Acta Numerica (1997), *pp.* 399–436

Numerical solution of multivariate polynomial systems by homotopy continuation methods

T. Y. Li *
Department of Mathematics
Michigan State University
East Lansing, MI 48824-1027
USA
E-mail: li@math.msu.edu

CONTENTS

1. Introduction

Let $P(x) = 0$ be a system of n polynomial equations in n unknowns. Denoting $P = (p_1, \ldots, p_n)$, we want to find all isolated solutions of

$$\begin{aligned} p_1(x_1, \ldots, x_n) &= 0, \\ &\vdots \\ p_n(x_1, \ldots, x_n) &= 0 \end{aligned} \tag{1.1}$$

for $x = (x_1, \ldots, x_n)$. This problem is very common in many fields of science and engineering, such as formula construction, geometric intersection problems, inverse kinematics, power flow problems with PQ-specified bases, computation of equilibrium states, *etc.* Elimination theory-based methods, most notably the Buchberger algorithm (Buchberger 1985) for constructing Gröbner bases, are the classical approach to solving (1.1), but their reliance on symbolic manipulation makes those methods seem somewhat unsuitable for all but small problems.

* This research was supported in part by the NSF under Grant DMS-9504953 and a Guggenheim Fellowship.

In 1977, Garcia and Zangwill (1979) and Drexler (1977) independently presented theorems suggesting that homotopy continuation could be used to find numerically the full set of isolated solutions of (1.1). During the last two decades, this method has been developed into a reliable and efficient numerical algorithm for approximating all isolated zeros of polynomial systems. Modern scientific computing is marked by the advent of vector and parallel computers and the search for algorithms that are to a large extent naturally parallel. A great advantage of the homotopy continuation algorithm for solving polynomial systems is that it is to a large degree parallel, in the sense that each isolated zero can be computed independently. This natural parallelism makes the method an excellent candidate for a variety of architectures. In this respect, it stands in contrast to the highly serial Gröbner bases method.

The homotopy continuation method for solving (1.1) is to define a trivial system $Q(x) = (q_1(x), \ldots, q_n(x)) = 0$ and then follow the curves in the real variable t which make up the solution set of

$$0 = H(x,t) = (1-t)Q(x) + tP(x). \tag{1.2}$$

More precisely, if $Q(x) = 0$ is chosen correctly, the following three properties hold:

Property 0 (*Triviality*). The solutions of $Q(x) = 0$ are known.

Property 1 (*Smoothness*). The solution set of $H(x,t) = 0$ for $0 \leq t < 1$ consists of a finite number of smooth paths, each parametrized by t in $[0,1)$.

Property 2 (*Accessibility*). Every isolated solution of $H(x,1) = P(x) = 0$ can be reached by some path originating at $t = 0$. It follows that this path starts at a solution of $H(x,0) = Q(x) = 0$.

When the three properties hold, the solution paths can be followed from the initial points (known because of Property 0) at $t = 0$ to all solutions of the original problem $P(x) = 0$ at $t = 1$ using standard numerical techniques; see Allgower and Georg (1990, 1993).

Several authors have suggested choices of Q that satisfy the three properties: *cf.* Chow, Mallet-Paret and Yorke (1979), Li (1983), Morgan (1986), Wright (1985) and Zulener (1988) for a partial list. A typical suggestion is

$$\begin{aligned} q_1(x) &= a_1 x_1^{d_1} - b_1, \\ &\vdots \\ q_n(x) &= a_n x_n^{d_n} - b_n, \end{aligned} \tag{1.3}$$

where $d_1, \ldots, d_n$ are the degrees of $p_1(x), \ldots, p_n(x)$ respectively, and a_i, b_i are random complex numbers (and therefore nonzero, with probability one).

So in one sense, the original problem posed is solved. All solutions of $P(x) = 0$ are found at the end of the $d_1 \cdots d_n$ paths that make up the solution set of $H(x,t) = 0, 0 \leq t < 1$.

In this article, we report on some recent developments that make this method more convenient to apply.

The reason the problem is not satisfactorily solved by the above considerations is the existence of *extraneous paths*. Although the above method produces $d = d_1 \cdots d_n$ paths, the system $P(x) = 0$ may have fewer than d solutions. We call such a system *deficient*. In this case, some of the paths produced by the above method will be extraneous paths.

More precisely, even though Properties 0–2 imply that each solution of $P(x) = 0$ will lie at the end of a solution path, it is also consistent with these properties that some of the paths may diverge to infinity as the parameter t approaches 1 (the smoothness property rules this out for $t \to t_0 < 1$). In other words, it is quite possible for $Q(x) = 0$ to have more solutions than $P(x) = 0$. In this case, some of the paths leading from roots of $Q(x) = 0$ are extraneous, and diverge to infinity when $t \to 1$ (see Figure 1).

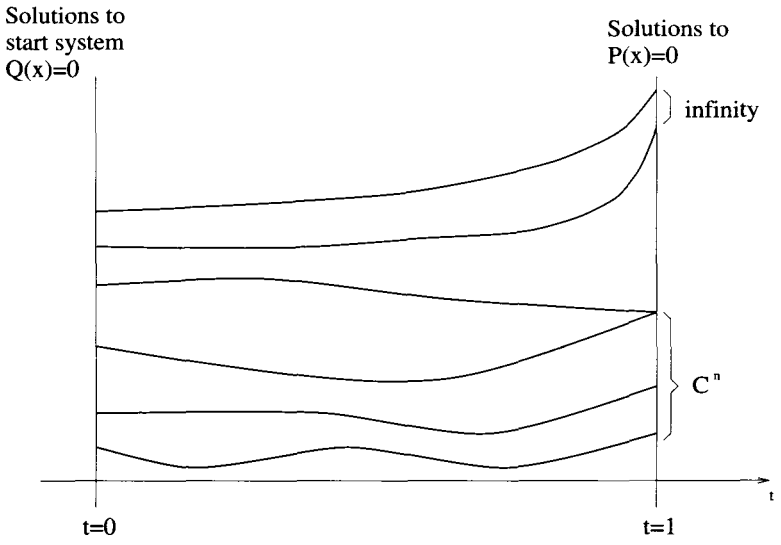

Fig. 1.

Empirically, we find that most systems arising in applications are deficient. A great majority of the systems have fewer than, and in some cases only a small fraction of, the 'expected number' of solutions. For a typical example

of this sort, let us look at the following Cassou–Nogues system

$$\begin{aligned}
p_1 &= 15b^4cd^2 + 6b^4c^3 + 21b^4c^2d - 144b^2c - 8b^2c^2e \\
&\quad -28b^2cde - 648b^2d + 36b^2d^2e + 9b^4d^3 - 120, \\
\\
p_2 &= 30b^4c^3d - 32cde^2 - 720b^2cd - 24b^2c^3e - 432b^2c^2 + 576ce \\
&\quad -576de + 16b^2cd^2e + 16d^2e^2 + 16c^2e^2 + 9b^4c^4 + 39b^4c^2d^2 \\
&\quad +18b^4cd^3 - 432b^2d^2 + 24b^2d^3e - 16b^2c^2de - 240c + 5184, \qquad (1.4) \\
\\
p_3 &= 216b^2cd - 162b^2d^2 - 81b^2c^2 + 1008ce - 1008de + 15b^2c^2de \\
&\quad -15b^2c^3e - 80cde^2 + 40d^2e^2 + 40c^2e^2 + 5184, \\
\\
p_4 &= 4b^2cd - 3b^2d^2 - 4b^2c^2 + 22ce - 22de + 261.
\end{aligned}$$

Since $d_1 = 7, d_2 = 8, d_3 = 6$ and $d_4 = 4$ for this system, the system $Q(x)$ in (1.3) will produce $d_1 \times d_2 \times d_3 \times d_4 = 7 \times 8 \times 6 \times 4 = 1344$ paths for the homotopy in (1.2). However, the system (1.4) has only 16 isolated zeros. Consequently, most of the paths are extraneous. Sending out 1344 paths in search of 16 solutions is a highly wasteful computation.

The choice of $Q(x)$ in (1.3) to solve the system $P(x) = 0$ requires an amount of computational effort proportional to $d_1 \cdots d_n$ and, roughly, proportional to the size of the system. We would like to derive methods for solving deficient systems for which the computational effort is instead proportional to the actual number of solutions.

To organize our discussion, we will at times use a notation that makes the coefficients and variables in $P(x) = 0$ explicit. Thus, when the dependence on coefficients is important, we will consider the system $P(c, x) = 0$ of n polynomial equations in n unknowns, where $c = (c_1, \ldots, c_M)$ are coefficients and $x = (x_1, \ldots, x_n)$ are unknowns. Two different problems can be posed:

Problem A Solve the system of equations $P(x) = 0$.

Problem B For each of several different choices of coefficients c, solve the system of equations $P(c, x) = 0$.

We divide our discussion on dealing with and eliminating extraneous paths for Problem A in Section 2, and for Problem B in Section 3. In Section 4, an algorithm is presented which, in some sense, uses the method for Problem B to treat Problem A. Some numerical considerations, the use of projective coordinates and real homotopies, are given in Section 5.

2. Methods for Problem A

Progress on Problem A has been the least satisfactory among the areas we discuss. For deficient systems, there are some partial results that use algebraic geometry to reduce the number of extraneous paths, with various degrees of success.

2.1. Random product homotopy

For a specific example that is quite simple, consider the system

$$\begin{aligned} p_1(x) &= x_1(a_{11}x_1 + \cdots + a_{1n}x_n) + b_{11}x_1 + \cdots + b_{1n}x_n + c_1 = 0, \\ &\vdots \\ p_n(x) &= x_1(a_{n1}x_1 + \cdots + a_{nn}x_n) + b_{n1}x_1 + \cdots + b_{nn}x_n + c_n = 0. \end{aligned} \tag{2.1}$$

This system has total degree $d = d_1 \cdots d_n = 2^n$. Thus the 'expected number' of solutions is 2^n, and the classical homotopy continuation method using the start system $Q(x) = 0$ in (1.3) sends out 2^n paths from 2^n trivial starting points. However, the system $P(x) = 0$ has only $n+1$ isolated solutions (even fewer for special choices of coefficients). This is a deficient system; at least $2^n - n - 1$ paths will be extraneous. It is never known from the start which of the paths will end up being extraneous, so they must all be followed to the end: wasteful computation.

The random product homotopy was developed in Li, Sauer and Yorke (1987*a*, 1987*b*) to alleviate this problem. According to that technique, a more efficient choice for the trivial system $Q(x) = 0$ is

$$\begin{aligned} q_1(x) &= (x_1 + e_{11})(x_1 + x_2 + \cdots + x_n + e_{12}), \\ q_2(x) &= (x_1 + e_{21})(x_2 + e_{22}), \\ &\vdots \\ q_n(x) &= (x_1 + e_{n1})(x_n + e_{n2}). \end{aligned} \tag{2.2}$$

Set

$$H(x,t) = (1-t)cQ(x) + tP(x).$$

It is clear by inspection that for a generic choice of the complex numbers e_{ij}, $Q(x) = 0$ has exactly $n+1$ roots. Thus there are only $n+1$ paths starting from $n+1$ starting points for this choice of homotopy. It is proved in Li, Sauer and Yorke (1987*b*) that Properties 0–2 hold for this choice of $H(x,t)$ for almost all complex numbers e_{ij} and c. Thus all solutions of $P(x) = 0$ are found at the end of the $n+1$ paths. The result of Li et al. (1987*b*) is then both a mathematical result (that there can be at most $n+1$ solutions to (2.1)) and the basis of a numerical procedure for approximating the solutions.

The reason this works is quite simple. The solution paths of (1.2) which do not proceed to a solution of $P(x) = 0$ in $\mathbb{C}^n$ diverge to infinity. If the system (1.2) is viewed in projective space

$$\mathbb{P}^n = \mathbb{C}^{n+1} \backslash \{(0,\ldots,0)\} / \sim,$$

where the equivalent relation '$\sim$' is given by $x \sim y$ if $x = cy$ for some nonzero $c \in \mathbb{C}$, the diverging paths simply proceed to a 'point at infinity' in $\mathbb{P}^n$.

For a polynomial $f(x_1, \ldots, x_n)$ of degree d, denote the associated homogeneous polynomial by

$$\tilde{f}(x_0, x_1, \ldots, x_n) = x_0^d f(\frac{x_1}{x_0}, \ldots, \frac{x_n}{x_0}).$$

The solutions of $f(x) = 0$ at infinity are those zeros of $\tilde{f}$ in $\mathbb{P}^n$ with $x_0 = 0$, and the remaining zeros of $\tilde{f}$ with $x_0 \neq 0$ are the solutions of $f(x) = 0$ in $\mathbb{C}^n$ when x_0 is set to be 1.

Viewed in projective space $\mathbb{P}^n$ the system $P(x) = 0$ in (2.1) has some roots at infinity. The roots at infinity make up a nonsingular variety, specifically the linear space $\mathbb{P}^{n-2}$ defined by $x_0 = x_1 = 0$. A Chern class formula from intersection theory (Fulton 1984, 9.1.1, 9.1.2) shows that the contribution of a linear variety of solutions of dimension e to the 'total degree' $(d_1 \times \cdots \times d_n)$, or the total expected number of solutions, of the system is at least s, where s is the coefficient of t^e in the Maclaurin series expansion of

$$(1+t)^{e-n} \prod_{i=1}^{n} (1 + d_i t).$$

In our case, $d_1 = \cdots = d_n = 2$, and $e = n - 2$, hence,

$$\frac{(1+2t)^n}{(1+t)^2} = \frac{\sum_{i=0}^{n} (1+t)^{n-i} t^i \binom{n}{i}}{(1+t)^2} = \sum_{i=0}^{n} (1+t)^{n-i-2} t^i \binom{n}{i}$$

and $s = \sum_{i=0}^{n-2} \binom{n}{i}$, meaning there are at least $\sum_{i=0}^{n-2} \binom{n}{i}$ solutions of $P(x) = 0$ at infinity. Thus there are at most

$$2^n - s = (1+1)^n - \sum_{i=0}^{n-2} \binom{n}{i} = n + 1$$

solutions of $P(x) = 0$ in $\mathbb{C}^n$. The system $Q(x) = 0$ is chosen to have the same nonsingular variety at infinity, and this variety stays at infinity as the homotopy progresses from $t = 0$ to $t = 1$. As a result, the infinity solutions stay infinite, the finite solution paths stay finite, and no extraneous paths exist.

