Resultants and Neighborhoods of a Polynomial

Valentina Marotta

Department of Mathematics and Computer Science
Viale A. Doria, 6 - 95125, Catania, Italy
marotta@dmi.unict.it *

Abstract. In this paper the concept of neighborhood of a polynomial is analyzed. This concept is spreading into Scientific Computation where data are often uncertain, thus they have a limited accuracy. In this context we give a new approach based on the idea of using resultant in order to know the common factors between an empirical polynomial and a generic polynomial in its neighborhood. Moreover given a polynomial, the Square Free property for the polynomials in its neighborhood is investigated.

1 Introduction

In Computer Algebra and Symbolic Computation we deal with exact data, but in many real life situations data come either from physical measurements or observations or they are the output of numerical computations. Moreover computers work with Floating Points, i.e. with finite arithmetic, so a real number is represented approximately because of the limited number of available digits. For these reasons it is growing up the necessity of combining methods in computer algebra and numerical analysis. This led to a new branch of the classic polynomial algebra, the *numerical polynomial algebra* [2, 11], whose aim is to accommodate the presence of inaccurate data and inexact computations. Since a polynomial is not exact, we consider a family of polynomials called *neighborhood*. For instance, given a polynomial $p(x) = 1.23x^3 + 5.78x^2 + 12x + 11.98$, if we suppose that the last decimal digit is off by one, then the coefficient 1.23 does not mean a precise rational number but rather any real number from a sufficiently small neighborhood, i.e. a real number belonging to the interval $[1.22, 1.24]$. Therefore computations with polynomial neighborhoods involve some notions of interval analysis [1].

The paper is organized as follows. In Section 2 the basic definition of polynomial neighborhood as in [9] is introduced, while the notion of pseudozero, as in [9, 11], is introduced in Section 3. In Section 4 the problem of finding the nearest polynomial with a given zero is briefly discussed. In Section 5 we give a new definition for the neighborhood imposing stronger conditions by means of the resultant. In Section 6 the Square Free property for the polynomials in the neighborhood is the object of our interest. Finally in Section 7 it is shown the geometric meaning of the problems previously discussed.

* Supported by University of Naples "Federico II" and INTAS grant n.99-1222.

Many definitions can be given for multivariate polynomials and the results in the paper can be extended in the future to multivariate polynomials case. The author has used two computer algebra packages: MAPLE 6 [3] and MATHEMATICA 4.1 [12]. The use of MATHEMATICA 4.1 is essential because this package works with interval arithmetic.

2 Polynomial Neighborhoods

Let $\mathbb{C}$ be the field of complex numbers and let $p(x) = p(x_1, \ldots, x_s)$ in $\mathbb{C}[x_1, \ldots, x_s]$ be a multivariate polynomial in the variables $x_1, \ldots, x_s$ with $\deg(p) = n$. Let $x^j = x_1^{j_1} \cdots x_s^{j_s}$ with $j = (j_1, \ldots, j_s) \in \mathbb{N}_0^s$. The polynomial $p(x)$ can be written as

$$p(x) = \sum_{j=0,\ldots,k} a_j x^j$$

with $a_j \in \mathbb{C}$, $k = \binom{n+s}{n}$ and $\sum_{h=1,\ldots,s} j_h \leq n$ for all j.

The *tolerance* e associated with $p(x)$ is a non negative vector $e = (e_k, \ldots, e_0)$ such that

$$e_j \in \mathbb{R} \quad e_j \geq 0 \ \text{ for } \ j = 0, \ldots, k$$

Let $p(x)$ be a polynomial with at least one coefficient having limited accuracy, i.e. a $p(x)$ such that there exists $j \in \{0, \ldots, k\}$ with $e_j > 0$. $p(x)$ will be called an *empirical* polynomial.

Let us give some different definitions for the polynomial neighborhood. We start introducing the definition given by Stetter as in [9].

Definition 1. *The* **neighborhood** $\mathcal{N}(p, e)$ *of an empirical polynomial p with tolerance e is the family of polynomials $\tilde{p} \in \mathbb{C}[x_1, \ldots, x_s]$, $\tilde{p} = \sum \tilde{a}_j x^j$, having vector of coefficients $\tilde{a} = (\tilde{a}_k, \ldots, \tilde{a}_0) \in \mathbb{R}^{k+1}$, such that*

$$\tilde{a}_j = a_j \ \text{ if } e_j = 0, \ \text{ i.e. } a_j \text{ is exact}$$

$$\|(\ldots, \frac{|\tilde{a}_j - a_j|}{e_j}, \ldots)\|^* \leq 1 \tag{1}$$

with $\| \cdot \|^*$ a generic norm over $\mathbb{R}^{k+1}$.

For instance if the polynomial $p(x) \in \mathbb{R}[x]$ and choosing $\|(\ldots, b_j, \ldots)\|^* = \max_j |b_j|$ as norm, the (1) becomes

$$|\tilde{a}_j - a_j| \leq e_j \ \text{ for } j = 0, \ldots, k \tag{2}$$

This means that the exact value $\tilde{a}_j \in [a_j - e_j, a_j + e_j]$, i.e. $\tilde{a}_j = a_j + \delta_j$ such that $|\delta_j| \leq e_j$. Each δ_j, for $j = 0, \ldots, k$, is called *perturbation* and the generic polynomial $\tilde{p}(x)$ in the neighborhood is called *perturbed polynomial.*
Another definition of polynomial neighborhood is the following as in [5].

Definition 2. *Let $p(x) = \sum a_j x^j \in \mathbb{C}[x_1, \ldots, x_s]$ be an empirical polynomial. The neighborhood $\mathcal{N}(p)$ of p is the family of polynomials $\tilde{p}$ in $\mathbb{C}[x_1, \ldots, x_s]$, $\tilde{p} = \sum \tilde{a}_j x^j$ such that $\|p - \tilde{p}\| \leq \epsilon$, with $\epsilon \geq 0$, $\epsilon \in \mathbb{R}$, and with $\|\cdot\|$ a generic coefficient vector norm.*

A special case of the Definition 1 is the following definition as in [8].

