RC 13287 (#59459) 11/18/87
Mathematics 22 pages

Research Report

Branch Switching in Homotopy Methods for Finding all Roots of a System of Polynomials

Michael E. Henderson

IBM Research Division
T.J. Watson Research Center
Yorktown Heights, N.Y. 10598

IBM Research Division
Almaden • Yorktown • Zurich

Copies may be requested from:

IBM Thomas J. Watson Research Center
Distribution Services F-11 Stormytown
Post Office Box 218
Yorktown Heights, New York 10598

Branch Switching in Homotopy Methods for Finding all Roots of a System of Polynomials

Michael E. Henderson

Summary. In 1978-79 three papers appeared which proposed using homotopy methods to find all the roots of a system of polynomials. A basic difference between the papers is how they avoid having to compute the tangents of the homotopy paths at singular points. In this paper we show how the tangents may be easily computed, using a theorem on the existence of families of related paths through the singular points. We then describe how this theorem may be used to construct a branch switching algorithm so that pseudo-arclength continuation [10] may be used to follow the homotopy paths.

Subject Classifications: AMS(MOS): 65H10; CR: G 1.5

I. Introduction

A system of polynomial equations is a set of N equations of the form

$$a_i(\underline{x}) = \sum_{j=1}^{M_i} c_{ij} x_1^{n_{ij1}} \ldots x_N^{n_{ijN}} = 0. \tag{1}$$

Here x_k are complex scalars, M_i is the number of terms in the i^{th} equation, c_{ij} is the coefficient of the j^{th} term of the i^{th} equation, and the n_{ijk} are positive integer exponents. The degree of equation i is the maximum of the sum of the exponents of its terms

$$d_i = \max_j \sum_{k=1}^{N} n_{ijk}.$$

Homotopy methods for finding the roots of a system of equations

$$A(\underline{x}) = 0$$

$$A: \ \mathbb{C}^N \rightarrow \ \mathbb{C}^N$$

use an initial system $A^0(\underline{x}) = 0$ whose roots $\{\underline{x}_i\}$ are known, and a system $H(\underline{x}, \lambda) = 0$, called a homotopy of A^0 into A, which depends on a homotopy parameter λ. The homotopy is such that

$$H(\underline{x}, 0) = A^0(\underline{x}), \qquad \text{and} \qquad H(\underline{x}, 1) = A(\underline{x}).$$

The homotopies considered here are all of the form

$$H(\underline{x}, \lambda) \equiv (1 - \lambda)A^0(\underline{x}) + \lambda A(\underline{x}).$$

To find the roots of A, a homotopy method computes the homotopy paths $\underline{x}_i(\lambda)$, satisfying $H(\underline{x}_i(\lambda), \lambda) = 0$, and $\underline{x}_i(0) = \underline{x}_i^0$. If all the paths exist for $\lambda = 1$, the roots of A have been found.

There are three components to a homotopy method:

1. An initial system A^0 must be specified,

2. It must be shown that the homotopy paths $\underline{x}_i(\lambda)$ exist for $\lambda = 1$,

3. A numerical method must be specified to compute the homotopy paths.

It is not easy to find polynomial systems which can be solved explicitly, so there is not much freedom in choosing the initial system. One class of systems whose roots are known is

$$a_i^0(\underline{x}) = x_i^{d_i} - \alpha_i.$$

The roots of this system are determined by the constants α_i, and it can be shown (Garcia and Li [9]) that this initial system always has at least as many solutions as does $A(\underline{x}) = 0$. It also has the desirable property that the homotopy paths are well separated in a neighborhood of $\lambda = 0$.

The existence of the homotopy paths has been proven in many different ways, see [3], [6], or [7] for example. Roughly speaking, Bézout's Theorem states that the number of solutions of H at a particular value of the homotopy parameter (counting multiple solutions and infinite solutions), depends only on the degree of H at that parameter value. Therefore, if the degree of H is independent of λ, the number of solutions is also independent of λ.

There are two types of numerical methods for following the homotopy paths; Simplicial Continuation methods (see Allgower and Georg [1]), and Predictor-Corrector type methods (e.g. Keller [10]). Simplicial Continuation methods partition the domain of H into small convex regions called simplices. The initial solution lies in one of these simplices, and by continuity the homotopy path must pass through an adjacent simplex. The faces of the initial simplex are checked to determine which the homotopy path crosses, the simplices adjacent to those faces are marked and take the place of the initial simplex.

Predictor-Corrector methods use the initial solution and an approximation to the homotopy path passing through that it to predict nearby points on the path. An iteration is used to locate a point on the homotopy path near the predicted point. Of the two methods predictor-corrector methods are simpler to program, and are faster for large systems, but may require expensive procedures to compute the predicted path near singular points on the homotopy paths. Simplicial Continuation methods can handle maps that are not smooth, and require no special routines to compute tangents near singular points, but are not well suited to large systems.

II. Existing Algorithms

In 1978 Chow, Mallet-Paret, and Yorke [3] introduced the idea of homotopy methods which are constructive with probability 1. They proposed to use a very simple predictor-corrector method, and avoided the difficulties caused by singular points on the homotopy paths by choosing the initial system so that the paths are regular. They showed that for almost every initial system A_0 a regular homotopy path exists which reaches $\lambda = 1$. The algorithm is essentially this:

1. Choose an initial system at random.
2. Follow the homotopy paths through each initial solution until $\lambda = 1$ is reached, or a singular point is found.
3. If the path reaches $\lambda = 1$ a root has been found.

   If a singular point is found, discard that path and initial system. Repeat starting at step 1 until all roots have been found.

This method finds all the roots with probability one.

Drexler [6] also specified that the homotopy paths be regular. The initial system was fixed, and the homotopy parameter was allowed to lie on a path in the complex plane between $\lambda = 0$ and $\lambda = 1$. He proved that the set of points $\lambda \in \mathbb{C}$ for which a homotopy path has a singular point is a set of isolated points. Thus a regular path exists in the complex λ plane which connects $\lambda = 1$ to the origin. Chow, Mallet-Paret, and Yorke [4] provided an alternate proof of this result. An algorithm based on this result is this:

1. Choose an initial system.
2. Follow the homotopy path through each initial solution until $\lambda = 1$ is reached or a singular point is found.
3. If the path reaches $\lambda = 1$ a root has been found.

   If a singular point is found, modify the path to avoid the singular point. Repeat starting at step 2 until $\lambda = 1$ is reached.

Garcia and Zangwill [7], [8] proposed using a simplicial continuation method to compute the homotopy paths. This eliminates the need to avoid singular points. The complex system of polynomials is first imbedded into $\mathbf{R}^{2N}$. They then showed that a real path exists for each initial solution which either reaches $\lambda = 1$, becomes infinite, or has an accumulation point at $\lambda = 1$. More recently, Brunovský and Meravý [2] and Wright [14] used a homogeneous parameter to map infinity to a finite point, thus eliminating the possibility of homotopy paths which contain infinite points.

Several recent papers report on implementations and refinements of the above algorithms. Wright [14], and Brunovský and Meravý [2] have implemented versions of Chow, Mallet-Paret, and Yorke's algorithm. Morgan [12], and Kojima and Mizuno [11] have implemented algorithms similar to those proposed by Zangwill and Garcia.

III. Pseudo-Arclength Continuation

Pseudo-arclength continuation [10] is a predictor-corrector method for following homotopy paths. It has tan advantage over other predictor-corrector methods in the way it handles paths with steep tangents, and its behavior near singular points. The particular algorithm we will describe is called Euler-Newton continuation. The tangent to the solution branch is used to predict the path (an Euler Predictor), and Newton's method is used as a corrector.

