Investigation of a Subdivision Based Algorithm for Solving Systems of Polynomial Equations

Jürgen Garloff* and Andrew P. Smith
University of Applied Sciences / FH Konstanz
Postfach 100543
D-78405 Konstanz
Germany
fax: +49-7531-206-559
email: {garloff, smith}@fh-konstanz.de

Abstract

A method for enclosing all solutions of a system of polynomial equations inside a given box is investigated. This method relies on the expansion of a multivariate polynomial into Bernstein polynomials and constitutes a domain-splitting approach. After a pruning step, a collection of subboxes remain which undergo an existence test provided by Miranda's theorem. In this paper, the complexity of this test is reduced from $O(n!)$ to nearly $O(n^7)$. Also, some observations on the effects of preconditioning of the system and results on the application of Bernstein expansion to the mean value form are presented.

Keywords: Polynomial equations, Bernstein polynomials, Miranda Theorem, mean value form

1 Introduction

Roughly speaking, methods for solving systems of polynomial equations can be divided into three classes: techniques based on elimination theory, e.g., Ch. 8 in Winkler [1], continuation, e.g., Allgower and Georg [2, 3], and subdivision. The first two classes frequently give us more information than we need since they determine all complex solutions of the system, whereas in applications often only the solutions in a given area of interest - typically a box - are sought. In the last category we collect all methods which apply a domain-splitting approach: Starting with a box of interest, such an algorithm sequentially splits it into subboxes, eliminating boxes which cannot contain a zero, and ending up with a union of boxes that contains all solutions of the system which lie within

*Corresponding author
the given box. Methods utilising this approach include interval computation
techniques, e.g., Van Hentenryck et al. [4], as well as methods which apply the
expansion of a multivariate polynomial into Bernstein polynomials (Sherbrooke
and Patrikalakis [5], Fausten and Luther [6] and Garloff and Smith [7]).

The approach presented in Garloff and Smith [7] eliminates infeasible boxes
by using bounds for the range of the polynomials under consideration over each
box. These bounds are provided by Bernstein expansion. The remaining boxes
undergo an existence test which is based on a theorem by Carlo Miranda [8].

The present paper aims to reduce the computational complexity of this test
from $O(n^l!)$ to nearly $O(n^{2l})$, where $n$ is the dimension of the problem. Also, a
closer look of the effects of preconditioning is presented.

The organisation of this paper is as follows: In the next section we briefly
recall the Bernstein expansion from Garloff [9] and Zettler and Garloff [10],
wherein the relevant references can be found. The solution method is presented
in Section 3. The reduction of the computational complexity as well as pre-
conditioning are discussed in Section 4. In the process of writing this paper
we were also investigating a new Bernstein form which is the application of the
usual Bernstein expansion to the mean value form, cf. Neumaier [11, Sect. 2.3].
Since this Bernstein form is not further used in our paper and we do not want to
belabour the presentation, we delegate our results to the Appendix. This part
requires a basic knowledge of interval arithmetic, e.g., Neumaier [11]. Numerical
results are given in Garloff and Smith [7].

2 Bernstein Expansion

For compactness, we will use multi-indices $I = (i_1, \ldots, i_l)$ and multi-powers
$x^I = x_1^{i_1} x_2^{i_2} \cdots x_l^{i_l}$ for $x \in \mathbb{R}^l$. Inequalities $I \leq N$ for multi-indices are meant
componentwise, where $0 \leq i_k, k = 1, \ldots, l$, is implicitly understood. With
$I = (i_1, \ldots, i_{r-1}, i_r, i_{r+1}, \ldots, i_l)$ we associate the index $I_{r,k}$ given by
$I_{r,k} = (i_1, \ldots, i_{r-1}, i_r + k, i_{r+1}, \ldots, i_l)$, where $0 \leq i_r + k \leq n_r$. Also, we write $\binom{n}{I}$ for $\binom{n_1}{i_1} \cdots \binom{n_l}{i_l}$.

We then write an $l$-variate polynomial $p$ in the form

$$p(x) = \sum_{I \leq N} a_I x^I, \quad x \in \mathbb{R}^l,$$

and refer to $N$ as the degree of $p$.

