Symbolic and Numerical Computation for Artificial Intelligence

edited by

Bruce Randall Donald
Department of Computer Science
Cornell University, USA

Deepak Kapur
Department of Computer Science
State University of New York, USA

Joseph L. Mundy
AI Laboratory
GE Corporate R&D, Schenectady, USA

Academic Press
Harcourt Brace Jovanovich, Publishers
London San Diego New York
Boston Sydney Tokyo Toronto

ACADEMIC PRESS LIMITED
24–28 Oval Road
London NW1

US edition published by
ACADEMIC PRESS INC.
San Diego, CA 92101

This book is printed on acid-free paper

A catalogue record for this book is available from the British Library

ISBN 0-12-220535-9

Printed and Bound in Great Britain by
The University Press, Cambridge

Solving Polynomial Constraints

Chapter 1

Polynomial Continuation and its Relationship to the Symbolic Reduction of Polynomial Systems

Alexander P. Morgan

Mathematics Department

General Motors Research Laboratories

Warren, Michigan 48090

morgan@gmr.com

Polynomial continuation is an approach to constructing numerical methods for computing the full list of geometrically isolated solutions to a system of polynomial equations. These methods have a solid theoretical basis (in algebraic geometry and numerical analysis) and a high degree of computational reliability (if properly implemented). They also require a significant amount of computational work, limiting the collection of solvable systems to those that are "small". A given polynomial system can often be made "smaller" via symbolic transformations, but this kind of "reduction" can also generate numerical difficulties. This complicates the development of symbolic reduction algorithms. This paper outlines basic and advanced forms of polynomial continuation. Included is a definition of the size of a system (from the point of view of polynomial continuation) and a discussion of the kinds of symbolic reduction which might be most useful to realize in computational algorithms. Examples from geometric modeling, chemical engineering, and mechanical engineering are used for illustration.

1. Introduction

Polynomial continuation is concerned with computing the complete list of geometrically isolated† solutions to polynomial systems. It can be made more efficient if augmented by a symbolic "reduction" of the systems to be solved. Although the theoretical and numerical aspects of polynomial continuation have been considerably developed, little has been done to create systematic reduction methods as an adjunct to it. In this expository paper we shall give an outline of some basic facts about polynomial continuation and then present some ideas on how effective tools for symbolic reduction might be structured for use with it. Readers whose main interest is symbolic methods may wish to read section 5 first.

† Defined at the beginning of section 3.

By way of introduction to polynomial continuation, let us consider a simple example. Suppose we wish to find all solutions to the system

$$\begin{aligned} x^2 + xy - 1 &= 0 \\ y^2 + x - 5 &= 0 \end{aligned} \tag{1.1}$$

The key "trick" of the method is to note that the related simplified system

$$\begin{aligned} x^2 - 1 &= 0 \\ y^2 - 5 &= 0 \end{aligned} \tag{1.2}$$

is easy to solve. The four solutions to (1.2) are $(x, y) = (\pm 1, \pm\sqrt{5})$. By defining the "homotopy"

$$h(x, y, t) = \begin{bmatrix} x^2 + & txy - 1 \\ y^2 + & tx - 5 \end{bmatrix}$$

the following strategy is developed: We can solve

$$h(x, y, 0) = \begin{bmatrix} x^2 + & 0xy - 1 \\ y^2 + & 0x - 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

This is, in fact, (1.2). Beginning with a solution to $h(x, y, 0) = 0$, we can modify it to be a solution to $h(x, y, t_1) = 0$, where t_1 is a small positive number, using some local solution method. Then we can modify the solution to $h(x, y, t_1) = 0$ to become a solution to $h(x, y, t_2) = 0$ where $t_2 > t_1$, and by proceeding in this way, finally derive a solution to $h(x, y, 1) = 0$, which is (1.1). Such a sequence of solutions is called a "continuation path". The idea is that each solution to (1.2) yields a solution to (1.1) via continuation.

Certain technical issues not made explicit in the example must be overcome in general. Thus, we must recast the problem in complex arithmetic, because in general a polynomial system will have complex solutions, even if its coefficients are real. We must take steps to assure that the continuation paths are well defined and do not bifurcate, turn back in t, "go singular", or get bogged down in some other way. We must devise fast and reliable numerical methods for path tracking.

Basically, we can distinguish two main steps:

1 Define a homotopy, $h(z, t)$.
2 Choose a numerical method for tracking the paths defined by $h(z, t) = 0$.

Step 1 is guided by results from algebraic geometry, while step 2 is based on methods for the numerical solution of ordinary differential equations and local methods for the solution of nonlinear systems. Both parts are nontrivial and important.

In the next section, we overview basic polynomial continuation, in which the number of paths to be tracked is equal to the *total degree* of the system, and we consider the theoretical and numerical issues that arise. Then, in sections 3 and 4, we touch on several advanced topics, defining *multi-homogeneous systems* and their *Bezout numbers* and *coefficient-parameter polynomial continuation* and *side conditions*. Much of this exposition is drawn from Morgan (1989) and Morgan and Sommese (1990). The book (Morgan, 1987) and the paper (Wampler *et al.*, 1990a) are recommended for further expository material. Section 5 puts forward some proposals for symbolic reduction tools that would complement polynomial continuation. The standard Gröbner basis algorithms typically

reduce a polynomial system to triangular form, which is often useless for fixed-precision floating point computations because of issues of numerical stability. However, less extreme reductions, which might be accomplished by the same basic algebra, could be very useful.

The idea of polynomial continuation was first suggested in Drexler (1977) and Garcia and Zangwill (1977). A number of papers have followed and there have also been significant physical applications of (see the references).

It is worth noting that polynomial continuation is especially efficient on parallel computers; see Morgan and Watson (1986, 1987, 1989). In Wampler *et al.* (1990b), better than mainframe speeds are obtained on an inexpensive parallel workstation. This feature of the approach greatly extends the range of solvable problems.

We generally say that only "small" problems are solvable in practice via polynomial continuation. However, *small* is not a precise concept. It depends in part on run-time expectations. Thus, a minute of run time might be acceptable for many applications, while others require run times in milliseconds. Most systems of two or three equations of low degree (less than six) can be solved in a minute or so on most workstations; in this sense, such systems are "small". On the other hand, a system of ten second-degree equations might take an hour on a typical mainframe, ten hours on a workstation, and five minutes on a 100-node parallel network. This still might be viewed as small. We have solved systems requiring the tracking of over half a million paths, involving hundreds of hours of mainframe time. Such systems seem unlikely candidates for the designation "small". However, parallel workstations are becoming available on which such a system would be solvable in a day. In any case, we will always be interested in making systems *smaller*, the observation which has inspired this paper.

2. Basic Polynomial Continuation

In this section we sketch the theory of polynomial continuation in the simplest and most straightforward case: the number of continuation paths is equal to the total degree of the system. The results in this section are stated more precisely and are generalized in section 3. The simple approach given here can be applied to any polynomial system, but the refinements can sometimes be much more efficient. The comments in this section on the numerical aspects of path tracking remain valid in the more sophisticated contexts presented in later sections.

2.1. Paths in Complex Euclidean Space, C^n

Let $f(z) = 0$ denote the system of n polynomial equations in n unknowns that we want to solve. Call this the *target system*. The target system can have real or complex coefficients. Let $d_j = \deg(f_j)$, the degree[†] of the j^{th} equation of the target system, for $j = 1$ to n. Let $d = d_1 \cdot d_2 \cdots d_n$, the *total degree* of f.

[†] The degree of a polynomial is the maximum of the degrees of its terms. The degree of a term is the sum of the exponents of the variables. Thus, the three terms of the polynomial $x^4 + x^2y^3 + 1$ have degrees 4, 5, and 0, respectively, and the degree of the polynomial is 5.

We choose a system $g(z) = 0$ of n polynomial equations in n unknowns, the *start system*. Since g is the start system, we must know what its solutions are. In addition, g must obey the following conditions:

- $\deg(g_j) = d_j$ for $j = 1$ to n.
- $g(z) = 0$ has d different isolated solutions.†

The system defined by $g_j = z_j^{d_j} - 1$ for $j = 1$ to n is an acceptable choice, but if we choose g to match more of the structure of f, then the method will generally be more efficient. For example, g might equal f with a few terms deleted.

Define the *homotopy* (or *continuation system*) by

$$h(z,t) = (1-t)\gamma g(z) + tf(z) \tag{2.1}$$

where $\gamma = re^{i\theta}$. Here $i = \sqrt{-1}$, e is the base of the natural logarithms, θ is a real number "chosen at random", and r is a positive real number. The theorem which justifies this homotopy (Theorem 3.2 in section 3) states that, for each choice of f and g, there are a finite number of θ that lead to singular continuation paths; i.e. paths that cross, bifurcate, or explode into higher-dimensional components. We want to avoid these unknown "bad" θ, because the associated "paths" are more difficult (or impossible) to track effectively. Since there are only a finite number of bad θ, there is a zero probability of choosing one of them at random.

