ABSTRACT
A local generic position method is proposed to isolate the real roots of a bivariate polynomial system
\[ \Sigma = \{ f(x, y), g(x, y) \}. \]
In this method, the roots of the system are represented as linear combinations of the roots of two univariate polynomial equations \( t(x) = 0 \) and \( T(X) = 0 \):
\[ x = \alpha, y = \beta - \alpha | \alpha \in V(t(x)), \beta \in V(T(X)), |\beta - \alpha| < S, \]
where \( s, S \) are constants satisfying certain conditions. The multiplicities of the roots of \( \Sigma = 0 \) are the same as that of the corresponding roots of \( T(X) = 0 \). This representation leads to an efficient and stable algorithm to isolate the real roots of \( \Sigma \).

Categories and Subject Descriptors
I.1.2 [SYMBOLIC AND ALGEBRAIC MANIPULATION]: Algorithms—Algebraic algorithms

General Terms
Algorithm

Keywords
Bivariate polynomial system, generic position, root isolation, root bound.

1. INTRODUCTION
Solving polynomial equation systems is a fundamental problem in symbolic computation. In this paper, we consider the problem of real root isolation for bivariate polynomial equation systems. Let \( f(x, y), g(x, y) \in \mathbb{Q}[x, y] \), where \( \mathbb{Q} \) is the field of rational numbers. We call
\[ \Sigma = \{ f(x, y), g(x, y) \}. \]

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
Copyright 2009 ACM 978-1-60558-609-0/09/07 ...$10.00.
the Gröbner basis method, the resultant method, the characteristic set method, and the subdivision based method. Here we compare our method with those that are devoted to bivariate equation systems.

In [7], Diochnos, Emiris, and Tsigaridas gave three algorithms to solve bivariate equation systems and analyzed their complexities. Algorithm GRID projects the roots of \( \Sigma \) to the \( x \) and \( y \) axes and checks whether a combination of the \( x \)- and \( y \)-coordinates is a solution of \( \Sigma \). Assuming the equation system in generic position, the algorithm MRUR uses signed subresultant sequences to compute an RUR like (2) and finds the solution of the system by estimating the value of the rational functions \( R(x) \) at the \( x \)-coordinates of the roots of \( \Sigma \). This method is quite similar to the method given in [13]. The GRUR method projects the roots of the system to the \( x \) and \( y \) axes, for each \( x \)-coordinate \( \alpha \) computes the GCD \( H(\alpha, y) \) of the square-free parts of \( f(\alpha, y) \) and \( g(\alpha, y) \), and isolates the roots of \( H(\alpha, y) = 0 \) based on computations of algebraic numbers and the RUR techniques. Among the three algorithms, GRUR has the lowest complexity and performs best in experiments. Our algorithm only uses resultant computation and root isolation for univariate polynomial equations with rational coefficients. Our algorithm totally avoids computation over algebraic numbers and is more efficient than GRUR as shown by experimental results in Section 5.

The method by Hong, Shan, and Zeng [14] projects the roots of \( \Sigma \) to the \( x \)-axis and \( y \)-axis respectively and uses a numerical iteration method to decide whether the boxes formed by the projection intervals contain a root of \( \Sigma \). The numerical method works for simple roots of \( \Sigma \) only. When the system has multiple roots, the RUR technique is used to isolate them. Comparing to this method, our method also computes two resultants of the same total degree. Our method is a complete one, while the method given in [14] needs to use the RUR technique to find multiple roots.

The rest of this paper is organized as follows. In Section 2, the theory behind the local generic position method is presented. In Section 3, we estimate the bounds needed in the algorithm. In Section 4, we give the local generic position algorithm. Experimental results are presented in Section 5 and conclusions are given in Section 6.

2. LOCAL GENERIC POSITION

In this section, we present the theory behind the local generic position algorithm. The idea is to do a shear transformation \( (x, y) \to (x + s y, y) \) so that the new equation system is in a “local generic position” with respect to the original equation system.

Let \( \pi \) be the projection map from the real plane to the \( x \)-axis:

\[
\pi : \mathbb{R}^2 \longrightarrow \mathbb{R}, \quad \text{such that } \pi(x, y) = x. \tag{3}
\]

For a zero-dimensional system \( \Sigma = \{f(x, y), g(x, y)\} \) defined in (1), let \( t(x) \in \mathbb{Q}[x] \) be the resultant of \( f(x, y) \) and \( g(x, y) \) w.r.t \( y \):

\[
t(x) = \text{Res}_y(f(x, y), g(x, y)). \tag{4}
\]

Since \( \Sigma \) is zero-dimensional, we have \( t(x) \neq 0 \). Then \( \pi(V(\Sigma)) \subseteq V(t(x)) \), where \( V(f_1, \ldots, f_n) \) is the set of common real zeros of \( f_1, \ldots, f_n \) is. Let the real roots of \( t(x) = 0 \) be

\[
\alpha_1 < \alpha_2 < \cdots < \alpha_m. \tag{5}
\]