2.1 Bernstein Transformation of a Polynomial

In this subsection we expand a given $l$-variate polynomial (1) into Bernstein
polynomials to obtain bounds for its range over an $l$-dimensional box. Without
loss of generality we consider the unit box $U = [0, 1]^l$ since any nonempty box
of $\mathbb{R}^l$ can be mapped affinely onto this box.
For $\mathbf{x} = (x_1, \ldots, x_l) \in \mathbb{R}^l$, the $I$th Bernstein polynomial of degree $N$ is defined as

$$B_{N,I}(\mathbf{x}) = b_{n_1,i_1}(x_1) b_{n_2,i_2}(x_2) \cdots b_{n_l,i_l}(x_l),$$

where for $i_j = 0, \ldots, n_j$, $j = 1, \ldots, l$

$$b_{n_j,i_j}(x_j) = \binom{n_j}{i_j} x_j^{i_j} (1 - x_j)^{n_j - i_j}.$$

The transformation of a polynomial from its power form (1) into its Bernstein form results in

$$p(\mathbf{x}) = \sum_{I \leq N} b_I(\mathbf{U}) B_{N,I}(\mathbf{x}),$$

where the Bernstein coefficients $b_I(\mathbf{U})$ of $p$ over $\mathbf{U}$ are given by

$$b_I(\mathbf{U}) = \sum_{J \leq I} \binom{I}{J} a_J, \quad I \leq N. \quad (2)$$

We collect the Bernstein coefficients in an array $B(\mathbf{U})$, i.e., $B(\mathbf{U}) = (b_I(\mathbf{U}))_{I \leq N}$. In analogy to Computer Aided Geometric Design we call $B(\mathbf{U})$ a patch. For an efficient calculation of the Bernstein coefficients, which does not use (2), see Garloff [12].

In the following lemma, we list some useful properties of the Bernstein coefficients, cf. Cargo and Shisha [13] and Farouki and Rajan [14].

**Lemma 1** Let $p$ be a polynomial of degree $N$. Then the following properties hold for its Bernstein coefficients $b_I(\mathbf{U})$:

i) Range enclosing property:

$$\forall \mathbf{x} \in \mathbf{U} : \min_{I \leq N} b_I(\mathbf{U}) \leq p(\mathbf{x}) \leq \max_{I \leq N} b_I(\mathbf{U}) \quad (3)$$

with equality in the left (resp., right) inequality if and only if $\min_{I \leq N} b_I(\mathbf{U})$ (resp., $\max_{I \leq N} b_I(\mathbf{U})$) is attained at a Bernstein coefficient $b_I(\mathbf{U})$ with $i_k \in \{0, n_k\}$, $k = 1, \ldots, l$.

ii) Partial derivative:

$$\frac{\partial p}{\partial x_r}(\mathbf{x}) = n_r \sum_{I \leq N \cap i_r = 1} [b_{I_{r,1}}(\mathbf{U}) - b_{I_{r,i}}(\mathbf{U})] B_{N_r - 1,I}(\mathbf{x}). \quad (4)$$

**Lemma 2** [7] Let $p$ be an $l$-variate polynomial and let $B(\mathbf{U})$ be the patch of its Bernstein coefficients on $\mathbf{U}$. Then the Bernstein coefficients of $p$ on the $m$-dimensional faces of $\mathbf{U}$ are just the coefficients on the respective $m$-dimensional faces of the patch $B(\mathbf{U})$, $0 \leq m \leq l - 1$. 

3
2.2 Sweep Procedure

The bounds obtained by the inequalities (3) can be tightened if the unit box $U$ is bisected into subboxes and Bernstein expansion is applied to the polynomial $p$ on these subboxes. i.e., to the polynomial shifted from each subbox back to $U$. A sweep in the $r$th direction $(1 \leq r \leq l)$ is a bisection perpendicular to this direction and is performed by recursively applying a linear interpolation. Let $D = [d_1, \bar{d}_1] \times \ldots \times [d_l, \bar{d}_l]$ be any subbox of $U$ generated by sweep operations (at the beginning, we have $D = U$). Starting with $B^{(0)}(D) = B(D)$ we set for $k = 1, \ldots, n_r$

$$b^{(k)}_i(D) = \begin{cases} b^{(k-1)}_i(D) : i_r < k \\ (b^{(k-1)}_{i_r+1}(D) + b^{(k-1)}_i(D))/2 : k \leq i_r. \end{cases}$$