Geometrically, the solutions $U = \{(z,t) \in C^n \times [0,1) \quad | \quad h(z,t) = 0\}$ form d distinct smooth nonintersecting paths in $C^n \times [0,1)$, each an embedding of a copy of $[0,1)$ in $C^n \times [0,1)$. These paths (with their righthand endpoints included, as discussed below) are called the *continuation paths* for the homotopy (2.1).

Let us consider one of these paths, $\alpha(s) = (z(s), t(s))$, parametrized by arc length, s. Thus $s \geq 0$ and, as s increases, t goes from 0 to 1. Further, t is strictly increasing as a function of s. Call $z^0 = z(0)$ the *lefthand end point* of the path. We have two cases:

- If the path has finite arc length, s^*, then $\lim_{s \to s^*} t(s) = 1$ and $\lim_{s \to s^*} z(s) = z^1$, the *righthand end point* of the path.
- If the path does not have finite arc length, then $\lim_{s \to \infty} t(s) = 1$ and $\lim_{s \to \infty} |z(s)| = \infty$. In this case, $f(z) = 0$ has a *solution at infinity*, noted below and discussed in the next section.

The proof that these cases exhaust the possibilities is based on the fact that $h^{-1}(0)$ is a 1-dimensional (real) smooth manifold for $t \in [0,1)$ and on the way complex Euclidean space embeds in complex projective space. (See, for example, Morgan, 1986a; Morgan and Sommese, 1987a.)

Now, the d lefthand endpoints of the d paths are the solutions to $g = 0$. Further, each isolated solution of $f(z) = 0$ shows up as a righthand end point of some path. Thus, if we track the paths numerically, beginning with the solutions to $g = 0$, we can find all the isolated solutions of $f(z) = 0$. (As a part of this process, the diverging paths will have

† Equivalently, all the solutions of $g(z) = 0$ are *nonsingular*; that is, if x^0 is a solution to $g(x) = 0$, then $dg(x^0)$ is nonsingular, where $dg(x^0)$ denotes the $n \times n$ Jacobian matrix of partial derivatives of g evaluated at x^0.

to be terminated. However, the projective transformation, described in (2.2), eliminates divergent paths.)

Path tracking is effected by solving a set of "initial value problems with energy function," as follows. Let $\dot{z}$ and $\dot{t}$ denote dz/ds and dt/ds, respectively. Denote by dh_z and dh_t the partial derivative matrices of h with respect to z and t, respectively (see Note 3 below). Then continuation paths are solutions to the initial value problem

$$\dot{z}(s) = -\delta[dh_z]^{-1}dh_t \qquad \dot{t}(s) = \delta \tag{2.2}$$

where δ is a positive constant chosen so that $|(\dot{z}, \dot{t})| = 1$, with initial conditions

$$z(0) = z^0 \qquad t(0) = 0 \tag{2.3}$$

for some solution z^0 to $g = 0$, having potential energy zero:

$$h(z(s), t(s)) = 0. \tag{2.4}$$

Notes:

1 Since $|(\dot{z}, \dot{t})| = 1$, the paths are parametrized by arc length.

2 The fact that (2.4) holds as well as (2.2) and (2.3) gives us some options for path tracking not available if this path were merely defined by an initial value problem. Thus we may augment standard ordinary differential equation solvers with special "path correction" devices (as described in Watson, 1979; Watson *et al.*, 1987), or we may use "prediction-correction" methods (as described in Allgower and Georg, 1980; Garcia and Zangwill, 1981; Morgan, 1987; Rheinboldt and Burkardt, 1983; Watson *et al.*, 1987) which require such additional information. Simplicial path tracking strategies also require it. See Allgower and Georg (1980) and Garcia and Zangwill (1981).

3 The dh_z is a $2n \times 2n$ real matrix, the "realification" of the $n \times n$ complex Jacobian matrix of partial derivatives of h, evaluated at $(z(s), t(s))$. If A is a $p \times q$ complex matrix, its *realification* is the $2p \times 2q$ matrix defined by replacing the entry $a_{i,j}$ by the 2×2 block $\begin{bmatrix} \mathrm{Re}(a_{i,j}) & -\mathrm{Im}(a_{i,j}) \\ \mathrm{Im}(a_{i,j}) & \mathrm{Re}(a_{i,j}) \end{bmatrix}$ where "Re" and "Im" denote the real and imaginary parts, respectively. Similarly, dh_t is the first column of the realification of the $n \times 1$ column matrix of partial derivatives of h with respect to t, evaluated at $(z(s), t(s))$. Further, for consistency with these conventions, we must interpret z and $\dot{z}$ as column vectors of real numbers; thus, $z = [Re(z_1), Im(z_1), \ldots, Re(z_n), Im(z_n)]^T$. This conversion of matrices via realification acknowledges that, while h is algebraically in complex arithmetic, we track continuation paths in real space. Thus, for n polynomials in n unknowns, we end up tracking paths in $(2n + 1)$-dimensional real space; see Morgan (1987, appendix 3) for more details. Note that t and $\dot{t}$ are real numbers. In fact, it is often convenient to implement polynomial continuation algorithms in complex computer arithmetic. The real space is conceptual.

4 In most cases, to generate the right hand side of (2.2), we will solve the linear system

$$dh_z\ \dot{z} = -dh_t \tag{2.5}$$

and not invert dh_z directly. Actually, (2.5) and (2.2) are derived from the $2n \times (2n + 1)$

linear system

$$[dh_z \mid dh_t] \begin{bmatrix} \dot{z} \\ \dot{t} \end{bmatrix} = 0. \tag{2.6}$$

Some researchers recommend alternative ways to solve for $\dot{z}$ and $\dot{t}$ using (2.6). See, for example, Watson (1986). However, we have had good results with (2.5).

5 The major computational challenge of the scheme we have outlined above is singular righthand endpoints. If all the solutions of $f(z) = 0$ are nonsingular, all the endpoints are nonsingular. This case is relatively routine to compute, but in practice it is the exception. Consequently, performance in the presence of singular endpoints is a major consideration for choosing a path tracker for polynomial continuation. Special methods for computing singular righthand endpoints are given in Morgan *et al.* (1991a, 1991b, 1991c).

6 We must also note the issue of *scaling*. The performance of a numerical method in solving a polynomial system is generally influenced by the scaling of the system. For example, poor scaling can lead to catastrophic overflows or underflows which degrade the performance of the solver. More subtle effects can also occur. Consider that the relative scaling of f and g will influence the geometry of the continuation paths associated with h. Little exact information on how to scale for polynomial continuation is available. The projective transformation generally has a helpful scaling effect (see the following subsection). The SCLGEN polynomial scaling algorithm (Meintjes and Morgan, 1987), originally created to tame the extreme scaling of chemical equilibrium systems, uses a simple optimization approach to minimize deviations in the coefficients from unity. Computational experience suggests that the performance of polynomial continuation is usually improved by using both of these together. See Morgan (1987) and Morgan *et al.* (1989) for more details.

2.2. Divergent Paths and the Projective Transformation

In the method described in section 2.1 above, paths may diverge. Since $|z(s)| \to \infty$ as $t(s) \to 1$ for these paths, a test must be implemented to terminate the paths when $|z|$ gets too large or when too many steps have been taken. While this can be satisfactory [even *recommended* in some cases (Morgan, 1987, Chapter 10)], usually it is better to make a simple change of context to eliminate divergent paths altogether. This eliminates the heuristic element in the numerical method resulting from programming the decision to terminate paths that appear to be diverging.

The approach we recommend is to replace the continuation system h by the *projective transformation* of h. We will sketch the mechanics of this substitution here. Section 3 gives more detail. An alternative mechanism for eliminating divergent paths is suggested in Wright (1985).

First, *homogenize* h via the substitutions $z_j \leftarrow y_j/y_0$ for $j = 1$ to n. This yields a new polynomial system (after we clear the powers of y_0 from the denominator): $\hat{h}(y_0, y_1, \ldots, y_n, t) = 0$ of n equations in $n + 2$ unknowns. We can identify the solutions to $h(z,t) = 0$ with the solutions to $\hat{h}(y,t) = 0$ obeying $y_0 = 1$. In addition, there may be solutions to $\hat{h}(y,t) = 0$ with $y_0 = y_1 = \ldots = y_{k-1} = 0$ and $y_k = 1$ for some

k between 1 and n. These do not correspond to solutions of $h(z,t) = 0$. (They, in fact, correspond to *solutions at infinity* of $h(z,t) = 0$. See section 3.)