Definition 3. *Let $p(x_1, \ldots, x_s) \in \mathbb{C}[x_1, \ldots, x_s]$ be a multivariate polynomial with tolerance e. Let $\prec$ be a term ordering on $PP(x_1, \ldots, x_s) = \{x_1^{t_1} \cdots x_s^{t_s} : (t_1, \ldots, t_s) \in \mathbb{N}_0^s\}$. We say that a polynomial $\tilde{p} \in \mathbb{C}[x_1, \ldots, x_s]$ is a* lower-order perturbed *of p if $\tilde{p} \in \mathcal{N}(p, e)$ and $HTerm_{\prec}(\tilde{p} - p)$ is less than $HTerm_{\prec}(p)$.*

In the sequel of the paper we will use the notion of neighborhood as in Definition 1 using the maximum norm and somewhere we will compare the results for different definitions.

3 Pseudozero

We are interested in finding meaningful solutions for algebraic problems with polynomial data of limited accuracy. The treatment of data with limited accuracy is based on the *backward error principle*. An approximate result is meaningful if it is the exact result for a problem within a tolerance neighborhood of the specified problem. Whenever we have an empirical polynomial the notion of zero must be replaced by the notion of "pseudozero".

Definition 4. *A value $z \in \mathbb{C}^s$ is a pseudozero of an empirical polynomial p with tolerance e if it is a zero of some polynomial $\tilde{p} \in \mathcal{N}(p, e)$.*

Definition 5. *A pseudozero domain of an empirical polynomial p with tolerance e is the set of all pseudozeros*

$$\mathcal{Z}(p, e) = \{z \in \mathbb{C}^s : \exists\, \tilde{p} \in \mathcal{N}(p, e) : \tilde{p}(z) = 0\}$$

For further details about pseudozeros see [9–11].

3.1 Univariate Case

Our aim is studying the behaviour of the pseudozero domain of a polynomial, i.e. where the zeros lie. A simple case occurs when we work in $\mathbb{R}[x]$ and we choose the maximum norm (obtaining (2)). Let $p(x) = \sum_{j=0}^{n} a_j x^j \in \mathbb{R}[x]$ be an empirical polynomial and let $e = (e_n, \ldots, e_0)$ be its tolerance vector. A perturbed polynomial in the neighborhood is $\tilde{p}(x) = \sum_{j=0}^{n} (a_j + \delta_j) x^j$; if we define the polynomial $p_{\Delta}(x) = \sum_{j=0}^{n} \delta_j x^j$, then $\tilde{p}(x) = p(x) + p_{\Delta}(x)$. It holds

$$\forall z : \tilde{p}(z) = 0 \Rightarrow p(z) + p_{\Delta}(z) = 0 \Rightarrow p(z) = -p_{\Delta}(z) \tag{3}$$

and then $|p_{\Delta}(z)| = |\sum_{j=0}^{n} \delta_j z^j| \leq \sum_{j=0}^{n} |\delta_j||z^j| \leq \sum_{j=0}^{n} e_j |z|^j$, i.e.

$$-\sum_{j=0}^{n} e_j |z|^j \leq p_{\Delta}(z) \leq \sum_{j=0}^{n} e_j |z|^j \tag{4}$$

We can distinguish 3 cases:

1. $z = 0$: trivially $a_0 + \delta_0 = 0$ and then $-e_0 \leq a_0 \leq e_0$;
2. $z > 0$: the inequalities (4) become $\sum_{j=0}^{n}(-e_j)z^j \leq p_\Delta(z) \leq \sum_{j=0}^{n} e_j z^j$, and from (3) z has to satisfy the following inequalities:

$$\sum_{j=0}^{n}(a_j + e_j)z^j \geq 0 \quad \sum_{j=0}^{n}(a_j - e_j)z^j \leq 0 \tag{5}$$

3. $z < 0$: from (3) and (4) z has to satisfy the following inequalities:

$$\sum_{h=0}^{n}(a_h + (-1)^h e_h)z^h \geq 0 \quad \sum_{h=0}^{n}(a_h + (-1)^{h+1} e_h)z^h \leq 0 \tag{6}$$

In this way we can find where all pseudozeros lie; the pseudozeros are distributed into disjoint components named *pseudozero components*.

Example 1. Let $p(x) = 1.1x^2 + 2.3x + 1$ be an empirical polynomial and let $e = (10^{-2}, 10^{-2}, 10^{-1})$ be the tolerance vector. We look for the pseudozero components. From equations (5) and (6), we have to solve these two systems:

$$\begin{cases} 1.11x^2 + 2.31x + 1.1 \geq 0 \\ 1.09x^2 + 2.29x + 0.9 \leq 0 \end{cases} \quad \begin{cases} 1.11x^2 + 2.29x + 1.1 \geq 0 \\ 1.09x^2 + 2.31x + 0.9 \leq 0 \end{cases}$$

By solving the former system we obtain the components $[-1.57, -1.34]$ and $[-0.73, -0.52]$; since $z > 0$ by hypothesis (case 2) we can say that there are not positive pseudozeros; from the latter system we obtain $[-1.60, -1.30]$ and $[-0.76, -0.51]$, that are the ranges where all pseudozeros lie.

There is a different way of checking where all pseudozeros lie. It is sufficient to consider a perturbed polynomial $\tilde{p}(x) = (1.1+\delta_2)x^2 + (2.3+\delta_1)x + 1 + \delta_0$ in the neighborhood. We substitute the interval $[-e_j, e_j]$ to each δ_j and we solve the equation $\tilde{p}(z) = 0$ in the variable z by using interval arithmetic. If we perform all these computations with MATHEMATICA 4.1, then we obtain the pseudozero components $[-1.60, -1.30]$ and $[-0.78, -0.49]$, i.e. components larger than the previous ones, that we have obtained by using formulas (5) and (6). It follows that the computation based on these formulas is more careful than the computation involving intervals (i.e. interval polynomials), which is affected by a *bad* interval arithmetic.

Roughly speaking, it seems that small perturbations in the coefficients of the polynomial give small variations of the roots, but it is not always true.

Example 2. Let $p(x) = (x-1)(x-2)\cdots(x-20) = x^{20} - 210x^{19} + \cdots + 20!$. This is a well-known polynomial in numerical analysis, named the *Wilkinson's polynomial*. If we change the coefficient -210 with $-210 + 10^{-7}$, then the roots $1, 2, \ldots, 20$ of the polynomial change dramatically (in fact there are a lot of complex roots).