With an arclength parameterization, the tangent to the solution path of $G(\underline{x}, \lambda) = 0$, through the point $(\underline{x}_0, \lambda_0)$ satisfies

$$G^0_{\underline{x}} \dot{\underline{x}}_0 + G^0_\lambda \dot{\lambda}_0 = 0 \tag{2}$$
$$\| \dot{\underline{x}}_0 \|^2 + \dot{\lambda}_0^2 = 1.$$

This has two solutions when $G^0_{\underline{x}}$ is nonsingular (the sign of the tangent is not determined). The sign is chosen so that the inner product with a tangent at a previous point on the homotopy path is positive. The predicted path is

$$\underline{x}_p(s) = \underline{x}_0 + s\dot{\underline{x}}_0$$

$$\lambda_p(s) = \lambda_0 + s\dot{\lambda}_0,$$

and the error in the prediction is bounded by

$$\| (\underline{x}_p(s), \lambda_p(s)) - (\underline{x}(s), \lambda(s)) \| \leq \frac{1}{2} \max_{0 \leq \zeta \leq s} \| (\ddot{\underline{x}}(\zeta), \ddot{\lambda}(\zeta)) \| \, s^2.$$

Newton's method is next used to solve the system

$$(4) \qquad \begin{aligned} &G(\underline{x}(s), \lambda(s)) = 0, \\ &\dot{\underline{x}}_0^*(\underline{x}(s) - \underline{x}_0) + \dot{\lambda}_0(\lambda(s) - \lambda_0) - s = 0, \end{aligned}$$

with the predicted path as an initial guess. The second constraint is called the Pseudo-arclength constraint, and requires that the solution at s lies in a hyperplane normal to the tangent at $(\underline{x}_0, \lambda_0)$, at a distance s from $(\underline{x}_0, \lambda_0)$. There exists a $\rho(s) > 0$ such that Newton's Method converges, provided that

$$\| (\underline{x}_p(s), \lambda_p(s)) - (\underline{x}(s), \lambda(s)) \| \leq \rho(s).$$

$\rho(s)$ can be given explicitly in terms of norms of G and its derivatives (see Decker and Keller [5]). In this ball the path is unique. Only one point on the path is computed ($s = \Delta s$), and the path is approximated by a line segment, or an Hermite interpolant, between $(\underline{x}(\Delta s), \lambda(\Delta s))$ and $(\underline{x}_0, \lambda_0)$. The step size Δs is chosen as large as possible so that the predicted solution lies within the ball of convergence for Newton's Method.

The Pseudo-arclength algorithm is this:

1. If G_x is non-singular, solve (2) for the tangent $(\dot{\underline{x}}, \dot{\lambda})$.

   Otherwise, switch branches using the algorithm described below.

2. Compute the predicted solution $(\underline{x}_p(\Delta s), \lambda_p(\Delta s))$ using (3).
3. Correct the predicted solution by using Newton's Method to solve (4), with the predicted solution as an initial guess.
4. Repeat.

IV. Singular Points At a point where $G_{\underline{x}}$ is singular the tangent to the path is not necessarily unique. Equation (2) only determines the tangent to within a component in the null space of $G_{\underline{x}}$. To overcome this difficulty it is usual to use a Lyapunov-Schmidt Decomposition (see Vainberg and Trenogin [13]). The equations $G = 0$ are first projected onto the range of G_x, and the Implicit Function Theorem is used to determine the component of the path in the complement of the null space of $G_{\underline{x}}$, as a function of the component in the null space. The equations are then projected onto the complement of the range of $G_{\underline{x}}$. These projected equations are called the bifurcation equations, and must be solved to determine the path near the singular point. They may usually be reduced to a set of algebraic equations for the tangent to the path by an appropriate scaling, and are then called the Algebraic Bifurcation Equations (the ABE's). Unfortunately, the problem of computing the tangents at a singular point is therefore equivalent to the problem we set out to solve, although the size of the system has been decreased.

We show below that there is a symmetry of the ABE's which allows the existence of a family of paths near the singular point to be inferred from a the existence of a single given path. We further show that one member of this family exists in a neighborhood of the singular parameter value which lies on the opposite side of the singular point from the given branch. This allows a path containing a singular point to be continued past the singular point without solving the ABE's. The resulting homotopy path is such that the parameter λ on the branch is monotone non-decreasing.

We assume that a branch of solutions of

$$(A)\qquad \begin{aligned} &G(\underline{x},\lambda)=0\\ &G:\mathbf{B}\times\mathbf{R}\rightarrow\mathbf{B}\end{aligned}$$

which passes through the singular solution $(\underline{x}_0,\lambda_0)$ is given. Here $\mathbf{B}$ is a complex Banach space. (In this paper we are concerned only with $\mathbf{C}^N$, but the result is more general.) We further assume that

$$(B)\qquad \begin{aligned} &G_{\underline{x}}(\underline{x}_0,\lambda_0) \text{ is Fredholm of index 0, and}\\ &\dim(\mathrm{Null}(G^0_{\underline{x}}))=d<\infty.\end{aligned}$$

(For $\mathbf{C}^N$ these are always satisfied.)

Let $\{\phi_i\}_1^d$ be an orthonormal basis for the null space of $G^0_{\underline{x}}$, $\{\psi_i\}_1^d$ an orthonormal basis for the null space of $G^{0*}_{\underline{x}}$, and Q be the projection of $\mathbf{B}$ onto the range of $G^0_{\underline{x}}$. By the Implicit Function Theorem there is a unique mapping $a(\underline{\xi},\eta)\in\mathrm{Range}(G^{0*}_{\underline{x}})$, such that

$$QG(\underline{x}_0+\sum_1^d \xi_i\phi_i+a(\underline{\xi},\eta),\lambda_0+\eta)=0$$

for all $\underline{\xi} \in \mathbb{C}^d$ and η near 0. The bifurcation equations for this branch are

$$\psi_i^* G(\underline{x}(s), \lambda(s)) = 0 \qquad 1 \leq i \leq d$$
$$\| \dot{\underline{x}}(s) \|^2 + |\dot{\lambda}(s)|^2 - 1 = 0,$$

where

(5)
$$\underline{x}(s) = \underline{x}_0 + \sum_{i=1}^{d} \xi_i(s)\phi_i + a(\underline{\xi}(s), \eta(s)),$$
and $\quad \lambda(s) = \lambda_0 + \eta(s).$

In the following discussion we will switch freely between the representations $(\underline{x}, \lambda)$ and $(\underline{\xi}, \eta)$, always subject to (5).

Let $F(\underline{x}, \lambda) = \psi_i^* G(\underline{x}(\underline{\xi}, \eta), \lambda(\eta))$, which is a mapping from $\mathbb{C}^d \times \mathbb{R} \to \mathbb{C}^d$. Expanding F in a power series in $\underline{x}$ and λ about $s = 0$ yields

$$F \sim \sum_{m=0}^{\infty} \frac{1}{m!} \sum_{n=0}^{m} \binom{m}{n} F^0_{\underline{x}^n \lambda^{m-n}} (\underline{x}(s) - \underline{x}_0)^n (\lambda(s) - \lambda_0)^{m-n}.$$

We will assume that near $s = 0$ the given branch behaves as

(6)
$$\underline{x}(s) \sim \underline{x}_0 + \frac{1}{M!} s^M \underline{x}_0^{(M)} + O(s^{M+1})$$
$$\lambda(s) \sim \lambda_0 + \frac{1}{K!} s^K \lambda_0^{(K)} + O(s^{K+1}),$$

for some positive integers M and K. Then

$$F \sim \sum_{m=0}^{\infty} \frac{1}{m!} \sum_{n=0}^{m} \left\{ \binom{m}{n} \frac{1}{M!^n K!^{m-n}} F^0_{\underline{x}^n \lambda^{m-n}} (\underline{x}_0^{(M)})^n (\lambda_0^{(K)})^{m-n} s^{nM+(m-n)K} (1 + O(s)) \right\}.$$

Or, grouping terms with like powers of s,

(7a)
$$F \sim \sum_{l=L}^{\infty} \sum_{m,n} \left\{ \frac{1}{m!} \binom{m}{n} \frac{1}{M!^n K!^{m-n}} F^0_{\underline{x}^n \lambda^{m-n}} (\underline{x}_0^{(M)})^n (\lambda_0^{(K)})^{m-n} s^l (1 + O(s)) \right\}.$$

Here the second sum is over positive integers m and n such that $nM + (m-n)K = l$ and $n < m$. The first two terms vanish, since $F^0 = \psi_i^* G^0 \equiv 0$, $F_{\underline{x}}^0 = \psi_i^* G_{\underline{x}}^0 \equiv 0$, and either $F_\lambda^0 = 0$ (a bifurcation point), or $F_\lambda^0 \neq 0 \Rightarrow \dot{\lambda}_0 = 0$ (a fold). We will assume that the first non-zero term is the L^{th}, for some $L \geq 2$.