This turns out to be a fairly typical situation. Even though the system $P(x) = 0$ to be solved has isolated solutions, when viewed in projective space there may be large number of roots at infinity, and quite often high-dimensional manifolds of roots at infinity. Extraneous paths are those that are drawn to the manifolds lying at infinity. If $Q(x) = 0$ can be chosen correctly, extraneous paths can be eliminated.

As another example, consider the algebraic eigenvalue problem

$$Ax = \lambda x,$$

where

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

is an $n \times n$ matrix. This problem is actually one of n polynomial equations in the $n+1$ variables $\lambda, x_1, \ldots, x_n$:

$$\begin{aligned} \lambda x_1 - (a_{11}x_1 + \cdots + a_{1n}x_n) &= 0, \\ &\vdots \\ \lambda x_n - (a_{n1}x_1 + \cdots + a_{nn}x_n) &= 0. \end{aligned}$$

Augmenting the system with a linear equation

$$c_1 x_1 + \cdots + c_n x_n + c_{n+1} = 0,$$

where $c_1, \ldots, c_{n+1}$ are chosen at random, we have a polynomial system of $n+1$ equations in $n+1$ variables. This system has total degree 2^n. However, it can have at most n isolated solutions. So, the system is deficient. But the system $Q(x)$ in random product form:

$$\begin{aligned} q_1 &= (\lambda + e_{11})(x_1 + e_{12}), \\ q_2 &= (\lambda + e_{21})(x_2 + e_{22}), \\ &\vdots \\ q_n &= (\lambda + e_{n1})(x_n + e_{n2}), \\ q_{n+1} &= c_1 x_1 + \cdots + c_n x_n + c_{n+1} \end{aligned}$$

has n isolated zeros for randomly chosen e_{ij}s. This $Q(x)$ will produce n curves for the homotopy in (1.3) that proceed to all solutions of the eigenvalue problem. Implicit in this is the fact that the algebraic eigenvalue problem has at most n solutions. Moreover, the generic eigenvalue problem has exactly n solutions.

To be more precise, we state the main random product homotopy result, Theorem 2.2 of Li et al. (1987*b*). Let $V_\infty(Q)$ and $V_\infty(P)$ denote the variety of roots at infinity of $Q(x) = 0$ and $P(x) = 0$ respectively.

Theorem 2.1 If $V_\infty(Q)$ is nonsingular and contained in $V_\infty(P)$, then Properties 1 and 2 hold.

Of course, Properties 1 and 2 are not enough. Without starting points, the path-following method cannot begin. Thus $Q(x) = 0$ should also be chosen to be of random product form, as in (2.2), these being trivial to solve.

This result was superseded by the result in Li and Sauer (1989). The complex numbers e_{ij} are chosen at random in Li et al. (1987*b*) to ensure Properties 1 and 2. In Li and Sauer (1989), it was proved that e_{ij} can be any fixed numbers; as long as the complex number c is chosen at random,

Properties 1 and 2 still hold. In fact, the result in Li and Sauer (1989) implies that the start system $Q(x) = 0$ in Theorem 2.1 need not be in product form. It can be any chosen polynomial system as long as its zeros in $\mathbb{C}^n$ are known or easy to obtain and its variety of roots at infinity $V_\infty(Q)$ is nonsingular and contained in $V_\infty(P)$.

Theorem 2.1 in Li and Wang (1991) goes one step further. Even when the set $V_\infty(Q)$ of roots at infinity of $Q(x) = 0$ has singularities, if the set is contained in $V_\infty(P)$ counting multiplicities, that is, containment in the sense of *scheme* theory of algebraic geometry, then Properties 1 and 2 still hold. To be more precise, let $I =< \tilde{q}_1, \ldots, \tilde{q}_n >$ and $J =< \tilde{p}_1, \ldots, \tilde{p}_n >$ be the homogeneous ideals spanned by homogenizations of q_is and p_is respectively. For a point p at infinity, if the *local rings* I_p and J_p satisfy

$$I_p \subset J_p,$$

then Properties 1 and 2 hold. However, this hypothesis can be much more difficult to verify than the singularity of the set. This limits the usefulness of this approach for practical examples.

2.2. m-homogeneous structure

In Morgan and Sommese (1987*b*), another interesting approach to Problem A is developed, using the concept of m-homogeneous structure.

The complex n-space $\mathbb{C}^n$ can be naturally embedded in $\mathbb{P}^n$. Similarly, the space $\mathbb{C}^{k_1} \times \cdots \times \mathbb{C}^{k_m}$ can be naturally embedded in $\mathbb{P}^{k_1} \times \cdots \times \mathbb{P}^{k_m}$. A point $(y_1, \ldots, y_m)$ in $\mathbb{C}^{k_1} \times \cdots \times \mathbb{C}^{k_m}$ with $y_i = (y_1^{(i)}, \ldots, y_{k_i}^{(i)})$, $i = 1, \ldots, m$, corresponds to a point $(z_1, \ldots, z_m)$ in $\mathbb{P}^{k_1} \times \cdots \times \mathbb{P}^{k_m}$ with $z_i = (z_0^{(i)}, \ldots, z_{k_i}^{(i)})$ and $z_0^{(i)} = 1$, $i = 1, \ldots, m$. The set of such points in $\mathbb{P}^{k_1} \times \cdots \times \mathbb{P}^{k_m}$ is usually called the *affine space* in this setting. The points in $\mathbb{P}^{k_1} \times \cdots \times \mathbb{P}^{k_m}$ with at least one $z_0^{(i)} = 0$ are called the *points at infinity*.

Let f be a polynomial in the n variables $x_1, \ldots, x_n$. If we partition the variables into m groups $y_1 = (x_1^{(1)}, \ldots, x_{k_1}^{(1)}), y_2 = (x_1^{(2)}, \ldots, x_{k_2}^{(2)}), \ldots, y_m = (x_1^{(m)}, \ldots, x_{k_m}^{(m)})$ with $k_1 + \cdots + k_m = n$ and let d_i be the degree of f with respect to y_i (more precisely, to the variables in y_i), then we can define its m-homogenization as

$$\tilde{f}(z_1, \ldots, z_m) = (z_0^{(1)})^{d_1} \times \cdots \times (z_0^{(m)})^{d_m} f(y_1/z_0^{(1)}, \ldots, y_m/z_0^{(m)}).$$

This polynomial is homogeneous with respect to each $z_i = (z_0^{(i)}, \ldots, z_{k_i}^{(i)})$, $i = 1, \ldots, m$. Here $z_j^{(i)} = x_j^{(i)}$, for $j \neq 0$. Such a polynomial is said to be m-*homogeneous*, and $(d_1, \ldots, d_m)$ is the m-homogeneous degree of f. To

illustrate this definition, let us consider the polynomial $p_i(x)$ in (2.1):

$$\begin{aligned} p_i(x) &= x_1(a_{i1}x_1 + \cdots + a_{in}x_n) + b_{i1}x_1 + \cdots + b_{in}x_n + c_i \\ &= a_{i1}x_1^2 + x_1(a_{i2}x_2 + \cdots + a_{in}x_n + b_{i1}) + b_{i2}x_2 + \cdots + b_{in}x_n + c_i. \end{aligned}$$

It is sufficient to set $y_1 = (x_1), y_2 = (x_2, \ldots, x_n)$ and $z_1 = (x_0^{(1)}, x_1)$, $z_2 = (x_0^{(2)}, x_2, \ldots, x_n)$. The degree of $p_i(x)$ is two with respect to y_1 and is one with respect to y_2. Hence, its 2-homogenization is

$$\begin{aligned} \widetilde{p}_i(z_1, z_2) &= a_{i1}x_1^2x_0^{(2)} + x_1x_0^{(1)}(a_{i2}x_2 + \cdots + a_{in}x_n + b_{i1}x_0^{(2)}) \\ &+ (x_0^{(1)})^2(b_{i2}x_2 + \cdots + b_{in}x_n + c_ix_0^{(2)}). \end{aligned}$$

which is homogeneous with respect to both z_1 and z_2. When the system (2.1) is viewed in $\mathbb{P}^n = \{(x_0, x_1, \ldots, x_n)\}$ with the homogenization

$$\begin{aligned} \widetilde{p}_1(x_0, x_1, \ldots, x_n) &= x_1(a_{11}x_1 + \cdots + a_{1n}x_n) \\ &\quad +(b_{11}x_1 + \cdots + b_{1n}x_n)x_0 + c_1x_0^2 = 0, \\ &\vdots \\ \widetilde{p}_n(x_0, x_1, \ldots, x_n) &= x_1(a_{n1}x_1 + \cdots + a_{nn}x_n) \\ &\quad +(b_{n1}x_1 + \cdots + b_{nn}x_n)x_0 + c_nx_0^2 = 0, \end{aligned}$$

its total degree, or Bézout number, is $d = d_1 \cdots d_n = 2^n$. However, when (2.1) is viewed in $\mathbb{P}^1 \times \mathbb{P}^{n-1} = \{(z_1, z_2) = ((x_0^{(1)}, x_1), (x_0^{(2)}, x_2, \ldots, x_n))\}$ with the 2-homogenization

$$\begin{aligned} \widetilde{p}_1(z_1, z_2) &= a_{11}x_1^2x_0^{(2)} + x_1x_0^{(1)}(a_{12}x_2 + \cdots + a_{in}x_n + b_{11}x_0^{(2)}) \\ &\quad +(x_0^{(1)})^2(b_{12}x_2 + \cdots + b_{1n}x_n + c_1x_0^{(2)}), \\ &\vdots \\ \widetilde{p}_n(z_1, z_2) &= a_{n1}x_1^2x_0^{(2)} + x_1x_0^{(1)}(a_{n2}x_2 + \cdots + a_{nn}x_n + b_{n1}x_0^{(2)}) \\ &\quad +(x_0^{(1)})^2(b_{n2}x_2 + \cdots + b_{nn}x_n + c_nx_0^{(2)}), \end{aligned} \tag{2.3}$$

the Bézout number d is different, and equals the coefficient of $\alpha_1^1\alpha_2^{n-1}$ in the product $(2\alpha_1 + \alpha_2)^n$. Thus, $d = 2n$. In general, for an m-homogeneous system

$$\begin{aligned} \widetilde{p}_1(z_1, \ldots, z_m) &= 0, \\ &\vdots \\ \widetilde{p}_n(z_1, \ldots, z_m) &= 0, \end{aligned} \tag{2.4}$$

in $\mathbb{P}^{k_1} \times \cdots \times \mathbb{P}^{k_m}$ with $\widetilde{p}_i$ having m-homogeneous degree $(d_1^{(i)}, \ldots, d_m^{(i)})$, $i = 1, \ldots, n$, with respect to $(z_1, \ldots, z_m)$, then the m-homogeneous Bézout

number d of the system with respect to $(z_1, \ldots, z_m)$ is the coefficient of $\alpha_1^{k_1} \times \cdots \times \alpha_m^{k_m}$ in the product

$$(d_1^{(1)}\alpha_1 + \cdots + d_m^{(1)}\alpha_m)(d_1^{(2)}\alpha_1 + \cdots + d_m^{(2)}\alpha_m) \cdots (d_1^{(n)}\alpha_1 + \cdots + d_m^{(n)}\alpha_m)$$

(Shafarevich 1977). The classical Bézout Theorem says the system (2.4) has no more than d isolated solutions, counting multiplicities, in $\mathbb{P}^{k_1} \times \cdots \times \mathbb{P}^{k_m}$. Applying this to our example in (2.3), the upper bound on the number of isolated solutions of (2.3), in affine space and at infinity, is $2n$. When solving the original system in (2.1), we may choose the start system $Q(x) = 0$ in the homotopy

$$H(x,t) = (1-t)cQ(x) + tP(x)$$

in random product form to respect the 2-homogeneous structure of $P(x)$. For instance, we may choose $Q(x) = 0$ to be

$$\begin{aligned} q_1(x) &= (x_1 + e_{11})(x_1 + e_{12})(x_2 + \cdots + x_n + e_{13}), \\ q_2(x) &= (x_1 + e_{21})(x_1 + e_{22})(x_2 + e_{23}), \\ &\vdots \\ q_n(x) &= (x_1 + e_{n1})(x_1 + e_{n2})(x_n + e_{n3}), \end{aligned} \tag{2.5}$$

which has the same 2-homogeneous structure as $P(x)$ with $y_1 = (x_1)$ and $y_2 = (x_2, \ldots, x_n)$. Namely, each $q_i(x)$ has degree two with respect to y_1 and degree one with respect to y_2. It is easy to see that for randomly chosen complex numbers e_{ij}, $Q(x) = 0$ has $2n$ solutions in $\mathbb{C}^n (= \mathbb{C}^1 \times \mathbb{C}^{n-1})$ (thus, no solutions at infinity when viewed in $\mathbb{P}^1 \times \mathbb{P}^{n-1}$). Hence there are $2n$ paths starting from $2n$ starting points for this choice of homotopy. It is shown in Morgan and Sommese (1987*b*) that Properties 1 and 2 hold for all complex numbers c, except those lying on a finite number of rays starting at the origin. Thus, all solutions of $P(x) = 0$ are found at the end of $n+1$ paths. The number of extraneous paths, $2n - (n+1) = n - 1$, is far less than the corresponding number, namely $2^n - n - 1$, arising via classical homotopy with $Q(x) = 0$ in (1.3).

More precisely, we state the main theorem in Morgan and Sommese (1987*b*).

Theorem 2.2 Let $Q(x)$ be a system of equations chosen to have the same m-homogeneous form as $P(x)$ with respect to a certain partition of the variables $(x_1, \ldots, x_n)$. Assume that $Q(x) = 0$ has exactly the Bézout number of nonsingular solutions with respect to this partition, and define

$$H(x,t) = (1-t)cQ(x) + tP(x),$$

where $t \in [0,1]$ and $c \in \mathbb{C}$. If $c = re^{i\theta}$ for some positive r, then, for all but finitely many θ, Properties 1 and 2 hold.