Using the notations in (1) and (5), let \( S, R, \) and \( s \) be rational numbers satisfying

\[
S < \frac{1}{2} \min(\alpha_{i+1} - \alpha_i, i = 1, \ldots, m - 1), \quad R > \max(|\beta|, \forall (\alpha, \beta) \in V(\Sigma)), \tag{6}
\]

\[
0 < s < \frac{S}{R}.
\]

For \( s \) satisfying (6), define an inversive linear map (a shear) from \( \mathbb{R}^2 \) to \( \mathbb{R}^2 \):

\[
\psi_s : (x, y) \longmapsto (X, Y) = (x + s y, y). \tag{7}
\]

We also define \( \psi_s(f(x, y)) = f(X - s Y, Y) \) for convenience. Geometrically, \( \psi_s \) maps a point \((x_0, y_0)\) to the intersection point of the lines \( y = y_0 \) and \( (x - x_0) = s y \). See Fig. 1 for an illustration.

![Figure 1: Map \( \psi_s \): the red squares are the roots of \( \Sigma = 0 \); the blue triangles are the roots of \( \psi_s(\Sigma) \); the black dots are the roots of univariate polynomial \( T(X) = 0 \).](image)

An equation system \( \Sigma \) is said to be in a generic position if different roots of \( \Sigma = 0 \) have different \( x \)-coordinates. We have:

**Lemma 2.1.** For \( S \) and \( R \) defined in (6) and \( \psi_s \) defined in (7), \( \psi_s(\Sigma) \) is in a generic position. Furthermore, a root \((\alpha, \beta)\) of \( \Sigma = 0 \) is mapped to \((\eta, \beta)\) where \( \eta \in (\alpha - S, \alpha + S) \). See Fig. 1 for an illustration.

**Proof.** For each \( \alpha_i \) in (5), let \( P_{i,j} = (\alpha_i, \beta_{i,j}) \) be the corresponding roots of \( \Sigma = 0 \). We have \( \psi_s(P_{i,j}) = (\alpha_i + s \beta_{i,j}, \beta_{i,j}) \). Then, for the same \( i \), different \( \psi_s(P_{i,j}) \) have different \( x \)-coordinates. Due to the conditions in (6), we have \( |\alpha_i + s \beta_{i,j} - \alpha_i| = |s \beta_{i,j}| < (S/R) \cdot R = S \). That is, \( \psi_s(P_{i,j}) \in R_i = (\alpha_i - S, \alpha_i + S) \times [-R, R] \). Since \( S < \frac{1}{2}(\alpha_{i+1} - \alpha_i) \), \( R_i \) are disjoint for different \( i \). Then different \( \psi_s(P_{i,j}) \) have different \( x \)-coordinates. This proves the lemma.

We project the roots of \( \psi_s(\Sigma) \) to the \( x \)-axis by computing the resultant \( T(X) \):

\[
T(X) = \text{Res}_y(\psi_s(f(x, y)), \psi_s(g(x, y))) = \text{Res}_y(f(X - s Y, Y), g(X - s Y, Y)). \tag{8}
\]

We hope that the zeros of \( \Sigma = 0 \) and the roots of \( T(X) \) are in a one-to-one correspondence. This may fail when

\[
h(X) = \text{gcd}(\text{LC}_Y(\psi_s(f(x, y))), \text{LC}_Y(\psi_s(g(x, y)))) \tag{9}
\]

has real roots, where \( \text{LC}_Y(f(X, Y)) \) is the leading coefficient of \( f(X, Y) \) w.r.t \( Y \).
By the property of the resultant, where $\psi(\Sigma) = 0$ and the real roots of $\Sigma = 0$ are in one-to-one correspondence. We can select the parameter $s$ properly so that $\text{LC}(\psi_s(f(x,y)))$ or $\text{LC}(\psi_s(g(x,y)))$ is a constant and $h(X) = 0$ has no real roots.

Write $f(x,y)$ and $g(x,y)$ as the sum of their homogeneous parts:

$$f(x,y) = f_p(x,y) + \cdots + f_0$$
$$g(x,y) = g_q(x,y) + \cdots + g_0$$

where $f_i$ and $g_i$ are homogeneous polynomials with total degree $i$. It is clear that when

$$f_p(-s,1) \neq 0 \text{ or } g_q(-s,1) \neq 0,$$  \hspace{1cm} (10)

$h(X)$ is a constant. It is always possible to choose an $s$ such that (6) is satisfied. Then we further have

**Lemma 2.2.** Let $s$ be a rational number satisfying (6) and (10). Then $\pi$ is a one-to-one and multiplicity preserving map between the roots of $\psi_s(\Sigma)$ and the roots of $T(X) = 0$, where $T(X)$ is defined in (8).

