Computing the Distance of a Point to an Algebraic Hypersurface and Application to Exclusion Methods

Xavier Gourdon
Algorithms Project, INRIA Rocquencourt

February 12, 1996

[summary by Pierre Nicodème]

Abstract

We compute lower bounds for the distance in \( \mathbb{C}^n \) from a point \( u \) to an algebraic surface \( \mathcal{Z} \). Such lower bounds or proximity tests give an approximation of \( \mathcal{Z} \). We present tests based on both Taylor’s formula and a generalization of the Dandelin-Graeffe process to the multivariate case, and their application to the exclusion method [2].

1. Introduction

Given a point \( a \) in \( \mathbb{C}^n \), and an algebraic hypersurface

\[
\mathcal{Z}(P) = \{(z_1, \ldots, z_n) \in \mathbb{C}^n | P(z_1, \ldots, z_n) = 0\},
\]

with \( P \in \mathbb{C}[z_1, \ldots, z_n] \), we want to evaluate the distance \( d(a, \mathcal{Z}) \) corresponding to the norm

\[
||z|| = \max_{1 \leq k \leq n} |z_k|.
\]

By shifting the variable \( z \), we can restrict to the case \( a = 0 \).

2. Univariate Polynomials

Let \( P(z) = \sum_{i=0}^{d} a_i z^i \in \mathbb{C}[z] \), \( a_d \neq 0 \), and \( \mathcal{Z}(P) = \{U_1, \ldots, U_d\} \). We want to evaluate \( d(0, \mathcal{Z}) = \min_i |U_i| \). In Henrici [4, vol. 1], Theorems 6.4.d and 6.4.i give the following classical bound for \( \mathcal{Z}(P) \):

**Proposition 1.** If \( \rho(P) \) is the nonnegative root of the equation \( |a_0| = \sum_{j=1}^{d} |a_j| \rho^j \), then

\[
\rho(P) \leq d(0, \mathcal{Z}) \leq \frac{1}{2^{1/d} - 1} \rho(P) \approx \frac{d}{\log 2} \rho(P).
\]

**Graeffe Iteration.** With \( P(z) = a_d \prod_{i=1}^{d} (z - U_i) \), we consider

\[
P(z)P(-z) = (-1)^{d} a_d^2 \prod_{i=1}^{d} (z^2 - U_i^2) = P^{(1)}(z^2).
\]

We note \( P^{(1)} \) the classical Graeffe iterate; the roots of \( P^{(1)} \) are the squares of those of \( P \), and \( d(0, \mathcal{Z}(P^{(1)})) = d(0, \mathcal{Z}(P)) \); we have

\[
\rho(P^{(1)}) \leq d(0, \mathcal{Z}(P^{(1)})) \leq \frac{\rho(P^{(1)})}{2^{1/d} - 1};
\]

\[
\rho(P^{(1)}) \approx \frac{d}{\log 2} \rho(P^{(1)}).
\]
so with $\rho_1 = \sqrt{\rho(P^{(1)})}$, we get
\[
\rho_1 \leq d(0, \mathcal{Z}(P)) \leq \frac{\rho_1}{(2^{1/d} - 1)^{1/2}}.
\]
Generally, we define $P^{(k)} = \text{Graeffe}(P^{(k-1)})$; then, we get $d(0, \mathcal{Z}(P^{(k)})) = d(0, \mathcal{Z}(P))^{2^k}$; with $\rho_k = \rho(P^{(k)})^{1/2^k}$, we have
\[
\rho_k \leq d(0, \mathcal{Z}(P)) \leq \frac{\rho_k}{(2^{1/d} - 1)^{1/2^k}}.
\]
The upper bound tends rapidly to the lower bound as $k$ increases, thus we have obtained an effective process to compute $d(0, \mathcal{Z})$.

**Computing the $P^{(k)}$.** With $A(z) = \sum_{i=0 \mod 2} a_i z^{i/2}$ and $B(z) = \sum_{i=1 \mod 2} a_i z^{(i-1)/2}$, we have
\[
P(z)P(-z) = A(z^2)^2 - z^2 B(z^2)^2,
\]
and therefore,
\[
\text{Graeffe}(P) = A(z)^2 - z B(z)^2.
\]
A practical problem is that the coefficient size doubles at each Graeffe iteration.

### 3. Multivariate Polynomials

In the multivariate case, the polynomial $P(z)P(-z)$ can not be written as $Q(z^2)$ where $Q(z)$ is a polynomial, thus we need to modify the definition. We generalize the Graeffe process to the multivariate case as follows:

**Definition 1.** We call the $N$-th Graeffe iterate of $P(z) \in \mathbb{C}[z_1, \ldots, z_n]$ the polynomial $P^{(N)}(z)$ defined by
\[
P^{(N)}(z) = \prod_{j=0}^{2^N-1} P(\omega^j z), \quad \omega = \exp\left(\frac{2i\pi}{2^N}\right), \quad i^2 = -1,
\]
where $\omega^j z$ denotes the point $(\omega^j z_1, \ldots, \omega^j z_n)$.

**Proposition 2.** For all non negative integer $N$, the $N$-th Graeffe iterate of $P(z)$ writes as
\[
P^{(N)}(z) = \sum_{j \geq 0} B_j^{(N)}(z),
\]
where the $B_j^{(N)}$’s are homogeneous polynomials of degree $2^N j$. The $(N+1)$-st Graeffe iterate can be computed from the $N$-th thanks to the formula
\[
P^{(N+1)}(z) = P_0^{(N)}(z)^2 - P_1^{(N)}(z)^2, \quad P_k^{(N)}(z) = \sum_{j \equiv k \mod 2} B_j^{(N)}(z).
\]

With the multivariate Graeffe process, we easily generalize the univariate algorithm to compute $d(0, \mathcal{Z})$ in the multivariate case.