In general, if $x = (x_1, \ldots, x_n)$ is partitioned into $x = (y_1, \ldots, y_m)$ where

$$y_1 = (x_1^{(1)}, \ldots, x_{k_1}^{(1)}), \quad y_2 = (x_1^{(2)}, \ldots, x_{k_2}^{(2)}), \quad \ldots, \quad y_m = (x_1^{(m)}, \ldots, x_{k_m}^{(m)}),$$

with $k_1 + \cdots + k_m = n$, and $p_i(x)$ has degree $(d_1^{(i)}, \ldots, d_m^{(i)})$ with respect to $(y_1, \ldots, y_m)$, $i = 1, \ldots, n$, then we may choose the start system $Q(x) = (q_1(x), \ldots, q_n(x))$,

$$q_i(x) = \prod_{j=1}^{m} \prod_{\ell=1}^{d_j^{(i)}} (c_{\ell 1}^{(j)} x_1^{(j)} + \cdots + c_{\ell k_j}^{(j)} x_{k_j}^{(j)} + c_{\ell 0}^{(j)}), \quad i = 1, \ldots, n. \quad (2.6)$$

Clearly, $q_i(x)$ has degree $(d_1^{(i)}, \ldots, d_m^{(i)})$ with respect to $(y_1, \ldots, y_m)$, the same degree structure of $p_i(x)$. Further, it is not hard to see that, for random coefficients, $Q(x)$ has exactly an m-homogeneous Bézout number, with respect to this particular partition $x = (y_1, \ldots, y_m)$, of nonsingular isolated solutions in $\mathbb{C}^n$. Those solutions are easy to obtain: the system $Q(x)$ in (2.5) is constructed according to this principle. In Wampler (1994), the product in (2.6) is modified along the same principle to be more efficient to evaluate.

In the example above, there are still $n - 1$ extraneous paths. This is because, even when it is viewed in $\mathbb{P}^1 \times \mathbb{P}^{n-1}$, $P(x)$ has zeros at infinity. One can see in (2.3) that

$$S = \{((x_0^{(1)}, x_1), (x_0^{(2)}, x_2, \ldots, x_n)) \in \mathbb{P}^1 \times \mathbb{P}^{n-1} : x_0^{(1)} = 0, x_0^{(2)} = 0\}$$

is a set of zeros of $P(x)$ at infinity. So, to lower the number of those extraneous paths further, we may choose the start system to have the same nonsingular variety of roots as $P(x) = 0$ at infinity, in addition to sharing the same 2-homogeneous structure of $P(x)$. For instance, the system $Q(x) = (q_1(x), \ldots, q_n(x))$ where

$$\begin{aligned} q_1(x) &= (x_1 + e_{11})(x_1 + x_2 + \cdots + x_n + e_{12}), \\ q_2(x) &= (x_1 + e_{21})(x_1 + x_2 + e_{22}), \\ &\vdots \\ q_n(x) &= (x_1 + e_{n1})(x_1 + x_n + e_{n2}) \end{aligned}$$

shares the same 2-homogeneous structure of $P(x)$ with $y_1 = (x_1)$ and $y_2 = (x_2, \ldots, x_n)$, namely, each $q_i(x)$ has degree two with respect to y_1 and degree one with respect to y_2. On the other hand, when viewed in $(z_1, z_2) \in \mathbb{P}^1 \times \mathbb{P}^{n-1}$ with $z_1 = (x_0^{(1)}, x_1)$ and $z_2 = (x_0^{(2)}, x_2, \ldots, x_n)$, this system has the same nonsingular variety S at infinity as $P(x)$. The system $Q(x) = 0$ also has $n+1$ solutions in $\mathbb{C}^n$ for generic e_{ij}s, and there are no extraneous paths. It can be shown (Li and Wang 1991, Morgan and Sommese 1987*a*) that if $Q(x) = 0$ in

$$H(x, t) = (1 - t)cQ(x) + tP(x)$$

is chosen to have the same m-homogeneous form as $P(x)$, and the set of zeros $V_\infty(Q)$ of $Q(x)$ at infinity is nonsingular and contained in $V_\infty(P)$, then Properties 1 and 2 hold for $c = re^{i\theta}$, $r > 0$, and for all but finitely many θ.

The zeros of an m-homogeneous polynomial system $\widetilde{P}(z_1, \ldots, z_m)$ at infinity in $\mathbb{P}^{k_1} \times \cdots \times \mathbb{P}^{k_m}$ may sometimes be difficult to obtain. Nevertheless, the choice of $Q(x) = 0$ in Theorem 2.2, assuming no zeros at infinity regardless of the structure of the zeros at infinity of $P(x)$, can still reduce the number of extraneous paths dramatically simply by sharing the same m-homogeneous structure of $P(x)$.

Let us consider the system

$$\begin{aligned} p_1(x) &= x_1(a_{11}x_1 + \cdots + a_{1n}x_n) + b_{11}x_1 + \cdots + b_{1n}x_n + c_1 = 0, \\ &\vdots \\ p_n(x) &= x_1(a_{n1}x_1 + \cdots + a_{nn}x_n) + b_{n1}x_1 + \cdots + b_{nn}x_n + c_n = 0 \end{aligned}$$

in (2.1) again. This time we partition the variables $x_1, \ldots, x_n$ into $y_1 = (x_1, x_2)$ and $y_2 = (x_3, \ldots, x_n)$. For this partition, the 2-homogeneous degree structure of $p_i(x)$ stays the same; namely, the degree of $p_i(x)$ is two with respect to y_1 and is one with respect to y_2. However, the Bézout number with respect to this partition becomes the coefficient of $\alpha_1^2\alpha_2^{n-2}$ in the product $(2\alpha_1 + \alpha_2)^n$. This number is

$$\binom{n}{2} \times 2^2 = 2n(n-1),$$

which is greater than the original Bézout number $2n$ with respect to the partition $y_1 = (x_1)$ and $y_2 = (x_2, \ldots, x_n)$ when $n > 2$. Apparently, the Bézout number is highly sensitive to the chosen partition: different ways of partitioning the variables produce different Bézout numbers. By using Theorem 2.2, we follow the Bézout number (with respect to the chosen partition) of curves to obtain all the isolated zeros of $P(x)$. To minimize the number of extraneous paths, it is certainly desirable to find a partition which provides the lowest Bézout number possible. In Wampler (1992), an algorithm to this end was given. By using this algorithm, one can determine, for example, the partition $\mathcal{P} = \{(b), (c, d, e)\}$ which gives the lowest possible Bézout number 368 for the Cassou–Nogues system in (1.4). Consequently, we may construct a random product start system $Q(x)$ to respect the degree structure of the system with respect to this partition. The start system $Q(x)$ will have 368 isolated zeros in $\mathbb{C}^n$, and, according to Theorem 2.2, only 368 homotopy curves need to be followed to obtain all 16 isolated zeros of the Cassou–Nogues system, in contrast to following the 1344 curves, 1344 being the total degree of the system.

The usefulness of the methods yet developed for Problem A is restricted to application on an *ad hoc* basis. The challenge is, in a specific case, to

find a $Q(x)$ that is simple to solve (Property 0) and also produces minimal extraneous paths.

3. Methods for Problem B

The situation for Problem B is different. A method called the 'cheater's homotopy' has been developed, which is, in some sense, an optimum solution procedure; see Li, Sauer and Yorke (1988) and Li, Sauer and Yorke (1989) (a similar procedure can be found in Morgan and Sommese (1989)). Problem B asks that the system $P(c, x) = 0$ be solved for several different values of the coefficients c. In other words, we think of $P(c, x) = 0$ as a system with the same structure or sparsity.

The idea of the method is to establish Properties 1 and 2 theoretically by deforming a sufficiently generic system (in a precise sense to be given later) and then to 'cheat' on Property 0 by using a preprocessing step. The amount of computation per preprocessing step may be large, but is shared among the several solving characteristics of Problem B.

We begin with an example. Let $P(x)$ be the system

$$\begin{aligned} p_1(x) &= x_1^3 x_2^2 + c_1 x_1^3 x_2 + x_2^2 + c_2 x_1 + c_3 = 0, \\ p_2(x) &= c_4 x_1^4 x_2^2 - x_1^2 x_2 + x_2 + c_5 = 0. \end{aligned} \tag{3.1}$$

This is a system of two polynomial equations in two unknowns x_1 and x_2. We want to solve Problem B, that is, we want to solve the system of equations several times, for various specific choices of $c = (c_1, \ldots, c_5)$.

It turns out that, for any choice of coefficients c, system (3.1) has at most 10 isolated solutions. More precisely, there is an open dense subset S of $\mathbb{C}^5$ such that, for $c \in S$, there are 10 solutions of (3.1). Moreover, 10 is an upper bound for the number of isolated solutions for all c in $\mathbb{C}^5$. The total degree of the system is $6 \times 5 = 30$, meaning that if we had taken a generic system of two polynomials in two variables of degree 5 and 6, there would be 30 solutions. Thus (3.1), with any choice of c, is a deficient system.

Classical homotopy using the start system $Q(x) = 0$ in (1.3) produces $d = 30$ paths, beginning at 30 trivial starting points. Thus there are (at least) 20 extraneous paths.

The cheater's homotopy continuation approach begins by solving (3.1) with *randomly chosen* complex coefficients $c^* = (c_1^*, \ldots, c_5^*)$; let X^* be the set of 10 solutions. No work is saved, since 30 paths need to be followed and 20 paths are wasted. However, the 10 elements of the set X^* are the seeds for the remainder of the process. Subsequently, for each choice of coefficients $c = (c_1, \ldots, c_5)$ for which the system (3.1) needs to be solved, we use the homotopy continuation method to follow a straight-line homotopy from the system with coefficient c^* to the system with coefficient c, and we follow the 10 paths beginning at the 10 elements of X^*. Thus Property 0,

the existence of trivial starting points, is satisfied. The fact that Properties 1 and 2 are also satisfied is the content of Theorem 3.1 below. Thus for each fixed c, all 10 (or fewer) isolated solutions of (3.1) lie at the end of 10 smooth homotopy paths beginning at the seeds in X^*. After the initial step of finding the seeds, the complexity of all further solvings of (3.1) is proportional to the number of solutions 10, rather than the total degree 30.

Furthermore, this method, unlike the method for Problem A, requires no *a priori* analysis of the system. The first preprocessing step of finding the seeds establishes a sharp upper bound on the number of isolated solutions as a by-product of the computation; further solving of the system uses the optimal number of paths to be followed.

We earlier characterized a successful homotopy continuation method as having three properties: triviality, smoothness, and accessibility (Properties 0, 1 and 2, respectively). Given an arbitrary system of polynomial equations, such as (3.1), it is not too hard (through generic perturbations) to find a family of systems with the last two properties. The problem is that one member of the family must be trivial to solve, or the path-following cannot begin. The idea of the cheater's homotopy is simply to 'cheat' on this part of the problem, and run a preprocessing step (the computation of the seeds X^*) which gives us Property 0 (triviality) in a roundabout way. Hence the name, the 'cheater's homotopy'.

A statement of the theoretical result we need follows. Let

$$\begin{aligned} p_1(c_1,\ldots,c_M,x_1,\ldots,x_n) &= 0, \\ &\vdots \\ p_n(c_1,\ldots,c_M,x_1,\ldots,x_n) &= 0, \end{aligned} \tag{3.2}$$

be a system of polynomial equations in the variables $c_1,\ldots,c_M,x_1,\ldots,x_n$. For each choice of $c=(c_1,\ldots,c_M)$ in $\mathbb{C}^M$, this is a system of polynomial equations in the variables $x_1,\ldots,x_n$. Let d be the total degree of the system for a generic choice of c.

Theorem 3.1 Let c belong to $\mathbb{C}^M$. There exists an open, dense, full-measure subset U of $\mathbb{C}^{n+M}$ such that for $(b_1^*,\ldots,b_n^*,c_1^*,\ldots,c_M^*)\in U$, the following holds.

(a) The set X^* of solutions $x=(x_1,\ldots,x_n)$ of

$$\begin{aligned} q_1(x_1,\ldots,x_n) &= p_1(c_1^*,\ldots,c_M^*,x_1,\ldots,x_n)+b_1^*=0, \\ &\vdots \\ q_n(x_1,\ldots,x_n) &= p_n(c_1^*,\ldots,c_M^*,x_1,\ldots,x_n)+b_n^*=0, \end{aligned} \tag{3.3}$$

consists of d_0 isolated points, for some $d_0\le d$.

(b) Properties 1 and 2 (smoothness and accessibility) hold for the homotopy

$$H(x,t) = P((1-t)c_1^* + tc_1, \ldots, (1-t)c_M^* + tc_M, x_1, \ldots, x_n) + (1-t)b^* \quad (3.4)$$

where $b^* = (b_1^*, \ldots, b_n^*)$. It follows that every solution of $P(x) = 0$ is reached by a path beginning at a point of X^*.

A proof of Theorem 3.1 can be found in Li et al. (1989). The theorem is used as part of the following procedure. Let $P(c,x) = 0$ as in (3.2) denote the system to be solved for various values of the coefficients c.

Cheater's homotopy procedure

(1) Choose complex numbers $(b_1^*, \ldots, b_n^*, c_1^*, \ldots, c_M^*)$ at random, and use the classical homotopy continuation method to solve $Q(x) = 0$ in (3.3). Let d_0 denote the number of solutions found (this number is bounded above by the total degree d). Let X^* denote the set of d_0 solutions.

(2) For each new choice of coefficients $c = (c_1, \ldots, c_M)$, follow the d_0 paths defined by $H(x,t) = 0$ in (3.4), beginning at the points in X^*, to find all solutions of $P(c,x) = 0$.

In step (1) above, for random complex numbers $(c_1^*, \ldots, c_M^*)$, using classical homotopy continuation methods to solve $Q(x) = 0$ in (3.3) may itself sometimes be computationally expensive. It is desirable that those numbers do not have to be random. For illustration, consider the linear system

$$\begin{aligned} c_{11}x_1 + \cdots + c_{1n}x_n &= b_1, \\ &\vdots \\ c_{n1}x_1 + \cdots + c_{nn}x_n &= b_n, \end{aligned} \quad (3.5)$$

which may be considered as a polynomial system with each equation having degree one. For generic c_{ij}s, (3.5) has a unique solution which is not available right away. However, if we choose $c_{ij} = \delta_{ij}$ (the Kronecker delta), the solution is obvious.