Furthermore, let the roots of $T(X) = 0$ in $(a_i - S, a_i + S)$ be

$$\beta_{i,1} < \beta_{i,2} < \cdots < \beta_{i,m}, i = 1, \ldots, m$$  \hspace{1cm} (11)

where $\alpha_i$ is defined in (5). Then the inversion of the $\pi$ is:

$$\pi^{-1}(\beta_{i,j}) = (\beta_{i,j} - \alpha_i)/s.$$  \hspace{1cm} (12)

**Proof.** By the property of the resultant, $\pi(V(\psi_s(\Sigma)) \subseteq V(T(X))$. By Lemma 2.1, we can derive that different roots of $\psi_s(\Sigma) = 0$ are mapped to different roots of $T(X) = 0$. Furthermore, by (10), the leading coefficient of $\psi_s(f)$ or $\psi_s(g)$ does not vanish. Then by the property of the resultant, $\pi(V(\psi_s(\Sigma)) = V(T(X))$. Hence $\pi$ is one-to-one. Based on the theory in Section 1.6 of [9], we can conclude that $\pi$ is also multiplicity preserving. $\pi^{-1}(\beta_{i,j})$ can be obtained as follows. By the proof of Lemma 2.1, a root $Q_{i,j} = (\beta_{i,j}, \gamma_{i,j})$ of $\psi_s(\Sigma) = 0$ is projected one-to-one to a root of $T(X) = 0$ in $(a_i - S, a_i + S)$. Then, from the definition of $\psi_s Q_{i,j}$ is on the line defined by $(x - \alpha_i) = sy$ (the skewed lines in Fig. 1). Then, we have $\gamma_{i,j} = (\beta_{i,j} - \alpha_i)/s$.

The following result shows how to recover the roots of $\Sigma = 0$ from the roots of two univariate polynomial equations $t(x) = 0$.

**Theorem 2.3.** Use the notations introduced in this section. If $\pi(\psi_s(\Sigma)) \subseteq V(T(X))$. Furthermore, the roots of $\Sigma = 0$ can be obtained by the inversion of $\pi$:

$$\theta^{-1}(\beta_{i,j}) = (\alpha_i, (\beta_{i,j} - \alpha_i)/s), |\beta_{i,j} - \alpha_i| < S, i = 1, \ldots, m, j = 1, \ldots, m_i$$  \hspace{1cm} (13)

where $\alpha_i \in V(t(x)), \beta_{i,j} \in V(T(X))$ are defined in (5) and (11) respectively, $t(x), T(X)$ are defined by (4) and (8), respectively.

**Proof.** Since $\psi_s$ is an inverse linear map, it is one-to-one and multiplicity preserving. Then by Lemma 2.2, $\theta$ is also one-to-one and multiplicity preserving. The inversion map $\theta^{-1} = \psi_s^{-1} \circ \pi^{-1}$ can be obtained directly from (12) and (7).

As corollaries of Theorem 2.3, we have

**Corollary 2.4.** Under the same condition of Theorem 2.3, we have

$$V(\Sigma) = \{(\alpha, (\beta - \alpha)/s) | (\alpha, \beta) \in V(t(x)), \beta \in V(T(X))$$

and $|\beta - \alpha| < S\}$.

Due to (14), if $|\alpha - \beta| < S$, we say that $\beta$ is associated with $\alpha$.

**Corollary 2.5.** If we separate the real roots of $t(x) = 0$ and $T(X) = 0$ with precisions $\rho_1$ and $\rho_2$ respectively, then the roots computed with (14) have precision $\max\{\rho_1, \rho_2\}$.

From Theorem 2.3, the four-tuple

$$\{t(x), T(X), s\}$$

provides a representation for the roots of $\Sigma = 0$, and from this representation, we can compute the roots of $\Sigma$ by solving two univariate equations. This method is called a local generic position method because the roots of $\Sigma = 0$ with the same $x$-coordinate $a$ are mapped to $(a - S, a + S)$ and can be recovered with a linear map (14). This makes the precision control much easier than the usual generic position method where the roots of $\Sigma = 0$ are represented as a univariate rational function of the roots of $T(X) = 0$.

**3. ESTIMATION OF PARAMETERS $S, R$**

From Section 2, we need to know the values of the parameters $S, R$, and $s$ defined in (6) in order to transform the equation system into a local generic position. In this section, we will show how to compute such parameters efficiently.

We can use the general root bounds for zero dimensional equation systems in [20, p. 341] to estimate $S$ and $R$. But the results obtained in this way is far from optimal. In this section, we will show how to obtain better estimations for $S, R$, and $s$.

We will use intervals to isolate the roots of a univariate equation. Let $I \subseteq \mathbb{Q}$ denote the set of intervals of the form $[a, b]$ where $a < b \in \mathbb{Q}$. The length of an interval $I = [a, b] \subseteq \mathbb{Q}$ is defined to be $|I| = b - a$. A set $\mathcal{B}$ of disjoint intervals is called isolation intervals for the roots of $t(x) = 0$ if each root of $t(x) = 0$ is in an interval in $\mathcal{B}$ and each interval in $\mathcal{B}$ contains one root of $t(x) = 0$.