To obtain the new coefficients, this is applied for $i_j = 0, \ldots, n_j$, $j = 1, \ldots, r - 1, r + 1, \ldots, l$. Then the Bernstein coefficients on $D_0$, where the subbox $D_0$ is given by $D_0 = [d_1, \bar{d}_1] \times \ldots \times [d_l, \bar{d}_l]$, with $\bar{d}_r$ denoting the midpoint of $[d_r, \bar{d}_r]$, are obtained as $B(D_0) = B^{(n_r)}(D)$. The Bernstein coefficients $B(D_1)$ on the neighbouring subbox $D_1$

$$D_1 = [d_1, \bar{d}_1] \times \ldots \times [d_r, \bar{d}_r] \times \ldots \times [d_l, \bar{d}_l]$$

are obtained as intermediate values in this computation, since for $k = 0, \ldots, n_r$ the following relation holds (Garloff [9]):

$$b_{i_1, \ldots, n_r-k, \ldots, i_l}(D_1) = b^{(k)}_{i_1, \ldots, n_r, \ldots, i_l}(D).$$

A sweep needs $O(n^{r+1})$ additions and multiplications, where $n = \max\{n_i : i = 1, \ldots, l\}$, cf. Zettler and Garloff [10]. Note that by the sweep procedure the explicit transformation of the subboxes generated by the sweeps back to $U$ is avoided. Fig. 1 illustrates the sweeping process for $l = 2$ and $r = 1$.

![Fig. 1. Two new patches are obtained by a sweep in the first direction](image-url)
3 The Method

Let \( n \) polynomials \( p_i, i = 1, \ldots, n \), in the real variables \( x_1, \ldots, x_n \) and a box \( Q \) in \( \mathbb{R}^n \) be given. We want to know the set of all solutions of the equations \( p_i(x) = 0, i = 1, \ldots, n \), within \( Q \). Without loss of generality we can assume that \( Q \) is the unit box.

Our procedure is very simple: We take away from \( Q \) all subboxes generated by sweeps for which there is a polynomial \( p_i \) being (strictly) positive or negative over the subbox. We check the sign of the polynomials by their Bernstein coefficients according to Lemma 1: If all Bernstein coefficients of a polynomial \( p_i \) are either positive or negative over a box, this box cannot contain a solution.

After this pruning step we end up with a set of boxes of sufficiently small volume. All these boxes now undergo an existence test. In a first attempt we exploit the existence test given by Miranda [8] which provides a generalisation of the fact that if a univariate continuous function \( f \) has a sign change at the endpoints of an interval then this interval contains a zero of \( f \):

\[ f_i(x) f_i(y) \leq 0 \mbox{ for all } x, y \in X_i^-, y = X_i^+, i = 1, \ldots, n, \]  

(5)

then the equation \( F(x) = 0 \) has a solution in \( X \).

We employ a heuristic sweep direction selection rule in an attempt to minimise the total number of subboxes which need to be processed. Such a rule may favour directions in which polynomials have large partial derivatives and in which the box edge lengths are larger, to avoid repetitive sweeps in a single direction. For a discussion of some selection rule variants and numerical examples see Garloff and Smith [7].

4 Reduction of Computational Cost and Preconditioning

The results of this section are valid for general continuous functions.

Kioustelidis [15], cf. Zuhle and Neumaier [16], argued that the system \( F(x) = 0 \) should be preconditioned, i.e., it should be replaced by \( AF(x) = 0 \) with a suitably chosen matrix \( A \). If \( F \) is differentiable on \( X \) and the Jacobian \( F' \) of \( F \) is nonsingular at \( x \), the midpoint of \( X \), i.e.,

\[ \hat{x}_i := (x_i + \bar{x}_i)/2, \quad i = 1, \ldots, n, \]

then

\[ F'(\hat{x}) \]  

is a well-conditioned approximation of the Jacobian.
a reasonable choice for $A$ is to take an approximation to $F'(\bar{x})^{-1}$. If we apply Miranda’s theorem to the given polynomial system and use Bernstein expansion we can then make use of the easy calculation of the Bernstein form of the partial derivatives of a polynomial from its Bernstein form, cf. (4). Furthermore, the test required in (5) costs nearly nothing since the Bernstein coefficients of $p$ on the faces of $X$ are known once the Bernstein coefficients of $p$ on $X$ are computed, cf. Lemma 2.

