ . For each  $\mathbf{s} = (s_1, \dots, \hat{s}_j, \dots, s_n) \in \{0, 1\}^{n-1}$ , define

$$E(\mathbf{s}) = \{(0, z_1, \dots, \hat{z}_j, \dots, z_n): s_k = 0 \Leftrightarrow z_k = 0\}.$$

If  $m(\mathbf{s})$  is the number of components of  $\mathbf{s}$  which are equal to one, then  $E(\mathbf{s})$  is an open subset of a coordinate subspace of  $C^n$  of complex dimension  $m(\mathbf{s})$ . Define

$$G: E(\mathbf{s}) \times (-\infty, 1) \times C^{2n} \rightarrow C^{m(\mathbf{s})+1}$$

by

$$\begin{aligned} \mathbf{G}_k(0, z_1, \dots, \hat{z}_j, \dots, z_n, t, \mathbf{a}) &= \mathbf{H}_k^j(0, z_1, \dots, \hat{z}_j, \dots, z_n, t, \mathbf{a}) \\ &= \mathbf{H}_k(0, z_1, \dots, z_{j-1}, 1, z_{j+1}, \dots, z_n, t, \mathbf{a}) \end{aligned}$$

for  $k = j$  and  $k$  such that  $s_k = 1$ .

The transversality theorem can be applied to  $G$  to show that  $\mathbf{0}$  is a regular value of  $\mathbf{G}_{\mathbf{a}}$  for almost all  $\mathbf{a} \in C^{2n}$ . But since  $E(\mathbf{s}) \times (-\infty, 1)$  has real dimension  $2m(\mathbf{s}) + 1$  and  $C^{m(\mathbf{s})+1}$  has real dimension  $2m(\mathbf{s}) + 2$ , this implies that  $\mathbf{G}_{\mathbf{a}}^{-1}(\mathbf{0}) = \emptyset$ . But  $Y_{\mathbf{a}}^j \cap (E(\mathbf{s}) \times [0, 1)) \subseteq \mathbf{G}_{\mathbf{a}}^{-1}(\mathbf{0})$ . Since

$$C^n - N = \bigcup \{E(\mathbf{s}): \mathbf{s} \in \{0, 1\}^{n-1}\},$$

$Y_{\mathbf{a}}^j \subseteq N \times [0, 1)$ . This proves assertions (1) and (2a) of the theorem. To prove assertion (2b), note that  $Y_{\mathbf{a}} = \bigcup_{j=0}^n \Phi^j(Y_{\mathbf{a}}^j)$ .

For the remainder of the proof we suppose that an  $\mathbf{a} \in C^{2n}$  has been chosen so that assertions (1) and (2) hold. For simplicity of notation, we denote  $\mathbf{H}_{\mathbf{a}}^0$  by  $\mathbf{H}^0$  and  $Y_{\mathbf{a}}^0$  by  $Y^0$ .

The Jacobian of  $\mathbf{H}^0$  is the  $2n$  by  $2n + 1$  real matrix

$$\begin{bmatrix} \frac{\partial \mathbf{H}^0}{\partial \mathbf{z}} & \frac{\partial \mathbf{H}^0}{\partial t} \end{bmatrix}.$$

When evaluated at a point  $(\mathbf{z}, t)$  such that  $\mathbf{H}^0(\mathbf{z}, t) = 0$ , this matrix must have rank  $2n$  by the assumption that  $\mathbf{0}$  is a regular value of  $\mathbf{H}^0$ . By Lemma 5.1,  $\partial \mathbf{H}^0 / \partial \mathbf{z}$  must have even rank. This implies that  $\partial \mathbf{H}^0 / \partial \mathbf{z}$  has rank  $2n$  and is nonsingular.