Let $t(x)$ be defined in (4) and the isolating intervals for the roots of $t(x) = 0$ be

$$\mathcal{B} = \{[a_1, b_1], \ldots, [a_m, b_m]\}.$$  \hspace{1cm} (16)

We can directly estimate $S$ from the isolating intervals for the roots of $t(x) = 0$:

$$S = \frac{1}{2}\min\{a_{i+1} - b_i, i = 1, \ldots, m - 1\}.$$  \hspace{1cm} (17)

Then we have

**Lemma 3.1.** Let $a_i$ be the roots of $t(x) = 0$ and $[a_i, b_i]$ the isolation interval for $a_i$. If $S$ is taken as (17), then the roots of $T(X) = 0$ associated with $a_i$ are in the intervals $(a_i - S, a_i + S), i = 1, \ldots, m$.

**Proof.** It is clear that the $S$ defined in (17) satisfies (6). By Lemma 2.1, roots of $T(X) = 0$ are in $(a_i - S, a_i + S)$ for some $i$. Since $a_i \in [a_i, b_i]$, the roots of $T(X) = 0$ associated with $a_i$ must be in $(a_i - S, a_i + S)$.

A simple way to estimate $R$ is as follows:

$$R = \text{RB}(h(y)), \text{ where } h(y) = \text{Res}_x(f(x, y), g(x, y))$$  \hspace{1cm} (18)

and $\text{RB}(h(y))$ is the root bound of $h(x)$. If $h(x) = c_0y^d + \cdots + c_d$, then $\text{RB}(h(y))$ can be taken as $1 + \max\{|c_1|, \ldots, |c_d|/|c_0|$ (page 322, [3]). In this method, we need to compute a resultant. When the degrees of $f$ and $g$ in $x$ are low, we can use this approach. Otherwise, we can avoid the resultant computation by using the concept of sieve functions (see [6, 16] for details). We will explain this approach below.

Given $f \in \mathbb{Q}[x, y]$, we decompose it uniquely as $f = f^+ - f^-$, where each $f^+, f^- \in \mathbb{Q}[x, y]$ has positive coefficients and with minimal number of monomials. Given an isolating interval $I = [a, b]$ for a root $\alpha$ of a univariate equation $t(x) = 0$, we define

$$f^+_I(y) = f^+(b, y) - f^-(a, y) \in \mathbb{Q}[y],$$
$$f^-_I(y) = f^+(b, y) - f^-(a, y) \in \mathbb{Q}[y],$$  \hspace{1cm} (19)
where \( \bar{C} \) o n s i d e r t h e c a s e t h a t \\
\( f(x,y) = f(a,y) \) for \( -a \) in \([-b, -a] \) when \( a, b < 0 \). When considering \( y \geq 0 \), the following result is clearly true (Fig. 2).

**Lemma 3.2.** We have \( f_y^i(y) \leq f(x,y) \leq f_y^i(y) \), or equivalently, \( f(x,y) \in \mathbb{R}(I,y) \). Furthermore, when \( |I| \) approaches zero, the interval \( \mathbb{R}(I,y) \) converges to \( f(x,y) \) for each \( y \).

We can use the sleeve to estimate the root bound \( R \).

**Lemma 3.3.** For \( \Sigma = \{ f(x,y), g(x,y) \} \), let \( I_i = [a_i, b_i], i = 1, \ldots, m \) be the intervals defined in (16). If \( a_i \geq 0, b_i \geq 0, f_y^i(y), f_{x}^i(y) \) have the same degree in \( y \), and their leading coefficients in \( y \) have the same sign, then we can take

\[
R = \max \{ RB(f_y^i), RB(f_{x}^i) \}, i = 1, \ldots, m \quad (20)
\]

where \( f = f(x,-y) \).

Proof. Consider the case that \( d_i = \deg(f_y^i(y), y) = \deg(f_{x}^i(y), y) \) is odd and the leading coefficients of \( f_y^i(y) \) and \( f_{x}^i(y) \) are positive. Other cases can be treated similarly. Then, there exists a positive number \( r_1 \) such that \( f_y^i(y) > 0 \) for \( y > r_1 \). By Lemma 3.2, we have \( f(x,y) > f_y^i(y) > 0 \) for \( y > r_1 \). Then, the largest positive root of \( f(x,y) = 0 \) is bounded by \( RB(f_y^i(y)) \). See Fig. 2 for an illustration.