For this purpose, an alternative is suggested in Li and Wang (1992). When a system $P(c,x) = 0$ with a particular parameter c^0 is solved, this c^0 may be chosen arbitrarily instead of being chosen randomly; then for any parameter $c \in \mathbb{C}^M$ consider the nonlinear homotopy

$$H(a,x,t) = P((1-[t-t(1-t)a])c^0 + (t-t(1-t)a)c, x) = 0. \quad (3.6)$$

It is shown in Li and Wang (1992) that for a randomly chosen complex a the solution paths of (3.6) emanating from the solutions of $P(c^0,x) = 0$ will reach the isolated solutions of $P(c,x) = 0$ under the natural assumption that, for generic c, $P(c,x)$ has the same number of isolated zeros in $\mathbb{C}^n$.

The most important advantage of the homotopy in (3.6) is that the parameter c^0 of the start system $P(c^0, x) = 0$ is arbitrary so long as $P(c^0, x) = 0$ has the same number of solutions as $P(c, x) = 0$ for generic c. Therefore, in some situations, when the solutions of $P(c, x) = 0$ are easily available for certain c^0, the system $P(c^0, x) = 0$ may be used as the start system (3.6) and the extra effort of solving $P(c, x) = 0$ for a randomly chosen c would be saved.

To finish, we give a more non-trivial example of the procedure described in this section.

Consider the indirect position problem for revolute-joint kinematic manipulators. Each joint is associated with a one-dimensional parametrization, namely the angular position of the joint. If all angular positions are known, then of course the position and orientation of the end of the manipulator (the hand) are determined. The indirect position problem is the inverse problem: given the desired position and orientation of the hand, find a set of angular parameters for the (controllable) joints which will place the hand in the desired state.

The indirect position problem for six joints is reduced to a system of eight nonlinear equations in eight unknowns in Tsai and Morgan (1985). The coefficients of the equations depend on the desired position and orientation, and a solution of the system (an eight-vector) represents the sines and cosines of the angular parameters. Whenever the manipulator's position is changed, the system needs to be resolved with new coefficients. The equations are too long to repeat here; see the appendix of Tsai and Morgan (1985). Suffice it to say that it is a system of eight degree-two polynomial equations in eight unknowns which is rather deficient. The total degree of the system is $2^8 = 256$, but there are at most 32 isolated solutions.

The nonlinear homotopy of (3.6) requires only 32 paths to solve the system with different sets of parameters (Li and Wang 1990, 1992). The system contains 26 coefficients, and a specific set of coefficients is chosen for which the system has 32 solutions. For subsequent solving of the system, for any choice of the coefficients $c_1, \ldots, c_{26}$, all solutions can be found at the end of exactly 32 paths, by using nonlinear homotopy in (3.6) with randomly chosen complex a.

4. Polyhedral homotopy

In the last few years, a major computational breakthrough has occurred in the solution of polynomial systems by the homotopy continuation method. The new method takes great advantage of the Bernshteĭn theory, which gives a much tighter bound, in general, for the number of isolated zeros of a polynomial system in the algebraic tori $(\mathbb{C}^*)^n$, where $\mathbb{C}^* = \mathbb{C}\backslash\{0\}$. In Huber and Sturmfels (1995), this root count was used to actually find

all the isolated zeros of the polynomial system by establishing polyhedral homotopies. For a given polynomial system, the new method solves a new polynomial system with the same monomials, but with randomly chosen coefficients. The new system is then used as the start system in the cheater's homotopy described in Section 3 to solve the original polynomial system. In a way, the new method uses the method for Problem B to solve Problem A. The new algorithm is very promising. In particular, for polynomial systems without special structure, the new algorithm substantially outperformed other methods.

We take the following example (Huber and Sturmfels 1995) as our point of departure. Setting $x = (x_1, x_2)$, consider the system $P(x) = (p_1(x), p_2(x))$, where

$$\begin{aligned} p_1 = \quad & c_{11}x_1x_2 + c_{12}x_1 + c_{13}x_2 + c_{14} \quad = 0, \quad \text{and} \\ p_2 = \quad & c_{21}x_1x_2^2 + c_{22}x_1^2x_2 + c_{23} \quad = 0. \end{aligned} \tag{4.1}$$

Here, $c_{ij} \in \mathbb{C}^* = \mathbb{C}\backslash\{0\}$. The monomials $\{1, x_1x_2, x_1, x_2\}$ in p_1 can be written as $x_1x_2 = x_1^1x_2^1$, $x_1 = x_1^1x_2^0$, $x_2 = x_1^0x_2^1$ and $1 = x_1^0x_2^0$. The set of their exponents

$$S_1 = \{(0,0), (1,0), (1,1), (0,1)\}$$

is called the *support* of p_1, and its convex hull $Q_1 = \text{conv}(S_1)$ is called the *Newton polytope* of p_1. Similarly, p_2 has support $S_2 = \{(0,0), (2,1), (1,2)\}$ with Newton polytope $Q_2 = \text{conv}(S_2)$. Using multi-index notation $x^q = x_1^{q_1}x_2^{q_2}$ where $q = (q_1, q_2)$, we may rewrite (4.1) as

$$p_1(x) = \sum_{q \in S_1} c_q x^q \quad \text{and} \quad p_2(x) = \sum_{q \in S_2} c_q x^q.$$

The *Minkowski sum* $R_1 + R_2$ of polytopes R_1 and R_2 is defined as

$$R_1 + R_2 = \{r_1 + r_2 : r_1 \in R_1 \text{ and } r_2 \in R_2\}$$

(polytopes Q_1, Q_2 and $Q_1 + Q_2$ for (4.1) are shown in Figure 2). Now, let us consider the area of the convex polytope $\lambda_1Q_1 + \lambda_2Q_2$ with non-negative variables λ_1 and λ_2 for the system (4.1). From elementary geometry, the area of a triangle on the plane with vertices u, v and w equals

$$\frac{1}{2}\left|\det\begin{pmatrix} u - v \\ w - v \end{pmatrix}\right|. \tag{4.2}$$

Thus, to compute the area $f(\lambda_1, \lambda_2)$ of $\lambda_1Q_1 + \lambda_2Q_2$, one may partition the polytope into a collection of triangles, $A_1, A_2, ..., A_k$. These triangles are mutually disjoint, and the vertices take the form $\lambda_1q_1 + \lambda_2q_2$, with $q_1 \in Q_1$ and $q_2 \in Q_2$. In other words, the vertices of these triangles coincide with the vertices of the polytope $\lambda_1Q_1 + \lambda_2Q_2$. It follows from (4.2) that the area of each A_i is a second-degree homogeneous polynomial in λ_1 and λ_2.

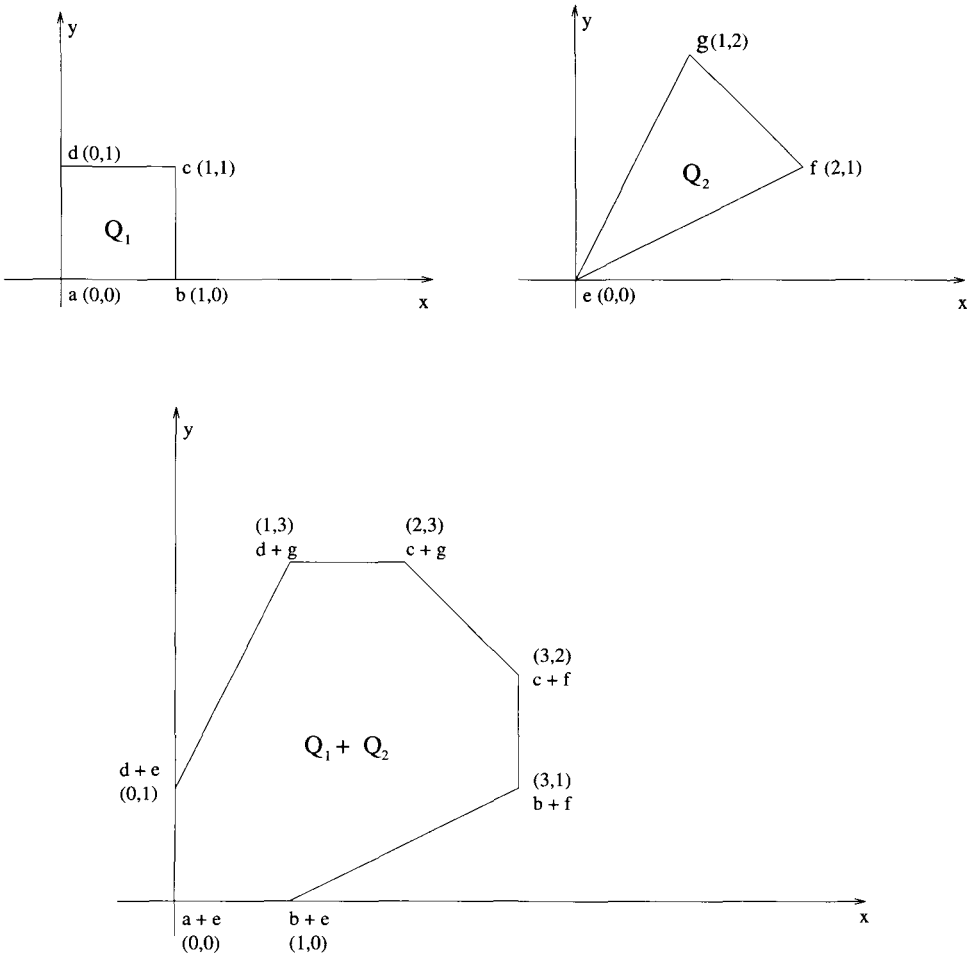

Fig. 2.

Therefore, $f(\lambda_1, \lambda_2)$, as a sum of the areas of $A_1, ..., A_k$, is also a second-degree homogeneous polynomial in λ_1 and λ_2. Writing

$$f(\lambda_1, \lambda_2) = a_1\lambda_1^2 + a_2\lambda_2^2 + a_{12}\lambda_1\lambda_2,$$

the coefficient a_{12} of $\lambda_1\lambda_2$ in f is called the *mixed volume* of the polytopes Q_1 and Q_2. We denote it by $\mathcal{M}(Q_1, Q_2)$, or $\mathcal{M}(S_1, S_2)$ when no ambiguity exists.

Clearly,

$$\begin{aligned} a_{12} &= f(1,1) - f(1,0) - f(0,1) \\ &= \text{area of } (Q_1 + Q_2) - \text{ area of } (Q_1) - \text{ area of } (Q_2). \end{aligned}$$

For (4.1), it is easy to see that the areas of $Q_1 + Q_2$, Q_1 and Q_2 are 6.5, 1 and 3.5 respectively. Therefore, $a_{12} = 6.5 - 1 - 1.5 = 4$. On the other hand, one can also easily see that system (4.1) has two zeros $(0, 0, 1)$ and $(0, 1, 0)$ at infinity in $\mathbb{P}^2$; hence it can have at most 4 isolated zeros in $\mathbb{C}^2$, or in $(\mathbb{C}^*)^2$ in particular. According to the Bernshteín theory, this is not a coincidence: the number of isolated zeros of (4.1) in $(\mathbb{C}^*)^2$, counting multiplicities, is bounded above by the mixed volume of its Newton polytopes. Further, when the coefficients in (4.1) are chosen generically, then these two numbers are exactly the same.

To state the Bernshteín theory in a more general form, we first allow monomials $x_1^{a_1} \cdots x_n^{a_n}$ to have negative exponents; such a polynomial is called a *Laurent polynomial*. With $x = (x_1, \ldots, x_n)$, let $p(x) = (p_1(x), \ldots, p_n(x))$ be a system of n Laurent polynomials with supports $S_1, \ldots, S_n$ respectively in $\mathbb{Z}^n$. The corresponding Newton polytopes are $Q_1, \ldots, Q_n$. Following reasoning similar to that described above, the n-dimensional volume of the polytope $\lambda_1 Q_1 + \cdots + \lambda_n Q_n$, with non-negative variables $\lambda_1, \ldots, \lambda_n$, is a homogeneous polynomial in $\lambda_1, \ldots, \lambda_n$ of degree n. The coefficient of $\lambda_1 \times \lambda_2 \times \cdots \times \lambda_n$ in this polynomial is defined as the *mixed volume* of $Q_1, \ldots, Q_n$, denoted by $\mathcal{M}(Q_1, \ldots, Q_n)$ or $\mathcal{M}(S_1, \ldots, S_n)$.

Theorem 4.1 (Bernshteín 1975) The number of isolated zeros, counting multiplicities, of $P(x) = (p_1(x), \ldots, p_n(x))$ in $(\mathbb{C}^*)^n$ is bounded above by the mixed volume $\mathcal{M}(S_1, \ldots, S_n)$. For generically chosen coefficients, the system $P(x) = 0$ has exactly $\mathcal{M}(S_1, \ldots, S_n)$ roots in $(\mathbb{C}^*)^n$.

In Canny and Rojas (1991), this bound was nicknamed the *BKK bound* after its inventors, Bernshteín (1975), Khovanskiĭ (1978) and Kushnirenko (1976). It turns out that this root count is very helpful in using the polyhedral homotopy to solve sparse polynomial systems, sparse in the sense that each polynomial in the system contains few terms. This sparseness is by no means a big restriction. After all, almost all the polynomial systems we encountered in application belong to this category.

An apparent limitation of the above theorem is that it counts only the roots of a polynomial system in $(\mathbb{C}^*)^n$, but not necessarily all roots in affine space $\mathbb{C}^n$. This problem was first attempted in Canny and Rojas (1991) and Rojas (1994) by introducing the notion of the *shadowed* sets, and a bound in $\mathbb{C}^n$ was obtained. Later, a significantly tighter bound was discovered in the following theorem.

Theorem 4.2 (Li and Wang 1996) The number of isolated zeros in $\mathbb{C}^n$, counting multiplicities, of a polynomial system $P(x) = (p_1(x), \ldots, p_n(x))$ with supports $S_1, \ldots, S_n$ is bounded above by the mixed volume

$$\mathcal{M}(S_1 \bigcup \{0\}, \ldots, S_n \bigcup \{0\}).$$

This theorem was further extended in several ways by Huber and Sturmfels (1997) and Rojas and Wang (1996). When $0 \in S_i$ for all $i = 1, \ldots, n$, so that each p_i has a nontrivial constant term, then Theorem 4.2 implies that the BKK bound of Theorem 4.1 gives the number of zeros of the polynomial system in $\mathbb{C}^n$. In fact, the proof of Theorem 4.2 uses the important fact that generic constant perturbations of a polynomial system can only have isolated zeros in $(\mathbb{C}^*)^n$, and all isolated zeros become nonsingular.