The arclength constraint becomes

$$(7b) \qquad \frac{1}{(M-1)!^2} s^{2(M-1)} \| \underline{x}_0^{(M)} \|^2 + \frac{1}{(K-1)!^2} s^{2(K-1)} |\lambda_0^{(K)}|^2 - 1 + O(s^{2M-1}, s^{2K-1}) = 0.$$

Because an arclength parameterization has been used, it is necessary that either $M = 1$ or $K = 1$.

The Algebraic Bifurcation Equations (ABE's) are obtained from $(7a)$ and $(7b)$ by scaling $(7a)$ by s^{-L}, substituting for $\underline{x}_0^{(M)}$ and $\lambda_0^{(K)}$ as functions of $\underline{\xi}$ and η, and then letting $s \to 0$. We have tabulated the first few ABE's for $M = 1$ in Table 1. Note that these are not the coefficients of a power series of $F(s)$ in s. (The coefficients of s^l for $l > L$ are modified by the $O(s^{k+1})$ terms in the coefficients of s^k, $L \leq k \leq l$.)

For each isolated root of the ABE's, the Implicit Function Theorem states that there is a corresponding isolated branch of solutions of the form (6) passing through the singular point. We shall assume that the given solution branch corresponds to such an isolated root. (If it does not, it may approach the singular point tangentially to another branch. If this occurs, the ansatz (6) must be modified to include terms which distinguish the branches.) One isolated root of the ABE's is therefore known, and we use this root to construct other isolated roots.

Theorem 1:

Let G be a mapping of type (A), satisfying (B). Assume that G has an infinite number of bounded derivatives. (A finite number are required, but the number depends on the degree of the ABE's.) If a branch of the form

$$\underline{x}(s) \sim \underline{x}_0 + s\dot{\underline{x}}_0 + O(s^2)$$

$$\lambda(s) \sim \lambda_0 + \frac{1}{K!} s^K \lambda_0^{(K)} + O(s^{K+1}),$$

exists in a neighborhood of the singular point $(\underline{x}_0, \lambda_0)$, and if that branch corresponds to an isolated *root of the algebraic bifurcation equations, then branches of the form*

(8)
$$\underline{x}(s) \sim \underline{x}_0 + \alpha s \underline{x}_0^{(1)} + O(s^2)$$
$$\lambda(s) \sim \lambda_0 + \beta \frac{1}{K!} s^K \lambda_0^{(K)} + O(s^{K+1})$$

also exist in a neighborhood of the singular point, provided that the pair (α, β) *satisfys*

$$\alpha = \beta^{1/K}$$
$$\beta = \pm 1.$$

Proof:

We must show that the branches (8) correspond to isolated roots of the ABE's. The ABE's for this type of branch are

(9)
$$\sum_{m-n=0}^{[L/K]} \binom{m}{n} F^0_{u^n \lambda^{m-n}} \frac{1}{K!^{m-n}} \dot{u}^n (\lambda^{(K)})^{m-n} = 0$$
$$\| \dot{u} \|^2 = 1,$$

where

$$\dot{u} = \sum_{i=1}^{d} \dot{\xi}_i(0) \phi_i$$
$$\lambda^{(K)} = \eta^{(K)}(0)$$

We look for other roots of (9) of the form $\alpha \dot{u}_0$ and $\beta \lambda_0^{(K)}$ ($\alpha \dot{\underline{\xi}}_0(0)$, and $\beta \eta_0^{(K)}(0)$.) Substituting, we must have

(10)
$$\alpha^L \sum_{m-n=0}^{[L/K]} (\alpha^{-K}\beta)^{(m-n)} \binom{m}{n} F^0_{u^n \lambda^{m-n}} \frac{1}{K!^{m-n}} \dot{u}_0^n (\lambda_0^{(K)})^{m-n} = 0,$$
$$|\alpha|^2 \| \dot{u}_0 \|^2 = 1.$$

Here the integers n and m are related by $n = L - K(m-n)$. When $\alpha^K = \beta$, and $|\alpha| = 1$ equation (10) is satisfied. λ is a real parameter, so β must be real, therefore $\beta = \pm 1$.

With this choice of α and β (10) is identical to (9). The given branch corresponds to an isolated solution $(\dot{\underline{\xi}}_0, \eta_0^{(K)})$ of equation (9), so the Jacobian J_0 of (9) at $(\dot{\underline{\xi}}_0, \eta_0^{(K)})$, is non-singular.

The Jacobian at the scaled root $(\alpha\dot{\xi}_0, \beta\lambda_0^{(K)})$ is a non-singular transformation of J_0, so the scaled root is also an isolated root of the algebraic bifurcation equations.

We have shown that for each pair $\alpha = (\pm 1)^{1/K}, \beta = \pm 1$, there is an isolated root of the ABE's, and hence a branch of solutions passing through the singular point. Near the singular point these branches behave as

$$u(s) \sim u_0 + (\pm 1)^{1/K} s\dot{u}_0 + O(s^2)$$

$$\lambda(s) \sim \lambda_0 \pm \frac{1}{K!}\lambda_0^{(K)} s^K + O(s^{K+1}). \blacksquare$$

If K is even, the given branch exists only in a neighborhood ton one side of λ_0. The branches with $\beta < 0$ exist on the other side, so if the path following algorithm switchs to one of these branches, the given branch can be continued past the singular point, with monotonically increasing parameter. Further, if one particular root of -1 is used to determine α, and the choice is consistent, distinct branches which approach the singular point non-tangentially will be switched to distinct branches. If K is odd, the branch switching algorithm does nothing, for in that case the given branch exists on both sides of λ_0.

This result is, in a local sense, the same as the result of [6]. The family of branches whose existance is given by Theorem 1 is the same as the family obtained by traversing an infinitesimal circular path in complex parameter space about the singular point, starting at a point on the given branch near the singular point. Each half circle traversed changes the sign of β, and each full circle multiplies α by $e^{2\pi i/K}$.

V. The Homogeneous Coordinate

To avoid difficulties with possibly infinite roots, we introduce a homogenous coordinate $t \in \mathbb{C}$, which maps infinity to a finite point. Let

$$u = (\underline{x}, t),$$

and scale each equation by t raised to the degree of the equation. This new system, $t^D H(\underline{x}/t, \lambda)$ will be used to compute the homotopy paths.

The solutions of this new system are not well defined. The system is homogenous in u, so any solution may be multiplied by a complex scalar. We introduce two different constraints to determine this scalar. For the purposes of applying Theorem 1 we will use the complex constraint

$$l^{*}u = 1, \tag{11}$$

where l is some fixed complex vector. For any particular choice of l it is possible for the solution to become unbounded, but there will be another choice of l for which that solution is bounded. For computational purposes it is better to have a single constraint which always guarentees a bounded solution. We therefore introduce a normalization constraint

$$\| \underline{x} \|^2 + |t|^2 = 1.$$

This determines the solution up to multiplication by a complex scalar of norm one. To determine this relative phase between $\underline{x}$ and t, we also impose the phase constraint

$$\mathrm{Im}(u)_k = 0.$$

The system we use for the homotopy is

$$G(u,\lambda) = \begin{Bmatrix} t^D H(\underline{x}/t, \lambda) \\ \| \underline{x} \|^2 + |t|^2 - 1 \\ \mathrm{Im}(u)_k \end{Bmatrix} = 0 \tag{12}$$

Theorem 1 cannot be directly applied to G. However, since solutions of $G(u, \lambda) = 0$ are isomorphic to solutions of the system with the normalization and phase conditions replaced by (11), we may apply Theorem 1 to this equivalent system, then use the isomorphism to bring the result back to G.

VI. The Branch Switching Algorithm

Suppose that a branch of solutions of equation (12) with a singular point (u_0, λ_0) is given, near which

$$u(s) = u_0 + s\dot{u}_0 + O(s^2)$$

$$\lambda(s) = \lambda_0 - O(s^K)$$

Theorem 1 (applied to an equivalent system), implies that there is also a branch near the singular point of the form

$$\tilde{u}(s) = u_0 + \alpha s \dot{u}_0 + O(s^2)$$

$$\tilde{\lambda}(s) = \lambda_0 + O(s^K),$$

where, $\alpha = (-1)^{1/K}$. This branch does not satisfy equation (12). The outgoing branch must be multiplied by an appropriate unitary scalar to make the k^{th} component of u real. Let

$$c_1 = (\overline{u}_0)_k / |(u_0)_k|,$$

and

$$c_2 = (\bar{\dot{u}}_0)_k / |(u_0)_k|.$$

The scalar is then $c_1 + \overline{\alpha} s c_2$. The rescaled branch is

$$\tilde{u}(s) = c_1 u_0 + s(\alpha c_1 \dot{u}_0 + \overline{\alpha} c_2 u_0) + O(s^2)$$

$$\tilde{\lambda}(s) = \lambda_0 + O(s^{K+1}),$$

which satisfies equation (12) and has real $\tilde{u}_k$ to first order in s. Because of the division it is desirable to choose k so that $|(u_0)_k|$ is as large as possible.