Note that \( c^a \) and \( c^d \) are the leading coefficients of \( f_y^i(y) \) and \( f_{x}^i(y) \), then the leading coefficients of \( f_y^i(y) \) and \( f_{x}^i(y) \) are \( c^a \) and \( c^d \), respectively, and \( c^d - c^a = 0 \). Then, there exists a positive number \( r_2 \) such that \( f_y^i(y) > 0 \) for \( y > r_2 \). By Lemma 3.2, we have \( f(x,y) > f_y^i(y) > 0 \) for \( y > r_2 \). Then, the largest positive root of \( f(x,y) = 0 \) is bounded by \( RB(f_y^i(y)) \).

2. If \( a_1 < 0 \), do a translation \( x := x + a_1 \) and still use \( t(x) \), \( f(x,y) \), and \( I_i = [a_i, b_i] \) to denote the translated polynomials and intervals.

3. Write \( f = f(x,y)^d + F_{d-1}(x,y)^{d-1} + \cdots + F_0(x,y) \). We assume that \( t(x) \) is not a factor of \( F(x) \); otherwise, we may remove \( F(x,y)^d \) from \( f \) since we have \( t(x) = 0 \).

4. For each root \( I \in \text{BS} \), let \( \alpha \in I \) be the root of \( t(x) = 0 \) in \( I \). Then \( F(\alpha) \neq 0 \). We assume that \( p = F_{d-1}^i - F_d^i \), and \( F_{d-1}^i \) and \( F_d^i \) are computed with (19); otherwise we repeatedly subdivide \( I \) and still denote \( I \) as the new interval containing \( \alpha^1 \) until \( p > 0 \).

5. As a consequence, \( f_y^i(y) \) and \( f_{x}^i(y) \) have same degree and their leading coefficients have the same sign. Then, by Lemma 3.3, we compute \( R \) according to (20).

**Proof of the correctness.** The correctness is obvious. We just need to show that Step 4 will terminate when we subdivide \( I \). By Lemma 2.2, the coefficients of \( f_y^i(y) \) and \( f_{x}^i(y) \) can approximate the coefficients of \( f(x,y) \) as close as we want. Since \( F(\alpha) \neq 0 \), when \( I \) is sufficiently subdivided, \( F_{d-1}^i - F_d^i > 0 \). And the program will terminate.

Now, we show how to compute \( s \) which satisfies (6) and (10). One way to do this is as follows.

**Lemma 3.5.** Let \( d = \deg(f(x,y)) \) and \( S, R \text{ rational numbers satisfying} \ (6). \text{ Then one of } s_i = \frac{(4d+2i+1)}{4(d+2)}, i = 1, \ldots, d+1 \text{ must satisfy} \ (10) \text{ and thus can be used as } s. \)

Proof. Each \( s_i \) satisfies (6). Since \( f_a(x,y) \) is homogeneous and is of total degree \( d \), \( f_a(x,1) = 0 \) can have at most \( d \) roots. Then, one of the \( s_i \) must satisfy \( f_a(x,-s_i) \neq 0. \)

Ideally, we want the bitsize of \( s \) to be as small as possible to make the computation of \( T(X) \) easier. For instance, \( \frac{1}{3} \) is much better than \( \frac{1}{1000004} \). As a heuristic, we may take \( s \) in \( Q \) satisfying (10) and with the smallest bitsize.

**4. ROOT ISOLATION OF BIVARIE POLYNOMIAL SYSTEMS**

In this section, we will present the local generic position method for real root isolation. We first find the parameters \( R, S, \) and \( s \), then obtain \( T(x) \) with (8), and finally isolate the real roots of the equation system with (14) by isolating the real roots of \( t(x) = 0 \) and \( T(X) = 0 \).

Let \( \mathbb{Q}^2 \) be the set of interval boxes of the form \( [a, b] \times [c, d] \) where \( [a, b], [c, d] \in \mathbb{Q} \). The length of an interval box \( B = [a, b] \times [c, d] \in \mathbb{Q}^2 \) is defined to be \( |B| = \max(b-a, d-c) \).

Let \( \Sigma = \{ f(x,y), g(x,y) \} \) and \( \xi = (\xi_1, \xi_2) \) be a root of \( \Sigma = 0 \). Then an interval box \( B = [a_1, b_1] \times [c_1, d_1] \in \mathbb{Q}^2 \) is called an isolation boxes of \( \xi \) if \( \xi \in (a_1, b_1) \times (c_1, d_1) \) and is the only root of \( \Sigma = 0 \) in \( B \). A set \( \text{BS} \) of disjoint interval boxes is called isolation boxes for \( \Sigma = 0 \) if each real root of \( \Sigma = 0 \) is in a box in \( \text{BS} \) and each box in \( \text{BS} \) contains one root of \( \Sigma = 0 \). A set of root isolation boxes of \( \Sigma = 0 \) is called \( \epsilon \)-isolation boxes if each box has size smaller than a given positive number \( \epsilon \).

In this section, we will present an algorithm to compute a set of \( \epsilon \)-isolation boxes for \( \Sigma = \{ f(x,y), g(x,y) \} \).