A polynomial as in Example 2 is said to be *sensitive*, which means that the relative error in the result is much greater than the relative error in the data, i.e. it is *ill-conditioned* (or similarly *ill-posed*). Pseudozeros searching analysis can be very useful for investigating more accurately this class of cases and for checking the sensitivity of a polynomial.

4 Nearest Polynomial with a Given Zero

Given an empirical polynomial with tolerance vector e and an approximate zero we want to investigate how we can check that it is a *pseudozero*, i.e. an exact zero of a polynomial that we can not distinguish from the specified one because of the indetermination in its data.

Assigned an empirical polynomial $p(x) \in \mathbb{R}[x]$, $p(x) = \sum_{j=0}^{n} a_j x^j$, with $\deg(p) = n$, and a value $z \in \mathbb{R}$, we look for a polynomial $\tilde{p}(x) = \sum_{j=0}^{n} \tilde{a}_j x^j$ in $\mathcal{N}(p, e)$ such that $\tilde{p}(z) = 0$. We can find more than one of such polynomials, but there is only one of these polynomials whenever we look for the *nearest* one, i.e. the polynomial with the smallest perturbations from the original coefficients.

The problem of finding the nearest polynomial with a given zero can be formulated in two different manners depending on the chosen definition for the polynomial neighborhood. The former case occurs when the distance between the two polynomials $p(x)$ and $\tilde{p}(x)$ is measured by using a norm as in Definition 2. In this case the concept of *nearest* depends on the chosen norm. The problem is the following: we want to compute

$$\delta_j \in \mathbb{R},\ j = 0, \ldots, k, \text{ such that } \|p - \tilde{p}\| \text{ is minimized for a given norm } \|\cdot\|$$

If we choose the Euclidean norm, i.e. $\|p - \tilde{p}\|_2 = (\sum (a_j - \tilde{a}_j)^2)^{\frac{1}{2}}$, then it is straightforward to find the δ_j's such that $\tilde{p}(z) = 0$ and $\|p - \tilde{p}\|_2$ is minimized. In fact we have the following theorem

Theorem 1. *Let $p(x)$ be an empirical polynomial with $\deg(p) = n$. Let $\|\cdot\|_2$ be the Euclidean norm and let $z \in \mathbb{R}$. $\|p - \tilde{p}\|_2$ is minimized whenever the δ_j's satisfy the following conditions:*

$$\delta_j = \frac{-z^j p(z)}{z^{2n} + z^{2n-2} + \ldots + z^2 + 1} \qquad j = 0, \ldots, n \tag{7}$$

Proof. The formula can be obtained minimizing a quadratic form as in [6].

If $p(x) \in \mathbb{C}[x]$ and $z \in \mathbb{C}$, then the formula for the δ_j's is the following

$$\delta_j = \frac{-\bar{z}^j p(z)}{\sum_{i=0}^{n} (z\bar{z})^i} \quad j = 0, \ldots, n \tag{8}$$

where $\bar{z}$ is the conjugate of z; a complete proof of formula (8) can be found in [4]. There are similar results using $\|\cdot\|_\infty$ [5].

The latter case is obtained using Definition 1 for the polynomial neighborhood. Whenever we impose that the perturbed polynomial $\tilde{p}(x) = \sum_{j=0}^{n} (a_j + \delta_j) x^j$ has a zero z, we locate a hyperplane in the space $(\delta_0, \ldots, \delta_n)$:

$$\delta_n z^n + \ldots + \delta_1 z + \delta_0 + p(z) = 0 \tag{9}$$

that is a linear manifold of codimension 1 in the space $(\delta_0, \ldots, \delta_n)$. The value for the δ_j's such that the distance between the origin $(0, \ldots, 0)$ and the hyperplane is minimum can be computed as follows. The straight line perpendicular to the hyperplane and crossing the origin has parametric equation

$$\begin{cases} \delta_n = z^n t \\ \vdots \\ \delta_1 = zt \\ \delta_0 = t \end{cases} \tag{10}$$

Computing the intersection between the hyperplane and the straight line, after simple manipulations, we obtain

$$t = \frac{-p(z)}{z^{2n} + z^{2n-2} + \ldots + z^2 + 1}$$

and then

$$\delta_j = \frac{-z^j p(z)}{z^{2n} + z^{2n-2} + \ldots + z^2 + 1} \quad j = 0, \ldots, n \tag{11}$$

Example 3. Given the polynomial $p(x) = 1.1x^2 + 2.3x + 1$ and its tolerance $e = (10^{-2}, 10^{-2}, 10^{-1})$ we look for the nearest polynomial in $\mathcal{N}(p, e)$ having $z = -1.5$ as a zero. Using formula (9) we obtain the hyperplane $2.25\delta_2 - 1.5\delta_1 + \delta_0 + 0.025 = 0$. The straight line as in formula (10) becomes

$$\begin{cases} \delta_2 = 2.25t \\ \delta_1 = -1.5t \\ \delta_0 = t \end{cases} \tag{12}$$

Using the method previously described we find the point $(\delta_2 = -0.0068, \delta_1 = 0.0045, \delta_0 = -0.003)$ belonging to the hyperplane and having minimum distance $d = 0.00866$ from the origin. Then the polynomial $\bar{p}(x) = 1.0932x^2 + 2.3045x + 0.997$ is the nearest polynomial in $\mathcal{N}(p, e)$ having $z = -1.5$ as a zero.

Remark 1. Actually the formula as in (11) is correct only if every coefficient of the empirical polynomial has a nonzero tolerance. For instance if we know by hypothesis that the leading coefficient of the empirical polynomial is exact, the equation (9) becomes $\delta_{n-1}\alpha^{n-1} + \ldots + \delta_1\alpha + \delta_0 + p(\alpha) = 0$, then, after computations as above, we find the following formula for the perturbations:

$$\delta_j = \frac{-\alpha^j p(\alpha)}{\sum_{k=0}^{n-1} \alpha^{2k}} \quad j = 0, \ldots, n-1$$

In a more general case, given a tolerance vector $e = (e_n, \ldots, e_0)$, let us denote by $\Omega(p) = \{j \in \{0, \ldots, n\} \mid e_j > 0\}$. Now it is simple to prove that the formula for the perturbations can be written as:

$$\delta_j = \frac{-\alpha^j p(\alpha)}{\sum_{k \in \Omega(p)} \alpha^{2k}} \quad j \in \Omega(p) \tag{13}$$

5 New Properties by Using Resultants

Sometimes beside the acceptable variations for the coefficients of the empirical polynomial $p(x)$, we know (often experimentally) more properties of the exact polynomial. For instance, we can know a zero of the exact polynomial we are looking for. So we construct $p(x)$, such that it has the same zero but its coefficients have limited accuracy. Similarly we can know more than one zero or a common factor between $p(x)$ and the generic $\tilde{p}(x)$. In these cases we use the notion of resultant.