Now define H, a system of n polynomials in $n+1$ unknowns, via $H(x_1,\ldots,x_n,t) = \hat{h}(x_0,x_1,\ldots,x_n,t)$ where $x_0 \equiv b_1x_1 + \ldots + b_nx_n + b_{n+1}$ for some fixed complex numbers $b_1,\ldots,b_{n+1}$. This H is called the *projective transformation* of h.

We have the following facts about H (see Morgan, 1986b; Morgan and Sommese, 1987b, 1989):

- For random choices of $b = (b_1,\ldots,b_{n+1}) \in R^{n+1}$ and $\theta \in R^1$, the homotopy $H(x,t)$ has no divergent continuation paths. (It also suffices to choose random $b \in C^{n+1}$ and/or random $\gamma \in C^1$.)
- If z^* is an isolated solution to $f(z) = 0$, then there is an isolated solution x^* to $H(x,1) = 0$ with $x_0^* \neq 0$, so that $z^* = \frac{1}{x_0^*}\, x^*$.

Thus, if we track the continuation curves defined by H, then we can recover all the isolated solutions of $f(z) = 0$ from those of $H(x,1) = 0$. The advantage is that we do not have to decide when to truncate paths that seem to be diverging, since no paths diverge. Note that H has the same total degree as h, so the number of continuation paths is the same. There are additional computational advantages to the projective transformation, in that it generally improves the scaling of the problem, as noted above.

3. Multi-homogeneous Polynomial Continuation

In this section we outline the multi-homogeneous method of constructing polynomial homotopies, introduced in Morgan and Sommese (1987a) as the "m-homogeneous" approach.

3.1. The multi-homogeneous context

First, we need the definitions of a *geometrically isolated* solution and the *multiplicity* of a solution.

Definition 3.1. *A solution to a polynomial system is called geometrically isolated (or simply isolated) if there is a ball around the solution that contains no other solution.*

A solution that is not geometrically isolated is singular, but an isolated solution can be singular also.

Definition 3.2. *Let z^* be a geometrically isolated solution to the polynomial system $f(z) = 0$. Let U be a closed ball about z^* containing no other solution. We can 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. For all sufficiently small coefficient perturbations, the perturbed systems have a constant number, m, of solutions in U. This m is (by definition) the multiplicity of z^*.*

A solution z^* to $f(z) = 0$ has multiplicity greater than one exactly when it is singular; that is, when the Jacobian matrix $df(z^*)$ is singular.

The algebraically proper context for generating the full solution list of a polynomial system is complex projective space rather than real or complex Euclidean space. This is because the structure of the solution set to $f(z) = 0$ is generic in projective space. (See, for example, Bezout's theorem below.) Homotopy continuation methods for generating the full solution list to $f(z) = 0$ have always implicitly acknowledged this by being formulated in complex Euclidean space and allowing paths to diverge to infinity. It is more numerically stable, however, to acknowledge projective space directly. We therefore follow the classical approach from algebraic geometry of homogenizing f and establishing our continuation process in projective space. In many cases it is advantageous to homogenize f so that it has an *multi-homogeneous* structure (m-homogeneous, for short). Then we view the solutions to $f(z) = 0$ as being in a Cartesian product of projective spaces. We will present our discussion based on this generality. Since 1-homogeneous systems are merely homogeneous systems, the m-homogeneous approach includes all polynomial systems and does not limit us to special cases.

Complex projective space, P^k, consists of the lines through the origin in C^{k+1}, denoted $[(z_0, \ldots, z_k)]$, where $(z_0, \ldots, z_k) \in C^{k+1} - \{0\}$; that is, $[(z_0, \ldots, z_k)]$ is the line through the origin that contains $(z_0, \ldots, z_k)$. It is natural to view P^k as a disjoint union of points $[(z_0, \ldots, z_k)]$ with $z_0 \neq 0$ (identified with Euclidean space via $[(z_0, \ldots, z_k)] \rightarrow (z_1/z_0, \ldots, z_k/z_0)$) and the "points at infinity," the $[(z_0, \ldots, z_n)]$ with $z_0 = 0$.

We partition the variables $\{z_1, \ldots, z_n\}$ into m nonempty collections. It will be notationally simpler here if we rename the variables with double subscripts. Thus, let $\{z_1, \ldots, z_n\} = \bigcup_{j=1}^{m} \{z_{1,j}, \ldots, z_{k_j,j}\}$, where $\sum_{j=1}^{m} k_j = n$. Now choose homogeneous variables $z_{0,j}$ for $j = 1$ to m and define $Z_j = \{z_{0,j}, z_{1,j}, \ldots, z_{k_j,j}\}$ for $j = 1$ to m. Then evoke the substitution $z_{i,j} \leftarrow z_{i,j}/z_{0,j}$ for $i = 1$ to k_j and $j = 1$ to m, generating a system $\hat{f} = 0$ of n equations in $n+m$ unknowns (after we clear the denominators of powers of the $z_{0,j}$). Now $\hat{f} = 0$ naturally has solutions in $P \equiv P^{k_1} \times P^{k_2} \times \ldots \times P^{k_m}$ (see Morgan and Sommese, 1987a). We say $\hat{f}$ is *multi-homogeneous* because the variables are partitioned into m collections, $Z_1, \ldots, Z_m$, so that $\hat{f}$ is homogeneous as a system in the variables of any one of the collections. We let $d_{j,l}$ denote the j^{th} degree of the l^{th} polynomial; that is, with all variables held fixed except those in Z_j, $\hat{f}_l$ has homogeneous degree $d_{j,l}$. Note that "1-homogeneous" is the same as "homogeneous", so theorems about m-homogeneous polynomial systems apply to all polynomial systems. Generally, we abuse the notation by not distinguishing f from its m-homogenization $\hat{f}$.

The *Bezout number*, d, of the m-homogeneous polynomial system f is defined to be the coefficient of $\prod_{j=1}^{m} \alpha_j^{k_j}$ in the product

$$D = \prod_{l=1}^{n} \sum_{j=1}^{m} d_{j,l} \alpha_j. \tag{3.1}$$

Then we have the classical:

THEOREM 3.1. (BEZOUT) *Let d denote the Bezout number of the system $f = 0$, and assume $f = 0$ does not have an infinite number of solutions in P. Then $f = 0$ has exactly d solutions in P, counting multiplicities.* †

† By "counting multiplicities" we mean that we count the multiplicities of the solutions rather than the solutions themselves.

The numerical significance of the Bezout number is that it is an upper bound on the number of homotopy continuation paths we will track in the space $P \times [0,1]$ (Theorem 3.2, below). The smaller d is, the better. Frequently, the m-homogenization of f for $m > 1$ has a (much) smaller Bezout number than the 1-homogenization. If $m = 1$, then $d = d_1 \cdots d_n$, the total degree of f, where $d_j = deg(f_j)$. This is the "traditional" number of paths to track in polynomial continuation.

We acknowledge $P = P^n$ and $P = P^{k_1} \times \ldots \times P^{k_m}$ computationally via *the projective transformation*, first proposed in Morgan (1986) and extended to the m-homogeneous case in Morgan and Sommese (1987b). We sketch what is involved here. Let us first consider the 1-homogeneous case. Thus h is a homogeneous system of n equations in the $n+1$ variables $z_0, \ldots, z_n$, and $h(z) = 0$ has solutions in P^n. Let constants $b_0, \ldots, b_n$ be given with $b_{k_0} \neq 0$. Define a (nonhomogeneous) system $h'(z)$ of $n+1$ equations in $n+1$ unknowns by

$$h'_l(z_0, \ldots, z_n) = h_l(z_0, \ldots, z_n)$$

for $l = 1$ to n, and

$$h'_0(z_0, \ldots, z_n) = L(z_0, \ldots, z_n) - 1$$

where

$$L(z) \equiv \sum_{j=0}^{n} b_j z_j .$$

Then the solutions to $h(z) = 0$ in $U_L = \{[z] \in P^n \mid L(z) \neq 0\}$ are in one-to-one correspondence with solutions to $h'(z) = 0$ in C^{n+1}. The multiplicity of solutions (and all other local properties) are preserved under this correspondence. Further, we may view $h'_1(z) = 0, \ldots, h'_n(z) = 0$ as a system in the n variables $z_0, \ldots, z_{k_0-1}, z_{k_0+1}, \ldots, z_n$ with

$$z_{k_0} = \frac{1 - (b_0 z_0 + \ldots + b_{k_0-1} z_{k_0-1} + b_{k_0+1} z_{k_0+1} + \ldots + b_k z_k)}{b_{k_0}} .$$

It is this system of n equations in n unknowns that we call "the projective transformation of h". We want to use this system for computations. Its solutions in C^n are in one-to-one correspondence with the solutions to $h'(z) = 0$ in C^{n+1}. In Theorem 3 of Morgan and Sommese (1987b) it is proven that the continuation paths are contained in U_L with probability one if the parameters are chosen at random. This theorem establishes the validity of the projective transformation.