The branch switching algorithm requires that K be estimated at the singular point. Suppose that an interval of arclength $[-a, b]$ is given which contains a singular point at $s = 0$, and that in the interval

$$\lambda(s) \sim c s^K + O(s^{K+1}).$$

Then in the interval, $\dot{\lambda}(s)$ is given by

$$\dot{\lambda}(s) \sim c K s^{K-1} + O(s^K).$$

If K is even $\dot{\lambda}(-a)$ and $\dot{\lambda}(b)$ will have opposite sign. The branch switching algorithm for odd K simply multiplies the incoming tangent by 1, so we will consider only the case in which $\dot{\lambda}$ changes sign on $[-a, b]$.

K is defined such that $\lambda^{(K+1)}(0)$ is the first non-zero derivative of $\lambda(s)$ at the singular point. We must therefore test the even derivatives of $\lambda(s)$ to see if they vanish at the singular point. The end points of the interval are given by

$$\begin{aligned} a &= \frac{\Delta s\dot{\lambda}(-a)}{\dot{\lambda}(-a) - \dot{\lambda}(b)} + O(\Delta s^2) \\ b &= \frac{\Delta s\dot{\lambda}(b)}{\dot{\lambda}(b) - \dot{\lambda}(-a)} + O(\Delta s^2), \end{aligned} \tag{13}$$

and

$$\lambda^{(N)}(0) = b\lambda^{(N)}(-a) + a\lambda^{(N)}(b) + E_N(0). \tag{14}$$

Here $E_N(s)$ is the error of the Hermite interpolant for $\lambda^{(N)}$ in the interval $[-a, b]$,

$$E_n(0) = \frac{1}{(n+1)!} \max_{-a \le \zeta \le b} |\lambda^{(n+1)}(\zeta)| ab.$$

So if

$$|b\lambda^{(N)}(-a) + a\lambda^{(N)}(b)| > E_N(0)$$

$\lambda^{(N)}(0)$ is non-zero.

The derivatives of λ at the end points may be computed by solving the same linear system used to perform the Newton corrector. Let

$$\begin{aligned} u_N(s) &= \sum_{n=0}^{N} \frac{1}{n!} s^n u_0^{(n)} \\ \lambda_N(s) &= \sum_{n=0}^{N} \frac{1}{n!} s^n \lambda_0^{(n)}. \end{aligned}$$

Then,

$$\begin{aligned} G_u(u_N(s), \lambda_N(s))u^{(N+1)}(s) + G_\lambda(u_N(s), \lambda_N(s))\lambda^{(N+1)}(s) &= -\frac{N!}{s^N} G(u_N(s), \lambda_N(s)) + O(s) \\ 2\dot{u}_0^* u^{(N+1)}(s) + 2\dot{\lambda}_0^* \lambda^{(N+1)}(s) &= -\frac{N!}{s^N} \left(\left\| \frac{\mathrm{d}}{\mathrm{ds}} u_N(s) \right\|^2 + \left| \frac{\mathrm{d}}{\mathrm{ds}} \lambda_N(s) \right|^2 - 1 \right) + O(s). \end{aligned}$$

Taking the limit as $s \to 0$,

$$(15)\qquad \begin{aligned} G_x^0 u_0^{(N+1)} + G_\lambda^0 \lambda_0^{(N+1)} &= -\frac{\mathrm{d}^N}{\mathrm{d}s^N} G(u_N(s), \lambda_N(s))\Big|_{s=0} \\ 2\dot{u}_0^* u_0^{(N+1)} + 2\dot{\lambda}_0^* \lambda_0^{(N+1)} &= -\frac{\mathrm{d}^N}{\mathrm{d}s^N}\left(\left\| \frac{\mathrm{d}}{\mathrm{d}s} u_N(s) \right\|^2 + \left| \frac{\mathrm{d}}{\mathrm{d}s} \lambda_N(s) \right|^2 - 1 \right)\Big|_{s=0} . \end{aligned}$$

Given an interval $[-a, b]$ containing a singular point, the algorithm for determining K is this:

1. If $\dot{\lambda}(-a)\dot{\lambda}(b) > 0$, K is odd, return.
2. Compute a and b using equation (13).
3. For K even, starting with $K = 2$;
   a. Compute $\lambda^{(K)}$ at the end points of the interval using equation (15).
   b. Use equation (14) to estimate $\lambda^{(K)}(0)$.
   c. If it is non-zero, return.
   If it is zero, $K \leftarrow K + 2$. Repeat from step 3a.

VII. Practical Considerations

For large problems the Jacobian $A_{\underline{x}}$ usually has some structure which may be exploited to speed the solution of the linear systems. The introduction of the homogeneous coordinate and homotopy parameter, and the normalization, phase and pseudo-arclength constraints destroys this structure. It is therefore desirable to use a bordering algorithm to solve the linear systems [10]. The systems which must be solved here may be decomposed into an $N \times N$ complex system, with three real borders. (The real and imaginary parts of the component of u which is real, and the homotopy parameter, and the normalization, phase and pseudo-arclength constraints.) These complex systems, bordered by real constraints, may be solved using a modified bordering algorithm. To solve

$$Ax + by = r$$

$$Re(c^* x) + dy = s,$$

where

$$x, r \in \mathbb{C}^N, \quad y, s \in \mathbb{R}^M,$$

$$A : \mathbb{C}^N \to \mathbb{C}^N, \quad b : \mathbb{R}^M \to \mathbb{C}^N, \quad c^* : \mathbb{C}^N \to \mathbb{R}^M, \quad d : \mathbb{R}^M \to \mathbb{R}^M,$$

let

$$Av = r$$

$$Az = -b.$$

Then

$$y = (d - Re(c^* z))^{-1}(s - Re(c^* v))$$

$$x = v + dy.$$

For small systems, the linear systems may be rewritten as equivalent $2N + 3 \times 2N + 3$ real systems. This has the advantage that a full pivoting strategy may be used.

It is impractical to compute $\rho(s)$ in order to determine the step size Δs. Instead, we require that between 2 and 3 corrector iterations be done at each step. If too many iterations are required the step size is decreased by a factor of 1.2 and the step is repeated. If too few are required the step size is increased by a factor of 1.2 for the next step.

This has the disadvantage that the solution may 'jump' from one path to another. In order to detect any jumps we require that

1. $|u - u_p| < c\Delta s^2$.

2. $|\Delta s| < \Delta s_{\max}$.

3. u may not be real if u_p is non-real.

If any of these requirements is violated the step size is decreased by a factor of 1.2 and the step is repeated.

VIII. Examples

We have applied the above algorithm to several polynomial systems from [12], [11], [14], and [2]. For each we tabulate the solutions found, and plot the quantity $\sum[(\mathrm{Re}(t) - \mathrm{Im}(t))(\mathrm{Re}(x_i) + \mathrm{Im}(x_i))]$ as a function of the homotopy parameter. Real homotopy paths are shown in solid line, non-real paths in broken line. As a measure of the complexity of the algorithm we count the average number of steps along each homotopy path. Each step requires at most three LU decompositions and three back-solves.

(1) From Morgan [12], Problem 3.

$$x + 10y = 0$$

$$\sqrt{5}(z - w) = 0$$

$$(x - 2z)^2 = 0$$

$$\sqrt{10}(x - w)^2 = 0.$$

There is only one solution, with multiplicity 4. It is

	w	x	y	z	t
1	0.0 0.0	0.0 0.0	0.0 0.0	0.0 0.0	1.0 0.0

The Homotopy paths for this problem are shown in Figure 1. An average of 43 steps per root were required along the homotopy paths. There are four singular points on the homotopy paths. Two are simple quadratic folds (K=2), one is a simple cubic fold (K=3), and the fourth is the singular point at $\lambda = 1$. The last is a multiple bifurcation point, with a two dimensional nullspace. The branches are paired, each pair with K=2. In [12], 36 initial paths were required to obtain 4 which reached $\lambda = 1$. For that algorithm the paths averaged 18 steps per root.