The resultant of a polynomial $p(x) = \sum_{j=0}^{n} a_j x^j$ and a generic perturbed polynomial in the neighborhood $\tilde{p}(x) = \sum_{j=0}^{n}(a_j + \delta_j)x^j$ is computed. By definition the resultant is the determinant of the Sylvester matrix of the two polynomials; in our case we work with the following $(n+n) \times (n+n)$ matrix:

$$S(p,\tilde{p}) = \begin{pmatrix} a_n & \dots & a_0 & 0 & \dots & 0 \\ 0 & a_n & \dots & a_0 & 0 & \dots \\ \vdots & \ddots & \ddots & \dots & \ddots & \dots \\ 0 & \dots & 0 & a_n & \dots & a_0 \\ a_n+\delta_n & \dots & a_0+\delta_0 & 0 & \dots & 0 \\ 0 & a_n+\delta_n & \dots & a_0+\delta_0 & 0 & \dots \\ \vdots & \ddots & \ddots & \dots & \ddots & \dots \\ 0 & \dots & 0 & a_n+\delta_n & \dots & a_0+\delta_0 \end{pmatrix}$$

and $\mathrm{Res}(p,\tilde{p}) = \det(S(p,\tilde{p}))$.

It is well known in linear algebra that the determinant of the matrix does not change (and the resultant too), when we substitute the $(n+1)$-th row with the difference between the $(n+1)$-th row and the first row, the $(n+2)$-th row with the difference between the $(n+2)$-th row and the second row and so on until the $2n$-th row that we substitute with the difference between the $2n$-th row and the n-th row. We obtain a new matrix:

$$S_1(p,\tilde{p}) = \begin{pmatrix} a_n & \dots & a_0 & 0 & \dots & 0 \\ 0 & a_n & \dots & a_0 & 0 & \dots \\ \vdots & \ddots & \ddots & \dots & \ddots & \dots \\ 0 & \dots & 0 & a_n & \dots & a_0 \\ \delta_n & \dots & \delta_0 & 0 & \dots & 0 \\ 0 & \delta_n & \dots & \delta_0 & 0 & \dots \\ \vdots & \ddots & \ddots & \dots & \ddots & \dots \\ 0 & \dots & 0 & \delta_n & \dots & \delta_0 \end{pmatrix}$$

Since $\det(S(p,\tilde{p})) = \det(S_1(p,\tilde{p}))$, we can write the resultant between $p(x)$ and $\tilde{p}(x)$ as the resultant between $p(x)$ and a new polynomial $p_{\Delta}(x) = \sum_{j=0}^{n} \delta_j x^j$.

Next theorems on resultants are well known in the literature and they are very useful for our purposes. They hold in the general case of polynomials with coefficients in a U.F.D. (Unique Factorization Domain).

Theorem 2. *Let R be a U.F.D., let $A(x) = \sum_{i=0}^{n} a_i x^i$, $B(x) = \sum_{j=0}^{m} b_j x^j$ be primitive polynomials in $R[x]$, then*

$$\mathrm{Res}(A, B) = 0 \Leftrightarrow \mathrm{GCD}(A, B) \neq 1$$

By Theorem 2 $p(x)$ and $\tilde{p}(x)$ have a common factor if and only if $\mathrm{Res}(p, \tilde{p}) = 0$. The resultant between $\tilde{p}$ and p is a polynomial in the indeterminates $\delta_0, \ldots, \delta_n$ hence $\mathrm{Res}(\tilde{p}, p) = 0$ is equivalent to give some conditions on the δ_j's.
We can give the following two new definitions.

Definition 6. *Let $p(x) \in \mathbb{R}[x]$ be an empirical polynomial and let e be a vector of tolerance. The polynomial $\tilde{p}(x) \in \mathbb{R}[x]$ is a* k common-factor perturbed *of $p(x)$ if $\tilde{p}(x) \in \mathcal{N}(p, e)$ and it has at least a common factor of degree k with $p(x)$.*

Definition 7. *Let $p(x) \in \mathbb{R}[x]$ be an empirical polynomial and let e be a vector of tolerance. A* k common-factor neighborhood *$\mathcal{N}_k(p, e)$ of a polynomial $p(x)$ is the set of all k common-factor perturbeds.*

Let us see two very important theorems about resultants [7]:

Theorem 3. *Let R be a U.F.D. and let $A(x) = \sum_{i=0}^{n} a_i x^i$ and $B(x) = \sum_{j=0}^{m} b_j x^j$ be primitive polynomials in $R[x]$. Let $D = \mathrm{GCD}(A, B)$. If the rank of the Sylvester matrix is $m + n - k$ then $\deg(D) = k$.*

Theorem 4. *Let R be a U.F.D. and let $A(x) = \sum_{i=0}^{n} a_i x^i$ and $B(x) = \sum_{j=0}^{m} b_j x^j$ be primitive polynomials in $R[x]$. Let $S_d(A, B)$ be the matrix obtained by the resultant matrix $S(A, B)$ when we put it in triangular echelon form by using row transformations only. The first nonzero row from the bottom of the matrix $S_d(A, B)$ gives the coefficients of a GCD of $A(x)$ and $B(x)$.*

Once the resultant matrix is given, then we can put it in triangular echelon form by using row transformations only. By using the Laidacker's result, the entry in the $2n$-th row and column is the resultant of the two polynomials. If the first nonzero row from the bottom is the $(2n - h)$-th, then the entries of such row are the coefficients of the polynomial $q(x) = \mathrm{GCD}(p, \tilde{p})$, having degree h. For example, if $\mathrm{Res}(p, \tilde{p}) = 0$, then the polynomials have, at least, a common linear factor. $\mathrm{Res}(p, \tilde{p})$ is a homogeneous polynomial of degree $2n$ in the δ_j's and a_j's and a homogeneous polynomial of degree n as polynomial in the only δ_j's.