Now consider the system (4.1) again. To compute the area of $Q_1 + Q_2$, we can certainly subdivide $Q_1 + Q_2$ as we wish. The subdivision may not consist of all triangles as before. However, the subdivision shown in Figure 3 – call it subdivision B – is of particular interest. By a *cell* of a subdivision we mean any member of the subdivision. It can be easily verified that all the cells in subdivision B have the following special properties.

Proposition 4.1

(a) Each one is a Minkowski sum of the convex hull of a subset C_1 in S_1 and the convex hull of a subset C_2 in S_2.
(b) For $i = 1, 2, \text{conv}(C_i)$ is a simplex of dimension $\#(C_i) - 1$, where $\#(C_i)$ is the number of points in C_i.
(c) Simplices $\text{conv}(C_1)$ and $\text{conv}(C_2)$ are complementary to each other in the sense that $\dim(\text{conv}(C_1)) + \dim(\text{conv}(C_2)) = \dim(\text{conv}(C_1) + \text{conv}(C_2))$.

In light of properties (a) and (b), each cell $C = \text{conv}(C_1) + \text{conv}(C_2)$ in B can be identified as a cell of type (l_1, l_2), where $l_1 = \dim(\text{conv}(C_1))$ and $l_2 = \dim(\text{conv}(C_2))$. Property (c) mainly says that simplices $\text{conv}(C_1)$ and $\text{conv}(C_2)$ are 'linearly independent', for otherwise their Minkowski sum would be lower dimensional.

In $\mathbb{R}^n$, consider the n-dimensional volume of the Minkowski sum of simplices $A_1, \ldots, A_n$ with dimensions $k_1, \ldots, k_n$, respectively, where $k_i \geq 0$ for $1 \leq i \leq n$ and $k_1 + k_2 + \cdots + k_n = n$. For $i = 1, \ldots, n$, let $A_i = \text{conv}\{q_0^{(i)}, \ldots, q_{k_i}^{(i)}\}$ and let V be the $n \times n$ matrix whose rows are $q_j^{(i)} - q_0^{(i)}$ for $1 \leq i \leq n$ and $1 \leq j \leq k_i$. Notice that any 0-dimensional simplex consists of only one point, and therefore contributes no rows to V. It can be shown that

$$\text{Vol}_n(A_1 + \cdots + A_n) = \frac{1}{k_1! \cdots k_n!} |\det V|. \tag{4.3}$$

Here, we use Vol_n to denote the n-dimensional volume; of course, $\text{Vol}_2(C)$ represents the area of C. Applying (4.3) to cell ①$= \text{conv}\{a, d\} + \text{conv}\{e, q\}$ in subdivision B, we have

$$\text{Vol}_2(\text{cell ①}) = \left| \det \begin{pmatrix} d - a \\ g - e \end{pmatrix} \right|.$$

Now, when Q_1 and Q_2 are scaled by λ_1 and λ_2, respectively, cell ① becomes $\text{conv}\{\lambda_1 a, \lambda_2 d\} + \text{conv}\{\lambda_2 e, \lambda_2 g\}$ and its volume becomes

$$\begin{aligned}\left|\det\begin{pmatrix}\lambda_1 d - \lambda_1 a\\ \lambda_2 g - \lambda_2 e\end{pmatrix}\right| &= \left|\det\begin{pmatrix}d-a\\ g-e\end{pmatrix}\right| \times \lambda_1\lambda_2\\ &= (\text{volume of cell ① before scaling}) \times \lambda_1\lambda_2.\end{aligned}$$

From the definition of the mixed volume, it follows that the volume of the original cell ① constitutes part of the mixed volume of Q_1 and Q_2. On the other hand, after scaling, cell ② in subdivision B becomes $\text{conv}\{\lambda_1 a, \lambda_1 c, \lambda_1 d\} + \{\lambda_2 g\}$ and its volume becomes, according to (4.3),

$$\begin{aligned}\frac{1}{2}\left|\det\begin{pmatrix}\lambda_1 c - \lambda_1 a\\ \lambda_1 d - \lambda_a a\end{pmatrix}\right| &= \frac{1}{2}\left|\det\begin{pmatrix}c-a\\ d-a\end{pmatrix}\right| \times \lambda_1^2\\ &= (\text{volume of cell ② before scaling}) \times \lambda_1^2.\end{aligned}$$

Apparently, the volume of the original cell ② has no contribution to the mixed volume of Q_1 and Q_2.

In summary, only cells of type (1,1) contribute to the mixed volume $\mathcal{M}(Q_1, Q_2)$ of Q_1 and Q_2 and, therefore,

$$\begin{aligned}\mathcal{M}(Q_1, Q_2) &= \text{the sum of the volumes of cells of type (1,1)}\\ &= \text{volume of cell ① + volume of cell ③ + volume of cell ⑤}\\ &= 1 + 2 + 1 = 4.\end{aligned}$$

The type of subdivisions of $Q_1 + Q_2$ that share the same special properties in Proposition 4.1 as subdivision B is called the *fine mixed subdivision*. To state a formal definition with less notation, we omit '+' and 'conv', except where absolutely necessary. For instance, $(S_1, \ldots, S_n)$ will replace $Q_1 + \cdots + Q_n (= \text{conv}(S_1) + \cdots + \text{conv}(S_n))$ as the key object.

Let $S = (S_1, \ldots, S_n)$ be a sequence of finite subsets of $\mathbb{Z}^n$, whose union affinely spans $\mathbb{R}^n$. By a *cell* of S we mean an n-tuple $C = (C_1, \ldots, C_n)$ of subsets $C_i \subset S_i$, for $i = 1, \ldots, n$. Define

$$\begin{aligned}\text{type}(C) &:= (\dim(\text{conv}(C_1)), \ldots, \dim(\text{conv}(C_n))),\\ \text{conv}(C) &:= \text{conv}(C_1) + \cdots + \text{conv}(C_n),\end{aligned}$$

and $\text{Vol}(C) := \text{Vol}(\text{conv}(C))$. A *face* of C is a subcell $F = (F_1, \ldots, F_n)$ of C where $F_i \subset C_i$ and some linear functional $\alpha \in (\mathbb{R}^n)^\vee$ attains its minimum over C_i at F_i, for $i = 1, \ldots, n$. We call such an α an *inner normal* of F. If F is a face of C then $\text{conv}(F_i)$ is a face of the polytope $\text{conv}(C_i)$ for $i = 1, \ldots, n$.

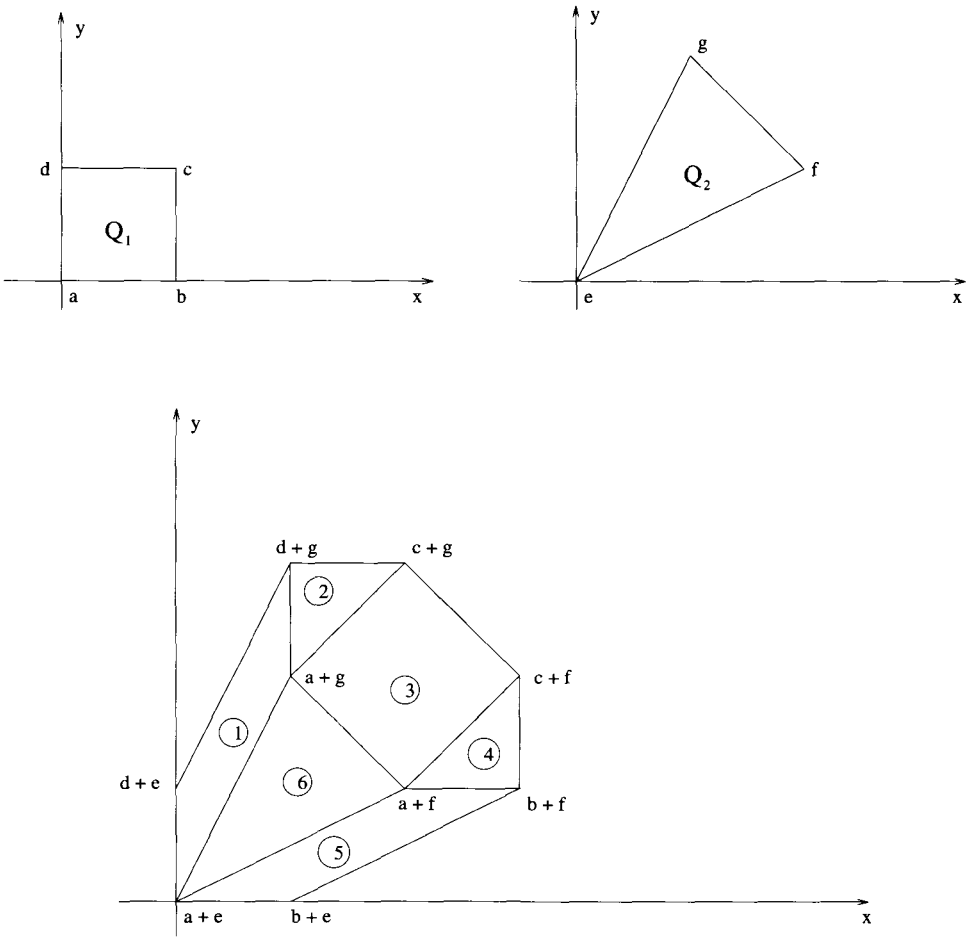

Fig. 3. Subdivision B for $Q_1 + Q_2$

Definition 4.1 A *fine mixed subdivision* of S is a set $\{C^{(1)}, \ldots, C^{(m)}\}$ of cells such that:

(a) for all $j = 1, \ldots, m$, $\dim(\text{conv}(C^{(j)})) = n$

(b) $\text{conv}(C^{(j)}) \cap \text{conv}(C^{(k)})$ is a proper common face of $\text{conv}(C^{(j)})$ and $\text{conv}(C^{(k)})$ when it is nonempty for $j \neq k$

(c) $\bigcup_{j=1}^{m} \text{conv}(C^{(j)}) = \text{conv}(S)$

(d) for $j = 1, \ldots, m$, write $C^{(j)} = (C_1^{(j)}, \ldots, C_n^{(j)})$. Then, each $\text{conv}(C_i^{(j)})$ is a simplex of dimension $\#C_i^{(j)} - 1$, and for each j,

$$\dim(\text{conv}(C_1^{(j)})) + \cdots + \dim(\text{conv}(C_n^{(j)})) = n.$$

As we have discussed for the special system (4.1), when a polynomial

system $P(x) = (p_1(x), \ldots, p_n(x))$ in $\mathbb{C}[x_1, \ldots, x_n]$ is given with support $S = (S_1, ..., S_n)$, where S_i is the support of p_i, and if we can find a fine mixed subdivision for S, then the mixed volume $\mathcal{M}(S_1, \ldots, S_n)$ will be the sum of the volumes of cells of type (1,...,1). Thus formula (4.3), together with condition (d) above, makes the volume computation of this type of cell quite easy.

A fine mixed subdivision for $S = (S_1, \ldots, S_n)$ can be found by the following standard process: choose real-valued functions $\omega^{(i)} : S_i \to \mathbb{R}$, for $i = 1, \ldots, n$; call the n-tuple $\omega = (\omega^{(i)}, \ldots, \omega^{(i)})$ a *lifting function* on S, and say that ω lifts S_i to its graph $\hat{S}_i = \{(q, \omega^{(i)}(q)) : q \in S_i\} \subset \mathbb{R}^{n+1}$. This notation is extended in the obvious way: $\hat{S} = (\hat{S}_1, \ldots, \hat{S}_n)$, $\hat{Q}_i = \text{conv}(\hat{S}_i)$, $\hat{Q} = \hat{Q}_1 + \cdots + \hat{Q}_n$, *etc.* Let S_ω be the set of cells $\{C\}$ of S which satisfy

(a) $\dim(\text{conv}(\hat{C})) = n$,
(b) $\hat{C}$ is a facet (an n-dimensional face) of $\hat{S}$ whose inner normal $\alpha \in (\mathbb{R}^{n+1})^\vee$ has positive last coordinate.

In other words, $\text{conv}(\hat{C})$ is a facet of the lower hull of $\hat{Q}$. The fact is that when the lifting function ω is chosen generically, S_ω always gives a fine mixed subdivision for S (Gel'fand, Kapranov and Zelevinskiĭ 1994, Lee 1991). The subdivision B in Figure 3 for system (4.1) is, in fact, induced by the lifting $\omega = ((0, 1, 1, 1), (0, 0, 0))$, that is,

$$\hat{S} = (\{(a, 0), (b, 1), (c, 1), (d, 1)\}, \{(e, 0), (f, 0), (g, 0)\})$$

(see Figure 4). While this lifting does not seem so generic, it is sufficient to give a fine mixed subdivision.

Let us return to our main issue: how can this Bernshteĭn theory help us to solve polynomial systems by homotopy continuation methods? Actually, the lifting function introduced above has already provided a nonlinear homotopy. This ingenious idea is due to Huber and Sturmfels (1995).

For a given polynomial system $P(x) = (p_1(x), \ldots, p_n(x))$ in $\mathbb{C}[x_1, \ldots, x_n]$, to find all isolated zeros of $P(x)$ in $\mathbb{C}^n$ instead of $(\mathbb{C}^*)^n$, we first, according to Theorem 4.2, augment the monomial $x^0 (= 1)$ to those p_is which do not have constant terms. We then choose the coefficients of all the monomials in $P(x)$ at random. For simplicity, we abuse notation and retain the name $P(x) = (p_1(x), \ldots, p_n(x))$ for this system. We wish to solve this system first, and then, by using the cheater's homotopy introduced in Section 3, it can be used as the start system for solution of the original system by linear homotopy.