(2) From Kojima and Mizuno [11], Problem 4-1

Determine x and y so that

$$\phi(x, y) = x^4 + y^4 - x^3 - x^2y + xy^2 - y^3 - x^2 + xy - y^2 + x - y.$$

is minimal. There are 9 extrema, given in the following table:

	x	y	t	ϕ
1	-0.531 0.000	0.733 0.000	0.425 0.000	-7.968 0.000
2	0.480 0.000	0.633 0.000	0.607 0.000	-0.987 0.000
3	0.577 0.000	0.577 0.000	0.577 0.000	-1.000 0.000
4	-0.401 0.084	-0.301 -0.316	0.801 0.000	-0.068 0.340
5	-0.401 -0.084	-0.301 0.316	0.801 0.000	-0.068 -0.340
6	0.293 0.012	-0.188 -0.305	0.886 0.000	0.330 0.155
7	0.293 -0.012	-0.188 0.305	0.886 0.000	0.330 -0.155
8	0.668 0.097	-0.111 0.327	0.652 0.000	0.118 -0.350
9	0.668 -0.097	-0.111 -0.327	0.652 0.000	0.118 0.350

The Homotopy paths for this system are shown in Figure 2. An average of 8 steps per root were required along the homotopy paths. There is only one singular point on the homotopy paths, a simple quadratic fold (K=2). In [11], each path required an average of 121 pivots. This is the number of simplex faces the path passed through. It therefore measures the number of function evaluations required.

(3) From Wright [14], Problem B

$$(x - y - 1)^2 = 0$$

$$x^2 - y^2 = 0$$

There are two solutions, both of multiplicity two.

	x	y	t
1	0.408 0.000	-0.408 0.000	0.826 0.000
2	0.707 0.000	0.707 0.000	0.000 0.000

The Homotopy paths for this system are shown in Figure 3. An average of 22 steps per root were required along the homotopy paths. In [14], approximately 431 steps were required per path.

We found four singular points on the homotopy paths. Three are simple quadratic folds (k=2), the last is at $\lambda = 1$ and has one pair of branches with K=1, and a second pair with K=2.

(4) From Brunovsky and Meravy [2], Problem 1

$$x + 10y = 20$$

$$x + 10y = -20$$

There is one solution,

	x	y	t
1	-0.995 0.000	0.100 0.000	0.000 0.000

The Homotopy paths for this system are shown in Figure 4. An average of 11 steps per root were required along the homotopy paths. Notice that the linear system is singular, and that the algorithm has found the null vector. There were no singular points on the homotopy path.

(5) Finally, we solve the discretized form of the two point boundary value problem

$$u'' + \tau^2\pi^2(u + \frac{1}{2}u^2) = 0,$$

$$u(0) = u(1) = 0,$$

This always has the trivial solution $u = 0$, and from this trivial solution other branches bifurcate at integer values of τ. For $\tau = 4.5$, and a discretization using six points on the interval, we found all 64 solutions. The first 8 are tabulated below.

	$u(h)$		$u(2h)$		$u(3h)$		$u(4h)$		$u(5h)$		$u(6h)$		t	
1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	0.000
2	-0.751	0.000	0.186	0.000	-0.063	0.000	0.019	0.000	-0.006	0.000	0.002	0.000	0.631	0.000
3	0.168	0.000	-0.717	0.000	0.180	0.000	-0.060	0.000	0.018	0.000	-0.005	0.000	0.649	0.000
4	-0.650	0.000	-0.627	0.000	0.148	0.000	-0.051	0.000	0.015	0.000	-0.004	0.000	0.400	0.000
5	-0.054	0.000	0.178	0.000	-0.713	0.000	0.179	0.000	-0.059	0.000	0.016	0.000	0.649	0.000
6	-0.596	0.000	0.237	0.000	-0.541	0.000	0.137	0.000	-0.045	0.000	0.012	0.000	0.525	0.000
7	0.140	0.000	-0.630	0.000	-0.629	0.000	0.149	0.000	-0.051	0.000	0.014	0.000	0.404	0.000
8	-0.527	0.000	-0.595	0.000	-0.506	0.000	0.119	0.000	-0.041	0.000	0.011	0.000	0.312	0.000

The Homotopy paths for this system are shown in Figure 5. An average of 39 steps per root were required along the homotopy paths.

IX. Conclusions

There are several features which a homotopy algorithm for finding the roots of a polynomial system should incorporate. It must use a homogeneous parameter to keep the homotopy path bounded, it must use a reliable continuation method, and it must have some mechanism for switching branches at singular points.

Of the existing algorithms, [7] uses simplicial continuation, which is reliable and incorporates an implicit branch switching technique, but is expensive for large systems. The algorithm of [3] uses a predictor-corrector continuation method with poor behavior near singular points, and a genericity argument to construct a homotopy which is free of singular points. The algorithm of [6] use a similar continuation method, and deforms the homotopy parameter into the complex plane to avoid singular points.

The algorithm we have described uses pseudo-arclength continuation, which is a reliable technique, and which does not fail near singular points [5]. At singular points we use a new branch switching result, which requires that the incoming tangent to the homotopy branch be multiplied by an appropriate complex scalar. This is, in a local sense, equivalent to [6], being an excursion along a small circle about the singular point, in the complex homotopy parameter space.

We have given five example problems. These illustrate the problem with genericity arguments, for although singular points are not generic, the initial system we chose exhibits singular points in three of the five examples. This is due mainly to out choice of an initial with real coefficients. Such systems have roots which occur in conjugate pairs, and as a consequence a complex path cannot become real without becoming singular. This ensures that a path will be singular if the system being solved does not have the same number of complex roots as the initial system. The singular paths are still of measure zero, but the set of homotopies has been restricted to the point that the set of singular homotopies has non-zero measure.

X. References

[1] Allgower, E. and Georg, K. *Simplicial and Continuation Methods for Approximating Fixed Points and Solutions to Systems of Equations*, SIAM Review **22** (Jan. 1980) pp. 28-85.

[2] Brunovský, P. and Meravý, P., *Solving Systems of Polynomial Equations by Bounded and Real Homotopy*, Numerische Mathematik, vol. 43, 1984 pp. 397-418. pp. 1-17.

[3] Chow, S.N., Mallet-Paret, J. and Yorke, J.A., *Finding Zeroes of Maps: Homotopy Methods that are Constructive with Probability One*, Mathematics of Computation, vol. 32 #143, July 1978 pp. 887-899.

[4] Chow, S.N., Mallet-Paret, J. and Yorke, J.A., *A Homotopy Method for Locating all Zeroes of a System of Polynomials*, in Functional Differential Equations and Approximation of Fixed points, H.O. Peitgen and H.O. Walther eds., Lecture Notes in Applied Mathematics 730, Springer-Verlag 1979, pp. 77-88.

[5] Decker, D.W. and Keller, H.B., *Path Following near Bifurcation*, Communications on Pure and Applied Mathematics, vol.34, pp. 149-175 (1981).

[6] Drexler, F.J., *A Homotopy Method for the Calculation of all Zeroes of Zero-Dimensional Polynomial Ideals*, in Continuation Methods, H. Wacker ed., Academic Press, 1978 pp. 69-93.

[7] Garcia, C.B. and Zangwill, W.I., *Determining all Solutions to Certain Systems of Nonlinear Equations*, Mathematics of Operations Research, vol. 4 #1, Feb. 1979 pp. 1-14.

[8] Garcia, C.B. and Zangwill, W.I., *Finding all Solutions to Polynomial Systems and Other Systems of Equations*, Mathematical Programming, vol. 16, 1979 pp. 159-176.

[9] Garcia, C.B. and Li, T.Y., *On the Number of Solutions to Polynomial Systems of Equations*, SIAM Journal of Numerical Analysis, vol. 17, #4, 1990 pp. 540-546.

[10] Keller, H.B., *Numerical Solution of Bifurcation and Nonlinear Eigenvalue Problems*, in Applications of

Bifurcation Theory, P. Rabinowitz ed. Academic Press, 1977.

