Journal of Symbolic Computation 44 (2009) 280–291

Contents lists available at ScienceDirect

Journal of Symbolic Computation

journal homepage: www.elsevier.com/locate/jsc

Solving polynomial systems via symbolic-numeric reduction to geometric involutive form

Greg Reid [a], Lihong Zhi [b]

[a] *Department of Applied Mathematics, University of Western Ontario, London, N6A 5B7, Canada*
[b] *Key Lab of Mathematics Mechanization, AMSS, Beijing 100080, China*

ARTICLE INFO

Article history:
Received 20 April 2005
Accepted 14 October 2007
Available online 26 September 2008

Keywords:
Numerical linear algebra
Numeric elimination
Partial differential equations
Jet spaces
Involutive bases
Numeric jet geometry

ABSTRACT

We briefly survey several existing methods for solving polynomial systems with inexact coefficients, then introduce our new symbolic-numeric method which is based on the geometric (Jet) theory of partial differential equations. The method is stable and robust. Numerical experiments illustrate the performance of the new method.

1. Introduction

Exact elimination methods for exactly given polynomial systems (e.g. Gröbner Bases), usually employ Gaussian Elimination (e.g. linear elimination of monomials). Such exact methods usually depend on the ordering of input (e.g. term ordering in the case of Gröbner Bases), and so are coordinate dependent. Since the order of elimination can force division by small leading entries, such methods are generally unstable, when used on approximate systems. In contrast, exact elimination methods from the geometric theory of PDE are coordinate independent (Kuranishi, 1957; Pommaret, 1978) and this motivated our study of numerical versions of such methods (Bonasia et al., 2004; Reid et al., 2002; Wittkopf and Reid, 2001; Reid et al., 2003) which is continued in this paper. See Tuomela and Arponen (2000) for applications of geometric methods to the numerical solution of ODE.

We exploit the well-known correspondence between polynomial systems and systems of constant coefficient linear homogeneous PDE. This equivalence has been extensively studied and exploited in the exact case by Gerdt (Gerdt and Blinkov, 1998) and his co-workers in their development of

E-mail addresses: reid@uwo.ca (G. Reid), lzhi@mmrc.iss.ac.cn (L. Zhi).

0747-7171/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jsc.2007.10.013

involutive bases (also see Saito et al. (2000)). We use this correspondence to write the polynomial system as a PDE system. The PDE system is brought to a geometric involutive form which is a variation of that of Kuranishi (1957) and Pommaret (1978) and much more distantly related to that of Gerdt and Blinkov (1998). Our new numerical methods are applied to this output involutive form.

The method depends on viewing the polynomial systems as matrix functions of their monomials and applying linear algebra to the null spaces of these maps (see Macaulay (1916) for an early example of this technique and especially see Emsalem (1978) and Mourrain (1996)). In our approach, we apply the Singular Value Decomposition (a fundamental technique of Numerical Linear Algebra) to the null spaces of these maps. We present a new method for computing the multiplication matrices from the null spaces of the involutive system and its geometric projections. This construction is based on a modification due to Bonasia et al. (2004), Wittkopf and Reid (2001) and Reid et al. (2003) of the classical criterion of involution (see Kuranishi (1957), Pommaret (1978) and Seiler (1994) for the classical criterion). The criterion is related to the one for zero dimensional systems given in Mourrain (1999) based on commutators which is closer to a Gröbner Basis formulation, with the commutators playing the rôle of S-polynomials. However, our criterion is not based on commutators, and for zero dimensional systems is coordinate independent. Our numerical criterion for output involutive form can be checked by computing dimensions (specifically of prolonged and projected systems). After the system is obtained in involutive form the solutions are found by applying eigenvalue–eigenvector techniques to a related eigen-problem constructed from the involutive form. In particular we give an eigen-problem formulation suitable for zero-dimensional involutive systems which is a modification of that of Auzinger and Stetter (1988), Möller and Stetter (1995), Corless et al. (1997) and Mourrain (1996).

The rest of the paper is organized as follows. In Section 2, we present the method based on symbolic prolongation and numeric projection. The algorithm and an illustrative example are given and discussed in Section 3. The performance on a set of well known examples is also given in Section 3. In Section 4 we outline strategies for larger systems based on exploiting their subsystem structure. In particular, we discuss the case where the subsystems are square and define complete intersections. In that case, the order for prolongation to involution, can be determined without any prolongations being first made. We also briefly describe relations to other approaches such as Mourrain and Trébuchet (2000), Bardet et al. (2005) and Faugère (2002).

2. Symbolic-numeric completion of polynomial systems

Consider a polynomial system S in $\mathbb{C}[x_1, \ldots, x_n]$ of degree q and its corresponding vector of monomials of degree less than or equal to q. The system can be written as

$$M_q \cdot [x_1^q, x_1^{q-1}x_2, \ldots, x_n^2, x_1, \ldots, x_n, 1]^{\mathrm{T}} = [0, 0, \ldots, 0, 0, \ldots, 0, 0]^{\mathrm{T}} \tag{1}$$

in terms of its coefficient matrix M_q. Here and hereafter, $[\ldots]^{\mathrm{T}}$ means the transposition. Further, $[\xi_1, \xi_2, \ldots, \xi_n]$ is one of the solutions of the polynomial system, if and only if

$$[\xi_1^q, \xi_1^{q-1}\xi_2, \ldots, \xi_n^2, \xi_1, \ldots, \xi_n, 1]^{\mathrm{T}} \tag{2}$$

is a null vector of the coefficient matrix M_q.

Since the number of monomials is usually much greater than the number of polynomials, the dimension of the null space can be large. Completion methods for polynomial ideals based on critical pairs (Lazard, 1983; Faugère, 2002; Gerdt and Blinkov, 1998; Möller and Sauer, 2000; Auzinger and Stetter, 1988; Möller and Stetter, 1995; Stetter, 2004; Mourrain, 1999; Mourrain and Trébuchet, 2000, 2002; Trébuchet, 2002) aim to include additional polynomials belonging to the ideal generated by S, until a (minimal) normal form is determined capable of deciding membership in the ideal. As a consequence, certain dimensions (including the dimension of the system's null space, and its projections) attain minimum values. Our method focuses on direct methods to calculate and minimize these dimensions without using critical pair techniques.