Let S_i be the support of p_i, so that

$$p_i(x) = \sum_{q \in S_i} c_q x^q, \qquad i = 1, \ldots, n,$$

where $q = (q_1, \ldots, q_n)$ and $x^q = x_1^{q_1} \cdots x_n^{q_n}$. Let t denote a new complex

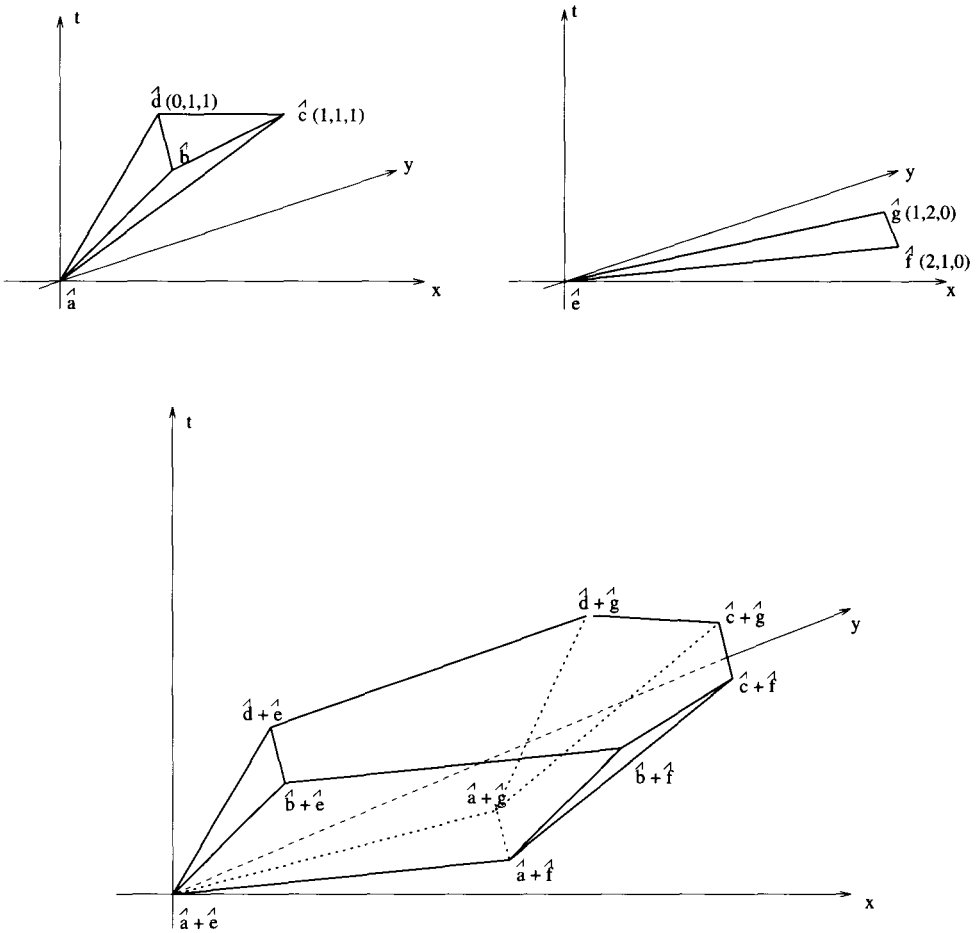

Fig. 4.

variable and consider the polynomials in $n+1$ variables given by

$$\hat{p}_i(x,t) = \sum_{q \in S_i} c_q x^q t^{\omega_i(q)}, \qquad i = 1, \ldots, n, \tag{4.4}$$

where each $\omega_i : S_i \to \mathbb{R}$ for $i = 1, \ldots, n$ is chosen generically. The support of $\hat{p}_i$ is now $\hat{S}_i = \{\hat{q} = (q, \omega_i(q)) : q \in S_i\}$ with Newton polytope $\hat{Q}_i = \text{conv}(\hat{S}_i)$. The function $\omega = (\omega_1, \ldots, \omega_n)$ can be viewed as a lifting function on $S = (S_1, \ldots, S_n)$ which lifts S_i to $\hat{S}_i$. The induced subdivision S_ω on S is then a fine mixed subdivision and the mixed volume $\mathcal{M}(S_1, \ldots, S_n)$ equals the sum of the volumes of cells of type (1,...,1) in S_ω. Recall that, for each $t \in (0, 1]$, the isolated zeros of the system

$$\hat{P}(x,t) = (\hat{p}_1(x,t), \ldots, \hat{p}_n(x,t))$$

are all nonsingular and, by the Bernshteĭn theory, the total number of those zeros is equal to $\mathcal{M}(S_1,\ldots,S_n)$. We may write these zeros as $x^1(t),\ldots,x^k(t)$ where $k = \mathcal{M}(S_1,\ldots,S_n)$, so $\hat{P}(x^j(t),t) = 0$ for each $t \in (0,1]$ and $j = 1,\ldots,k$.

Let $C = (C_1,\ldots,C_n)$ be a cell of type (1,...,1) in S_ω. For $i = 1,\ldots,n$, let $C_i = \{q_i^{(0)}, q_i^{(1)}\} \subset S_i$ and $v_i = q_i^{(1)} - q_i^{(0)}$. Since S_ω is a fine mixed subdivision, $\{v_1,\ldots,v_n\}$ is linearly independent; otherwise, $\dim(\mathrm{conv}(C_1))+\cdots+\dim(\mathrm{conv}(C_n)) < n$. So,

$$\mathrm{Vol}_n(C) = \left|\det\begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}\right|.$$

On the other hand, $\hat{C} = (\hat{C}_1,\ldots,\hat{C}_n)$ is a facet of $\hat{S} = (\hat{S}_1,\ldots,\hat{S}_n)$ whose inner normal $\hat{\alpha} \in (\mathbb{R}^{n+1})^\vee$ has positive last coordinate. Let $\hat{\alpha} = (\alpha_1,\ldots,\alpha_n,1)$ and $\alpha = (\alpha_1,\ldots,\alpha_n)$, so $\hat{\alpha} = (\alpha,1)$. Let $x(t)$ represent general solution curves $x^1(t),\ldots,x^k(t)$ of $\hat{P}(x,t) = 0$. Setting $x(t) = (x_1(t),\ldots,x_n(t))$, let

$$\begin{aligned} t^{\alpha_1} y_1(t) &= x_1(t), \\ &\vdots \\ t^{\alpha_n} y_n(t) &= x_n(t). \end{aligned}$$

Or, simply, $t^\alpha y(t) = x(t)$. Substituting this into (4.4) yields

$$\begin{aligned} \hat{p}_i &= \sum_{q\in S_i} c_q y^q t^{\alpha q} t^{\omega_i(q)} \\ &= \sum_{q\in S_i} c_q y^q t^{\langle(\alpha,1),(q,\omega_i(q))\rangle} \\ &= \sum_{q\in S_i} c_q y^q t^{\langle\hat{\alpha},\hat{q}\rangle}, \qquad i = 1,\ldots,n. \end{aligned} \tag{4.5}$$

Let $\beta_i = \min_{q\in S_i}\langle\hat{\alpha},\hat{q}\rangle$. Since $\hat{C}$ is a facet of $\hat{S}$, $\hat{C}_i = \{\hat{q}_i^{(0)}, \hat{q}_i^{(1)}\}$ is a face of $\hat{S}_i$ and $\hat{\alpha} = (\alpha,1)$ also serves as an inner normal of $\hat{C}_i$. It follows that $\langle\hat{\alpha},\hat{q}_i^{(0)}\rangle = \langle\hat{\alpha},\hat{q}_i^{(1)}\rangle = \beta_i$ and $\langle\hat{\alpha},\hat{q}\rangle > \beta_i$ for $\hat{q} \in \hat{S}_i\backslash\hat{C}_i$. Hence, factoring out t^{β_i} in (4.5), we have

$$\hat{p}_i = t^{\beta_i}(c_{i0}y^{q_i^{(0)}} + c_{i1}y^{q_i^{(1)}} + R_i(y,t)), \qquad i = 1,\ldots,n,$$

where $c_{i0} = c_{q_i^{(0)}}, c_{i1} = c_{q_i^{(1)}}$ and

$$R_i(y,t) = \sum_{q\in S_i\backslash C_i} c_q y^q t^{\langle\hat{\alpha},\hat{q}\rangle-\beta_i}.$$

Evidently, $R_i(y,0) = 0$ for each i, since $\langle\hat{\alpha},\hat{q}\rangle - \beta_i > 0$ for $q \in S_i\backslash C_i$. Now,

consider the homotopy $H(y,t) = (h_1(y,t), \ldots, h_n(y,t)) = 0$ where

$$h_i(y,t) = c_{i0}y^{q_i^{(0)}} + c_{i1}y^{q_i^{(1)}} + R_i(y,t), \qquad i = 1, \ldots, n. \tag{4.6}$$

The solutions $(y(t),t)$ of this homotopy satisfy

$$c_{j0}y^{q_j^{(0)}} + c_{j1}y^{q_j^{(1)}} = 0, \qquad j = 1, \ldots, n, \tag{4.7}$$

at $t = 0$. For $t \neq 0$, they agree with the zeros of (4.5) and, since $t^\alpha y(t) = x(t)$ for $y(t)$ in (4.5), they also agree with the zeros of (4.4) at $t = 1$. In other words, $y(1)$ of (4.6) are solutions of $P(x) = 0$. So, by following the solution curves $(y(t),t)$ of the homotopy $H(y,t) = 0$ defined by (4.6), we may reach the solutions of $P(x) = 0$, at $t = 1$. Of course, we need to solve the system (4.7) at $t = 0$ to begin with. It can be shown that for randomly chosen c_{ij}, for $i = 1, \ldots, n$ and $j = 0, 1$, system (4.7) has

$$\left| \det \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \right| = \text{ the volume of } C$$

solutions in $(\mathbb{C}^*)^n$; recall that $v_i = q_i^{(1)} - q_i^{(0)}$ for $i = 1, \ldots, n$. To see how to solve (4.7) in $(\mathbb{C}^*)^n$, we rewrite (4.7) as

$$\begin{aligned} y^{v_1} &= b_1, \\ &\vdots \\ y^{v_n} &= b_n, \end{aligned} \tag{4.8}$$

where $b_1 \cdot b_2 \cdot \ldots \cdot b_n \neq 0$, and let

$$V = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}.$$

For brevity, write $y^V = (y^{v_1}, \ldots, y^{v_n})$ and $b = (b_1, \ldots, b_n)$. Then (4.8) becomes

$$y^V = b. \tag{4.9}$$

With this notation, one can easily check that for an $n \times n$ integer matrix U, the following holds:

$$(y^U)^V = y^{(VU)}.$$

When V is a lower nonsingular triangular integer matrix

$$V = \begin{bmatrix} v_{11} & & & \\ v_{21} & v_{22} & & 0 \\ \vdots & \vdots & \ddots & \\ v_{n1} & v_{n2} & \cdots & v_{nn} \end{bmatrix},$$

(4.8) becomes

$$\begin{aligned} y_1^{v_{11}} &= b_1, \\ y_1^{v_{21}} y_2^{v_{22}} &= b_2, \\ &\vdots \\ y_1^{v_{n1}} y_2^{v_{n2}} \cdots y_n^{v_{nn}} &= b_n. \end{aligned} \tag{4.10}$$

Obviously, by forward substitution, (4.10) has $|v_{11}| \times \cdots \times |v_{nn}| = |\det V|$ solutions. In general, we may lower triangularize V by multiplying on the right by an integer matrix U with $|\det U| = 1$, which can be found by the following procedure. Firstly, the greatest common divisor d of two integers a and b is

$$d = \gcd(a, b) = ka + lb, \qquad \text{for certain } k, l \in Z.$$

Let

$$M = \begin{bmatrix} k & l \\ -\frac{b}{d} & \frac{a}{d} \end{bmatrix}.$$

Then $\det(M) = 1$ and

$$M \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} k & l \\ -\frac{b}{d} & \frac{a}{d} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} d \\ 0 \end{bmatrix}.$$

In view of this, a series of $n \times n$ matrices like M can be used to produce zeros in matrices in a similar way to the use of Givens rotations for the QR factorization. For instance, if a and b are the ith and jth components of an n-dimensional vector v, that is,

$$v = \begin{bmatrix} \vdots \\ a \\ \vdots \\ b \\ \vdots \end{bmatrix} \begin{matrix} \\ \rightarrow i\text{th} \\ \\ \rightarrow j\text{th}, \\ \end{matrix}$$

then we set

$$U(i,j,a,b) = \begin{array}{c} \begin{matrix} & & & i & & & j & & \\ 1 & & & & & & & & \\ & 1 & & & & & & & \\ & & \ddots & & & & & & \\ & & & k & & & l & & \\ & & & & 1 & & & & \\ & & & & & \ddots & & & \\ & & & & & & 1 & & \\ & & & -\frac{b}{d} & & & \frac{a}{d} & & \\ & & & & & & & 1 & \\ & & & & & & & & \ddots \\ & & & & & & & & & 1 \end{matrix} \end{array} \begin{matrix} \\ \\ \\ \\ i \\ \\ \\ \\ j \\ \\ \\ \\ \\ \end{matrix},$$

so that $\det(U(i,j,a,b)) = 1$ and the jth component of $U(i,j,a,b)v$ will vanish. Thus, a product of a series of matrices in the form of $U(i,j,a,b)$ can be chosen to upper triangularize a matrix from the left. To lower triangularize V, let U be an integer matrix with $|\det U| = 1$ such that $U^T V^T$ is upper triangular; hence, VU is lower triangular.

Now, let $z^U = y$ and substitute it into (4.9); we have

$$y^V = (z^U)^V = z^{VU} = b. \tag{4.11}$$

Since VU is lower triangular, $z = (z_1, \ldots, z_n)$ in (4.11) can be solved and the number of solutions is equal to $|\det(VU)| = |\det(V)| \cdot |\det(U)| = |\det(V)|$. Consequently, we have as many solutions of $y = (y_1, \ldots, y_n)$ in (4.9).

In summary, to find all the isolated zeros of a polynomial system $P(x) = (p_1(x), \ldots, p_n(x))$ in $\mathbb{C}[x_1, \ldots, x_n]$, we augment x^0 to those p_is without constant terms first, then equip all the monomials in $P(x)$ with generic coefficients. In the same notation, we construct $\hat{P}(x,t) = (\hat{p}_1(x,t), \ldots, \hat{p}_n(x,t))$, where

$$\hat{p}_i(x,t) = \sum_{q \in S_i} c_q x^q t^{\omega_i(q)}, \qquad i = 1, \ldots, n.$$

Here S_i is the support of p_i and the lifting function $\omega = (\omega_1, \ldots, \omega_n)$ is chosen at random. Then each cell C of type (1,...,1) in the induced fine mixed subdivision S_ω provides a set of k starting points for the homotopy $H(y,t) = 0$ defined by (4.6), where k denotes the volume of C. Following the solution curves of this homotopy with those k starting points from 0 to 1, we reach k of the solutions of $P(x) = 0$. By Bernshteĭn's theory, the total number of isolated zeros of $P(x)$ equals the sum of the volumes of all cells of this type. We are thus able to find all the isolated zeros of $P(x)$, and this modified system can then be used as a start system of the linear homotopy to find all the isolated zeros of the original system.