[11] Kojima, M. and Mizuno, S., *Computation of all Solutions to a System of Polynomial Equations*, Mathematical Programming 25 (1983) pp. 131-157.

[12] Morgan, A.P., *A Method for Computing all Solutions to Systems of Polynomial Equations*, ACM Transactions on Mathematical Software, vol. 9 #1, Mar. 1983 pp. 1-17.

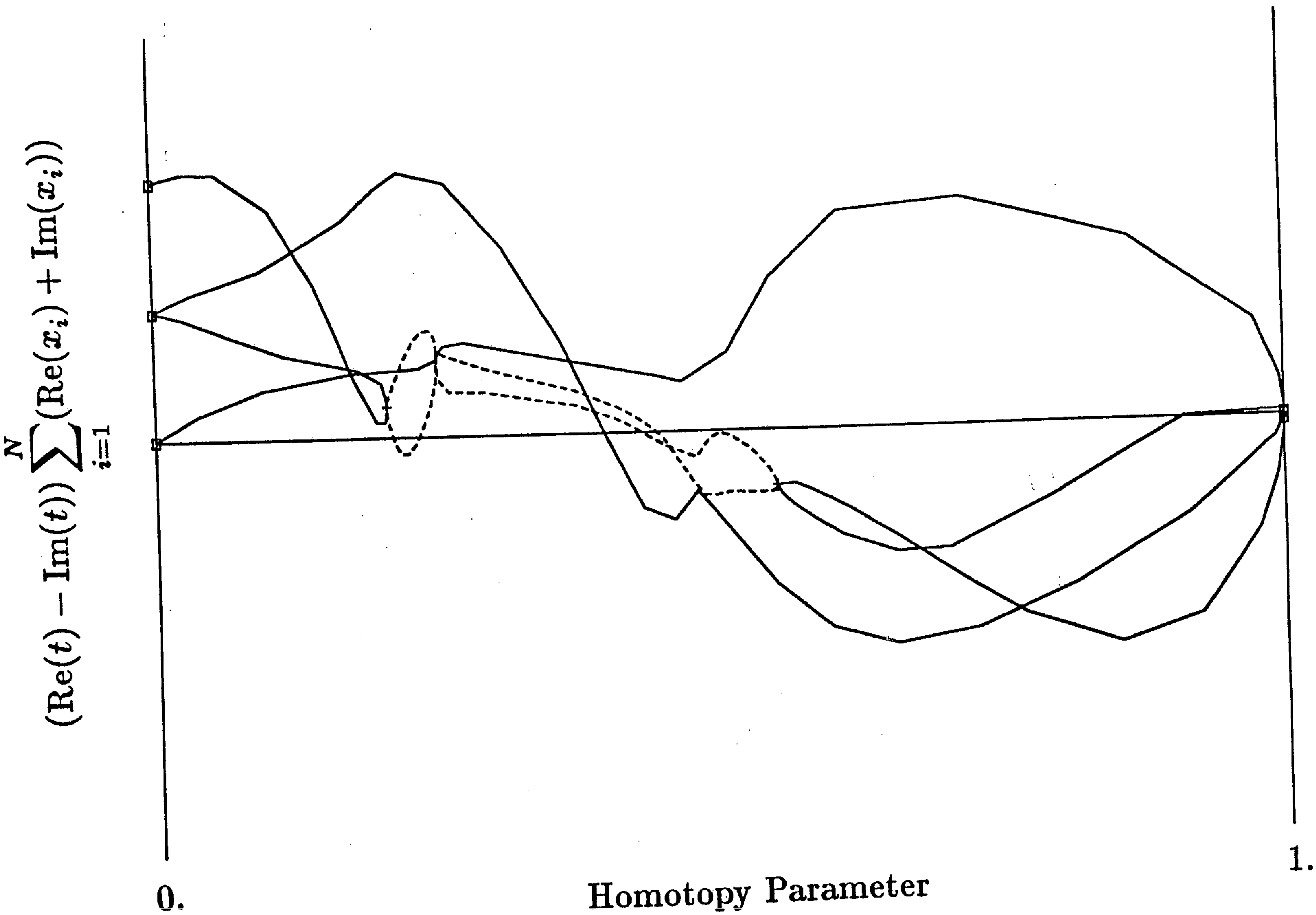
[13] Vainberg, M.M. and Trenogin, V.A. *The Methods of Lyapunov and Schmidt in the Theory of Non-Linear Equations and their Further Development*, Russian Math Surveys 17 #2 (1962) pp. 1-60.

[14] Wright, A.H., *Finding all Solutions to a System of Polynomial Equations*, Mathematics of Computation, vol. 44 #169, Jan. 1985 pp. 125-133. pp. 1-17.

L	K	Algebraic Bifurcation Equations
2	1	$F^0_{xx}\dot{x}_0\dot{x}_0 + 2F^0_{x\lambda}\dot{x}_0\dot{\lambda}_0 + F^0_{\lambda\lambda}\dot{\lambda}_0\dot{\lambda}_0 = 0$ $\Vert \dot{x}_0 \Vert^2 + \vert\dot{\lambda}_0\vert^2 = 1$
	2	$F^0_{xx}\dot{x}_0\dot{x}_0 + \frac{1}{2}F^0_{\lambda}\ddot{\lambda}_0 = 0$ $\Vert \dot{x}_0 \Vert^2 = 1$
3	1	$F^0_{xxx}\dot{x}_0\dot{x}_0\dot{x}_0 + 3F^0_{xx\lambda}\dot{x}_0\dot{x}_0\dot{\lambda}_0 + 3F^0_{x\lambda\lambda}\dot{x}_0\dot{\lambda}_0\dot{\lambda}_0 + F^0_{\lambda\lambda\lambda}\dot{\lambda}_0\dot{\lambda}_0\dot{\lambda}_0 = 0$ $\Vert \dot{x}_0 \Vert^2 + \vert\dot{\lambda}_0\vert^2 = 1$
	2	$F^0_{xxx}\dot{x}_0\dot{x}_0\dot{x}_0 + F^0_{x\lambda}\dot{x}_0\ddot{\lambda}_0 = 0$ $\Vert \dot{x}_0 \Vert^2 = 1$
	3	$F^0_{xxx}\dot{x}_0\dot{x}_0\dot{x}_0 + \frac{1}{6}F^0_{\lambda}\lambda_0^{(3)} = 0$ $\Vert \dot{x}_0 \Vert^2 = 1$
4	1	$F^0_{xxxx}\dot{x}_0\dot{x}_0\dot{x}_0\dot{x}_0 + 4F^0_{xxx\lambda}\dot{x}_0\dot{x}_0\dot{x}_0\dot{\lambda}_0 + 6F^0_{xx\lambda\lambda}\dot{x}_0\dot{x}_0\dot{\lambda}_0\dot{\lambda}_0 + 4F^0_{x\lambda\lambda\lambda}\dot{x}_0\dot{\lambda}_0\dot{\lambda}_0\dot{\lambda}_0 + F^0_{\lambda\lambda\lambda\lambda}\dot{\lambda}_0\dot{\lambda}_0\dot{\lambda}_0\dot{\lambda}_0 = 0$ $\Vert \dot{x}_0 \Vert^2 + \vert\dot{\lambda}_0\vert^2 = 1$
	2	$F^0_{xxxx}\dot{x}_0\dot{x}_0\dot{x}_0\dot{x}_0 + \frac{3}{2}F^0_{xx\lambda}\dot{x}_0\dot{x}_0\ddot{\lambda}_0 + \frac{1}{4}F^0_{\lambda\lambda}\ddot{\lambda}_0\ddot{\lambda}_0 = 0$ $\Vert \dot{x}_0 \Vert^2 = 1$
	3	$F^0_{xxxx}\dot{x}_0\dot{x}_0\dot{x}_0\dot{x}_0 + \frac{1}{3}F^0_{x\lambda}\dot{x}_0\lambda_0^{(3)} = 0$ $\Vert \dot{x}_0 \Vert^2 = 1$
	4	$F^0_{xxxx}\dot{x}_0\dot{x}_0\dot{x}_0\dot{x}_0 + \frac{1}{24}F^0_{\lambda}\lambda_0^{(4)} = 0$ $\Vert \dot{x}_0 \Vert^2 = 1$

Table 1. The Algebraic Bifurcation Equations for M=1

Figure 1.