In the worst case, the test required in (5) has to be performed $n!$ times. To reduce the computational cost of the Miranda test we start with the following partial uniqueness result: Suppose that the Miranda test for the component function $f_i$, $i \in \{1, \ldots, n\}$, is successful for some $k \in \{1, \ldots, n\}$, i.e.,

$$f_i(x)f_i(y) \leq 0 \quad \forall x \in X_k^-, \forall y \in X_k^+. \quad (6)$$

Assume that there exist $j \in \{1, \ldots, k-1, k+1, \ldots, n\}$ and $X_k^+ \in \{X_k^-, X_k^+\}$ with

$$f_i(x) \neq 0 \quad \text{on } X_k^+ \cap (X_j^- \cup X_j^+). \quad (7)$$

In the polynomial case, this condition can easily be checked by inspection of the Bernstein coefficients on two $(n-2)$-dimensional subpatches of $B(X)$, cf. Lemma 2. It follows then from (6) that

$$\exists v \in X_k^+ \cap X_j^-, \exists w \in X_k^+ \cap X_j^+ : f_i(v)f_i(w) > 0,$$

so that the Miranda test fails for $f_i$ on the pair $(X_j^-, X_j^+)$. So, if condition (7) is fulfilled for all other pairs $(X_j^-, X_j^+)$, with $j \neq k$, then the pair $(X_k^-, X_k^+)$ is the only one for which the Miranda test will be successful.

However, condition (7) may not be fulfilled. A simple example is provided by the two polynomials

$$p_1(x_1, x_2) = x_1^2 + x_2^2 - 1,$$

$$p_2(x_1, x_2) = x_1 - x_2.$$

In Fig. 2 the zero sets of both polynomials are displayed. Choosing $X = U$, $p_1$ has a sign change on both pairs of opposite faces, but attains the value zero on all four face intersections (7).

The preconditioned system is

$$q_1(x_1, x_2) = \frac{1}{2} \left\{ (x_1 + \frac{1}{2})^2 + (x_2 - \frac{1}{2})^2 - \frac{3}{2} \right\},$$

$$q_2(x_1, x_2) = \frac{1}{2} \left\{ (x_1 - \frac{1}{2})^2 + (x_2 + \frac{1}{2})^2 - \frac{3}{2} \right\}.$$

The zero sets of the two polynomials are displayed in Fig. 3. Both polynomials have a sign change on one pair of opposite faces and condition (7) is now satisfied for both.
If condition (7) is fulfilled — which is likely the case after preconditioning — the computational cost of the Miranda test is reduced to an $O(n^2)$ complexity: Define an $n \times n$ matrix $M = (m_{ij})$ by

$$m_{ij} := \begin{cases} 1 & \text{if the test of } p_i \text{ on } (X^-_j, X^+_j) \text{ succeeds} \\ 0 & \text{fails.} \end{cases}$$

As soon as we find any zero row or column in $M$, we may terminate since no successful permutation is possible. If each row and column of $M$ contain a 1, we may terminate since there is a successful permutation, and we can conclude that the box under consideration contains a solution of the system.

We conclude this section with a comment on preconditioning. Let $G(x) := A F(x)$ be the preconditioned function, where $A := F'(\hat{x})^{-1}$ and $\hat{x}$ is an approximation to $x^*$, a regular solution to $F(x) = 0$ and therefore also to $G(x) = 0$. Zuhre and Neumaier [16] have shown that then

$$G(x) = x - x^* + o(\varepsilon)$$

holds for all $x \in \mathbb{R}^n$ with $\|x - x^*\| \leq \varepsilon$, $\varepsilon$ small. This implies that

$$g_i(x) < 0 \quad \forall x \in X^-_i \quad \text{and} \quad g_i(x) > 0 \quad \forall x \in X^+_i, \quad i = 1, \ldots, n,$$

(8)

holds for any box $X$ with $[\underline{x}_i, \overline{x}_i] = [\hat{x}_i - d_i, \hat{x}_i + d_i]$, $i = 1, \ldots, n$, where $d$ is a vector with

$$|\hat{x}_i - x^*_i| < \frac{d}{2} \quad i = 1, \ldots, n.$$

(9)
||d|| sufficiently small.