Remark 2. A trivial case $\mathrm{Res}(p, \tilde{p}) = 0$ occurs when the coefficients of the polynomials are the same up to a nonzero constant. This case occurs when we have the following choice for the perturbations:

$$\delta_n \in [-e_n, e_n], \quad \delta_j = a_j \frac{\delta_n}{a_n} \quad j = 0, \ldots, n-1 \tag{14}$$

where $a_n \neq 0$, since $n = \deg(p)$. In order to guarantee that for any choice of δ_n the corresponding values of each δ_j belong to their intervals, we have to compute

$$\tau = \min_{j=0,\ldots,n-1} \left\{ \left| \frac{e_j a_n}{a_j} \right| \right\} \quad \text{with} \quad a_j \neq 0$$

Now if $\delta_n \in [-\tau, +\tau] \cap [-e_n, e_n]$ by applying (14) we can find all polynomials in the neighborhood, whose coefficients are proportional to the coefficients of $p(x)$.

In the general case we have the following: if $p(x)$ and $\tilde{p}(x)$ have at least a k degree common factor, with $1 \leq k \leq n$, then all entries from the $(2n-k+1)$-th row to the $2n$-th row of the resultant matrix $S_d(p, \tilde{p})$ must vanish. Furthermore, denoted by $S_d[2n-k, 2n-k]$ the entry in the $(2n-k)$-th row and $(2n-k)$-th column of the matrix $S_d(p, \tilde{p})$, if $S_d[2n-k, 2n-k]$ is different from zero, then $p(x)$ and $\tilde{p}(x)$ have exactly a k degree common factor and $S_d[2n-k, 2n-k]$ is its leading coefficient. We find a system of $\frac{k(k+1)}{2}$ homogeneous equations in the δ_j's. We observe that, by construction of the matrix S_d, if the entries in the $(2n-k+1)$-th row vanish then all entries from the $(2n-k)$-th row to the $2n$-th row vanish too; then the system of $\frac{k(k+1)}{2}$ equations in the δ_j's is equivalent to a system of only k equations. If $k = n$ then the only solution is given by (14), i.e. $p(x)$ and $\tilde{p}(x)$ have proportional coefficients. Otherwise we find the conditions on the δ_j's in order to satisfy that $\tilde{p}(x)$ is a k common-factor perturbed of $p(x)$.

5.1 An Example: Polynomials of Degree Two

Let $p(x) = a_2x^2 + a_1x + a_0$ be a generic polynomial of degree two and let $\tilde{p}(x) = (a_2+\delta_2)x^2 + (a_1+\delta_1)x + (a_0+\delta_0)$ be a perturbed polynomial in $\mathcal{N}(p, e)$, for a given tolerance e. Their Sylvester matrix in triangular form (using the Fraction Free Gaussian Elimination) is

$$\begin{pmatrix} a_2 & a_1 & a_0 & 0 \\ 0 & a_2^2 & a_1a_2 & a_0a_2 \\ 0 & 0 & \delta_1a_2^2 - a_1a_2\delta_2 & \delta_0a_2^2 - a_2a_0\delta_2 \\ 0 & 0 & 0 & R \end{pmatrix}$$

By Laidacker's theorem $R = \mathrm{Res}(p, \tilde{p}) = -a_2^2\delta_0^2 - a_0a_2\delta_1^2 + 2a_2a_0\delta_0\delta_2 + a_1a_2\delta_0\delta_1 + a_0a_1\delta_1\delta_2 - a_0^2\delta_2^2 - a_1^2\delta_0\delta_2$. If we impose the condition $\mathrm{Res}(p, \tilde{p}) = 0$ then we have to satisfy

$$(a_1\delta_2 - a_2\delta_1)(a_0\delta_1 - a_1\delta_0) - (a_2\delta_0 - a_0\delta_2)^2 = 0 \tag{15}$$

Moreover if $\delta_1a_2^2 - a_1a_2\delta_2 \neq 0$, i.e. $\delta_1a_2 - a_1\delta_2 \neq 0$ since $a_2 \neq 0$ because $\deg(p) = 2$, then p and $\tilde{p}$ have exactly the following common linear factor

$$(\delta_1a_2^2 - a_1a_2\delta_2)x + \delta_0a_2^2 - a_2a_0\delta_2 \tag{16}$$

If $\tilde{p}(x)$ is a lower-order perturbed as in Definition 3, with $\deg(\tilde{p} - p) = 1$, then we impose $e_2 = 0$, and in (15) and (16) we have to substitute the value $\delta_2 = 0$.

Example 4. We compute the resultant between the empirical polynomial $p(x) = 1.1x^2+2.3x+1$ and a generic polynomial in its neighborhood $\tilde{p}(x) = (1.1+\delta_2)x^2 + (2.3+\delta_1)x + (1+\delta_0)$. We choose the tolerance vector $e = (10^{-2}, 10^{-2}, 10^{-1})$ and we assume to work with 3 digits after the point. $\mathrm{Res}(p, \tilde{p}) = (2.3\delta_2 - 1.1\delta_1)(\delta_1 - 2.3\delta_0) - (1.1\delta_0 - \delta_2)^2 = 0$ in order to have at least a common factor. This condition gives the following choice of the perturbations: $\delta_1 \in [-e_1, e_1], \delta_2 \in [-e_2, e_2], \delta_0 = \frac{23}{22}\delta_1 - \frac{309}{242}\delta_2 \pm \frac{1}{242}\sqrt{89}(-23\delta_2 + 11\delta_1)$. For instance if we choose $\delta_0 = -0.025, \delta_1 = -0.002, \delta_2 = 0.01$, then we find the polynomial $1.11x^2 + 2.298x + 0.975$ in the neighborhood having the common root -1.474 with the initial polynomial.

Remark 3. In the paper the notion of polynomial neighborhood is introduced using the standard basis $\{1, x, x^2, \dots, x^n\}$ for polynomials, but other bases can be used, e.g. Bernstein polynomials, Lagrangian polynomials, etc. Unfortunately some problems can be found in order to define the resultant with respect to a different basis. Since there is a transformation matrix from the standard basis to a new basis, then a procedure to define a corresponding resultant can be given.