In creating a computer code to implement the projective transformation, our usual procedure is to let $k_0 = 0$ and write a subroutine for h as a system of n equations in the $n+1$ variables $z_0, \ldots, z_n$ but include the formula

$$z_0 = \sum_{j=1}^{n} \beta_j z_j \; + \; \beta_0 ,$$

which makes z_0 an implicitly defined function of the other variables. The partial derivatives of the projective transformation with respect to $z_1, \ldots, z_n$ are then generated from those of h with respect to $z_0, \ldots, z_n$ using the chain rule. To make use of Theorem 3 from Morgan and Sommese (1987b), the $\beta_0, \ldots, \beta_n$ are chosen at random.

If h is m-homogeneous, we may evoke the projective transformation on each component of P. Thus, with m-homogeneous h in the variables $z_{i,j}$ for $i = 0$ to n and $j = 1$ to m, we define

$$z_{0,j} = \sum_{i=1}^{k_j} \beta_{i,j} z_{i,j} + \beta_{0,j}$$

for $j = 1$ to m. [Theorem 3 of Morgan and Sommese (1987b) is proven in this generality.] The finite solutions of $f(z) = 0$ are recovered via $z_{i,j} \leftarrow z_{i,j}/z_{0,j}$ for $i = 1$ to k_j and $j = 1$ to m. If any $z_{0,j} = 0$, then the solution is at infinity.

Example 1. Consider the system:

$$\begin{aligned} z_1^2 + z_2^2 - 25 &= 0 \\ z_1^2 + z_2^2 - 16z_2 + 39 &= 0 \end{aligned} \tag{3.2}$$

This is the intersection of two circles of radius 5 with centers at the origin and at $(0, 8)$. We homogenize (that is, 1-homogenize) via the substitutions $z_1 \leftarrow z_1/z_0$ and $z_2 \leftarrow z_2/z_0$, yielding

$$\begin{aligned} z_1^2 + z_2^2 - 25z_0^2 &= 0 \\ z_1^2 + z_2^2 - 16z_2z_0 + 39z_0^2 &= 0 \end{aligned} \tag{3.3}$$

We obtain the solutions at infinity by solving the system with $z_0 = 0$ and $z_1 = 1$. This reduces (3.3) to $1 + z_2^2 = 0$. Thus, $z_2 = \pm i$, with $i = \sqrt{-1}$. (In finding solutions in projective space, we always set one of the variables equal to 1.) The projective transformation of (3.2) is (3.3) with $z_0 \equiv \beta_1 z_1 + \beta_2 z_2 + \beta_0$, where the β_j are chosen at random.

Example 2. Consider the following system:

$$\begin{aligned} z_1z_2z_3z_4 + 1 &= 0 \\ z_1z_3 + z_2z_4 + z_1z_4 &= 0 \\ 4z_1z_3z_4 - 2z_2z_3z_4 + 1 &= 0 \\ z_1 + z_2 &= 0 \end{aligned} \tag{3.4}$$

By grouping the variables of (3.4) into different sets, we create different m-homogeneous structures and Bezout numbers. Normally, we would want to solve such a system with the m-homogeneous structure that gives the smallest Bezout number. For each grouping of variables, we will form the combinatorial product, D, defined in (3.1) above, and then pick out the distinguished coefficient that gives the Bezout number, d. Thus:

Example 2.1. Group variables as: $\{z_1, z_2\} \cup \{z_3, z_4\}$. Then, $D = (2\alpha_1 + 2\alpha_2)(\alpha_1 + \alpha_2)(\alpha_1 + 2\alpha_2)(\alpha_1 + \alpha_2)$, and $d = Coef[D, \alpha_1^2\alpha_2^2]$ (i.e. the coefficient of the $\alpha_1^2\alpha_2^2$ term of D). Thus, $d = 10$.

Example 2.2. Group variables as: $\{z_1, z_2\} \cup \{z_3\} \cup \{z_4\}$. Then, $D = (2\alpha_1 + \alpha_2 + \alpha_3)(\alpha_1 + \alpha_2 + \alpha_3)^2(\alpha_1 + 0\alpha_2 + 0\alpha_3)$, and $d = Coef[D, \alpha_1^2\alpha_2\alpha_3] = 8$.

Example 2.3. Group variables as: $\{z_1\} \cup \{z_2\} \cup \{z_3, z_4\}$. Then, $D = (\alpha_1 + \alpha_2 + 2\alpha_3)^2(\alpha_1 + \alpha_2 + \alpha_3)(\alpha_1 + \alpha_2 + 0\alpha_3)$, and $d = Coef[D, \alpha_1\alpha_2\alpha_3^2] = 16$.

We see that 2.2 gives the smallest Bezout number. Thus, while the 1-homogeneous (traditional) polynomial continuation yields a 24 path homotopy (i.e. the number of the total degree), we can (easily) find a 3-homogeneous 8 path homotopy. Such a savings in computer work (i.e. by a factor of 1/3) can be significant in some applications. The

projective transformation for 2.2 is easy to construct: first, we 3-homogenize (3.4) via $z_1 \leftarrow z_1/z_5$, $z_2 \leftarrow z_2/z_5$, $z_3 \leftarrow z_3/z_6$, and $z_4 \leftarrow z_4/z_7$, where we take z_5, z_6, and z_7 to be the three homogeneous variables. This yields

$$\begin{aligned} z_1z_2z_3z_4 + z_5^2z_6z_7 &= 0 \\ z_1z_3 + z_2z_4 + z_1z_4 &= 0 \\ 4z_1z_3z_4 - 2z_2z_3z_4 + z_5z_6z_7 &= 0 \\ z_1 + z_2 &= 0 \end{aligned} \tag{3.5}$$

with $z_5 \equiv \beta_1 z_1 + \beta_2 z_2 + \beta_5$, $z_6 \equiv \beta_3 z_3 + \beta_6$, and $z_7 \equiv \beta_4 z_4 + \beta_7$, where the β_j are chosen at random.

3.2. A BASIC THEOREM

Consider the homotopy (2.1), where now we take g to be an m-homogeneous system of n polynomials in $n+m$ variables having exactly the Bezout number of nonsingular solutions, where g is chosen so that its m-homogeneous structure matches that of f; that is, $Z_1, \ldots, Z_m$ are specified and the $d_{j,l}$ for g are exactly the same as those for f. Many such g will exist. In Wampler *et al.* (1990a, section 4.1, p. 63), a scheme for generating such a g for any f is given.

Then the following holds.

THEOREM 3.2. *For any positive r and for all but a finite number of angles, θ, if $\gamma = re^{i\theta}$, then $h^{-1}(0)$ consists of smooth paths over $[0,1)$ and every geometrically isolated solution of $f(z) = 0$ has a path converging to it. In fact, if m_0 is the multiplicity of a geometrically isolated solution, z^0, then z^0 has exactly m_0 paths converging to it. Further, the paths are strictly increasing in t, and $dt/ds > 0$, where s denotes arc length.*

This theorem is stated and proven in Morgan and Sommese (1987a).

4. Parameter Continuation and Side Conditions

4.1. THE CONTEXT OF COEFFICIENT-PARAMETER POLYNOMIAL CONTINUATION

Let us expand the context of our considerations somewhat, as follows. Let

$$f_j(C,z) = \sum_{k=1}^{r_j} C_{j,k} \prod_{l=1}^{n} z_l^{\delta_{j,l,k}} \tag{4.1}$$

denote a polynomial system in the n variables $z_1, \ldots, z_n$ with the j-th equation having r_j terms, where the k-th term has (complex number) coefficient $C_{j,k}$ and in which z_l is raised to the $\delta_{j,l,k}$-th power. Here $C = (C_{j,k})$. We will generally consider systems of $n+N$ equations where $N \geq 0$. That is, we let $j = 1, \ldots, n+N$. Sometimes we call the equations indexed by $j = n+1, \ldots, N$ *side conditions.*

Traditionally, polynomial continuation has not considered side conditions, focusing on systems of n equations in n unknowns. Further, the traditional goal of polynomial continuation has been to compute all geometrically isolated solutions to $f(z) = 0$. However,

we wish to refine and make more flexible the goal of polynomial continuation, allow more equations than unknowns and compute a distinguished subset of the solutions to $f(z) = 0$.

In polynomial continuation, homotopies are typically constructed by letting the $C_{j,k}$ in (4.1) vary via a homotopy parameter t. Thus $C_{j,k}(t) : [0,1] \rightarrow C^r$ where $r = \sum_{j=1}^{n} r_j$, so that (4.1) with coefficients $C_{j,k}(0)$ defines the start system and (4.1) with coefficients $C_{j,k}(1)$ defines the target system.