The bijection

$$\phi : x_i \leftrightarrow \frac{\partial}{\partial x_i}, \quad 1 \le i \le n, \tag{3}$$

maps the system S to an equivalent system of linear homogeneous PDE denoted by R. Jet space approaches are concerned with the study of the jet variety

$$V(R) = \left\{ \left(\underset{q}{u}, \underset{q-1}{u}, \ldots, \underset{1}{u}, u \right) \in J^q : R\left(\underset{q}{u}, \underset{q-1}{u}, \ldots, \underset{1}{u}, u \right) = 0 \right\}, \tag{4}$$

where $\underset{j}{u}$ denotes the formal jet coordinates corresponding to derivatives of order exactly j. Here $J^q \approx \mathbb{C}^{N_q}$ where $N_q = \binom{q+n}{q}$, is the number of jet variables of derivative order less than or equal to q (or equivalently the number of monomials of total degree less than or equal to q).

A single prolongation of a system R of order q consists of augmenting the system with all possible derivatives of its equations, so that the resulting augmented system, denoted by $\mathbf{D}R$, has order $q+1$. Under the bijection ϕ, the equivalent operation for polynomial systems is to multiply by monomials, so that the resulting augmented system has degree $q+1$. Successive prolongations of the system yield R, $\mathbf{D}R$, $\mathbf{D}^2R$, ..., and a sequence of corresponding linear homogenous constant matrix systems:

$$M_q \mathbf{v_q} = \mathbf{0}, M_{q+1}\mathbf{v_{q+1}} = \mathbf{0}, M_{q+2}\mathbf{v_{q+2}} = \mathbf{0}, \ldots \tag{5}$$

where $\mathbf{v_i} = \left(\underset{i}{u}, \underset{i-1}{u}, \ldots, \underset{1}{u}, u \right)^{\mathrm{T}}$. A single geometric projection is defined as

$$\boldsymbol{\pi}(R) := \left\{ \left(\underset{q-1}{u}, \ldots, \underset{1}{u}, u \right) \in J^{q-1} : \exists\, \underset{q}{u}, R\left(\underset{q}{u}, \underset{q-1}{u}, \ldots, \underset{1}{u}, u \right) = 0 \right\}. \tag{6}$$

The projection operator $\boldsymbol{\pi}$ maps a point in J^q to one in J^{q-1} by removing the jet variables of order q (i.e. eliminating $\underset{q}{u}$). For polynomial systems of degree q, by the bijection ϕ, the projection is equivalent to eliminating the monomials of the highest degree q. We have adopted an abbreviated notation for projection here. To avoid cumbersome notation, we have omitted the traditional indices in the projection operator (e.g. $\boldsymbol{\pi}^q_{q-1}$ to indicate that the projection acts from J^q to J^{q-1}). Our convention is that the projection is determined by the space on which it acts. Thus $\boldsymbol{\pi}^3(R) \equiv \boldsymbol{\pi}^{q-2}_{q-3}\boldsymbol{\pi}^{q-1}_{q-2}\boldsymbol{\pi}^q_{q-1}(R)$. A similar convention has been followed for the prolongation operator $\mathbf{D}$.

To implement an approximate involutive form method, we proposed in Bonasia et al. (2004), Wittkopf and Reid (2001) and Reid et al. (2003) a numeric projection operator $\hat{\boldsymbol{\pi}}$ based on singular value decomposition (SVD). We first find the SVD of M_{q+k}:

$$M_{q+k} = U \cdot \Sigma \cdot V.$$

Here, U and V are unitary matrices. Σ is a diagonal matrix whose diagonal entries are real decreasing non-negative numbers. The approximate rank r is the number of singular values bigger than a fixed tolerance. The tolerance is chosen close to the number of correct digits for the coefficients of the input polynomials. Deleting the first r rows of V yields an approximate basis for the null space of M_{q+k}. This yields an estimate for $\dim(\mathbf{D}^kR)$. To estimate $\dim(\hat{\boldsymbol{\pi}}(\mathbf{D}^kR))$, the components of the approximate basis for $\mathbf{D}^kR$ corresponding to the highest order ($(q+k)$th order) derivatives are deleted. This projected basis yields an approximate spanning set for $\hat{\boldsymbol{\pi}}(\mathbf{D}^kR)$. Proceeding in the same way, deleting components corresponding to the highest order jet variables from the approximate spanning set just obtained, yields an approximate spanning set for $\hat{\boldsymbol{\pi}}^2(\mathbf{D}^kR)$, and then for $\hat{\boldsymbol{\pi}}^3(\mathbf{D}^kR)$, etc. Application of the SVD to each of these approximate spanning sets yields the approximate dimensions of $\hat{\boldsymbol{\pi}}(\mathbf{D}^kR)$, $\hat{\boldsymbol{\pi}}^2(\mathbf{D}^kR)$, $\hat{\boldsymbol{\pi}}^3(\mathbf{D}^kR)$, ..., required for the application of our approximate involutive form test.

Throughout this article, we confine ourselves to polynomial systems, and consequently by the bijection, to linear homogeneous systems of PDE with constant coefficients. As a result, many of the more complicated phenomena that occur for non-constant coefficient systems and nonlinear PDE systems do not occur. Such phenomena include splitting into components of different dimensions, singular points, etc.

The symbol of a system of PDE (and by the bijection of a system of polynomials) is central to our approach.

Definition 2.1 (*Symbol*). The symbol matrix of a system of PDE is the Jacobian matrix of the system with respect to its highest order jet coordinates:

$$\text{Symbol } R := \frac{\partial R}{\partial \underset{q}{u}}. \tag{7}$$

Roughly speaking, this is the geometric generalization, of the leading term, of a single polynomial. In case of polynomials of degree q (and their equivalent linear PDE) the symbol matrix is simply the submatrix of the coefficient matrix M_q of the system corresponding to highest degree (q) monomials. Another central object of our paper is that of an involutive system. One of the most important requirements of involutive systems, is that their symbols are involutive. Involution of the symbol is intrinsically defined in terms of the vanishing of certain Homology groups (Pommaret, 1978). Involutivity of the symbol is equivalent to its Mumford regularity (Malgrange, 2003).