6 Square Free Property

Let $p(x)$ be a univariate polynomial in $\mathbb{R}[x]$. $p'(x)$ will denote the usual derivative of $p(x)$ with respect to x.

Definition 8. *A polynomial $p(x)$ is square free if* $\mathrm{GCD}(p, p') = 1$, *i.e. it has no factors with multiplicity greater than one.*

For instance, square free property is important in order to use the Sturm's algorithm for finding the zeros of a polynomial in a given real interval (a, b). Given an empirical polynomial $p(x) = \sum_{j=0}^{n} a_j x^j \in \mathbb{R}[x]$, we suppose that it is square free (respectively not square free). Our aim is to investigate if this property holds for any perturbed polynomial in the neighborhood of $p(x)$. We compute the resultant between $p(x)$ and its derivative $p'(x)$. If $p(x)$ is square free then, by Theorem 2, $\mathrm{Res}(p, p') \neq 0$. Similarly given $\tilde{p}(x)$ we compute $\mathrm{Res}(\tilde{p}, \tilde{p}')$. $\tilde{p}(x)$ is square free if $\mathrm{Res}(\tilde{p}, \tilde{p}') \neq 0$.

Theorem 5. *Let $p(x) = \sum_{j=0}^{n} a_j x^j \in \mathbb{R}[x]$ be an empirical polynomial with $\deg(p) = n$ and let e be its tolerance vector. Let $\tilde{p}(x) = \sum_{j=0}^{n} (a_j + \delta_j) x^j$ be a generic perturbed in the neighborhood of $p(x)$. We have*

$$\mathrm{Res}(\tilde{p}, \tilde{p}') = \mathrm{Res}(p, p') + \phi(\delta_0, \dots, \delta_n) \tag{17}$$

where $\phi(\delta_0, \dots, \delta_n)$ is a polynomial in the indeterminate δ_j's of degree less or equal than $2n - 1$.

Proof. By definition, the resultant $\mathrm{Res}(\tilde{p}, \tilde{p}')$ is the determinant of the following $(2n-1) \times (2n-1)$ matrix

$$\begin{pmatrix}
a_n + \delta_n & \dots & a_1 + \delta_1 & a_0 + \delta_0 & 0 & \dots & 0 \\
0 & a_n + \delta_n & \dots & a_1 + \delta_1 & a_0 + \delta_0 & 0 & \dots \\
\vdots & \ddots & \dots & \ddots & \dots & \ddots & \dots \\
0 & \dots & 0 & a_n + \delta_n & \dots & a_1 + \delta_1 & a_0 + \delta_0 \\
n(a_n + \delta_n) & \dots & a_1 + \delta_1 & 0 & 0 & \dots & 0 \\
0 & n(a_n + \delta_n) & \dots & a_1 + \delta_1 & 0 & 0 & \dots \\
\vdots & \ddots & \ddots & \dots & \ddots & \dots & \vdots \\
0 & \dots & 0 & 0 & n(a_n + \delta_n) & \dots & a_1 + \delta_1
\end{pmatrix}$$

Applying a well known property of the matrices to the first row, and iterating this process, we obtain

$$\operatorname{Res}(\tilde{p},\tilde{p}') = \det \begin{pmatrix} a_n & \dots & a_1 & a_0 & 0 & \dots & 0 \\ 0 & a_n & \dots & a_1 & a_0 & 0 & \dots \\ \vdots & \ddots & \dots & \ddots & \dots & \ddots & \dots \\ 0 & \dots & 0 & a_n & \dots & a_1 & a_0 \\ na_n & \dots & a_1 & 0 & 0 & \dots & 0 \\ 0 & na_n & \dots & a_1 & 0 & 0 & \dots \\ \vdots & \ddots & \ddots & \dots & \ddots & \dots & \vdots \\ 0 & \dots & 0 & 0 & na_n & \dots & a_1 \end{pmatrix} + \dots$$

i.e. $\operatorname{Res}(\tilde{p},\tilde{p}') = \operatorname{Res}(p,p') + r$, where r is the sum of the other $2^{2n-1}-1$ determinants and it is a polynomial function of the δ_j's.

Since by hypothesis $|\delta_j| \le e_j$ for $j = 0,\dots,n$, then we can identify each δ_j with an interval $\delta_j = [-e_j, e_j]$. By using interval arithmetic let $[\rho,\eta]$ be the interval obtained by substituting $\delta_j = [-e_j, e_j]$ as interval in the formula of ϕ. If we look at $\operatorname{Res}(p,p')$ as a degenerate interval $[\operatorname{Res}(p,p'), \operatorname{Res}(p,p')]$, then

$$\operatorname{Res}(\tilde{p},\tilde{p}') = [\operatorname{Res}(p,p') + \rho, \operatorname{Res}(p,p') + \eta]$$

By working with intervals, whenever we ask $\operatorname{Res}(\tilde{p},\tilde{p}') \neq 0$, it must hold that $0 \notin [\operatorname{Res}(p,p') + \rho, \operatorname{Res}(p,p') + \eta]$, i.e. we have to guarantee that either

$$\begin{cases} \operatorname{Res}(p,p') + \rho > 0 \\ \operatorname{Res}(p,p') + \eta > 0 \end{cases} \quad \text{or} \quad \begin{cases} \operatorname{Res}(p,p') + \rho < 0 \\ \operatorname{Res}(p,p') + \eta < 0 \end{cases} \tag{18}$$

On the other hand if we want that $\tilde{p}(x)$ is not square free, we must impose

$$\begin{cases} \rho + \operatorname{Res}(p,p') < 0 \\ \eta + \operatorname{Res}(p,p') > 0 \end{cases} \tag{19}$$

so $0 \in [\operatorname{Res}(p,p') + \rho, \operatorname{Res}(p,p') + \eta]$. If we know that $p(x)$ is square free and $\operatorname{Res}(p,p') > 0$ (respectively $\operatorname{Res}(p,p') < 0$), then we can evaluate (17) in $\delta_j = 0$ for all $j = 0,\dots,n$. We find $\operatorname{Res}(\tilde{p},\tilde{p}')_{|\{\delta_0=0,\dots,\delta_n=0\}} = \operatorname{Res}(p,p') + \phi(\delta_0,\dots,\delta_n)_{|\{\delta_0=0,\dots,\delta_n=0\}} = \operatorname{Res}(p,p') > 0$. By using Bolzano's theorem, there is a neighborhood of the point $(0,\dots,0)$ in the space of $(\delta_0,\dots,\delta_n)$ where all perturbed polynomials in the neighborhood are square free.