However, when polynomial systems arise in engineering, the *coefficients* are usually not the *physical parameters*. To mirror this reality in the mathematics, we assume the coefficients are functions on a space Q, and the continuation will be generated in "parameter space," Q, rather than "coefficient space," C^r. Thus

$$C_{j,k} = c_{j,k}[q], \tag{4.2}$$

where we call (4.2) the *coefficient parameter formulas*. The homotopy will be defined by the composition

$$C_{j,k}(t) = c_{j,k}[q(t)] \tag{4.3}$$

where $q(t) : [0,1] \rightarrow Q$ and $c_{j,k} : Q \rightarrow C^1$. Generally, the dimension of the parameter space will be much less than that of the coefficient space. The special structure of the solutions of (4.1) induced by (4.2) is naturally acknowledged with a homotopy like (4.3). The result, we will see, is that fewer paths have to be tracked to solve the system, significantly reducing in some cases the total numerical cost. We further allow $c_{j,k}[q]$ to be a general analytic function of q, rather than merely a polynomial in q. The need for this generality arises, for example, in mechanical engineering, where it is common for trigonometric functions to arise.

Consider the following simple example:

$$\begin{aligned} z_1 z_2 + q_2 z_1 + q_1 z_2 + q_1 q_2 &= 0 \\ z_1^2 + (q_1 + q_3) z_1 + q_1 q_3 &= 0 \end{aligned} \tag{4.4}$$

Here z_1 and z_2 are the variables and q_1, q_2, q_3 are the coefficient parameters. Thus (4.4) can be written

$$\begin{aligned} c_{1,1} z_1 z_2 + c_{1,2} z_1 + c_{1,3} z_2 + c_{1,4} &= 0 \\ c_{2,1} z_1^2 + c_{2,2} z_1 + c_{2,3} &= 0 \end{aligned}$$

where

$$c_{1,1} = 1 \quad c_{1,2} = q_2 \quad c_{1,3} = q_1 \quad c_{1,4} = q_1 q_2 \quad c_{2,1} = 1 \quad c_{2,2} = q_1 + q_3 \quad c_{2,3} = q_1 q_3.$$

Now, construct a parameter homotopy so that: when $t = 0$, $q_1 = 0$, $q_2 = 1$, and $q_3 = 1$, and when $t = 1$, $q_1 = 1$, $q_2 = 2$, and $q_3 = 3$; for example, $q(t) = [(1-t), (1+t), (1+2t)]$. Then, the coefficients will continue via the formulas:

$$c[q(t)] = [1, 1+t, 1-t, (1-t)(1+t), 1, (1-t)+(1+2t), (1-t)(1-t)(1+2t)].$$

Thus the coefficient parameter continuation is not an affine transformation with respect to the homotopy parameter t.

The polynomial equations f defined by (4.1) naturally can be viewed as a map

$$f : Q \times C^n \to C^{n+N}$$

with the solution set

$$X_+ \equiv f^{-1}(0) \subseteq Q \times C^n.$$

However, X_+ accounts only for the *finite* solutions of $f = 0$, whereas it is convenient to expand the context of the system in such a way that additional *solutions at infinity* are defined. It is further convenient to be able to do this in a variety of ways. What we have in mind is to find algebraic compactifications P of C^n and extensions ϕ of f with $\phi : Q \times P \to E$, for appropriately defined E. It will follow that $X_+ \subseteq \phi^{-1}(0) \equiv X$, where X therefore depends on the choice of compactification and the extension of f, but X_+ does not. See Morgan and Sommese (1989, 1990) for details.

The two most common ways of defining P and ϕ are:

- $P = P^n$, n-dimensional complex projective space; ϕ the "homogenization of f"
- $P = P^{k_1} \times \ldots \times P^{k_m}$; ϕ the "m-homogenization of f".

4.2. A COEFFICIENT-PARAMETER THEOREM

Now we state Theorem 4.1, a corollary to the very general theorem cited in Morgan and Sommese (1989). We focus on this special case because:

- It highlights an important practical result.
- It illustrates the more general theory without requiring a background in algebraic geometry.

Recall that we are considering the system $f(c[q], z) = 0$ of $n + N$ equations in n unknowns with $q \in Q$ and $z \in P$. We make the following assumptions:

- $Q = C^s$, for some s
- $c : Q \to C^r$ is polynomial
- $P = P^{k_1} \times \ldots \times P^{k_m}$

We emphasize that these assumptions are not necessary, but are merely taken to allow a more elementary presentation. We need the following definition.

DEFINITION 4.1. *The solution z to $f(c[q], z) = 0$ is generically nonsingular if there is a dense open $Q_0 \subseteq Q$ such that for every neighborhood B about z in P there is a neighborhood B' about q in Q so that if $q' \in B' \cap Q_0$, then $f(c[q'], z) = 0$ has a nonsingular solution z' with $z' \in B$.*

Thus a solution z is generically nonsingular if almost all nearby systems have a nonsingular solution near z. Now we have:

THEOREM 4.1. *Given $q^1 \in Q$, there is a dense open full measure $Q_0 \subseteq Q$ such that if $q^0 \in Q_0$ and S is the (finite) set of nonsingular solutions to $f(c[q^0], z) = 0$ in P, then the homotopy*

$$h(z, t) = f(c[(1 - t)q^0 + tq^1], z)$$

with start points in S and $(z,t) \in P \times [0,1]$ will have well defined homotopy paths in P that are smooth and strictly increasing in t as a function of arc length, and the endpoints will include all the geometrically isolated generically nonsingular solutions of $f(c[q^1], z) = 0$.

The proof of this theorem is given in Morgan and Sommese (1989).

We sketch the idea of the proof here. Define $X = \{(q,z) \in Q \times P \mid f(c[q],z) = 0\}$. Consider the natural map $\pi : X \to Q$ defined by $\pi(q,z) = q$. We then can prove that there is an open dense $Q_0 \subseteq Q$ and a smooth manifold $X_0 \subseteq X$ such that $\pi_0 \equiv \pi \mid X_0 : X_0 \to Q_0$ is a finite-to-one *covering space*, and X_0 can be identified with the geometrically isolated solutions of $f = 0$ when $q \in Q_0$. Further, $X_0 = X_{ns} \cup X_s$, where X_{ns} and X_s denote the nonsingular and singular solutions, respectively, and $\pi \mid X_{ns} : X_{ns} \to Q_0$ and $\pi \mid X_s : X_s \to Q_0$ are both covering spaces. In addition we show that if $K \subseteq Q$ is a polynomial (complex) curve and $K_0 \equiv K \cap Q_0$, then $K - K_0$ is finite. This means that all but a finite number of points in K are in Q_0. For example, K might be taken to be $\{(1-t)q^0 + tq^1 \mid t \in C^1\}$. To define a standard homotopy, we choose K a complex curve, and then choose $\alpha : [0,1] \to K$ such that $\alpha : [0,1) \to K_0$, which is easy to do because all but a finite number of points in K are in K_0. (Note that the image of α is one-real-dimensional in the two-real-dimensional K.) Now consider the diagram:

$$\begin{array}{ccccccc}
\pi^{-1}(\alpha[0,1]) & \subseteq & \pi^{-1}(K) & \subseteq & X & \subseteq & Q \times P \\
\downarrow & & \downarrow & & \pi\downarrow & & \\
\alpha[0,1] & \subseteq & K & \subseteq & Q & &
\end{array}$$

The image of α, which is a path in the parameter space Q, is lifted via π to $Q \times P$. This lifted path is a collection of paths in the "parameter–solution" space $Q \times P$. These paths are the continuation paths that we will track numerically.

What this theorem says is that if we choose almost any system (given by $q^0 \in Q_0$) and solve it somehow (say, by a traditional m-homogeneous continuation) and pick out the nonsingular solutions (the set S), then we will find the generically nonsingular solutions that are geometrically isolated of any other system (given by q^1), via a path tracking approach with fewer paths (because we can restrict the choice of start points to S). There are three key observations about the "real" world that make this theorem important:

- In engineering and scientific practice, small polynomial systems tend to arise in families (as indexed by Q). It makes sense to solve one (given by q^0) at some expense, if the rest can then be solved cheaply.
- The number of points in the set S is often much less than the full Bezout number for a general system.
- The physically meaningful solutions tend to be included among the generically nonsingular solutions.

In our experience, polynomial models with physically meaningful solutions that are not generically nonsingular are degenerate, often because a special singular case has arisen in a general model. We note also that when physically meaningful solutions are not geometrically isolated, a similar type of model degeneracy has usually arisen. If such degeneracies are unavoidable, then using "random real" parameters as noted in section

4.3 can help, because then polynomial continuation will tend to find some real solutions, which is usually what is being sought. We cannot "prove" these observations about reality. We cite the examples given in Morgan and Sommese (1989), Morgan *et al.* (1990), Morgan and Wampler (1990), Wampler and Morgan (1991) and Wampler *et al.* (1990a, 1990b) for supporting evidence.