A determination of involutivity of the symbol can be made using explicit coordinates (Seiler, 2002, 1994) as follows. First define the class of an order q jet variable (or equivalently a degree q monomial $x^J := x_1^{j_1} \dots x_n^{j_n}$ where $q = j_1 + \cdots + j_n$) as follows. The class of x^J is first nonzero k such that j_k is nonzero in the list $J = [j_1, j_2, \dots, j_n]$. Next, order the columns of Symbol R from higher to lower class, and row reduce the Symbol R. Then, define $\beta_k^{(q)}$ to be the number of pivots in the row reduced form of Symbol R corresponding to class k jet variables. Finally, we have:

Definition 2.2 (*Involutivity Test for Symbol*). The symbol of a qth order system $R = 0$ is involutive in a generic system of coordinates if

$$\text{rank Symbol}(\mathbf{D}R) = \sum_{j=1}^{n} j\beta_j^{(q)}. \tag{8}$$

To numerically implement (8) the $\beta_j^{(q)}$ are determined by dimensions of projections of subspaces corresponding to different classes of monomials.

Definition 2.3 (*Involutive System*). A system of linear homogeneous PDE $R = 0$ with constant coefficients is involutive if $\dim\ \boldsymbol{\pi}(\mathbf{D}R) = \dim\ R$ and the symbol of R is involutive.

As a special case of the Cartan–Kuranishi prolongation theorem (Kuranishi, 1957; Bryant et al., 1991; Seiler, 2002) we have:

Theorem 2.4 (*Cartan–Kuranishi Prolongation Theorem*). *A linear homogeneous system of PDE order q with constant coefficients becomes involutive after a finite number of projections and prolongations. In particular, given an input system $R = 0$ there exist $C_r, \dots C_2, C_1$ with $C_j = \boldsymbol{\pi}$ or $C_j = \mathbf{D}$ for each j such that $C_r \dots C_2 C_1(R)$ is an involutive system of order $\geq q$.*

Also, algorithms and computer algebra implementations exist for carrying out the above completion process for exact input (Seiler, 1994, 2002).

Our symbolic-numeric completion method requires that we compute prolongations $\mathbf{D}^k R$, and then compute $\boldsymbol{\pi}^\ell \mathbf{D}^k R$. This led us to introduce the concept of projected involutive systems in Bonasia et al. (2004), Wittkopf and Reid (2001) and Reid et al. (2003):

Definition 2.5 (*Projected Involutive System*). The system of linear homogenous PDE $R = 0$ with constant coefficients is said to be projectively involutive at prolongation order $k \geq 0$ and projected order ℓ such that $0 \leq \ell \leq k$, if $\boldsymbol{\pi}^\ell(\mathbf{D}^k R)$ satisfies the projected elimination test

$$\dim\ \boldsymbol{\pi}^\ell \left(\mathbf{D}^k R\right) = \dim\ \boldsymbol{\pi}^{\ell+1} \left(\mathbf{D}^{k+1} R\right) \tag{9}$$

and the symbol of $\boldsymbol{\pi}^\ell(\mathbf{D}^k R)$ is involutive.

Theorem 3.4 of Bonasia et al. (2004) states that:

Theorem 2.6. *A system is projectively involutive if and only if it is involutive.*

Theorem 3.3 of Bonasia et al. (2004) ensures that such systems exist. It states that given a linear homogeneous PDE system $R = 0$ there exist finite integers $k \geq 0$ and ℓ with $0 \leq \ell \leq k$ such that the system $R = 0$ is projectively involutive at prolongation order $k \geq 0$ and projected order ℓ.

In this article, we are concerned with systems of polynomials with finitely many solutions (zero dimensional systems) which correspond to linear homogeneous PDE with finitely many parameters in their solutions (finite type PDE). Under this hypothesis, we have the following simple and computationally convenient characterization of projectively involutive (equivalently involutive) systems. This result is essentially well-known and appears in many different guises (e.g. see Seiler (2002) for the differential case), but we include its proof for completeness.

Theorem 2.7 (*Criterion of Involution for Zero-dimensional Polynomial Systems*). *A q-th order system of linear homogeneous PDE R corresponding to a zero dimensional polynomial system S is projectively involutive at order k and projected order ℓ if and only if, $\boldsymbol{\pi}^{\ell}(\mathbf{D}^k R)$ satisfies the projected elimination test* (9) *and*

$$\dim \boldsymbol{\pi}^{\ell}\left(\mathbf{D}^k R\right) = \dim \boldsymbol{\pi}^{\ell+1}\left(\mathbf{D}^k R\right). \tag{10}$$

Proof of Theorem 2.7: The definition of the symbol space implies that

$$\dim\left(\text{Symbol } \boldsymbol{\pi}^{\ell}\left(\mathbf{D}^k R\right)\right) = \dim \boldsymbol{\pi}^{\ell}\left(\mathbf{D}^k R\right) - \dim \boldsymbol{\pi}^{\ell+1}\left(\mathbf{D}^k R\right). \tag{11}$$

We need to show under the hypotheses of the Theorem that:

$$\dim \boldsymbol{\pi}^{\ell}\left(\mathbf{D}^k R\right) = \dim \boldsymbol{\pi}^{\ell+1}\left(\mathbf{D}^k R\right) \Longleftrightarrow \text{Symbol } \boldsymbol{\pi}^{\ell}(\mathbf{D}^k R) \quad \text{is involutive.} \tag{12}$$

Suppose that (10) holds then by (11) $\dim(\text{Symbol } \boldsymbol{\pi}^{\ell}(\mathbf{D}^k R)) = 0$. In this case, it is easily shown that the symbol of $\boldsymbol{\pi}^{\ell}(\mathbf{D}^k R)$ is involutive (Seiler, 2002).

Suppose that Symbol $\boldsymbol{\pi}^{\ell}(\mathbf{D}^k R)$ is involutive under the hypotheses of the Theorem. Since S is a zero dimensional system, it has finitely many zeros and the Hilbert function of the ideal generated by the polynomials S as a function of its degree d, is zero for $d \geq d^*$ for sufficiently large d^* (Cox et al., 1992).