$(\mathrm{Re}(t) - \mathrm{Im}(t)) \sum_{i=1}^{N} (\mathrm{Re}(x_i) + \mathrm{Im}(x_i))$
0.
1.
Homotopy Parameter

2.

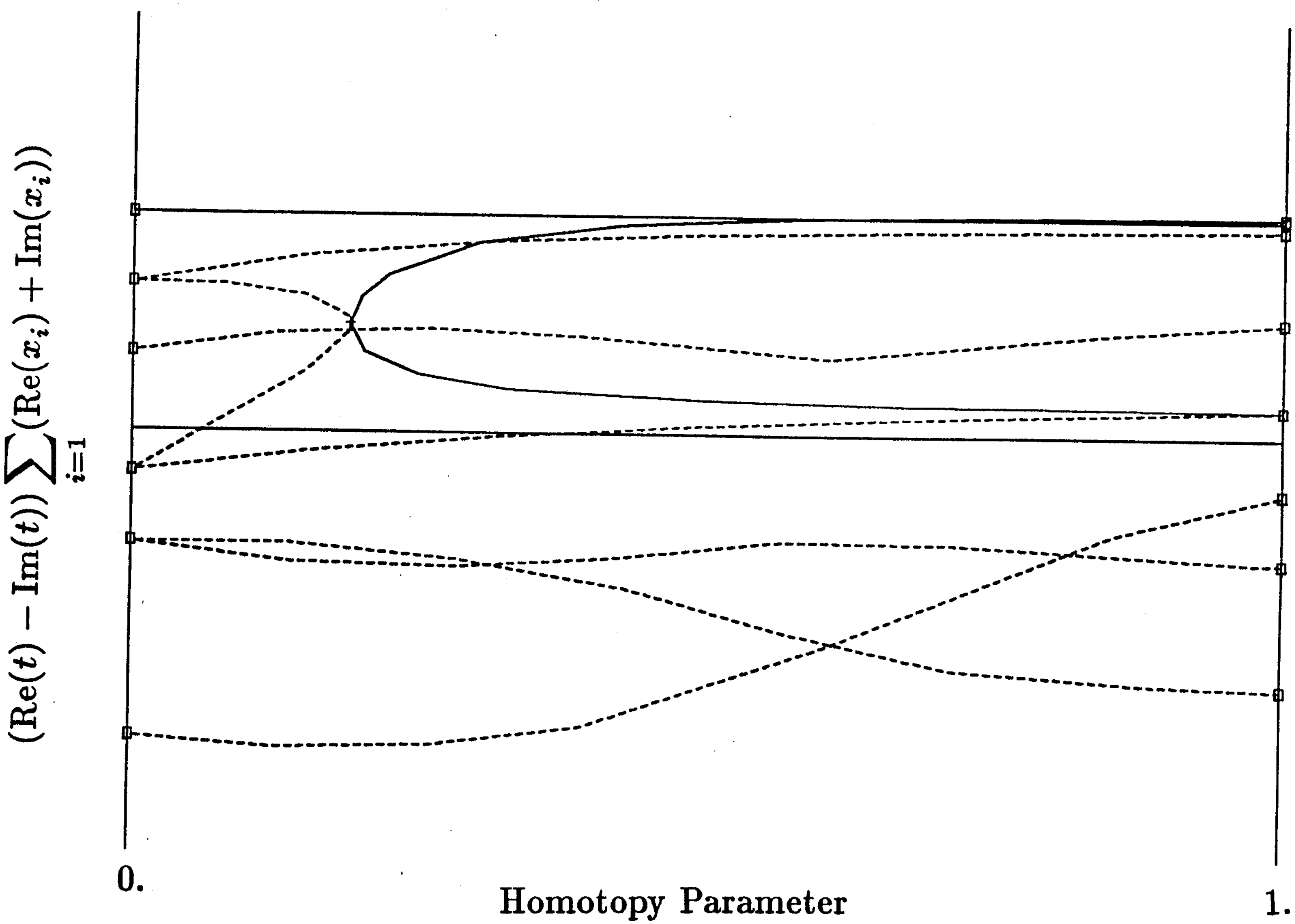
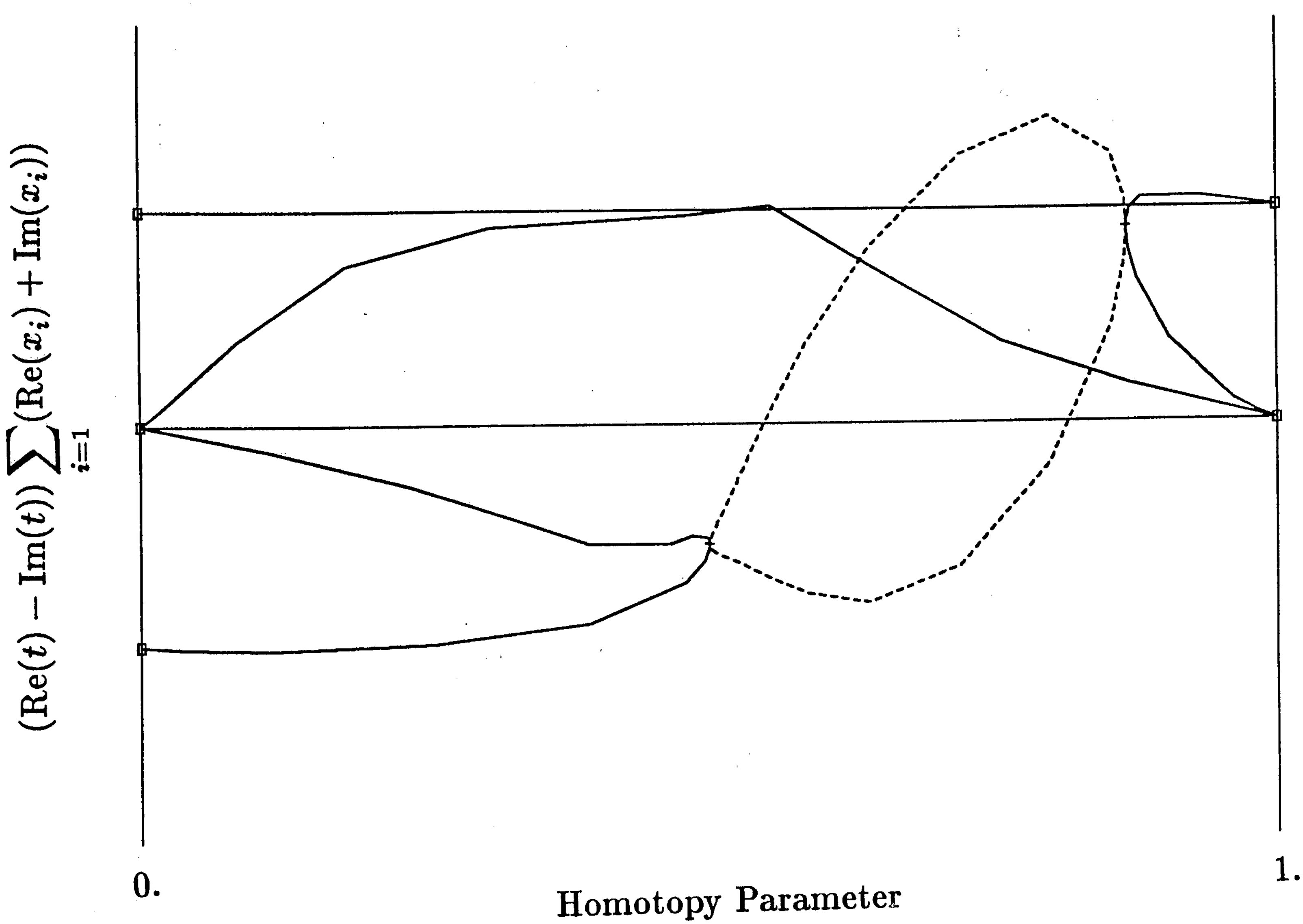

$(\mathrm{Re}(t) - \mathrm{Im}(t)) \sum_{i=1}^{N} (\mathrm{Re}(x_i) + \mathrm{Im}(x_i))$
0.
1.
Homotopy Parameter

Figure 2.

Figure 3.