What seems to be missing in the process described above is a constructive way of finding cells of type (1,...,1) in the induced fine mixed subdivision S_ω corresponding to the lifting ω. This issue was discussed in Emiris (1994), Verschelde (1996) and Verschelde, Gatermann and Cools (1996), papers which provided different ways to deal with this problem. At present, the most efficient technique for finding those cells is still undetermined.

The algorithm has been implemented with remarkable success. Recall that the Cassou–Nogues system in (1.4) has total degree 1344 and optimal m-homogeneous Bézout number 368. This system has 16 isolated zeros and its mixed volume equals 24. So, by using polyhedral nonlinear homotopies, one need only follow 24 paths to reach all isolated zeros of the system.

Originally, a more general version of the above process was presented in Huber and Sturmfels (1995). If some of the p_is have the same supports, then cells of the 'appropriate' types, instead of cells of type $(1,\ldots,1)$, can serve the same purpose. The method can be made much more efficient by taking this special structure into consideration. For simplicity, we describe here only the special, and more common, case where the supports of the p_is are all different.

Polyhedral homotopies have been applied to solve symmetric polynomial systems by means of constructing symmetric polyhedral homotopies (Verschelde and Cools 1994, Verschelde and Gatermann 1995). On the other hand, the Bernshteĭn theory is also used for constructing random product start systems for linear homotopies with various degrees of success (Li, Wang and Wang 1996, Li and Wang 1994).

5. Numerical considerations

5.1. Projective coordinates

As described in Section 1, solution paths of (1.2) that do not proceed to a solution of $P(x) = 0$ in $\mathbb{C}^n$ diverge to infinity: a very poor state of affairs for numerical methods. However, there is a simple idea from classical mathematics which improves the situation. If the system (1.2) is viewed in $\mathbb{P}^n$, the diverging paths are simply proceeding to a 'point at infinity' in projective space. Since projective space is compact, we can force all paths, including the extraneous ones, to have finite length by using projective coordinates.

For $P(x) = (p_1(x_1,\ldots,x_n),\ldots,p_n(x_1,\ldots,x_n)) = 0$, consider the system of $n+1$ equations in $n+1$ unknowns after homogenization,

$$\widetilde{P}: \quad \begin{cases} \widetilde{p}_1(x_0,\ldots,x_n) & = 0, \\ & \vdots \\ \widetilde{p}_n(x_0,\ldots,x_n) & = 0, \\ a_0x_0+\cdots+a_nx_n-1 & = 0, \end{cases}$$

where $a_0, \ldots, a_n$ are complex numbers. When a start system

$$Q(x) = (q_1(x_1, \ldots, x_n), \ldots, q_n(x_1, \ldots, x_n)) = 0$$

is chosen, we also homogenize $Q(x)$ and consider the system

$$\widetilde{Q}: \quad \begin{cases} \widetilde{q}_1(x_0, \ldots, x_n) & = 0, \\ & \vdots \\ \widetilde{q}_n(x_0, \ldots, x_n) & = 0, \\ a_0 x_0 + \cdots + a_n x_n - 1 & = 0. \end{cases}$$

We then use the classical homotopy continuation procedure to follow all the solution paths of the homotopy

$$\widetilde{H}(x_0, x_1, \ldots, x_n, t) = (1-t)c\widetilde{Q}(x_0, \ldots, x_n) + t\widetilde{P}(x_0, \ldots, x_n).$$

For almost every choice of $a_0, \ldots, a_n$, the paths stay in $\mathbb{C}^{n+1}$. It only remains to ignore solutions with $x_0 = 0$. Of the remaining solutions with $x_0 \neq 0$, it is easy to see that $x = (x_1/x_0, \ldots, x_n/x_0)$ is the corresponding solution of $P(x) = 0$.

A similar technique is described in Morgan and Sommese (1987*a*), where it is called a 'projective transformation'. It differs from the above in the following way. Instead of increasing the size of the problem from $n \times n$ to $(n+1) \times (n+1)$, they implicitly consider solving the last equation for x_0 and substituting in the other equations, essentially retaining n equations in n unknowns. Then the chain rule is used for the Jacobian calculations needed for path following. In many cases, it seems that this may create extra work. Suppose, for example, that the tenth equation in the system is $p_{10}(x) = x_7^3 - x_1 x_2$; its homogeneous version is $\widetilde{p}_{10}(x) = x_7^3 - x_0 x_1 x_2$. Since x_0 is now considered as a function of all other variables, the partial derivative of p_{10} with respect to every variable is suddenly nonzero. This results in added computation for each Jacobian evaluation, and is particularly problematic if the original problem is large and/or sparse.

A more advanced technique, the *projected Newton method*, was suggested in Shub and Smale (1993). A typical step to follow a solution curve in $\mathbb{C}^n$ of homotopy $H(x,t) = 0$, a system of n equations in $n+1$ variables, consists of two major steps: *prediction* and *correction*. The prediction step locates a point $(x^{(0)}, t_0)$. For fixed t_0, $H(x, t_0) = 0$ is a system of n equations in n unknowns. With starting points $x^{(0)}$, Newton's iteration,

$$x^{(m+1)} = x^{(m)} - [H_x(x^{(m)}, t_0)]^{-1} H(x^{(m)}, t_0), \qquad m = 0, 1, \ldots$$

can be applied to find the solution of $H(x, t_0) = 0$. If $x^{(0)}$ is suitably chosen by the prediction step, the iteration will converge to a solution of $H(x, t_0) = 0$ close to $x^{(0)}$. This is called the correction step. To follow the solution curve in projective space $\mathbb{P}^n$ after homogenization, $H(x,t) = 0$ becomes $\widetilde{H}(\widetilde{x}, t) = 0$ and, for fixed t_0, $\widetilde{H}(\widetilde{x}, t_0) = 0$ is now a system of n

equations in $n+1$ variables: $x_0, \ldots, x_n$. It is, therefore, unsuitable for the classical Newton iteration at the correction step. However, for any nonzero constant $c \in \mathbb{C}$, $\widetilde{x}$ and $c\widetilde{x}$ in $\mathbb{C}^{n+1}$ are considered to be equal in $\mathbb{P}^n$, whence the magnitude of $\widetilde{x}$ in $\mathbb{C}^{n+1}$ is no longer significant in $\mathbb{P}^n$. Therefore it is reasonable to project every step of the Newton iteration onto the hyperplane perpendicular to the current point in $\mathbb{C}^{n+1}$. At $\widetilde{x}^{(m)} \in \mathbb{C}^{n+1}$, we now have $n+1$ equations in $n+1$ unknowns, namely

$$\bar{H}(\widetilde{x}, t_0) = \begin{cases} \widetilde{H}(\widetilde{x}, t_0) = 0 \\ (\widetilde{x} - \widetilde{x}^{(m)}) \cdot \widetilde{x}^{(m)} = 0 \end{cases}$$

and one step of Newton's iteration for this system can be used to obtain

$$\widetilde{x}^{(m+1)} = \widetilde{x}^{(m)} - [\bar{H}_{\widetilde{x}}(\widetilde{x}^{(m)}, t_0)]^{-1} \bar{H}(\widetilde{x}^{(m)}, t_0).$$

The efficiency of this strategy, known as the projected Newton iteration, when applied to following the homotopy curve in $\mathbb{P}^n$, is intuitively clear. See Figure 5. It frequently allows a bigger step size at the prediction stage.

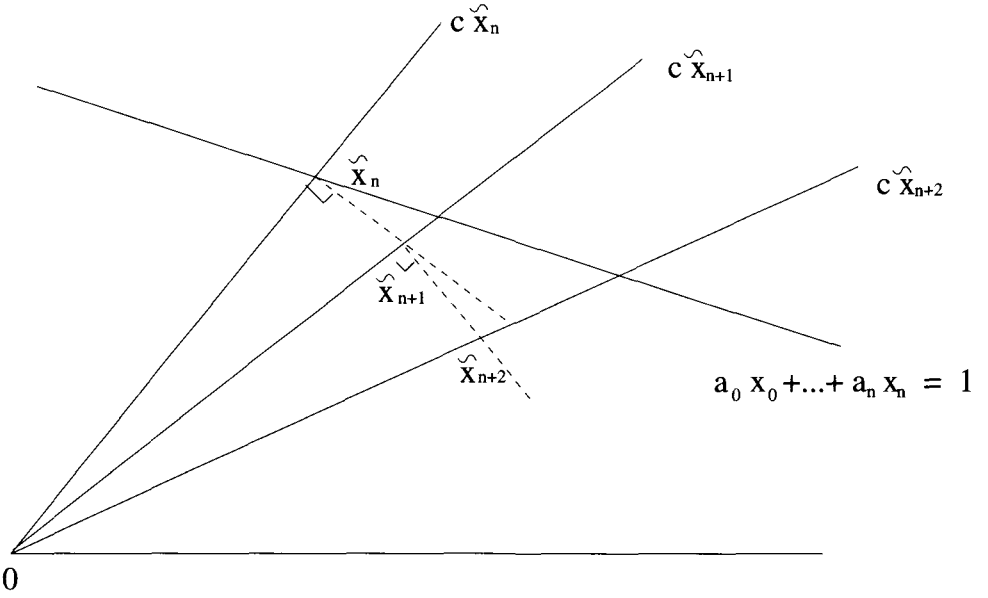

Fig. 5.

For practical considerations, we revise the above procedure as follows. At $t = t_1$, let

$$\widetilde{x}(t_1) = (x_0(t_1), x_1(t_1), \ldots, x_n(t_n)) \in \mathbb{P}^n$$

be the corresponding point on a homotopy curve $(\widetilde{x}(t), t)$ of $\widetilde{H}(\widetilde{x}(t), t) = 0$.

Let

$$|x_i(t_1)| = \max(|x_0(t_1)|, \ldots, |x_n(t_1)|).$$

We then fix the variable x_i in $\widetilde{H}(\widetilde{x}, t) = 0$ by the number $x_i(t_1)$, and thereafter, $\widetilde{H}(x_0, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n, t) = 0$ becomes a system of n equations in $n+1$ variables. A standard prediction–correction procedure can now be applied to arrive at a new point $(x_0(t_2), \ldots, x_{i-1}(t_2), x_{i+1}(t_2), \ldots, x_n(t_2))$, which satisfies

$$\widetilde{H}(x_0, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n, t_2) = 0.$$

Letting $\widetilde{x}(t_2) = (x_0(t_2), \ldots, x_{i-1}(t_2), x_i(t_1), x_{i+1}(t_2), \ldots, x_n(t_2))$, the point on the curve $(\widetilde{x}(t), t)$ for $t = t_2$ is obtained. A major advantage of this revision is that the size of the problem remains $n \times n$ throughout the procedure.

5.2. Real homotopy

Most polynomial systems arising in applications consist of polynomials with real coefficients, and most often the only desired solutions are real solutions. This suggests the use of real homotopies. That is, when the coefficients of the target polynomial system $P(x) = 0$ we want to solve are all real, we may choose a start system $Q(x) = 0$ with real coefficients, ensuring that the homotopy $H(x, t) = 0$ has real coefficients for all t. Thus, for fixed t, if x is a solution of $H(x, t) = 0$, so is its conjugate $\bar{x}$. Accordingly, a major advantage of real homotopy is that following a complex homotopy path $(x(s), t(s))$ provides its conjugate homotopy path $(\bar{x}(s), t(s))$ as a by-product without any further computation. On the other hand, although the homotopy $H(x, t)$ is still a map from $\mathbb{C}^n \times [0, 1]$ to $\mathbb{C}^n$, when a real homotopy path is traced, we may consider $H(x, t)$ as a map from $\mathbb{R}^n \times [0, 1]$ to $\mathbb{R}^n$, and hence the computation can be achieved in real arithmetic. In this way, a considerable reduction in computation is achieved.

There are numerous computational problems associated with the path following algorithms of real homotopies. In particular, when real homotopies are used, in contrast to the complex homotopy, bifurcation of some of the homotopy paths is inevitable. Hence, efficient algorithms must be developed to identify the bifurcation points and to follow the path after bifurcation. We can no longer parametrize the homotopy path of $H(x, t) = 0$ by t conventionally. Instead, the arclength s can be used as a parameter, and both x and t are considered to be independent variables. We now have

$$H(x(s), t(s)) = 0$$

and

$$H_x \dot{x} + H_t \dot{t} = 0,$$

where $\dot{x} \equiv \dfrac{dx}{ds}$, $\dot{t} \equiv \dfrac{dt}{ds}$ and $\|\dot{x}\|^2 + |\dot{t}|^2 = 1$. It is easy to see that bifurca-

tions can only occur at *turning points*, points (x^*, t^*) for which $\dot{t} = 0$ and $H_x(x^*, t^*) = 0$ is singular. To identify the bifurcation point, let $a_0 = (x^{(0)}, t_0)$ be a point on the homotopy path Γ with $\dot{t}(a_0) > 0$. After a standard Euler prediction with step size h_0 and Newton corrections (Allgower and Georg 1990, 1993), we obtain a point $a_1 = (x^{(1)}, t_1)$ on Γ. When the tangent vector $(\dot{x}, \dot{t})$ is calculated at a_1 with $\dot{t}(a_1) < 0$, a turning point $a^* = (x^*, t^*)$ apparently exists in this situation (see Figure 6).

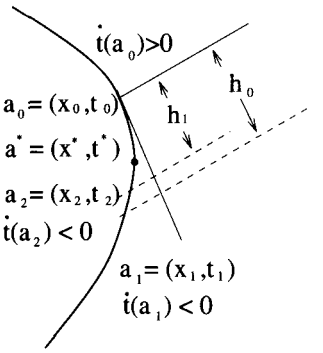

Fig. 6.

To identify a^*, we take the following procedure.

(1) Let h_1 be the solution of the equation ...

$$\frac{h}{h_0}\dot{t}(a_0) + \frac{h_0 - h}{h_0}\dot{t}(a_1) = 0.$$

Taking the Euler prediction at a_0 with step size h_1 followed by Newton corrections, we obtain a new point a_2 on Γ.