Seiler (2002, Def. 4.2.4) gives the definition of the Hilbert function in the differential case. By our isomorphism, this leads to the same Hilbert function (i.e. zero) as that for S. This is the key link justifying the equivalence between the PDE and polynomial case. Also see Robin Scott's thesis (Scott, 2006), for a more algebraically oriented approach. In particular d^* can be taken at least to be the order of $\boldsymbol{\pi}^{\ell}(\mathbf{D}^k R)$, that is $d^* = q + k - \ell$. But, then by (4.10) of Seiler (2002), we have $\dim \boldsymbol{\pi}^{\ell}\left(\mathbf{D}^k R\right) = \dim \boldsymbol{\pi}^{\ell+1}\left(\mathbf{D}^k R\right)$ and so (10) holds completing the proof. ■

We briefly discuss the case where the dimension of the symbol space is not zero. When there are 2 variables, then it is easily shown that:

$$\text{Symbol } \boldsymbol{\pi}^{\ell}(\mathbf{D}^k R) \text{ is involutive} \Longleftrightarrow \dim \text{Symbol } \boldsymbol{\pi}^{\ell}\left(\mathbf{D}^{k+1} R\right) = \dim \text{Symbol } \boldsymbol{\pi}^{\ell}\left(\mathbf{D}^k R\right)$$

and this gives a computationally easy characterization by using (11). However, when the number of variables is ≥ 3 a finer analysis of the structure of the symbol space and its Cartan characters is required (Seiler, 2002).

Suppose that finite type PDE R is involutive at prolonged order k and projected order ℓ, and by the bijection ϕ has corresponding system of polynomials S. Then, the dimension of $\hat{\boldsymbol{\pi}}^{\ell}(\mathbf{D}^k R)$ allows us to determine the number of approximate solutions of S up to multiplicity. In particular, these solutions approximately generate the null space of $\hat{\boldsymbol{\pi}}^{\ell}(\mathbf{D}^k R)$. Moreover, from

$$d = \dim \hat{\boldsymbol{\pi}}^{\ell}\left(\mathbf{D}^k R\right) = \dim \hat{\boldsymbol{\pi}}^{\ell+1}\left(\mathbf{D}^k R\right), \tag{13}$$

we can form the multiplication matrices from the null space of $\hat{\boldsymbol{\pi}}^{\ell}\left(\mathbf{D}^k R\right)$ and $\hat{\boldsymbol{\pi}}^{\ell+1}\left(\mathbf{D}^k R\right)$. Instead of choosing monomials to form a normal set of size d, we compute the SVD of the approximate basis of the null space of $\hat{\boldsymbol{\pi}}^{\ell+1}(\mathbf{D}^k R)$. Since the first d left singular vectors permit a stable representation of the other rows in the approximate basis of the null space of $\hat{\boldsymbol{\pi}}^{\ell}\left(\mathbf{D}^k R\right)$, a polynomial basis formed from these singular vectors leads to a stable representation of multiplicative structure of the quotient ring $\mathbb{C}[x_1, \ldots, x_n]/I$. Here, I is the zero-dimensional ideal generated by the polynomials in S. This is the key step which significantly improves the stability of our method. The small errors shown on Fig. 2 also reveal the stability of our method on a set of benchmarks.

The details for computing the multiplication matrix are described in the following three steps:

(1) We compute an approximate basis of the null space of $\hat{\boldsymbol{\pi}}^{\ell}\left(\mathbf{D}^k R\right)$, denoted by a $N_{q+k-\ell} \times d$ matrix B. The $N_{q+k-\ell-1} \times d$ submatrix B_1 of B by deleting entries corresponding to the $q+k-\ell$ degree monomials is a basis of null space of $\hat{\boldsymbol{\pi}}^{\ell+1}\left(\mathbf{D}^k R\right)$ due to the condition (13).

(2) Let $\mathcal{N}_{q+k-\ell-1} = \left[x_1^{q+k-\ell-1}, x_1^{q+k-\ell-2}x_2, \ldots, x_n, 1\right]$ be the set of all monomials of degree less than or equal to $q+k-\ell-1$. We compute the SVD of $B_1 = U \cdot S \cdot V$. The first d columns of U form the submatrix U_s, and guarantee a stable polynomial set $\mathcal{N}_p = U_s^{\mathrm{T}} \cdot \mathcal{N}^{\mathrm{T}}$.

(3) The multiplication matrices of x_j with respect to $\mathcal{N}_p$ can be formed as $M_{x_j} = U_s^{\mathrm{T}} \cdot B_{x_j} \cdot V^{\mathrm{T}} \cdot S_i$ where B_{x_j} are the rows of B corresponding to monomials $x_j \cdot \mathcal{N}_{q+k-\ell-1}$ respectively, and S_i is a diagonal matrix with elements which are reciprocals of the first d elements of S. The matrix S_i is well-conditioned since all elements of S_i are bounded by the reciprocal of the fixed tolerance.

Finally, the solutions can be obtained by computing eigenvalues and eigenvectors of the multiplication matrices.

3. Algorithm and examples

The following algorithm solves zero dimensional polynomial systems based on the symbolic-numeric elimination method we discussed above.

Algorithm SNEPSolver

Input:

- a degree q polynomial system S in $\mathbb{C}[x_1, \ldots, x_n]$ with a finite number of zeros
- a small tolerance ϵ.

Output: All numerical solutions of S

(1) Form the PDE system R by the bijection ϕ. Let $\ell = 0;\ k = 0$.

(2) Applying the symbolic-numeric completion method to R with tolerance ϵ, we obtain the table of dimensions of $\dim \hat{\boldsymbol{\pi}}^{\ell}(\mathbf{D}^k R)$.

(3) We seek the smallest k such that there exists an ℓ with $\hat{\boldsymbol{\pi}}^{\ell}(\mathbf{D}^k R)$ approximately involutive, i.e., satisfying the conditions (9) and (10). We choose the largest such ℓ if there are several such values for the given k. (From Definition 2.5 recall that $0 \le \ell \le k$).

(4) Form the multiplication matrices of $x_1, \ldots, x_n$ from the null vectors of $\hat{\boldsymbol{\pi}}^{\ell}\left(\mathbf{D}^k R\right)$ and $\hat{\boldsymbol{\pi}}^{\ell+1}\left(\mathbf{D}^k R\right)$, we can compute eigenvalues and eigenvectors to find solutions of S (Auzinger and Stetter, 1988; Möller and Stetter, 1995; Corless et al., 1997).