Let us see some numerical examples performed with MATHEMATICA 4.1.

Example 5. Let $p(x) = 1.1x^2 + 2.3x + 1$ be an empirical polynomial with vector of tolerance $e = (10^{-2}, 10^{-2}, 10^{-1})$. $p(x)$ is square free, in fact $\operatorname{Res}(p,p') = -0.979 \neq 0$. A perturbed polynomial in $\mathcal{N}(p,e)$ is $\tilde{p}(x) = (1.1+\delta_2)x^2 + (2.3+\delta_1)x + (1+\delta_0)$, with $|\delta_2| \le 10^{-2}$, $|\delta_1| \le 10^{-2}$ and $|\delta_0| \le 10^{-1}$. By working with the δ_j's as intervals we compute $\operatorname{Res}(\tilde{p},\tilde{p}') = [-1.852, -0.087]$, and then we can assert that every polynomial in the neighborhood of $p(x)$ is square free.

Example 6. Let $q(x) = 0.03x^2 - 0.4x + 1.4$ be an empirical polynomial with vector of tolerance $e = (10^{-2}, 10^{-1}, 10^{-1})$. Let us work with four digits after the point. $q(x)$ is square free because $\mathrm{Res}(q, q') = 0.0002$, but for a perturbed $\tilde{q}(x) = (0.03 + \delta_2)x^2 + (-0.4 + \delta_1)x + 1.4 + \delta_0 \in \mathcal{N}(q, e)$ we have $\mathrm{Res}(\tilde{q}, \tilde{q}') = [-0.016, 0.016]$. So there exist non-square free polynomials in the neighborhood; for instance if we choose $\delta_2 = 0.01$, $\delta_1 = -0.08$ and $\delta_0 = 0.04$ we obtain the polynomial $\bar{p} = 0.04x^2 - 0.48x + 1.44 = (0.2x - 1.2)^2$, i.e. a non-square free polynomial.

6.1 Nearest Non-Square Free Polynomial

Given an empirical polynomial $p(x) = \sum_{j=0}^{n} a_j x^j$ and its tolerance vector e, we suppose that $p(x)$ is square free. We are interested in computing the minimal values $(\bar{\delta}_n, \ldots, \bar{\delta}_0)$ for the perturbations δ_j's with $|\bar{\delta}_j| \leq e_j$, for $j = 0, \ldots, n$, such that the polynomial $\bar{p}(x) = \sum_{j=0}^{n} (a_j + \bar{\delta}_j)x^j \in \mathcal{N}(p, e)$ is not square free, i.e. $\mathrm{Res}(\bar{p}, \bar{p}') = 0$. Let $\tilde{p}(x) = \sum_{j=0}^{n} (a_j + \delta_j)x^j$ be a generic perturbed polynomial in the neighborhood. The equation $\mathrm{Res}(\tilde{p}, \tilde{p}') = 0$ gives an hypersurface in the space $(\delta_0, \ldots, \delta_n)$. In order to find the minimum values for the perturbations we have to minimize the distance between the origin $(0, \ldots, 0)$ and a generic point $(\delta_0, \ldots, \delta_n)$ belonging to the hypersurface, then we have to minimize the value $\delta_0^2 + \ldots + \delta_n^2$ under the constraint $\mathrm{Res}(\tilde{p}, \tilde{p}') = 0$. A well-known method to determine the minima and maxima of a function under given constraints is the *Lagrange Multipliers Method.* In our case, we have the function $\mathcal{D} = \delta_0^2 + \ldots + \delta_n^2$ to minimize and the constraint $\mathrm{Res}(\tilde{p}, \tilde{p}') = 0$. We construct a new function in the indeterminates $\delta_0, \ldots, \delta_n$ with a new indeterminate λ:

$$F(\delta_0, \ldots, \delta_n, \lambda) = \delta_0^2 + \ldots + \delta_n^2 + \lambda \,\mathrm{Res}(\tilde{p}, \tilde{p}')$$

and we compute the solution of the following system:

$$\begin{cases} \frac{\partial F(\delta_0, \ldots, \delta_n, \lambda)}{\partial \delta_0} = 0 \\ \vdots \\ \frac{\partial F(\delta_0, \ldots, \delta_n, \lambda)}{\partial \delta_n} = 0 \\ \frac{\partial F(\delta_0, \ldots, \delta_n, \lambda)}{\partial \lambda} = 0 \end{cases} \tag{20}$$

Example 7. For the polynomial $q(x) = 0.03x^2 - 0.4x + 1.4$ as in Example 6, using the Lagrange Multipliers Method, we find the following assignment for the perturbations: $\delta_2 = -0.0014$, $\delta_1 = -0.0002$, $\delta_0 = -0.00003$. The correspondent polynomial $\bar{q}(x) = 0.286x^2 - 0.4002x + 1.39997 = (0.1691x - 1.183203)^2$ is not square free.

Let us give a final example [1] where we use all concepts till now discussed.

Example 8. Let $p(x) = x^4 - 2.83088x^3 + 0.00347x^2 + 5.66176x - 4.00694$ be an empirical polynomial and let $e = (0, 10^{-5}, 10^{-5}, 10^{-5}, 10^{-5})$ be its tolerance vector. We look for the zeros of $p(x)$. Using the method as described in Sect. 3.1, we

[1] Found in [11].

find the interval $[-1.4142168, -1.4142104]$, that gives the value -1.41421 as a good approximation of a zero of $p(x)$. For the other three zeros of the polynomial the situation becomes more complicated, in fact we can find three distinct real zeros or a real zero and a complex conjugate pair. In this case it is not possible to attribute individual pseudozero sets to each of the three zeros, but it must be considered as a single cluster.
Computing the resultant between a generic perturbed $\tilde{p} \in \mathcal{N}(p,e)$ and its derivative $\tilde{p}'$, we obtain $\mathrm{Res}(\tilde{p},\tilde{p}') = [-18.8098, 18.8098]$ and $\mathrm{Res}(p,p') = 9.96352 \times 10^{-16}$. This means that $p(x)$ is square free by an algebraic point of view, but "numerically" $p(x)$ is not square free because $\mathrm{Res}(p,p')$ is zero with respect to the precision we are working to. For instance the polynomial $\tilde{p}_1(x) = (x-\sqrt{2})^3(x+\sqrt{2})$ belongs to $\mathcal{N}(p,e)$ and it is not square free. This is why we cannot assign individual pseudozero sets to the other three zeros of $p(x)$.