In Theorem 2.3, roots of \( \Sigma = 0 \) are represented by algebraic numbers. In the following, we will give an interval version of this result, which leads to an algorithm directly.

Let the isolation boxes for \( \alpha_{i1} \) in (5) and \( \beta_{i1} \) in (11) be

\[
B = \{ [\alpha_1, \beta_1], \ldots, [\alpha_m, \beta_m] \} \quad (21)
\]

\[
B_i = \{ [c_{i1}, d_{i1}], \ldots, [c_{im}, d_{im}] \}, i = 1, \ldots, m,
\]

respectively. Theorem 4.1 shows how to compute isolation boxes for \( \Sigma = 0 \).

3This can be easily done since \( t(x) \) is irreducible.
Theorem 4.1. Let $\epsilon$ be a positive number and $s$ a number satisfying (6) and (10). If the intervals in (21) satisfy

\[ b_i - a_i < \epsilon \quad \text{and} \quad b_i - a_i + d_{i,j} - c_{i,j} < s, \]

\[ c_{i,j+1} - d_{i,j} = 1, \ldots, m, \]

then a set of $\epsilon$-isolation boxes for the roots $P_{i,j} = (a_{i,j}, b_{i,j})$ of $\Sigma = 0$ are

\[ B_{i,j} = [a_{i,j}, b_{i,j}] \times [(c_{i,j} - b_i)/s, (d_{i,j} - a_i)/s]. \]  

Proof. Since $[a_{i,j}, b_{i,j}]$ is an isolation interval of $a_i$ and $[c_{i,j}, d_{i,j}]$ is an isolation interval of $b_i$, from the basic rules of interval computation, we have $P_{i,j} = (a_{i,j}, b_{i,j})$. See Fig. 3 for an illustration. We need only to show that $B_{i,j} < \epsilon$ and $B_{i,j}$ are disjoint.

From (22), we have $B_{i,j} = [\max(b_i - a_i, d_{i,j} - a_i)/s - (c_{i,j} - b_i)/s, \max(\epsilon, b_i - a_i + d_{i,j} - c_{i,j})/s] < \epsilon$. Finally, we will show that different $B_{i,j}$ are disjoint. The isolating box for a root $P_{i,j}$ of $\Sigma = 0$ is in a unique $\prod_{\beta}$. The isolating box for a root $P_{i,j}$ of $\Sigma = 0$ is formed based on (24).

Remark 4.4. 1. In the algorithm, we may need to isolate the roots of $t(x) = 0$ twice. When isolating their roots in the second time, we need only subdivide the existing intervals. It is not necessary to start the isolation procedure from scratch.

2. An advantage of this method is that we need only to isolate the roots of $T(X) = 0$ associated with a root $\alpha_i$ of $r(x) = 0$ in $(a_i - S, b_i + S)$. From Steps 6 and 7, we know that condition (23) is also valid. From Steps 3 and 4, it is clear that conditions (6) and (10) are satisfied. Then, Theorem 4.1 can be used to compute the isolation boxes. What we need to do is to choose those $[p_j, q_j]$ which are associated with a given $[a_i, b_i]$, which is the purpose of Step 8.

3. The most time consuming step of the algorithm is Step 5 and Step 6. There are two reasons for this. First, the shear transformation changes a spare polynomial into a dense one. Second, if the bitsize of $s$ is large, the coefficients of $T(X)$ could be very large.

Example 4.5. We use a simple example to illustrate the algorithm. Let $\Sigma = \{x^2 - y^2 - 1, 2x^2 + 3y^2 - 6\}$, and $\epsilon = 10^{-3}$.

1. $t(x) = (-5 \times x^2 + 9)^2$.
2. $p_1 = 10^{-3}$ and $B = \{-687 \div 2747 \div \\{3137 \div 1024 \div \{1374 \div 687\}\}}.$
3. Compute $R$, we get $R = 1$.
4. $S = \frac{1374}{2747}$. Since $S > 2\epsilon$, we choose 1 to replace $S$ in the computation of $s$. We obtain $s = 1$.
5. $T(X) = 5 \times X^2 - 26 \times X^2 + 5$.
6. $p_2 = 2 \epsilon / 2 = 10^{-3}/2$ and $T = \{[-1145 \div 4572], [-2289 \div 917], [4572 \div 1145], [-3137 \div 2747]\}$. The multiplicity of all the roots are one.
7. $\rho = \frac{4572}{1374}, \theta = \min\{\frac{4572}{1374}, 1 \times 10^{-3}/2\} = 10^{-3}/2$. Since $p_2 > \epsilon$, refine $B$ with $p_1 = \theta$ and derive $B = \{[-687 \div 2747], [-3137 \div 1374], [-1145 \div 4572]\}.$
8. For each element of $T$, recover the isolation box of the corresponding root of $\Sigma = 0$. Consider the first element $T_1 = [-1145 \div 4572]$. It is easy to check that $T_1$ is associated with $B_1 = [-687 \div 1374].$ Then, the corresponding isolation box can be computed with (24), which is $[-687 \div 2747, -2747 \div 2048].$ The multiplicity of the root of the system is in one. In a similar way, we can find other isolation boxes.
5. IMPLEMENTATION AND EXPERIMENTAL RESULTS

We have implemented Algorithm 4.3 as a software package LGP in Maple, which is available at http://www.mmrc.iss.ac.cn/~xgao/software.html. Extensive experiments with this package show that this approach is efficient and stable, especially for bivariate equation systems with multiple roots.