$(\mathrm{Re}(t) - \mathrm{Im}(t)) \sum_{i=1}^{N} (\mathrm{Re}(x_i) + \mathrm{Im}(x_i))$
0.
1.
Homotopy Parameter

4

Figure 4.

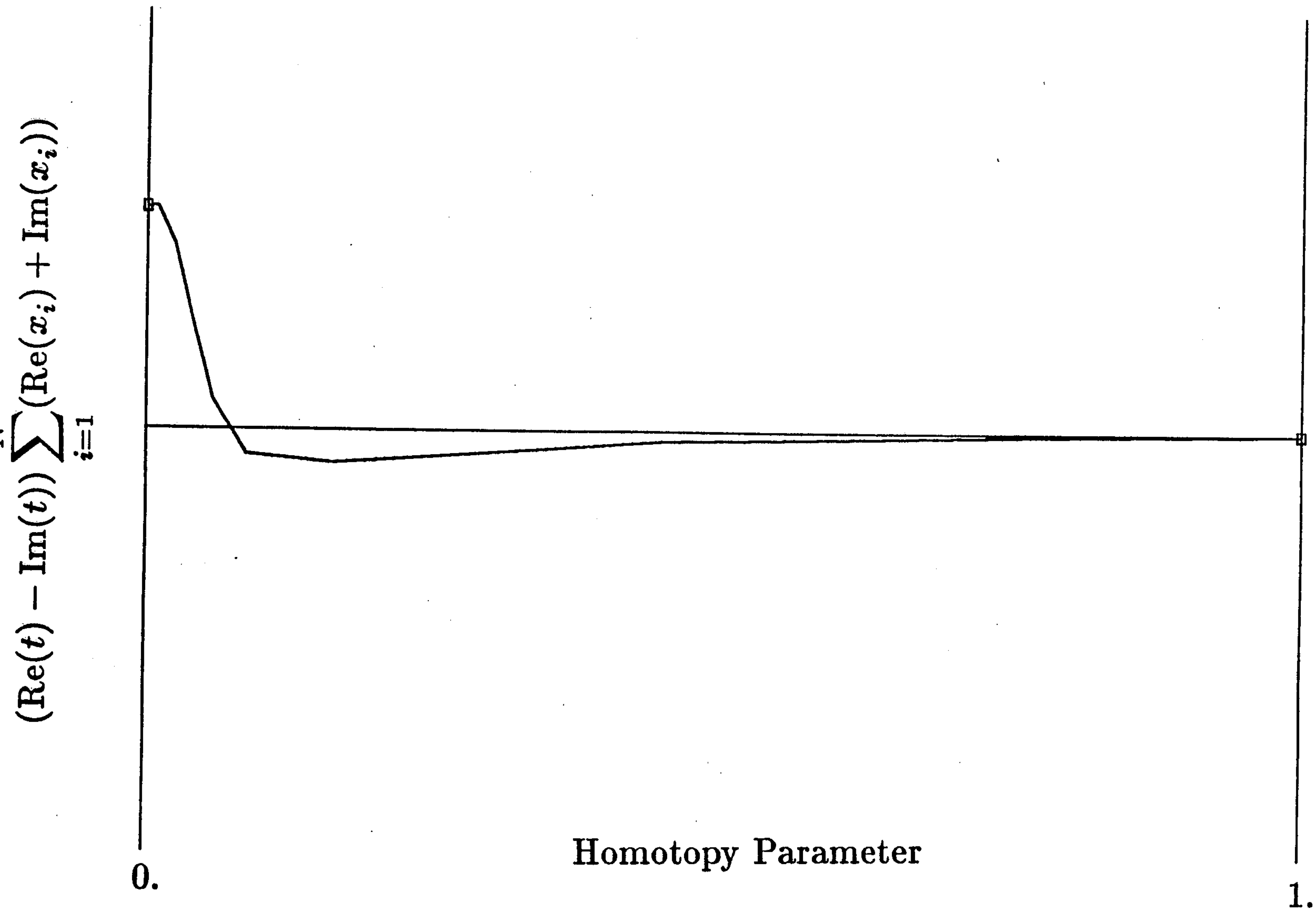
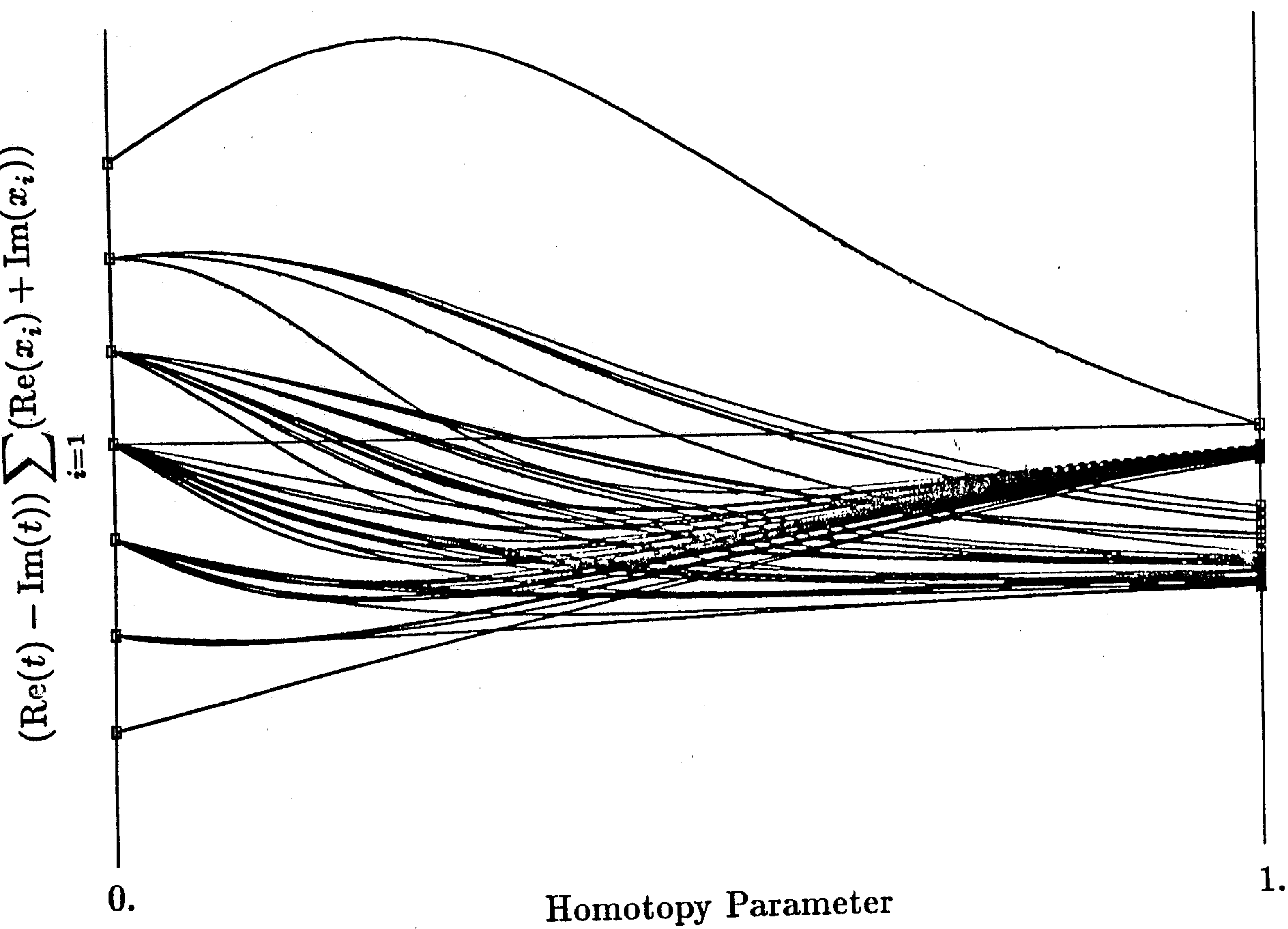

$(\mathrm{Re}(t) - \mathrm{Im}(t)) \sum_{i=1}^{N} (\mathrm{Re}(x_i) + \mathrm{Im}(x_i))$
0.
1.
Homotopy Parameter

Figure 5.



$(\mathrm{Re}(t) - \mathrm{Im}(t)) \sum_{i=1}^{N} (\mathrm{Re}(x_i) + \mathrm{Im}(x_i))$
0.
1.
Homotopy Parameter