7 Geometric Interpretation

In this section we analyze the geometric meaning of the concept of polynomial neighborhood and we focus *geometrically* the conditions discussed in the previous sections.

Let $p(x) = \sum_{j=0}^{n} a_j x^j$ be an empirical polynomial, let e be its tolerance vector and let $\tilde{p}(x) = \sum_{j=0}^{n} (a_j+\delta_j)x^j \in \mathcal{N}(p,e)$ be a perturbed polynomial. We have two possibilities.

(i). Let us consider the affine (n+1)-dimensional space over $\mathbb{R}$ given by the (n+1)-tuples $(\tilde{a}_0, \ldots, \tilde{a}_n)$. This means that we identify the initial polynomial $p(x)$ with the point $(a_0, \ldots, a_n)$ and the neighborhood with the polytope having points $(a_0+\delta_0, \ldots, a_n+\delta_n)$, with $|\delta_j| \leq e_j$ for all $j = 0, \ldots, n$.

(ii). Let us consider the affine (n+1)-dimensional space over $\mathbb{R}$ given by the (n+1)-tuples $(\delta_0, \ldots, \delta_n)$. This means that we identify the initial polynomial $p(x)$ with the point $(0, \ldots, 0)$ and the neighborhood with the polytope having points $(\delta_0, \ldots, \delta_n)$, with $|\delta_j| \leq e_j$ for all $j = 0, \ldots, n$.

We obtain case (ii) from case (i) by the linear isomorphism of affine spaces defined by $T_j = t_j - a_j$ for $j = 0, \ldots, n$, where t_j is a coordinate system in the space as in (i) and T_j is the coordinate system of the space as in (ii). Let denote by Δ^{n+1} the space of (n+1)-tuples $(\delta_0, \ldots, \delta_n)$. In our case we prefer the space Δ^{n+1} because, as in Sections 5 and 6, we find algebraic conditions on the δ_j's.

In Section 5 if we impose $\mathrm{Res}(p,\tilde{p}) = 0$, then we require more conditions on the δ_j's. If $\deg(p) = n$ then the resultant is a homogeneous polynomial of degree n in the δ_j's. It represents a hypersurface in the space Δ^{n+1}. The points where the polytope intersects this hypersurface are in one to one correspondence with the perturbed polynomials having at least a linear common factor with $p(x)$. Similarly when we ask for a k common-factor neighborhood, with $1 \leq k \leq n$, we have k homogeneous equations in the δ_j's. All common points between these hypersurfaces and the polytope are the $(n+1)$-tuple $(\delta_0, \ldots, \delta_n)$'s, that are in one to one correspondence with the $\tilde{p}(x)$ and have at least a common factor of degree k with $p(x)$.

Finally, in Section 6 we study the case when the polynomial $p(x)$ is square free. If the polynomial $\tilde{p}$ has no multiple factors, then we must have $\mathrm{Res}(\tilde{p}, \tilde{p}') \neq 0$. If $\deg(p) = n$ then $\mathrm{Res}(\tilde{p}, \tilde{p}')$ is a polynomial of degree $2n-1$ in the δ_j's. The condition $\mathrm{Res}(\tilde{p}, \tilde{p}') = 0$ locates a hypersurface in the space Δ^{n+1}. If the hypersurface intersects the polytope, then we can find the values of δ_j's, with $|\delta_j| \leq e_j$, such that $\mathrm{Res}(\tilde{p}, \tilde{p}') = 0$, i.e. such that the polynomial $\tilde{p}$ is not square free.

Acknowledgement

I would like to thank Prof. Giuseppa Carrà Ferro (University of Catania) and Prof. Hans J. Stetter (Tech. Univ. Vienna) for their useful comments.

References

1. Alefeld, G., Claudio, D. (1998): The basic properties of interval arithmetic, its software realization and some applications. Computer and Structures, **67**, 3–8
2. Emiris, I. Z. (1998): Symbolic-numeric algebra for polynomials. In: Encyclopedia of Computer Science and Technology, A. Kent and J.G. Williams, editors, vol. 39, 261–281, Marcel Dekker, New York
3. Geddes, K.O., Monogan, M.B., Heal, K.M., Labahn, G., Vorkoetter, S.M., McCarron, S. (2000): Maple 6, Programming Guide. Waterloo Maple Inc., Canada
4. Hitz, M.A., Kaltofen, E. (1998): Efficient algorithms for computing the nearest polynomial with constrained root. In: Proceed. ISSAC'98 (Ed. O.Gloor), 236–243
5. Hitz, M.A., Kaltofen, E. (1999): Efficient algorithms for computing the nearest polynomial with a real root and related problems. In: Proceed. ISSAC'99 (Ed. S.Dooley), 205–212
6. Karmarkar, N., Lakshman, Y.N. (1996): Approximate polynomial greatest common divisors and nearest singular polynomials. In: Proceed. ISSAC'96, 35–39
7. Laidacker, M.A. (1969): Another theorem relating Sylvester's matrix and the greatest common divisors. Mathematics Magazine **42**, 126–128
8. Jia, R. (1996): Perturbation of polynomial ideals. Advances in Applied Mathematics **17**, 308–336
9. Stetter, H.J. (1999): Polynomial with coefficients of limited accuracy. In: Computer Algebra in Scientific Computing (Eds. V.G.Ganzha, E.W.Mayr, E.V.Vorozhtsov), Springer, 409–430
10. Stetter, H.J. (1999): Nearest polynomial with a given zero, and similar problems. SIGSAM Bull. **33** n. 4, 2–4
11. Stetter, H.J. (2000): Numerical polynomial algebra: concepts and algorithms. In: ATCM 2000, Proceed. 5th Asian Technology Conf. in Math. (Eds.: W.-Ch. Yang, S.-Ch. Chu, J.-Ch. Chuan), 22–36, ATCM Inc. USA
12. Wolfram, S. (1999): The Mathematica Book. 4th ed., Wolfram Media, Cambridge University Press