The details are discussed in the following example given by Stetter (2004).

$$\begin{aligned} p_1 &:= -3.8889 + 0.078524\,x + 0.66203\,y + 2.9722\,x^2 - 0.46786\,xy + 1.0277\,y^2, \\ p_2 &:= -3.8889 + 0.66203\,x - 0.078524\,y + 1.0416\,x^2 + 0.70179\,xy + 3.9584\,y^2. \end{aligned} \tag{14}$$

Using the methods of Stetter (2004), this is a difficult problem which required about 30 Digits of precision to obtain 10 correct digits for the y-component if we are using a generic normal set $\{1, x, x^2, x^3\}$.

The method we now describe does not use a normal set, and only needs Digits $= 10$ for success in Maple 9. Under the bijection ϕ, the system is equivalent to the PDE system R. Applying the symbolic-numeric completion method to R with tolerance 10^{-9}, we obtain the table of dimensions of $\hat{\boldsymbol{\pi}}^{\ell}(\mathbf{D}^k R)$. Fig. 1 is the table of dimensions of the prolonged and projected PDE systems of R, R is equivalent to polynomial system (14) by bijection ϕ defined on (3).

	$k=0$	$k=1$	$k=2$	$k=3$
$\ell=0$	4	4	4	4
$\ell=1$	3	4	4	4
$\ell=2$	1	3	4	4
$\ell=3$		1	3	4
$\ell=4$			1	3
$\ell=5$				1

Fig. 1. Table of dim $\hat{\boldsymbol{\pi}}^{\ell}(\mathbf{D}^k R)$ for (14).

Applying the approximate version of the involutive test to the example shows that the system is involutive after one prolongation and no projection, i.e. $k = 1$, $\ell = 0$, yielding $\mathbf{D}R$ as the sought approximately involutive system.

The involutive system has $\dim(\mathbf{D}R) = 4$ and so by the bijection the polynomial system has 4 solutions up to multiplicity, and the monomial bases for these spaces should include the second degree monomials in order to recover all solutions. In the following, we show how to find the solutions without computing normal set w.r.t. a specified order of variables. It is a key improvement on Reid et al. (2003) since there a type of normal set was used.

- Compute an approximate basis of the null space of $\mathbf{D}R$, denoted by a 10×4 matrix B. The 6×4 submatrix B_1 of B by deleting entries corresponding to the third degree monomials is a basis of null space of $\hat{\boldsymbol{\pi}}(\mathbf{D}R)$ since $\dim(\hat{\boldsymbol{\pi}}(\mathbf{D}R)) = \dim(\mathbf{D}R) = 4$.
- Let $\mathcal{N} = \left[x^2, xy, y^2, x, y, 1\right]$ be the set of all monomials of degree less than or equal to 2. For numerical stability, instead of selecting four monomials as the normal set from $\mathcal{N}$, we compute the SVD of $B_1 = U \cdot S \cdot V$. The first four columns of U form the 6×4 submatrix U_s, and guarantee a stable polynomial set $\mathcal{N}_p = U_s^{\mathrm{T}} \cdot \mathcal{N}^{\mathrm{T}}$ (including four quadratic polynomials) for computing multiplication matrices accurately.
- The multiplication matrices of x, y with respect to $\mathcal{N}_p$ can be formed as $M_x = U_s^{\mathrm{T}} \cdot B_x \cdot V^{\mathrm{T}} \cdot S_i$ and $M_y = U_s^{\mathrm{T}} \cdot B_y \cdot V^{\mathrm{T}} \cdot S_i$, where B_x, B_y are the 1, 2, 3, 5, 6, 8 and 2, 3, 4, 6, 7, 9 rows of B corresponding to monomials $x^3, x^2y, xy^2, x^2, xy, x$ and $x^2y, xy^2, y^3, xy, y^2, y$ respectively, and S_i is a well-conditioned diagonal matrix with elements which are reciprocals of the first four elements of S: 0.99972, 0.95761, 0.64539, 0.58916.
- Compute the eigenvectors v_p of $M_x - M_y$ (or any random linear combination of M_x, M_y), and recover the eigenvector corresponding to the monomial set $\mathcal{N}$ by $v = U_s \cdot v_p$. Since $x, y, 1$ appear as the last three components in $\mathcal{N}$, the solutions of p_1, p_2 can be obtained as $x = v[4, i]/v[6, i]$, $y = v[5, i]/v[6, i]$:

$$\{x = 1.04972, y = -0.80689\}; \qquad \{x = 1.04972, y = 0.64062\};$$
$$\{x = -1.20441, y = -0.78652\}; \qquad \{x = -0.76039, y = 1.05888\}.$$

Substituting these solutions back to p_1, p_2, we found that the errors are smaller than 10^{-8}. It should be noticed that, for this example, although the first two solutions have the same x values, they do not correspond to a multiple root. So the last step is successful. Otherwise, we apply a reordered Schur factorization method in Corless et al. (1997) to the multiplication matrices $M_x, M_y, \ldots$ to recover all roots including the multiplicities.

We have implemented the SNEPSolver in Maple 9. In the following table, we show the performance for some well known examples on Pentium 4 at 2.0 GHz for Digits = 10 and tolerance being 10^{-9} in Maple 9 under Windows. Here, *vars* is the number of variables; d is the number of solutions; k and ℓ denote the numbers of prolongation and projection for system become involutive respectively; *error* indicates the maximal absolute value after submitting the solutions to the original polynomial system.

Ex.	$\deg(f_i)$	$vars$	d	k	ℓ	*error*
1	2, 2	2	3	1	1	0.52e-17
2	2, 2	2	4	1	0	0.3e-8
3	2, 2, 2, 2, 2, 2	4	4	2	1	0.44e-8
4	2, 2	2	4	1	0	0.7e-9
5	2, 2	2	2	1	1	0.3e-9
6	2, 3, 2	3	12	2	0	0.7e-8
7	8, 7	2	49	6	1	0.10e-6
8	3, 3, 3	3	27	4	0	0.18e-14
9	3, 3, 3	3	14	4	3	0.22e-9
10	3, 3, 3	3	21	4	2	0.82e-9
11	1, 1, 2, 2, 3, 2	6	10	2	2	0.34e-8
12	3, 3, 3	3	27	4	0	0.24e-8

Fig. 2. Algorithm performance on twelve examples.