4.3. Implementation considerations

Let us continue with the special assumption that $Q = C^s$, but remind the reader that the theory developed in Morgan and Sommese (1989) allows a more general context.

We take the implementation problem to be: Given the coefficient parameter polynomial system $f(c[q], z)$, we want to solve a sequence of systems $f(c[q^k], z) = 0$ for $k = 1, 2, 3, \ldots$. We recognize two steps:

1 Choose $q^0 \in Q$, solve $f(c[q^0], z) = 0$, pick out the set of nonsingular solutions, S.
2 Track the solution paths of $h(z,t) = f(c[(1-t)q^0 + tq^k], z) = 0$, beginning at the points in S.

There are two ways to choose q^0:

- Choose q^0 "at random" from Q.
- Choose q^{00} "at random" from $Q \cap R^s$, choose γ "at random" from C^1, and take $q^0 = \gamma q^{00}$. (This "random-real" alternative has advantages for certain physical problems.)

The coefficient parameter formulas for a particular problem may be complicated and involve many transcendental functions. This makes the function and Jacobian matrix evaluations expensive, and typically polynomial continuation involves thousands if not millions of these evaluations. Sometimes variants of the coefficient parameter polynomial homotopies have a similar small number of paths and are much more efficient in implementation. We sketch some of these variants here. We have found particularly useful the *secant homotopy*. What these variants amount to is that from the given coefficient parameter formulas we can define corollary coefficient parameter formulas that have a simpler form but admit a larger class of systems. Generally, this means that there will be more generically nonsingular solutions; therefore, "weakening" the coefficient parameter formulas in this way would generally seem to be a bad idea. However, often the trade-off between efficiencies makes a variant preferable.

Let us consider a *hierarchy of homotopies*:

- **Traditional:** $h(z,t) = (1-t)g(z) + tf(c[q^k], z)$ where $g(z)$ has d solutions.
- **Coefficient:** $h(z,t) = f((1-t)C_0 + tc[q^k], z)$ where $C_0 \in C^r$ is a random choice of coefficients.
- **Secant:** $h(z,t) = f((1-t)c[q^0] + tc[q^k], z)$ where q^0 is a random choice of parameters. Note that here the start system is defined by coefficient parameter formulas but the intermediate systems (for $0 < t < 1$) are not given by coefficient parameters.
- **Parameter:** $h(z,t) = f(c[(1-t)q^0 + tq^k], z)$ where q^0 is a random choice of parameters. Note that taking the convex combination of q^0 and q^k is natural when $Q = C^s$, but almost any path in Q from q^0 to q^k will do.

We can easily see that the traditional, coefficient, and secant homotopies are "coefficient parameter" homotopies for appropriately defined coefficient parameter formulas. These variant coefficient parameter formulas are derived from the "original" coefficient parameter formulas, which we conceptualize as having been given by physically meaningful relations.

4.4. Side conditions

Here we explore the surprising and useful fact that *algebraic relations persist along homotopy paths (generically)*. This becomes apparent from the fact that the theory is developed for $n + N$ equations in n unknowns, where $N \geq 0$. The theory of polynomial continuation until now has taken $N = 0$, but we can recognize considerable implementation advantages from the more general context.

We have called the equations indexed by j for $j = n + 1$ to $n + N$ "side conditions", although in the structure of the theory they are not distinguished from the other equations. Now, however, it will be better for the exposition if we call the side conditions and the rest of the system by different names. Thus let $f'(c[q], z)$ denote a system of n polynomials in n unknowns, with coefficient parameters $q \in Q$, and let $s(c[q], z)$ denote a system of N polynomials in the same n unknowns with the same coefficient parameter set Q. The result we want to point out is as follows. Suppose we choose q^0 at random from Q and suppose z^0 is a geometrically isolated solution to $f'(c[q^0], z) = 0$ that also satisfies $s(c[q^0], z) = 0$. Let $(q(t), z(t))$ denote the homotopy path of $h(z, t) = f'(c[(1-t)q^0+q^1], z)$ with start point (q^0, z^0). Then $s(q(t), z(t)) = 0$ for $0 \leq t \leq 1$.

In other words, if the side conditions hold for a solution to the start system $f'(c[q^0], z) = 0$, then the side conditions will hold along the associated path and, in particular, at the corresponding solution to the target system $f'(c[q^1], z) = 0$. Therefore, if the side conditions represent conditions that we *do not want* to hold, we can omit the start points that *satisfy* the side conditions, and only track paths from the resulting smaller set of start points. Thus, we can track fewer paths than we would otherwise need to track. (Without the side condition result, we might track all the paths whose start points are given by solutions to $f'(c[q^0], z) = 0$ and then omit the resulting endpoints that satisfy $s(c[q^1], z) = 0$.) For example, solutions at infinity of $f'(c[q], z) = 0$ are distinguished by obeying the additional relation $s(z) = 0$, where $s(z)$ denotes the product of the homogeneous coordinates of z. If we are not interested in solutions that are generically at infinity, we can omit any start points that are at infinity, since we know that they will end up at infinity.

The converse idea is generally true, but there are some subtleties. That is, if we *want* the side conditions to hold, it is a reasonable approach to omit the paths whose start points are solutions to $f'(c[q^0], z) = 0$ that do not also obey $s(c[q^0], z) = 0$. This will certainly yield only those solutions that obey the side conditions. However, certain types of solutions may be missed. In many cases, this is not a difficulty. However, for completeness, let us consider exactly what is involved.

- It might happen, for a choice of $q^1 \notin Q_0$, that $s(c[q^1], x) = 0$ might be satisfied in the limit of a continuation path as $t \to 1$ but not along the path. Generally, we can argue that these *unstable solutions* are nonphysical.

- $f'(c[q^0], z) = 0$ could have some positive-dimensional solutions sets that yield isolated solutions when the side conditions are included. We will miss such solutions if we find merely the isolated solutions of $f'(c[q^0], z) = 0$ and then eliminate those not satisfying $s(c[q^0], z) = 0$. But usually we are not able to find the positive-dimensional solutions sets of $f'(c[q^0], z) = 0$. So, to be absolutely rigorous, we can find instead the isolated solutions of

$$f'_j(c[q^0], z) + \sum_{k=1}^{N} c_{j,k} s_k(c[q^0], z) = 0 \qquad \text{for } j = 1 \text{ to } n$$

that also satisfy

$$s_j(c[q^0], z) = 0 \qquad \text{for } j = 1 \text{ to } N,$$

where the $c_{j,k}$ are random numbers.

See Morgan and Sommese (1989, section 3.2) for a further discussion of side conditions.

5. Symbolic Reduction

Let $f(z) = 0$ be a system we want to solve. By *reduction* we mean the process of generating a new system $\bar{f}(z)$ which is "smaller" than $f(z)$ but whose solutions yield easily those solutions of $f(z) = 0$ that are of interest. How should $\bar{f}$ look compared to f? It might have

1 fewer variables;
2 fewer paths for the associated polynomial continuation (smaller total degree or Bezout number);
3 a simpler solution set; and
4 shorter paths and/or better conditioned Jacobian matrices along paths.

By "a simpler solution set" we have in mind that the reduction will eliminate (extraneous) solutions cut (out or in) by side conditions. For example, it might eliminate a known generic positive-dimensional solution set or eliminate some generic solutions at infinity. Item 4 is a bit special and comes more under the category of "scaling" than reduction. (The SCLGEN scaling algorithm cited in Note 6 of section 2.1 and the projective transformation are examples.)

However, reduction can have a devastating effect on the numerical characteristics of a system which is to be solved in fixed-precision floating point arithmetic. Consider the following example from geometric modeling.

$$\begin{aligned} 1.6 \times 10^{-3} x_1^2 + 1.6 \times 10^{-3} x_2^2 - 1 &= 0 \\ 5.3 \times 10^{-4} x_1^2 + 5.3 \times 10^{-4} x_2^2 + 5.3 \times 10^{-4} x_3^2 + 2.7 \times 10^{-2} x_1 - 1 &= 0 \\ -1.4 \times 10^{-4} x_1 + 10^{-4} x_2 + x_3 - 3.4 \times 10^{-3} &= 0 \end{aligned}$$

This is the intersection of a cylinder, a sphere, and a plane. It has two real solutions of norm about 25, and a complex conjugate pair of norm about 10^9. In the original geometric context, the two real solutions had physical meaning and the complex pair did not. By standard elimination [here effected by a Gröbner basis algorithm (Buchberger,

1985)] the system is reduced to a 4^{th} degree polynomial equation:

$$6.38281970398352x_1^4 - 7.12554854545301 \times 10^9 x_1^3 + 1.89062308408416 \times 10^{19} x_1^2 \\ +9.36558635415069 \times 10^{20} x_1 - 1.15985845720186 \times 10^{22} = 0,$$

where the coefficients have been rounded to 15 digits. Note that in the original system there is a range of 4 orders of magnitude in the coefficients, but in the single polynomial there are 22. When this polynomial was submitted to a code which implemented the quartic formula, an overflow was generated and the system crashed. The point is not that these numerical difficulties are insurmountable, but simply that a rather mild problem was made much more numerically unstable by reduction.