**Theorem 1.** Let $P(z)$ be a polynomial in $\mathbb{C}[z_1, \ldots, z_n]$ of total degree $d$. Let $P^{(N)}(z) = \sum_{j \geq 0} B_j^{(N)}(z)$ be its $N$-th Graeffe iterate and $R_N$ the non-negative solution of the equation in $R$

\[
|P^{(N)}(0)| = \sum_{j \geq 1} \|B_j^{(N)}\|_{\infty} R^j,
\]
<table>
<thead>
<tr>
<th>d</th>
<th>$r_0/d$</th>
<th>$r_1/d$</th>
<th>$r_2/d$</th>
<th>$r_3/d$</th>
<th>$r_4/d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.7673</td>
<td>0.9725</td>
<td>0.9996</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>5</td>
<td>0.6525</td>
<td>0.9479</td>
<td>0.9973</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>7</td>
<td>0.6325</td>
<td>0.9400</td>
<td>0.9960</td>
<td>0.9999</td>
<td>1.0000</td>
</tr>
<tr>
<td>15</td>
<td>0.6067</td>
<td>0.9271</td>
<td>0.9938</td>
<td>0.9999</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

**Table 1.** Some values of $r_N/d(0, Z_{n,d})$ for $n = 2$.

where $\|B_j^{[N]}\|_\infty = \sup_{\|z\|_1} \|B_j(z)\|$. Then we have

$$r_N \leq d(0, Z) \leq \left( \frac{1}{2^{1/d} - 1} \right)^{2^{-N}} r_N, \quad r_N = B_N^{2^{-N}}.$$  

Computing $\|B_j^{[N]}\|_\infty$ raises a difficult practical problem; therefore, we make use of the norm $\| \sum a_n z^n \| = \sum |a_n|$, easy to compute. Our main result is stated using this norm; one demonstrates the equivalence of the norms $\| \cdot \|_\infty$ and $\| \cdot \|$ by combination of the Parseval identity and of the Cauchy-Schwarz inequality.

**Theorem 2.** Let $\rho_N$ be the unique nonnegative solution of

$$|P_N^{[N]}(0)| = \sum_{j=1}^{d} \|B_j^{[N]}\| \rho^j$$

The distance from 0 to $Z$ satisfies

$$r_N \leq d(0, Z) \leq \kappa_N r_N,$$

where

$$r_N = \rho_N^{2^{-N}} \quad \text{and} \quad \kappa_N = \left( \frac{1}{2^{1/d} - 1} \right)^{\frac{2^{N} + n - 1}{n-1}}.$$  

Moreover $\lim_{N \to \infty} \kappa_N = 1$, which implies $\lim_{N \to \infty} r_N = d(0, Z)$.

4. **Examples**

We take a polynomial of degree $d$ in $n$ variables: $P_{n,d} = \sum_{j=1}^{n} (1 - z_j)^d - 1$. With $Z_{n,d} = Z(P_{n,d})$, we have $d(0, Z_{n,d}) = 1 - \frac{1}{n^{1/2}}$.

Tables 1 and 2 give the value of the ratio $r_N/d(0, Z_{n,d})$ of Theorem 3 for several values of $n$, $d$, and $N$. The computations were performed in Maple. These examples show that the bound is quite good for a small value $N$ of Graeffe iterates.

5. **Exclusion methods**

We give the principle of the method for a polynomial of one variable $P(z) \in \mathbb{C}[z]$.

- Let the **exclusion function** be: $z_0 \mapsto \rho(z_0)$, with $\rho$ given by theorem 2 after a proper shift of the variable, and
  
  (1) $\rho(z_0) = 0 \iff P(z_0) = 0$,
  
  (2) $P$ has no zero in $|z - z_0| < \rho(z_0)$, which is equivalent to $\rho(z_0) \leq d(z_0, Z)$;

- then, the **exclusion test** is: let $C$ be a square of centre $z_0$ and half-side $a > 0$. If $\rho(z_0) \geq \sqrt{2a}$, $C$ contains no zero of $P$.

49
Exclusion algorithm.

- Consider the reciprocal polynomial \( R(z) \) of \( P(z) \); compute by Graeffe a lower bound of the smallest root of \( R(z) \), which gives an upper bound \( b_u \) of the largest root of \( P(z) \);
- Start from a big square centred at the origin, with side \( 2b_u \), which contains all the roots of \( P(z) \);
- Recursively split the square in four squares of equal size, discarding by the exclusion test squares containing no zeros;
- Stop the recursion when the desired precision is reached (the surface of the area covering the zeros decreases exponentially fast to zero).

Figure 1 shows an application of the exclusion method to localize an algebraic curve in \( \mathbb{R}^2 \).

For an algebraic variety \( \mathcal{Z} = \mathcal{Z}(P_i) \) and \( \mathcal{Z} = \bigcap_i \mathcal{Z}(P_i) \), with \( P_1, \ldots, P_m \in \mathbb{C}[z_1, \ldots, z_n] \), let \( \rho_i(z_0) \) be an exclusion function defined by theorem 2 for \( P_i \), \( 1 \leq i \leq m \); we can define an exclusion function for the variety as \( \rho(z_0) = \sup_{1 \leq i \leq m} \rho_i(z_0) \).

An application of exclusion method to localize the intersection of two curves in \( \mathbb{R}^2 \) is given in Figure 2.

Bibliography