Examples are cited from Lazard (1983), Mourrain (1996), Reid et al. (2003), Stetter (2004), Corless et al. (1997), Sturmfels (2002) and Gatermann (1990). The systems 3, 6, 7, 8, 9, 12 have multiple roots.

4. Subsystem strategies for larger systems

We describe a strategy suitable for systems with subsystems that are square and define complete intersections.

Let $f(x) = (f_1(x), \ldots, f_m(x)) = 0$ be a system of $m = \#f$ polynomials in $\mathbb{C}[x]$ where $x = (x_1, \ldots, x_n)$. The total (Bézout) degree of $f(x)$ is $\deg(f) := \prod_i \deg f_i$, the product of degrees of its polynomials. For generic square systems, $\deg(f)$ is the number of isolated roots of $f(x) = 0$ and a fundamental parameter for measuring the complexity of polynomial solving methods. A key disadvantage of our method is that as the prolongation order grows, the total degrees of the prolonged systems can increase explosively.

It is natural to develop strategies based on the structure of f to lessen this growth. Here, we discuss the case where f has a subsystem structure $f(x) = (g(y), h(y, z)) = 0$ corresponding to a partition of the variables of the form $x = (y, z)$. Already a number of symbolic and numeric solution approaches exploit such subsystem structure (e.g. Gröbner Bases and Triangular Decomposition methods can exploit such structure by choosing appropriate orderings of the variables). The idea of such approaches is to first solve subsystem $g(y) = 0$, then substitute the solutions $y = \hat{y}$ into $h(y, z)$ and finally solve $h(\hat{y}, z) = 0$. In the case where the systems $g(y), f(\hat{y}, z)$ and $f(y, z)$ are generic and square comparing their total degrees helps highlight the advantages of such approaches.

We consider the case where $g(y) = 0$ is a square system of polynomials (i.e. $\#(g) = \#(y)$) of the same degree (i.e. $\deg(g_i) = d = \text{constant}$). Let g_{hom} be the system of polynomial functions obtained from g by removing all of its monomials of degree strictly less than d (i.e. g_{hom} is simply obtained from Symbol (g)). The variety of the leading homogeneous part of g is given by

$$V_{\text{hom}}(g) := \{\tilde{y} : g_{\text{hom}}(\tilde{y}) = 0\} \tag{15}$$

In Möller and Sauer (2000) it is shown that:

Theorem 4.1. *Suppose that $g(y) = 0$ is a square polynomial system of n equations in $\mathbb{C}[y]$ with each polynomial having the same degree d. If $V_{\hom}(g)$ only contains the zero solution, that is*

$$V_{\hom}(g) = \{0\}, \tag{16}$$

then:

(1) $g(y) = 0$ *is a complete intersection;*
(2) $g(y)$ *is an H Basis;*
(3) $g(y)$ *is a zero-dimensional ideal with Hilbert Function* $\mathrm{HF}(r) = d^n$ *for* $r \geq (n-1)(d-1)$*;*

Further in Reid et al. (2005) under the hypotheses of Theorem 4.1 it is easy to show that $\mathbf{D}^{(n-1)(d-1)}R$ is involutive; and that the PDE system R corresponding to $g(y)$ by the mapping ϕ is formally integrable. In addition, it is easy to show that if $g(y)$ is generic and square then $V_{\hom}(g) = \{0\}$ (e.g. see Reid et al. (2005)).

Thus, Theorem 4.1 applies to the large class of square generic systems. Square generic systems are a strict subset of systems with $V_{\hom}(g) = \{0\}$ (e.g. just consider $g(y) = y^2$ which has $V_{\hom}(g) = \{0\}$ but which is not generic).

An interesting feature of Theorem 4.1, is that for a large class of systems, it does not require the checking of S-polynomials, or other objects related to higher degree polynomials, obtained by prolonging its polynomials (e.g. by multiplying them by monomials). It only requires information coming from a part of the system ($V_{\hom}(g) = \{0\}$). This yields fundamental information on the ideal generated by the system such as its Hilbert function.

We now comment on a method to numerically check the hypothesis (16) in Theorem 4.1; which by Reid et al. (2005) yields the order of prolongation at which the system becomes involutive.

At first sight, the computational expense of checking (16) by solving $g_{\hom}(\tilde{y}) = 0$ is similar to solving $g(y) = 0$ since both systems consist of $\#(g)$ polynomials of degree d. This motivates us to use more refined polyhedral methods which exploit the structure of the problem. Such structure includes the fact that $g_{\hom}(\tilde{y}) = 0$, unlike $g(y)$ is homogeneous, and is obtained from $g(y)$ by removing all its lower ($< d$) degree monomial terms. However direct computation of the mixed volume, $M_{\mathrm{vol}}(g_{\hom}(\tilde{y}))$, does not yield useful information since being homogeneous it does not have nonzero isolated solutions.

Since any homogeneous system has its solution set left invariant by the map $\psi_\mu : \tilde{y} \mapsto \mu\tilde{y}$ there is no loss in scaling one of the variables to zero or one to yield overdetermined systems which are respectively homogeneous and non-homogeneous. The main idea here is that these systems can be analyzed by numerical homotopy methods, which enable the testing of the condition $V_{\hom}(g) = \{0\}$. The subtleties are that one has to use Numerical Algebraic Geometry, to rule out the possible existence of positive dimensional components in $V_{\hom}(g)$. The cost, even in the generic case, is less than that of numerically solving the whole system $g(y)$, and a tree of cases, where y_j are equal to zero or one, has to be analyzed. Eigenvalue methods or homotopy methods can be then used to solve the system $g(y) = 0$, and its solutions $y = \hat{y}$ substituted into $h(y, z) = 0$. We note that, if $g_{\hom}$ is sparse then input system g may not necessarily sparse. Specialization of the values of the variables as described above to values of 0 or 1 will not destroy the sparsity of $g_{\hom}$. Thus, the mixed volume of a specialized form of $g_{\hom}(y)$ could be much smaller than that of $g(y)$. So for such input systems, the cost to determine $V_{\hom}(g) = \{0\}$ can be even cheaper. The degree of $h(\hat{y}, z) = 0$ is potentially much smaller than $h(y, z) = 0$ making it much more amenable to our SNEPSolver, which can solve non-generic non-square zero dimensional systems.