Now let  $(\mathbf{z}(\lambda), t(\lambda))$  be a local parametrization of a curve of  $Y^0$  with respect to a real parameter  $\lambda$ . We can assume that  $d\mathbf{z}/d\lambda$  and  $dt/d\lambda$  are not both zero at any  $\lambda$  for which  $(\mathbf{z}(\lambda), t(\lambda))$  is defined. Then  $\mathbf{H}^0(\mathbf{z}(\lambda), t(\lambda)) = 0$ , so

$$\frac{\partial \mathbf{H}^0}{\partial \mathbf{z}} \frac{d\mathbf{z}}{d\lambda} + \frac{\partial \mathbf{H}^0}{\partial t} \frac{dt}{d\lambda} = \mathbf{0}.$$

Since  $\partial \mathbf{H}^0 / \partial \mathbf{z}$  is nonsingular,  $dt/d\lambda = 0$  implies that  $d\mathbf{z}/d\lambda = 0$ . Thus,  $dt/d\lambda$  is always nonzero. This means that  $z$  can be parametrized by  $t$  so that  $z(t) \in C^n \times \{t\}$ . This proves assertion (3a).

Since each component of  $Y = Y_a$  is closed in  $CP^n \times [0, 1)$ , it must intersect  $CP^n \times \{0\}$ . But  $Y \cap (CP^n \times \{0\})$  is the set of solution classes of the equation  $\mathbf{Q}(\mathbf{z}) = \mathbf{0}$ . This set consists of the  $d_1 d_2 \cdots d_n$  points described earlier. Thus,  $Y$  has  $d_1 d_2 \cdots d_n$  components, and assertion (3b) is proved.

Assertion (4a) follows from the continuity of  $\mathbf{H}_a$ . To prove (4b), choose  $j$  so that  $z_j^* \neq 0$ . In [2] it is shown that  $H^j | C^n \times \{1\}$  has positive degree at  $\varphi_j^{-1}([z^*])$ . Let  $V$  be a ball neighborhood of  $\varphi_j^{-1}([z^*])$  in  $C^n$  which contains no other root of  $H^j | C^n \times \{1\}$ . Choose  $\varepsilon > 0$  sufficiently small that  $\mathbf{0} \notin H^j(\text{Bd}(V) \times [1 - \varepsilon, 1])$ . Then, by the homotopy invariance of degree,  $H^j | V \times \{1 - \varepsilon\}$  has positive degree, and so  $Y^j \cap (V \times \{1 - \varepsilon\}) \neq \emptyset$ . Thus,  $\varphi_j^{-1}([z^*])$  is a limit point of  $Y^j$ , and  $[z^*]$  is a limit point of  $Y$ .

**7. Following the Solution Curves Numerically.** As before, we assume that  $\mathbf{a} \in C^{2n}$  is chosen so that the conclusions of the main theorem hold. To solve the vector equation  $\mathbf{P}(\mathbf{z}) = \mathbf{0}$ , the arcs of  $Y$  are followed from the known solutions of  $\mathbf{Q}(\mathbf{z}) = \mathbf{0}$  in  $PC^n \times \{0\}$  to the solutions of  $\mathbf{P}(\mathbf{z}) = \mathbf{0}$  in  $PC^n \times \{1\}$ . We can start by following the curves of  $Y^0$  in  $C^n \times [0, 1]$ . If the modulus of some component  $z_j$  of  $\mathbf{z}$  becomes large as the curve is followed, we can switch to the corresponding point  $(1/z_j, z_1/z_j, \dots, \hat{z}_j/z_j, \dots, z_n/z_j)$  of  $Y^j$ . In this way, numerical difficulties of following curves that go to infinite roots can be avoided, and infinite roots can be computed explicitly.

One way to follow the solution curves is to use an initial value differential equation solver. Let  $\mathbf{z}(t)$  be a parametrization of a solution curve. Then  $\mathbf{H}(\mathbf{z}(t), t) = \mathbf{0}$ . Taking the derivative of both sides by the chain rule, we obtain the differential equation

$$(3) \quad \frac{\partial \mathbf{H}}{\partial \mathbf{z}} \frac{d\mathbf{z}}{dt} + \frac{\partial \mathbf{H}}{\partial t} = \mathbf{0}$$

or

$$\frac{d\mathbf{z}}{dt} = - \left( \frac{\partial \mathbf{H}}{\partial \mathbf{z}} \right)^{-1} \frac{\partial \mathbf{H}}{\partial t}.$$

Thus, each function evaluation for the differential equation solver involves solving the linear system of equations (3) for  $d\mathbf{z}/dt$ . This method of following the solution curves was implemented and is reported on in Section 8.