(2) If $\dot{t}(a_2) > 0$, we replace a_0 by a_2 and replace h_0 by the real part of the inner product of $(a_1 - a_2)$ and the unit tangent vector at a_2. If $\dot{t}(a_2) < 0$, we replace a_1 by a_2 and h_0 by h_1.

(3) (3) Repeat step 1 until $\dot{t}(a_2)$ is sufficiently small. Then, a_2 will be taken as a bifurcation point $a^* = (x^*, t^*)$.

When the bifurcation point a^* is identified, in order to follow the bifurcation branches, tangent vectors of the branches need to be characterized. It turns out that for the following special kind of turning point the bifurcation phenomenon is rather simple.

Definition 5.1 A singular point $(x^*, t^*) \in \mathbb{C}^n \times [0, 1]$ is said to be a quadratic turning point of $H(x, t) = 0$ if

(1) $\text{Rank}_R H_x(x^*, t^*) = 2n - 2$

(2) $\text{Rank}_R[H_x(x^*, t^*), H_t(x^*, t^*)] = 2n - 1$

(3) For $y \in \mathbb{C}^n \setminus \{0\}$ satisfying $H_x(x^*, t^*)y = 0$, we have

$$\text{Rank}_R[H_x(x^*, t^*), H_{xx}(x^*, t^*)yy) = 2n.$$

Here, Rank_R denotes the real rank.

Proposition 5.1 (Li and Wang 1994) Let (x^*, t^*) be a quadratic turning point. Then, there are only two branches of solution paths Γ and Γ' passing through (x^*, t^*). If ϕ is the tangent vector of the path Γ at (x^*, t^*), then the tangent vector of Γ' is the direction of $i\phi$ (see Figure 7).

When Γ is a real path, the assertion of this proposition can be considered as a special case of Allgower (1984) and Henderson and Keller (1990). The most general version, where Γ and x^* may both be complex, was proved in Li and Wang (1993).

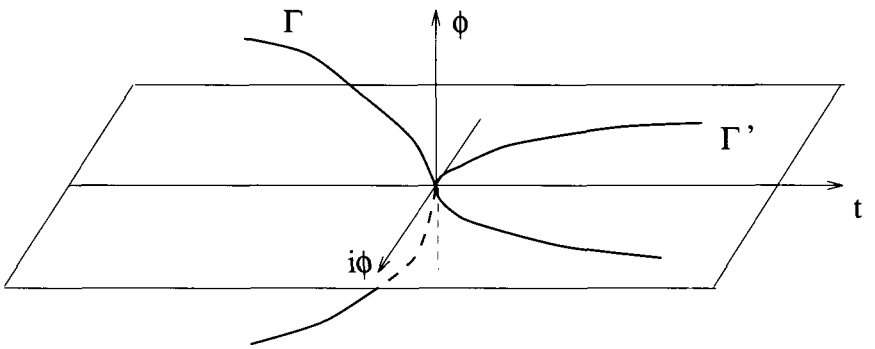

Fig. 7.

To follow the bifurcation branch Γ' at a quadratic turning point, we consider the following three situations.

(1) Γ *is a real path.*
Then, ϕ is real and $i\phi$ is pure imaginary. Apparently, the bifurcation branch Γ' consists of a complex path and its complex conjugacy. We need only to follow one of them with tangent vector $i\phi$ or $-i\phi$.

(2) Γ *is a complex path and* (x^*, t^*) *is real.*
Then, Γ consists of complex conjugate pairs (x, t) and $(\bar{x}, t)$ for each $t < t^*$. The tangent vector at (x^*, t^*) is

$$\phi = \lim_{s_1 - s_2) \to 0} \frac{(x(s_1), t(s_1)) - (x(s_2), t(s_2))}{s_1 - s_2},$$

where $x(s_2) = \bar{x}(s_1), t(s_1) = t(s_2)$ is clearly pure imaginary. Hence, $i\phi$ is real. Consequently, the bifurcation branch Γ' consists of two real paths. We may follow them in real space $\mathbb{R}^n \times [0, 1]$ with real tangent vectors $i\phi$ and $-i\phi$ respectively.

(3) Γ *is a complex path and* x^* *is complex.*
The bifurcation branch Γ', in this case, consists of two complex solution paths. They are not conjugate to each other. We may follow them with tangent vector $i\phi$ and $-i\phi$ respectively.

It was conjectured in Brunovský and Meravý (1984) and proved in Li and Wang (1993) that, generically, real homotopies contain no singular points other than a finite number of quadratic turning points.

5.3. Software

Several software packages dedicated to solving polynomial systems by homotopy continuation are publicly available. HOMPACK (Morgan, Sommese and Watson 1989) and CONSOL (Morgan 1987) are written in FORTRAN 77. HOMPACK is a general package for homotopy continuation with a polynomial driver. It has been parallelized to various architectures (Allison, Chakraborty and Watson 1989, Harimoto and Watson 1989). The code for CONSOL is contained in Morgan (1987), Appendix 6. The programs pss (Malajovich, software) and Pelican (Huber, software) are written in C. The pss contains facilities for parallel continuation and Pelican provides the polyhedral methods. The package PHC and MVC (Verschelde 1995) is written in Ada and compiled on three different hardware platforms, for which executables are available on request. Two main features of this package are the wide variety of homotopy methods and the powerful facilities for mixed volume computation.

Nonetheless, a more efficient and user-friendly code including all the features described in this article is still under development. In particular, a better understanding of the convex geometry with a clever use of linear programming techniques will make the polyhedra homotopy method described in Section 3 much more powerful.

REFERENCES

E. L. Allgower (1984), Bifurcation arising in the calculation of critical points via homotopy methods, in *Numerical Methods for Bifurcation Problems* (T. Kupper, H. D. Mittelman and H. Weber, eds), Birkhäuser, Basel, pp. 15–28.

E. L. Allgower and K. Georg (1990), *Numerical Continuation Methods, an Introduction*, Springer, Berlin. Springer Series in Computational Mathematics, Vol. 13.

E. L. Allgower and K. Georg (1993), Continuation and path following, in *Acta Numerica*, Vol. 2, Cambridge University Press, pp. 1–64.

D. C. S. Allison, A. Chakraborty and L. T. Watson (1989), 'Granularity issues for solving polynomial systems via globally convergent algorithms on a hypercube', *J. Supercomputing* **3**, 5–20.

D. N. Bernshteín (1975), 'The number of roots of a system of equations', *Functional Anal. Appl.* **9**, 183–185. Translated from *Funktsional. Anal. i Prilozhen.*, **9**, 1–4.

P. Brunovský and P. Meravý (1984), 'Solving systems of polynomial equations by bounded and real homotopy', *Numer. Math.* **43**, 397–418.

B. Buchberger (1985), Gröbner basis: An algorithmic method in polynomial ideal theory, in *Multidimensional System Theory* (N. Bose, ed.), D. Reidel, Dordrecht, pp. 184–232.

J. Canny and J. M. Rojas (1991), An optimal condition for determining the exact number of roots of a polynomial system, in *Proceedings of the 1991 International Symposium on Symbolic and Algebraic Computation*, ACM, pp. 96–101.

S. N. Chow, J. Mallet-Paret and J. A. Yorke (1979), Homotopy method for locating all zeros 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 Mathematics, Vol. 730, Springer, Berlin, pp. 77–88.

F. J. Drexler (1977), 'Eine Methode zur Berechnung sämtlicher Lösungen von Polynomgleichungssystemen', *Numer. Math.* **29**, 45–58.

I. Emiris (1994), Sparse Elimination and Applications in Kinematics, PhD thesis, University of California at Berkeley.

I. Emiris and J. Canny (1995), 'Efficient incremental algorithms for the sparse resultant and the mixed volume', *J. Symb. Computation* **20**, 117–149.

W. Fulton (1984), *Intersection Theory*, Springer, Berlin.

C. B. Garcia and W. I. Zangwill (1979), 'Finding all solutions to polynomial systems and other systems of equations', *Mathematical Programming* **16**, 159–176.

I. M. Gel'fand, M. M. Kapranov and A. V. Zelevinskiĭ (1994), *Discriminants, Resultants and Multidimensional Determinants*, Birkhäuser, Boston.

S. Harimoto and L. T. Watson (1989), The granularity of homotopy algorithms for polynomial systems of equations, in *Parallel Processing for Scientific Computing* (G. Rodrigue, ed.), SIAM, Philadelphia.

M. E. Henderson and H. B. Keller (1990), 'Complex bifurcation from real paths', *SIAM J. Appl. Math.* **50**, 460–482.

B. Huber (software), *Pelican Manual.* Available via the author's web page, `http://math.cornell.edu/~birk`.

B. Huber and B. Sturmfels (1995), 'A polyhedral method for solving sparse polynomial systems', *Math. Comp.* **64**, 1541–1555.

B. Huber and B. Sturmfels (1997), 'Bernstein's theorem in affine space', *Discrete Comput. Geom.* To appear.

A. G. Khovanskiĭ (1978), 'Newton polyhedra and the genus of complete intersections', *Functional Anal. Appl.* **12**, 38–46. Translated from *Funktsional. Anal. i Prilozhen.*, **12**, 51–61.

A. G. Kushnirenko (1976), 'Newton polytopes and the Bézout theorem', *Functional Anal. Appl.* **10**, 233–235. Translated from *Funktsional. Anal. i Prilozhen.*, **10**, 82–83.

C. W. Lee (1991), Regular triangulations of convex polytopes, in *Applied Geometry and Discrete Mathematics – The Victor Klee Festschrift, DIMACS Series Vol. 4* (P. Gritzmann and B. Sturmfels, eds), American Mathematical Society, Providence, RI, pp. 443–456.

T. Y. Li (1983), 'On Chow, Mallet-Paret and Yorke homotopy for solving systems of polynomials', *Bulletin of the Institute of Mathematics, Acad. Sin.* **11**, 433–437.

T. Y. Li and T. Sauer (1989), 'A simple homotopy for solving deficient polynomial systems', *Japan J. Appl. Math.* **6**, 409–419.

T. Y. Li and X. Wang (1990), A homotopy for solving the kinematics of the most general six- and-five-degree of freedom manipulators, in *Proc. of ASME Conference on Mechanisms, D1-Vol. 25*, pp. 249–252.

T. Y. Li and X. Wang (1991), 'Solving deficient polynomial systems with homotopies which keep the subschemes at infinity invariant', *Math. Comp.* **56**, 693–710.

T. Y. Li and X. Wang (1992), 'Nonlinear homotopies for solving deficient polynomial systems with parameters', *SIAM J. Numer. Anal.* **29**, 1104–1118.

T. Y. Li and X. Wang (1993), 'Solving real polynomial systems with real homotopies', *Math. Comp.* **60**, 669–680.

T. Y. Li and X. Wang (1994), 'Higher order turning points', *Appl. Math. Comput.* **64**, 155–166.

T. Y. Li and X. Wang (1996), 'The BKK root count in $\mathbb{C}^n$', *Math. Comp.* **65**, 1477–1484.

T. Y. Li, T. Sauer and J. A. Yorke (1987*a*), '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 (1987*b*), 'The random product homotopy and deficient polynomial systems', *Numer. Math.* **51**, 481–500.

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.

T. Y. Li, T. Wang and X. Wang (1996), Random product homotopy with minimal BKK bound, in *Proceedings of the AMS-SIAM Summer Seminar in Applied Mathematics on Mathematics of Numerical Analysis: Real Number Algorithms, Park City, Utah*, pp. 503–512.

G. Malajovich (software), *pss 2.alpha, polynomial system solver, version 2.alpha, README file.* Distributed by the author through gopher, `http://www.labma.ufrj.br/~gregorio`.

A. P. Morgan (1986), 'A homotopy for solving polynomial systems', *Appl. Math. Comput.* **18**, 173–177.

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

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

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

A. P. Morgan and A. J. Sommese (1989), 'Coefficient-parameter polynomial continuation', *Appl. Math. Comput.* **29**, 123–160. Errata: *Appl. Math. Comput.* **51**, 207 (1992).

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.

J. M. Rojas (1994), 'A convex geometric approach to counting the roots of a polynomial system', *Theoret. Comput. Sci.* **133**, 105–140.

J. M. Rojas and X. Wang (1996), 'Counting affine roots of polynomial systems via pointed Newton polytopes', *J. Complexity* **12**, 116–133.

I. R. Shafarevich (1977), *Basic Algebraic Geometry*, Springer, New York.

M. Shub and S. Smale (1993), 'Complexity of Bézout's theorem I: Geometric aspects', *J. Amer. Math. Soc.* **6**, 459–501.

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 Journal of Mechanics, Transmissions, and Automation in Design* **107**, 189–200.

J. Verschelde (1995), PHC and MVC: two programs for solving polynomial systems by homotopy continuation, Technical report. Presented at the PoSSo workshop on software, Paris. Available by anonymous ftp to `ftp.cs.kuleuven.ac.be` in the directory /pub/NumAnal-ApplMath/PHC.

J. Verschelde (1996), Homotopy continuation methods for solving polynomial systems, PhD thesis, Katholieke Universiteit Leuven, Belgium.

J. Verschelde and R. Cools (1993), 'Symbolic homotopy construction', *Applicable Algebra in Engineering, Communication and Computing* **4**, 169–183.

J. Verschelde and R. Cools (1994), 'Symmetric homotopy construction', *J. Comput. Appl. Math.* **50**, 575–592.

J. Verschelde and K. Gatermann (1995), 'Symmetric Newton polytopes for solving sparse polynomial systems', *Adv. Appl. Math.* **16**, 95–127.

J. Verschelde, K. Gatermann and R. Cools (1996), 'Mixed volume computation by dynamic lifting applied to polynomial system solving', *Discrete Comput. Geom.* **16**, 69–112.

C. W. Wampler (1992), 'Bézout number calculations for multi-homogeneous polynomial systems', *Appl. Math. Comput.* **51**, 143–157.

C. W. Wampler (1994), 'An efficient start system for multi-homogeneous polynomial continuation', *Numer. Math.* **66**, 517–523.

A. H. Wright (1985), 'Finding all solutions to a system of polynomial equations', *Math. Comp.* **44**, 125–133.

W. Zulener (1988), 'A simple homotopy method for determining all isolated solutions to polynomial systems', *Math. Comp.* **50**, 167–177.