We refer the reader to the interesting paper of Mourrain and Trébuchet (2000), where they show that they can efficiently solve square polynomial systems, provided that certain matrices in their method are well-conditioned, and this occurs if the systems are generic enough. Then, they can avoid dealing with significant numbers of multiples of polynomials in the prolongation. However, they do not give tests for their input systems being generic, complete intersections, or zero-dimensional. The numerical procedures we describe above test for these properties and thus are a valuable supplement to their methods. Their methods allow the practical processing of large sufficiently generic systems.

5. Conclusion

In this paper, we present a symbolic-numeric elimination method to find the numeric solutions for zero-dimensional polynomial systems. The matrices appearing in our approach can be large in comparison with those occurring in other approaches such as Mourrain (1999), Mourrain and Trébuchet (2000), Mourrain and Trébuchet (2002) and Trébuchet (2002). Future work involves reducing the size of the matrices by exploiting the structure of the problems and making use of structured singular value decomposition. That is, we plan to make our method more efficient, while maintaining its stability. An analytical backward error analysis in terms of an appropriate error metric is also an important future task, made feasible by the backward error analyses, that exist for the SVD.

Our differential-algebraic method is easily reformulated and implemented by the bijection ϕ in terms of pure linear algebra on monomials. Such an implementation would be more efficient than our current differential method in Maple since it does not have the additional overhead for differentiation. Our method is related to that of Mourrain (1999), Mourrain and Trébuchet (2000), Mourrain and Trébuchet (2002) and Trébuchet (2002). They also gave a necessary and sufficient condition for a projection onto a set of polynomials to be a normal form modulo an ideal, and present a new algorithm for constructing the multiplicative structure of a zero-dimensional algebra. However, during their algorithm, numeric reductions with respect to some set of polynomials are performed. While in our algorithm, we only need to check the dimensions of the prolonged and projected differential system; no reduction is needed. So there is no new error introduced during the completion to involutivity. Although the matrices in our approach can be bigger, they only consist of the coefficients of the original input polynomial system. The threshold value we used for computing the numeric rank indicates how near the coefficient matrix is to a matrix of lower rank. It would be interesting to investigate how close our polynomial system is to the system with the given dimensional table. We can apply structured total least norm method (Lemmerling, 1999; Lemmerling et al., 2000; Rosen et al., 1996) to compute nearby projectively involutive polynomial systems (Scott, 2006; Reid et al., 2005). This should make our SVD based on method more reliable.

The algebraic method which also is related to our method is the method of H-bases (Möller and Sauer, 2000), which also focuses on the dimensions of the vector spaces of generated by monomials. Example 2.4 of Möller and Sauer (2000) is an H-basis of degree 4, but can be shown to become involutive only after prolongation to degree 7 (when it also becomes a Gröbner Basis). However, this H-basis is *minimally formally integrable*, in the sense defined in Reid et al. (2001, Appendix A). In future work we will investigate the relation between H-bases and minimal formal integrability which unlike H-bases applies to the more general case of differential systems. Indeed, the current paper is part of a more general symbolic-numeric investigation of approximate systems of differential equations (Reid et al., 2002). The fact that the methods apply not only to zero dimensional polynomial systems, but also to positive dimensional ones, and even more generally to systems of partial differential equations, is a favourable aspect of this approach.

Under-determined systems (i.e. positive dimension systems) can have their involutivity checked by the use of their Cartan characters (e.g. see Pommaret (1978) and Seiler (2002)); which is numerically accomplished in generic coordinates by making projections between different classes of jet variables in their symbol space. An alternative is to attempt to interpolate their Hilbert polynomial from the values $\dim\ \boldsymbol{\pi}^{\ell}\left(\mathbf{D}^{k}R\right)$. This allows the determination of the top dimensional positive dimensional components of polynomial systems. Subsequently, an appropriate random linear subspace of complementary dimension, when intersected with these components, cuts out generic points on those components (i.e. by using a variation of the methods of Numerical Algebraic Geometry (Sommese and Wampler, 2005)). Such generic points can then be calculated with the eigen-method of this paper. We plan to extend our implementation to such positive dimensional systems.

The discussion in Section 4 also can be extended to square systems of polynomials of differing degrees (Möller and Sauer (2000) consider this case). Most notably, the order of involution can be detected before prolongation. Undetermined systems defining complete intersections admit a similar treatment in the positive dimensional case, and this is worthy of detailed investigation. It is interesting to note that complete intersections, as a way of avoiding redundant polynomial calculations, has been

a key ingredient in the fastest known improvements on Gröbner type algorithms (Bardet et al., 2005; Faugère, 2002). We will investigate these connections in future work.

The methods of Wu (1991) and Sommese and Wampler (2005) are concerned with radicals of ideals, and more directly with the set of zeros of polynomial systems, than the methods we present here. Our approach maintains the structure of the ideal, and hence the multiplicity structure, in contrast to those of Wu (1991) and Sommese and Wampler (2005) which change and lose multiplicity information during their execution.

The method has been applied successfully to solve some over-determined problems such as camera pose determination in singular positions (Reid et al., 2003). Our test suite and Maple implementation are available by request.

Acknowledgements

Greg Reid was supported in part by a Canadian NSERC Grant. Lihong Zhi was supported in part by NKBRPC 2004CB318000 and Chinese NSF under Grant 10401035. We gratefully acknowledge the very helpful comments of Wenyuan Wu, especially on Section 4. We gratefully acknowledge the helpful comments of Robin Scott and the Referees.