We compare our method with Discoverer [19], GRUR [7], Hybrid [14], and Isolate [17]. Discoverer is a tool for solving problems about polynomial equations and inequalities. GRUR is a tool to solve bivariate equation systems. Hybrid is a numeric and symbolic hybrid algorithm for solving bivariate equation systems. Isolate is a tool to solve general equation systems based on the Realsolving C library by Rouillier.

We did three sets of experiments. All the results are collected on a PC with a 3.2GHz CPU, 2.00G memory, and running Microsoft Windows XP. We use Maple 12 in the experiments. The precision in the experiments is set to be $10^{-3}$. In these experiments, $f$ and $g$ are generated as follows.

- Both $f$ and $g$ are randomly generated dense polynomials with the same degree and with integer coefficients between $-99$ and $99$. The results are given in Fig. 5. In order to give more details about the results, we show the timings of Isolate, Hybrid, and LGP in Fig. 6 with a smaller time scaling.
- Both $f$ and $g$ are randomly generated sparse polynomials in the same degree, with sparsity $10\%$, and with integer coefficients between $-99$ and $99$. The results are given in Fig. 7 and Fig. 8.
- The third set of experiments is done with polynomial systems with multiple roots. We randomly generate a polynomial $h(x, y, z)$ and take $f(x, y) = \text{Res}_z(h, h_z)$, $g(x, y) = f_y(x, y)$. Since $f(x, y)$ is the projection of a space curve to the $xy$-plane, it most probably has singular points and $f = g = 0$ is an equation system with multiple roots. The results are given in Fig. 9 and Fig. 10.

For each possible degree, we generate ten examples and the results are the average values for the ten examples. According to Figures 5, 7, and 9, we have the following observations.

- In all cases, GRUR and Discoverer generally work for equation systems with degrees not higher than ten within reasonable time.
- In the first two cases, the equations are randomly generated and hence have no multiple roots. For systems without multiple roots, Hybrid is the fastest method, which is significantly faster than LGP and Isolate. Both Hybrid and LGP compute two resultants and isolate their real roots. LGP is slow, because the polynomials obtained by the shear map are usually dense and with large coefficients.
- We also observe that all methods spend more time with sparse polynomials than with dense polynomials in the same high degree. This phenomenon needs further exploration.

- For systems with multiple roots, LGP is the fastest method, which is significantly faster than Hybrid and Isolate. Note that our method is quite stable for equation systems with and without multiple roots. Isolate is also quite stable, but slower than LGP for bivariate equation systems.

Of course, we should mention that Discoverer and Isolate can be used to solve general polynomial equations and even inequalities. Our comparison is limited to the bivariate case.