The above method does not utilize all of the available information. In contrast to the usual situation in solving an initial value problem, we can use Newton's method to get closer to the solution curve. If we are at a point  $(\mathbf{z}, t)$ , one step of Newton's method would be:

$$\text{Solve } \mathbf{s} \frac{\partial \mathbf{H}}{\partial \mathbf{z}}(\mathbf{z}, t) = -\mathbf{H}(\mathbf{z}, t) \quad \text{for } \mathbf{s} \in C^n.$$

Let  $\mathbf{z} = \mathbf{z} + \mathbf{s}$  (and let  $t$  remain unchanged).

Notice that the coefficient matrix of the system of linear equations to be solved is the same as for a function evaluation of the differential equation solution. In both cases the computation of a decomposition (such as the  $LU$  decomposition) of this

Jacobian matrix will be the major part of the computation of a step. Thus, it should be possible to combine these steps in some way.

In addition, one does not need to follow the solution curves closely at the beginning. All that is required is that one remain within the domain of convergence for Newton's method for the solution curve one is trying to follow.

Clearly, more research is needed on techniques for following the solution curves.

**8. An Implementation.** This method was implemented using the initial value differential equations package LSODE, available from the Lawrence Livermore Laboratory, Livermore, California 94550, to follow the curves of  $Y$ . (LSODE is a package based on the GEAR and GEARB packages and on the Oct. 23, 1978 version of the tentative ODEPACK user interface standard, with minor modifications.) As mentioned in the last section, there should be much more efficient methods for following these curves, so the computation times reported here should not be used to judge the eventual efficiency of the algorithm. However, the results presented here do show that all roots of small systems of polynomials can be computed.

The following systems of equations were solved.

$$(A) \quad (2z_1 - z_2 - z_0)^2 = 0, \quad z_1^2 - z_2^2 = 0.$$

Solutions:

$$(z_0, z_1, z_2) = \begin{cases} (1, 1, 1) & \text{(multiplicity 2),} \\ (\frac{1}{3}, -\frac{1}{3}, 1) & \text{(multiplicity 2).} \end{cases}$$

$$(B) \quad (z_1 - z_2 - z_0)^2 = 0, \quad z_1^2 - z_2^2 = 0.$$

Solutions:

$$(z_0, z_1, z_2) = \begin{cases} (0, 1, 1) & \text{(multiplicity 2),} \\ (2, 1, -1) & \text{(multiplicity 2).} \end{cases}$$

$$(C) \quad z_k^2 + \sum_{i=1}^5 z_0 z_i - 2z_0 z_k - 10z_0^2 = 0 \quad \text{for } k = 1, 2, \dots, 5.$$

Note that this system is invariant under permutation of the coordinates  $z_1, z_2, \dots, z_5$ . In the following list of solutions, the coordinates  $z_1, z_2, \dots, z_5$  should be permuted to obtain the remaining solutions. All solutions have multiplicity 1. This system comes from [10].

Solutions:

$$(z_0, z_1, z_2, z_3, z_4, z_5) = \begin{cases} (1, 2, 2, 2, 2, 2) & \text{(1 solution),} \\ (1, -5, -5, -5, -5, -5) & \text{(1 solution),} \\ (1, -a, 2 + a, 2 + a, 2 + a, 2 + a) & \text{(5 solutions),} \\ (1, 5 + a, -3 - a, -3 - a, -3 - a, -3 - a) & \text{(5 solutions),} \\ (1, -1, -1, 3, 3, 3) & \text{(10 solutions),} \\ (1, -2, -2, 4, 4, 4) & \text{(10 solutions),} \end{cases}$$

where  $a = (-5 + \sqrt{33})/2$ .

$$(D) \quad z_k^2 + \sum_{i=1}^5 z_0 z_i - 2z_0 z_k - 4z_0^2 = 0 \quad \text{for } k = 1, 2, \dots, 5.$$

Again, this system is invariant under permutation of the coordinates  $z_1, z_2, \dots, z_5$ .  
Solutions:

$$(z_0, z_1, z_2, z_3, z_4, z_5) = \begin{cases} (1, 1, 1, 1, 1, 1) & \text{(1 solution of multiplicity 16),} \\ (1, 2, 2, 0, 0, 0) & \text{(10 solutions of multiplicity 1),} \\ (-1, -4, 2, 2, 2, 2) & \text{(5 solutions of multiplicity 1),} \\ (-1, 4, 4, 4, 4, 4) & \text{(1 solution of multiplicity 1).} \end{cases}$$

These problems were solved on a Digital Equipment PDP-10 computer with dual KL-10 processors. A FORTRAN program was written which:

(1) Computed the parameters  $a_{kj}$ ,  $k = 1, 2, \dots, n$ ;  $j = 1, 2$  with  $-2 \leq \text{Re}(a_{kj}) \leq 2$  and  $-2 \leq \text{Im}(a_{kj}) \leq 2$  using the random number generator supplied with the FORTRAN system.