References

Auzinger, W., Stetter, H., 1988. Polynomial factorization 1987–1991. Internat. Ser. Numer. Math. 86, 11–30.
Bardet, M., Faugère, J., Salvy, B., 2005. Asymptotic expansion of the degree of regularity for semi-regular systems of equations. In: Proc. MEGA 2005.
Bonasia, J., Lemaire, F., Reid, G., Scott, R., Zhi, L., 2004. Determination of approximate symmetries of differential equations. In: CRM Proceedings and Lecture Notes, pp. 233–249.
Bryant, R., Chern, S., Gardner, R., Goldschmidt, H., Griffiths, P., 1991. Exterior Differential Systems, vol. 18. Springer-Verlag, New York.
Corless, R.M., Gianni, P.M., Trager, B.M., 1997. The singular value decomposition for polynomial systems. In: Küchlin (Ed.), Proc. 1997 Internat. Symp. Symbolic Algebraic Comput. ISSAC'97. ACM Press, New York, pp. 96–103.
Cox, D., Little, J., O'Shea, D., 1992. Ideals, Varieties, and Algorithms. In: Undergraduate Texts in Mathematics, Springer-Verlag, New York.
Emsalem, J., 1978. Géométrie des points épais. Bull. Soc. Math. 106, 399–416.
Faugère, J., 2002. A new efficient algorithm for computing Gröbner bases without reduction to zero (F5). In: Mora, T. (Ed.), Proc. 2002 Internat. Symp. Symbolic Algebraic Comput. ISSAC'02. ACM Press, New York, pp. 75–83.
Gatermann, K., 1990. Symbolic solution of polynomial equation systems with symmetry. In: Watanabe, S., Nagata, M. (Eds.), Proc. 1990 Internat. Symp. Symbolic Algebraic Comput. ISSAC'90. ACM Press, New York, pp. 112–119.
Gerdt, V., Blinkov, Y., 1998. Involutive bases of polynomial ideals. Mathematics and Computers in Simulation 45, 519–541.
Kuranishi, M., 1957. On E.Cartan's prolongation theorem of exterior differential systems. Amer. J. Math 79, 1–47.
Lazard, D., 1983. Gröbner bases, Gaussian elimination and resolution of systems of algebraic equations. In: Proc. European Conf. Comput. Algebra. In: Lect. Notes in Comp. Sci., vol. 162. Springer, pp. 146–157.
Lemmerling, P., 1999. Structured total least squares: Analysis, algorithms and applications. Dissertation, Katholieke Universiteit Leuven, Belgium.
Lemmerling, P., Mastronardi, N., Van Huffel, S., 2000. Fast algorithm for solving the Hankel/Toeplitz structured total least squares problem. Numer. Algorithms 23, 371–392.
Macaulay, F., 1916. The Algebraic Theory of Modular Systems. In: Cambridge tracts in Math. and Math. Physics, vol. 19.
Malgrange, B., 2003. Cartan involutiveness = mumford regularity. Contemp. Math. 331, 193–205.
Möller, H., Sauer, T., 2000. H-bases for polynomial interpolation and system solving. Adv. Comput. Math. 12, 23–35.
Möller, H., Stetter, H., 1995. Multivariate polynomial equations with multiple zeros solved by matrix eigenproblems. Numer. Math. 70, 311–329.
Mourrain, B., 1996. Isolated points, duality and residues. J. Pure Appl. Algebra 117 & 118, 469–493.
Mourrain, B., 1999. A new criterion for normal form algorithms. In: Fossorier, M., Imai, H., Lin, S., Poli, A. (Eds.), AAECC, vol. 1719. Springer, Berlin, pp. 430–443.
Mourrain, B., Trébuchet, P., 2000. Solving projective complete intersection faster. In: Traverso, C. (Ed.), Proc. 2000 Internat. Symp. Symbolic Algebraic Comput. ISSAC'00. ACM Press, New York, pp. 430–443.
Mourrain, B., Trébuchet, P., 2002. Algebraic methods for numerical solving. In: Proceedings of the 3rd International Workshop on Symbolic and Numeric Algorithms for Scientific Computing, pp. 42–47.
Pommaret, J., 1978. Systems of Partial Differential Equations and Lie Pseudogroups. Gordon and Breach Science Publishers.
Reid, G., Lin, P., Wittkopf, A., 2001. Differential elimination–completion algorithms for DAE and PDAE. Stud. Appl. Math. 106 (1), 1–45.
Reid, G., Scott, R., Wu, W., Zhi, L., 2005. Algebraic and geometric properties of nearby projectively involutive polynomial systems, manuscript.
Reid, G., Smith, C., Verschelde, J., 2002. Geometric completion of differential systems using numeric-symbolic continuation. SIGSAM Bull. 36 (2), 1–17.
Reid, G., Tang, J., Zhi, L., 2003. A complete symbolic-numeric linear method for camera pose determination. In: Sendra, J. (Ed.), Proc. 2003 Internat. Symp. Symbolic Algebraic Comput. ISSAC'03. ACM Press, New York, pp. 215–223.

Rosen, J.B., Park, H., Glick, J., 1996. Total least norm formulation and solution for structured problems. SIAM J. Matrix Anal. Appl. 17 (1), 110–128.
Saito, M., Sturmfels, B., Takayama, N., 2000. Gröbner Deformations of Hypergeometric Differential Equations. Springer.
Scott, R.J., 2006. Approximate Gröbner bases. Master thesis, University of Western Ontario, Canada.
Seiler, W., 1994. Analysis and application of the formal theory of partial differential equations. Dissertation, Lancaster University, Germany.
Seiler, W., 2002. Involution — the formal theory of differential equations and its applications in computer algebra and numerical analysis. Habilitation thesis, Univ. of Mannheim, Germany.
Sommese, A.J., Wampler, C.W., 2005. The Numerical Solution of Systems of Polynomials Arising in Engineering and Science. World Scientific.
Stetter, H., 2004. Numerical Polynomial Algebra. SIAM, Philadelphia.
Sturmfels, B., 2002. Solving Systems of Polynomial Equations. In: CBMS Regional Conference Series in Mathematics, vol. 97. AMS.
Trébuchet, P., 2002. Vers une résolution stable et rapide des équations algébriques. Dissertation, Université Pierre et Marie Curie, France.
Tuomela, J., Arponen, T., 2000. On the numerical solution of involutive ordinary differential systems. IMA J. Numer. Anal. 20, 561–599.
Wittkopf, A., Reid, G., 2001. Fast differential elimination in C: The CDiffelim Environment. Comput. Phys. Comm. 139 (2), 192–217.
Wu, W., 1991. On the construction of Gröbner basis of a polynomial ideal based on Riquier-Janet theory. Syst. Sci. Math. Sci. 4, 193–207.