\[ f = 16 + 40x - 72y - 119x^{12}y^{2} - 4755x^{6}y^{4} - 1803x^{8}y^{5} - 983x^{8}y^{7} + 4582x^{4}y^{9} + 153x^{7}y^{6} + 402x^{10}y^{2} + 201x^{12}y + 391x^{11}y^{2} + 6221x^{3}y^{3} - 1692x^{7}y^{3} - 3216x^{5}y^{3} + 4922x^{9}y^{3} + 585x^{11}y^{3} - 2301x^{10}y^{3} + 5749x^{7}y^{4} + 6066x^{10}y^{4} - 19x^{13}y^{4} + 1141x^{7}y^{5} + 1516x^{10}y^{5} - 2349x^{11}y^{5} - 3315x^{8}y^{5} + 1153x^{7}y^{6} + 2775x^{10}y^{6} + 784x^{10}y^{6} + 5439x^{9}y^{6} - 818x^{8}y^{6} - 131x^{12}y^{6} - 625x^{8}y^{7} - 1174x^{8}y^{7} - 1244x^{9}y^{7} - 16x^{6}y^{7} - 965x^{7}y^{7} + 345x^{4}y^{3} + 337x^{7}y^{3} + 7x^{7}y^{3} + 2276x^{9}y^{3} - 2276x^{8}y^{3} - 420x^{7}y^{3} - 348x^{9}y^{3} + 309x^{10}y^{3} + 3264x^{9}y^{3} + 1597x^{7}y^{3} + 1055x^{9}y^{3} + 98x^{8}y^{3} - 234x^{10}y^{3} - 401x^{10}y^{3} + 659x^{8}y^{3} - 84x^{7}y^{3} + 559x^{8}y^{3} - 154x^{8}y^{3} - 1075x^{7}y^{3} - 486x^{8}y^{3} + 665x^{11}y^{2} + 847x^{10}y^{2} + 294x^{9}y^{2} - 616x^{8}y^{2} - 5413x^{7}y^{2} + 1441x^{6}y^{2} - 2176x^{9}y^{2} - 2641x^{8}y^{2} - 2670x^{7}y^{2} + 3845x^{6}y^{2} + 626x^{7}y^{2} + 902x^{5}y^{2} - 880x^{7}y - 983x^{7}y^{3} + 190x^{7}y^{4} - 5890x^{5}y^{7} + 75x^{7}y^{7} + 826x^{6}y^{7} - 617x^{7}y^{7} - 8154x^{9}y^{7} + 3832x^{9}y^{7} + 3728x^{7}y^{7} + 918x^{9}y^{7} - 2759x^{8}y^{7} - 3400x^{7}y^{7} - 984x^{9}y^{7} + 2641x^{4}y^{8} + 879x^{7}y^{8} - 1269x^{7}y^{8} + 1450x^{8}y^{8} + 640x^{7}y^{8} + 829x^{8}y^{8} + 44x^{9}y^{8} + 992x^{7}y^{9} - 284x^{8}y^{9} - 919x^{7}y^{9} - 6516x^{7}y^{9} + 5401x^{7}y^{9} - 24x^{9}y^{9} + 2896x^{7}y^{9} - 256y^{9} - 116x^{9} - 56x^{9} - 113x^{9} + 172y^{9} + 238y^{9} - 134y^{9} + 305y^{9} - 107x^{9} + 222x^{8} - 94x^{9} + 105y^{9} + \]

![Figure 5: Σ consists of dense polynomials and has no multiple roots.](image)

![Figure 6: Same as Fig. 5, with a smaller time scaling.](image)

![Figure 7: Σ consists of sparse polynomials and has no multiple roots.](image)

![Figure 8: Same as Fig. 7, with a smaller time scaling.](image)

![Figure 9: Σ is a system with multiple roots.](image)

![Figure 10: A smaller time scaling of Fig. 9.](image)

**EXAMPLE 5.1.** A critical step to compute the topology for a plane curve is to determine its $x$-critical points. In this example, we compute all $x$-critical points of an algebraic curve $f = 0$ (the solid curve in Figure 11), which is defined below. The polynomial $f(x, y)$ is actually one of the examples used in the third set of experiments shown in Fig. 9. We need to solve the equation system $\Sigma = \{ f, \frac{\partial f}{\partial y} \}$. The method LGP uses 9.6 seconds, Isolate uses 80.1 seconds, Hybrid uses 326.4 seconds, and GRUR and Discover both give no results in 1200 seconds with the same PC as mentioned before.
55 x^7 + 993 y^8 - 1564 y^9 + 81 x^{12} + 157 x^{11} + 1497 y^{10} + 250 x^{10} + 179 y^{13} - 96 y^{14} + 50 y^{12} + 34 x^{13} - 766 y^{11} - 380 y^7 + 3 x^{15} + 19 y^{15} - 2014 y^2 x - 324 y^3 x y^4 + 281 x^{12} y^3 - 76 x^{13} y^4 + 201 x^{11} y^3 - 41 x^8 y^5 + 111 x^4 y^6 + 6 x^{15} y^6 - 64 x^{14} y^5 + 49 x^{13} y^5 + 60 x^8 y^6 + 66 x^7 y^6 + 26 x^6 y^7 + 8 x^5 y^7 - 136 x^4 y^7 - 175 x^3 y^8 + 17 x^2 y^8 + 138 x^4 y^8 - 53 x^3 y^9 + 1580 x^2 y^9 - 882 x^6 y^9 + 183 x^5 y^9 + 144 x^4 y^9 + 832 x^3 y^9 - 355 x^2 y^{10} + 236 x^2 y^8 + 245 x^4 y^{10} + 16 x^5 y^{10}.

Figure 11: The intersection of \( f = 0 \) and \( \partial f / \partial y = 0 \).

6. CONCLUSION

In this paper, we propose a local generic position method to solve bivariate polynomial equation systems. The method can be used to represent the roots of a bivariate equation system as the linear combination of the roots of two univariate equations. As a result, root isolation for bivariate systems is reduced easily to root isolation of univariate equations. The multiplicities of the roots are also derived.

The results of this paper can be extended to isolate the real roots of bivariate equation systems with more than two polynomials by using the resultant systems for several polynomials given in [18]. It is also possible to extend the method to multivariate equation systems. But the procedure is very complicated. It is an interesting problem to give a simple and effective algorithm for multivariate equation solving based on the idea of local generic position.

Acknowledgement

The authors would like to thank Prof. Hoon Hong and Ms. Mei-jiang Shan for providing the Maple code of their method.

7. REFERENCES