Condition (9) means that we require

\[ x^* \in X_{cent} := \frac{3x_1 + x_1}{4} \times \frac{x_1 + 3x_1}{4} \times \ldots \times \frac{x_n + 3x_n}{4} \times \frac{x_n + 3x_n}{4}. \]

In general, we have \( x^* \notin X_{cent} \). Although a sequence of sweep directions which eventually yields a box \( X \) with \( x^* \in X_{cent} \) is achievable, given knowledge of \( x^* \), we generally do not choose sweep directions in this fashion.

If (8) holds, then the (problematic) case that \( G \) vanishes on the boundary of \( X \) is avoided, so that condition (7) is fulfilled. Then the Miranda test is not only guaranteed to succeed, but succeed for the identity permutation. In this case of course we do not need to check all permutations, yet the complexity is not reduced, since the preconditioning requires a matrix inversion.

**Acknowledgement**

Support from the Ministry of Education, Science, Research, and Technology of the Federal Republic of Germany under contract no.1706998 is gratefully acknowledged.

**Appendix: The Mean Value Bernstein Form**

In interval computations, often the mean value form, e.g., Neumaier [11, Sect. 2.3], is used to find an enclosure for the range \( p(X) \) of a function \( p \) over an
interval $X$ (for simplicity, we consider here only the univariate case). In deriving this form, the mean value theorem is used which tells us that

$$p(x) = p(\bar{x}) + p'(\xi)(x - \bar{x}).$$

where $\xi \in X$ lies between $x$ and the midpoint $\bar{x}$ of $X$. To compute an enclosure for $p(X)$ we replace $x$ on the right hand side of (10) by $X$ and have to find an enclosure for $p'(X)$. If $p$ is a polynomial we can apply Bernstein expansion to $p'$, where we profit again from the easy calculation of the Bernstein form of $p'$ from the Bernstein form of $p$, cf. (4). We call the resulting form (specified here with $X = [0, 1]$, for simplicity)

$$p(\frac{1}{2}) + RF[p'][-\frac{1}{2}, \frac{1}{2}]$$

the mean value Bernstein form which encloses $p([0, 1])$. Here $RF[p']$ denotes the interval spanned by the minimum and maximum of the Bernstein coefficients of $p'$, cf. (3). However, the mean value Bernstein form does not yield any improvement on the usual Bernstein form:

**Theorem 2** Let $p$ be a univariate polynomial of degree $n$ with its Bernstein coefficients $b_i$, $i = 0, \ldots, n$. If the lower or upper bound for $p([0, 1])$ provided by the Bernstein coefficients is not sharp, then the width of this enclosure is strictly less than that of the mean value Bernstein form.

**Proof:** Using (4), we can write (11) as

$$p(\frac{1}{2}) + \frac{n}{2} \max_{i=0, \ldots, n-1} |b_{i+1} - b_i| [-1, 1].$$

Thus the width of the mean value Bernstein form is

$$n \max_{i=0, \ldots, n-1} |b_{i+1} - b_i|.$$  \hspace{1cm} (12)

Let

$$b_l := \min_{i=0, \ldots, n} \{b_i\} \quad \text{and} \quad b_u := \max_{i=0, \ldots, n} \{b_i\}.$$ 

Assume w.l.o.g. that $l < u$. If either bound is non-sharp, we have $n - l \leq n - 1$. The width of the enclosure provided by the usual Bernstein form can be bounded as follows:

$$b_u - b_l = (b_u - b_{u-1}) + (b_{u-1} - b_{u-2}) + \ldots + (b_{l+2} - b_{l+1}) + (b_{l+1} - b_l) \leq |b_u - b_{u-1}| + |b_{u-1} - b_{u-2}| + \ldots + |b_{l+2} - b_{l+1}| + |b_{l+1} - b_l| \leq (u - l) \max_{i=0, \ldots, n-1} |b_{i+1} - b_i| \leq (n - 1) \max_{i=0, \ldots, n-1} |b_{i+1} - b_i|.$$ 

Comparison with (12) concludes the proof.
References