(2) Computed the  $\prod_{k=1}^n d_k$  starting points.

(3) For each starting point, called LSODE to follow the solution curves.

The relative tolerance for the LSODE package was set at  $10^{-10}$  and the absolute tolerance was set at  $10^{-12}$ . All computations were done in DOUBLE PRECISION (machine epsilon  $\approx 2 \times 10^{-19}$ ).

The computation times and approximate maximum absolute errors are given in Table 1. It appears that going from a single root to a multiple root approximately quadruples the number of function evaluations, and approximately halves the number of accurate significant digits. Increasing the dimension did not seem to affect the number of function evaluations required.

When problem B was run *without* the recoordinating feature, the Jacobian matrix became somewhat ill-conditioned (condition number greater than  $10^{10}$ ). After about 1300 steps (1400 function evaluations) the norms of the points were approximately  $10^6$  and  $t \approx 1 - 10^{-12}$ . This contrasts with an average of 431 steps (526 function evaluations) to compute the infinite root using recoordinating.

TABLE 1

Problem	$n$	Sum of degrees	CPU time (seconds)	Approximate maximum absolute error (single roots)	Approximate maximum absolute error (multiple roots)	Average number of function eval. per single root	Average number of function eval. per multiple root
A	2	4	26		$1.5 \times 10^{-5}$		631
B	2	4	32		$2.5 \times 10^{-5}$		589
C	5	32	151	$5 \times 10^{-10}$		136	
D	5	32	449	$1.6 \times 10^{-9}$	$2 \times 10^{-5}$	158	631

1. R. ABRAHAM & J. ROBBIN, *Transversal Mappings and Flows*, Benjamin, New York, 1967.
2. S. N. CHOW, J. MALLET-PARET & J. A. YORKE, "A homotopy method for locating all zeros of a system of polynomials," *Functional Differential Equations and Approximation of Fixed Points* (Proceedings, Bonn, 1978), Lecture Notes in Math., Vol. 730, Springer-Verlag, Berlin and New York, 1979.
3. S. N. CHOW, J. MALLET-PARET & J. A. YORKE, "Finding zeros of maps: homotopy methods that are constructive with probability one," *Math Comp.*, v. 32, 1978, pp. 887–899.
4. F. J. DREXLER, "A homotopy method for the calculation of all zero-dimensional polynomial ideals," *Continuation Methods* (H. Wacker, ed.), Academic Press, New York, 1978, pp. 69–93.
5. F. J. DREXLER, "Eine Methode zur Berechnung sämtlicher Lösungen von Polynomgleichungssystemen," *Numer. Math.*, v. 29, 1977, pp. 45–58.
6. C. B. GARCIA & T. Y. LI, "On the number of solutions to polynomial systems of equations," *SIAM J. Numer. Anal.*, v. 17, 1980, pp. 540–546.
7. C. B. GARCIA & W. I. ZANGWILL, "Finding all solutions to polynomial systems and other systems of equations," *Math. Programming*, v. 16, 1979, pp. 159–176.
8. M. HIRSCH, *Differential Topology*, Springer-Verlag, New York, 1976.
9. R. W. KLOPFENSTEIN, "Zeros of non-linear functions," *J. Assoc. Comput. Mach.*, v. 8, 1961, pp. 366–373.
10. M. KOJIMA & S. MIZUNO, "Computation of all solutions to a system of polynomial equations," *Math. Programming*, v. 25, 1983, pp. 131–157.
11. B. L. VAN DER WAERDEN, "Die Alternative bei nichtlinearen Gleichungen," *Nachrichten der Gesellschaft der Wissenschaften zu Göttingen*, Math. Phys. Klasse, 1928, pp. 77–87.