This phenomenon is quite common. Here is a more extreme example. The following polynomial system comes from the study of chemical equilibria. [This is the Model A combustion chemistry example from Morgan (1987, Chapter 9), for a temperature of 3000°.]

$$\begin{aligned} x_2 + 2x_6 + x_9 + 2x_{10} &= 10^{-5} \\ x_3 + x_8 &= 3 \times 10^{-5} \\ x_1 + x_3 + 2x_5 + 2x_8 + x_9 + x_{10} &= 5 \times 10^{-5} \\ x_4 + 2x_7 &= 10^{-5} \\ 0.5140437 \times 10^{-7} x_5 &= x_1^2 \\ 0.1006932 \times 10^{-6} x_6 &= x_2^2 \\ 0.7816278 \times 10^{-15} x_7 &= x_4^2 \\ 0.1496236 \times 10^{-6} x_8 &= x_1 x_3 \\ 0.6194411 \times 10^{-7} x_9 &= x_1 x_2 \\ 0.2089296 \times 10^{-14} x_{10} &= x_1 x_2^2 \end{aligned}$$

The following reduction (as given in the above reference) to a system of two equations is not much more extreme than the original problem, in terms of range of coefficients or powers of variables:

$$\begin{aligned} &0.9572608 \times 10^{15} x_1 x_2^2 + 0.1614359 \times 10^8 x_1 x_2 + 0.1986230 \times 10^8 x_2^2 + x_2 - .3 \times 10^{-4} = 0 \\ &0.5200670 \times 10^{15} x_1^3 + 0.1078946 \times 10^{15} x_1^2 x_2 + 0.9118119 \times 10^8 x_1^2 \\ &\quad +0.8245123 \times 10^{15} x_1 x_2^2 + 0.2560374 \times 10^8 x_1 x_2 - 0.1985030 \times 10^3 x_1 - 0.8 \times 10^{-4} = 0 \end{aligned}$$

This system has total degree 9 and Bezout number 8. Its solutions (including the physical solution) are easy to compute to full double precision via polynomial continuation implemented in double precision, in spite of the range in the coefficients (19 orders of magnitude).

However, the reduction of this system to a single polynomial (via a Gröbner basis algorithm) is

$$2.47638543974253 + 173863854.837537x_1 + 3.64001686812513 \times 10^{15} x_1^2 \\ +2.03096667862772 \times 10^{22} x_1^3 + 2.2580193833473 \times 10^{28} x_1^4 \\ -9.05543678075695 \times 10^{34} x_1^5 - 1.74762562120847 \times 10^{41} x_1^6 \\ +9.44753095267844 \times 10^{46} x_1^7 + 2.43085098985668 \times 10^{53} x_1^8 = 0$$

Here the coefficients are rounded to 15 digits. There are 53 orders of magnitude in the coefficients, and a standard bisection-Newton's method solver did not obtain a single solution to a single digit of accuracy. Again the point is not that this polynomial is

impossible to solve, but rather than the numerical stability of the two variable system is made much worse by the final reduction to triangular form.

These examples illustrate that often symbolic reduction has the effect of degrading the numerical stability of a system, especially when this reduction is carried beyond a certain point. Let us consider the possible bad effects of reduction, which we would like to minimize:

- Numerical instability, signaled by
  - The variables of $\bar{f}$ typically have higher powers than those of f.
  - The coefficients of $\bar{f}$ typically are given by complicated formulas in terms of the coefficients of f. The numerical range of the coefficients of $\bar{f}$ is typically greater than those of f.
- Cost, including
  - The cost of the reduction: cpu time and memory requirements.
  - The cost of evaluating the system: extra terms and higher powers.

We need a concept of a reduction which improves some aspect of a system without disturbing the system more than is needed. Let us refer to a *small reduction* as a reduction that accomplishes a limited reduction goal without changing the system any more than it has to. Conceptually, a small reduction is less likely to devastate the numerical stability of the system than reduction to triangular form.

We would like software tools which allow us to make specified small reductions to improve the characteristics of a system which we will then solve by polynomial continuation. What would a good reduction algorithm package look like?

- It would generate a sequence of intermediate systems, each a small reduction of the previous, in contrast to the typical Gröbner basis reduction to triangular form. (The goal of each of these small reductions would be one or more of the above items 1–4.)
- It would offer the option of "one-step" and "two-step" reductions, with human guidance. Thus a few small steps would be taken, with the user able to offer suggestions on the most fruitful directions. Alternatively, perhaps an AI or expert system implementation would be possible.
- It might offer the option of finding the best conditioned sequence of reductions. (Note that pivoting provides a means of doing this in Gaussian elimination.) For example, the system

$$\begin{aligned} x^2 + y^2 - 1 &= 0 \\ x + \epsilon y - 0.5 &= 0 \end{aligned}$$

  reduces stably to a polynomial in y but not in x as $\epsilon \to 0$.
- It does not have to be optimal. A 1/12 reduction in a year-long computation makes it take only a month.

We complete this section with a set of questions, projects, and problems. They are not equally difficult. Even partial success with some of them would amount to a breakthrough (e.g. question 3). Most should be investigated with specific problems in mind; for example, the problems discussed in Meintjes and Morgan (1985, 1987, 1989, 1990), Morgan and Wampler (1990), Richter and De Carlo (1984), Safonov (1984), Wampler and Morgan (1991), Wampler *et al.* (1990b) and Watson and Morgan (1991).

1 Create an efficient algorithm to find all the partitions of variables that yield the minimal Bezout number for a given polynomial system.
2 Given a polynomial system with Bezout number d, find a small reduction with total degree d. (Classical reduction to a single polynomial accomplishes this, but it is not small.)
3 Find a small reduction strategy for the Lagrange multiplier formulation of the constrained optimization problem with polynomial constraints and objective function. [This would yield global minima. See Watson and Morgan (1991).]
4 Given a known positive-dimensional solution set T for $f(z) = 0$, find a small reduction $\bar{f}(z)$ so that T is not contained in the solution set of $\bar{f}(z) = 0$. (T might still intersect it.) Consider the case that T is given by linear equations. The systems discussed in Morgan and Wampler (1990) and Wampler *et al.* (1990b) are examples.
5 Given a polynomial side condition $s(z) = 0$ so that the solutions of interest of $f(z) = 0$ always obey $s(z) \neq 0$, find a small reduction, $\bar{f}(z)$, so that $\bar{f}(z) = 0$ has no (or fewer) solutions with $s(z) = 0$. For example, if $s(z)$ equals the homogeneous coordinate, then classical reduction to a single polynomial accomplishes this (by eliminating solutions at infinity), but this reduction is not small.
6 Define and implement a "pivoting" strategy that guides elimination algorithms to minimize numerical instability.

6. Summary

Polynomial continuation is a well developed numerical method for computing the isolated solution to polynomial systems. This method would benefit from flexible tools to symbolically reduce systems, to make them smaller with minimum degradation of their numerical characteristics. Section 5 sketches some features these tools might possess and proposes some related research problems and directions.

Acknowledgments

I acknowledge the continuing support, encouragement, and inspiration of my colleagues Professor Andrew Sommese and Dr. Charles Wampler.

Dedication

To Alex Dejanikus with love, respect, and appreciation.

References

E.L. Allgower and K. Georg (1980), "Simplicial and continuation methods for approximating fixed points and solutions to systems of equations", *SIAM Review*, **22**, 28–85.
B. Buchberger (1985), "Gröbner bases: an algorithmic method in polynomial ideal theory", *Recent Trends in Multidimensional Systems Theory*, N.K. Bose, ed., D. Reidel, NY, 184–232.
P. Brunovský and P. Meravý (1984), "Solving systems of polynomial equations by bounded and real homotopy", *Numer. Math.*, **43**, 397–418.
S.N. Chow, J. Mallet-Paret and J.A. Yorke (1978), "Finding zeros of maps: homotopy methods that are constructive with probability one", *Math. Comput.*, **32**, 887–899.

S.N. Chow, J. Mallet-Paret and J.A. Yorke (1979), "A homotopy method for locating all zeros of a system of polynomials", *Functional Differential Equations and Approximation of Fixed Points*, Lecture Notes in Math. 730, H.O. Peitgen and H.O. Walther, eds., Springer-Verlag, NY, 228–237.

F.J. Drexler (1977), "Eine methode zur berechung sämtlicher lösunger von polynomgleichungessystemen", *Numer. Math.*, **29**, 45–58.

F.J. Drexler (1978), "A homotopy method for the calculation of zeros of zero-dimensional polynomial ideals", *Continuation Methods*, H.G. Wacker, ed., Academic Press, NY, 69–93.

C.B. Garcia and W.I. Zangwill (1977), "Global continuation methods for finding all solutions to polynomial systems of equations in N variables", *Center for Math. Studies in Business and Economics*, Report 7755, University of Chicago, Chicago, IL.

C.B. Garcia and W.I. Zangwill (1979a), "Finding all solutions to polynomial systems and other systems of equations", *Math. Prog.*, **16**, 159–176.

C.B. Garcia and W.I. Zangwill (1979b), "An approach to homotopy and degree theory", *Math. Oper. Res.*, **4**, 390–405.

C.B. Garcia and W.I. Zangwill (1979c), "Determining all solutions to certain systems of nonlinear equations", *Math. Oper. Res.*, **4**, 1–14.

C.B. Garcia and W.I. Zangwill (1981), *Pathways to Solutions, Fixed Points, and Equilibria*, Prentice-Hall, Englewood Cliffs, NJ.

T.Y. Li (1983), "On Chow, Mallet-Paret and Yorke homotopy for solving system of polynomials", *Bull. Inst. Math. Academica Sinica*, **11**, 433–437.

T.Y. Li, T. Sauer and J.A. Yorke (1987), "Numerical solution of a class of deficient polynomial systems", *SIAM J. Numer. Anal.*, **24**, 435–451.

T.Y. Li, T. Sauer and J.A. Yorke (1988), "Numerically determining solutions of systems of polynomial equations", *Bull. Amer. Math. Soc.*, **18**, 173–177.

T.Y. Li, T. Sauer and J.A. Yorke (1989), "The cheater's homotopy: an efficient procedure for solving systems of polynomial equations", *SIAM J. Numer. Anal.*, **26**, 1241–1251.

K. Meintjes and A.P. Morgan (1985), "Performance of algorithms for calculating the equilibrium composition of a mixture of gases", *J. Comput. Phys.*, **60**, 219–234.

K. Meintjes and A.P. Morgan (1987), "A methodology for solving chemical equilibrium systems", *Appl. Math. Comput.*, **22**, 333–361.

K. Meintjes and A.P. Morgan (1989), "Element variables and the solution of complex chemical equilibrium problems", *Combust. Sci. Tech.*, **68**, 35–48.

K. Meintjes and A.P. Morgan (1990), "Chemical equilibrium systems as numerical test problems", *ACM Trans. Math. Software*, **16**, 143–151.

A.P. Morgan (1983), "A method for computing all solutions to systems of polynomial equations", *ACM Trans. Math. Software*, **9**, 1–17.

A.P. Morgan (1986a), "A homotopy for solving polynomial systems", *Appl. Math. Comput.*, **18**, 87–92.

A.P. Morgan (1986b), "A transformation to avoid solutions at infinity for polynomial systems", *Appl. Math. Comput.*, **18**, 77–86.

A.P. Morgan (1987), *Solving Polynomial Systems Using Continuation for Scientific and Engineering Problems*, Prentice-Hall, Englewood Cliffs, NJ.

A.P. Morgan (1989), "Polynomial continuation", *Impact of Recent Computer Advances on Operations Research*, R. Sharda, ed., Publications in Operations Research Series, North Holland.

A.P. Morgan and R.F. Sarraga (1982), "A method for computing three surface intersection points in GMSOLID", ASME paper 82-DET-41. Also available as Res. Pub. GMR-3964, Math. Dept., G.M. Research Labs., Warren, MI 49090, 1981.

A.P. Morgan and A.J. Sommese (1987a), "A homotopy for solving general polynomial systems that respects m-homogeneous structures", *Appl. Math. Comput.*, **24**, 101–113.

A.P. Morgan and A.J. Sommese (1987b), "Computing all solutions to polynomial systems using homotopy continuation", *Appl. Math. Comput.*, **24**, 115–138.

A.P. Morgan and A.J. Sommese (1989), "Coefficient-parameter polynomial continuation", *Appl. Math. Comput.*, **29**, 123–160.

A.P. Morgan and A.J. Sommese (1990), "Generically nonsingular polynomial continuation", *Lectures in Appl. Math. Vol. 26: Computational Solution of Nonlinear Systems of Equations*, E. Allgower and K. Georg, eds., American Math. Soc., 467–493.

A.P. Morgan, A.J. Sommese and C.W. Wampler (1990), "Polynomial continuation for mechanism design problems", *Lectures in Appl. Math. Vol. 26: Computational Solution of Nonlinear Systems of Equations*, E. Allgower and K. Georg, eds., American Math. Soc., 495–517.

A.P. Morgan, A.J. Sommese and C.W. Wampler (1991a), "Computing singular solutions to nonlinear analytic systems", *Numer. Math.*, **58**, 669–684.

A.P. Morgan, A.J. Sommese and C.W. Wampler (1991b), "Computing singular solutions to polynomial systems", *Advances in Appl. Math.*, to appear.

A.P. Morgan, A.J. Sommese and C.W. Wampler (1991c), "A power series method for computing singular

solutions to nonlinear analytic systems", submitted for publication.
A.P. Morgan, A.J. Sommese and L.T. Watson (1989), "Finding all isolated solutions to polynomial systems using HOMPACK", *ACM Trans. Math. Software*, **15**, 93–122.
A.P. Morgan and C.W. Wampler (1990), "Solving a planar four-bar design problem using continuation", *ASME J. Mechanical Design*, **112**, 544–550.
A.P. Morgan and L.T. Watson (1986), "Solving nonlinear equations on a hypercube", *Super and Parallel Computers and Their Impact on Civil Engineering*, M.P. Kamat, ed., ASCE Structures Congress '86, New Orleans, LA, 1–15.
A.P. Morgan and L.T. Watson (1987), "Solving polynomial systems of equations on a hypercube", *Hypercube Multiprocessors 1987*, M.T. Heath, ed., SIAM, Philadelphia, PA, 501–511.
A.P. Morgan and L.T. Watson (1989), "A globally convergent parallel algorithm for zeros of polynomial systems", *Nonlinear Analysis, Theory, Methods and Applications*, **13**, 1339–1350.
W.C. Rheinboldt and J.V. Burkardt (1983), "Algorithm 596: a program for a locally parameterized continuation process", *ACM Trans. Math. Software*, **9**, 236–241.
S. Richter and R. De Carlo (1984), "A homotopy method for eigenvalue assignment using decentralized state feedback", *IEEE Trans. Auto. Control*, **AC-29**, 148–158.
M.G. Safonov (1984), "Exact calculation of the multivariable structured-singular-value stability margin", *IEEE Control Decision Conf.*, Las Vegas, NV.
L.-W. Tsai and A.P. Morgan (1985), "Solving the kinematics of the most general six- and five-degree-of-freedom manipulators by continuation methods", *ASME J. Mechanisms, Transmissions Automation in Design*, **107**, 189–200.
C.W. Wampler and A.P. Morgan (1991), "Solving the 6R inverse position problem using a generic-case solution methodology", *Mechanism Mach. Theory*, **26**, 91–106.
C.W. Wampler, A.P. Morgan and A.J. Sommese (1990a), "Numerical continuation methods for solving polynomial systems arising in kinematics", *ASME J. Mechanical Design*, **112**, 59–68.
C.W. Wampler, A.P. Morgan and A.J. Sommese (1990b), "Complete solution of the nine-point path synthesis problem for four-bar linkages", *ASME J. Mechanical Design*, to appear.
L.T. Watson (1979), "A globally convergent algorithm for computing fixed points of C^2 maps", *Appl. Math. Comput.*, **5**, 297–311.
L.T. Watson (1986), "Numerical linear algebra aspects of globally convergent homotopy methods", *SIAM Review*, **28**, 529–545.
L.T. Watson, S.C. Billups, and A.P. Morgan (1987), "HOMPACK: a suite of codes for globally convergent homotopy algorithms", *ACM Trans. Math. Software*, **13**, 281–310.
L.T. Watson and A.P. Morgan (1991), "Global optimization for polynomial programming problems using m-homogeneous polynomial homotopies", *J. Computational Appl. Math.*, to appear.
A.H. Wright (1985), "Finding all solutions to a system of polynomial equations", *Math. Comput.*, **44**, 125–133.
W. Zulehner (1988a), "A simple homotopy method for determining all isolated solutions to polynomial systems", *Math. Comput.*, **50**, 167–177.
W. Zulehner (1988b), "On the solutions to polynomial systems obtained by homotopy methods", *Numer. Math.*, **54**, 303–317.
W. Zulehner (1988c), "Numerical solution of polynomial equation systems by curve tracing", *Zeitschrift fur Angewandte Mathematik und Mechanik*, **68**, T504–T505.