Numerical Polynomial Algebra: Concepts and Algorithms

Hans J. Stetter
Tech.Univ.Vienna
stetter@aurora.anum.tuwien.ac.at

Abstract

In this expository lecture, we explain how problems of classical polynomial algebra may be understood and numerically solved when they appear in the context of Scientific Computing, i.e. with data of limited accuracy. A suitable conceptual framework is introduced and its use explained in a number of situations.

1 Introduction

We consider polynomials in their function as a *modelling tool* of Scientific Computing; throughout, coefficients and variables take values in $\mathbb{R}$ or $\mathbb{C}$. Moreover, we assume that some coefficients are only known to a *limited accuracy*. This may reflect the limited accuracy of measured or observed data or the effect of previous numerical computation. This assumption that algebraic objects may have a "tolerance" has far-reaching consequences for their mathematical and computational treatment. But *linear algebra* has accepted this assumption a long time ago; it has led to the enormous growth of *numerical linear algebra* into one of the supporting pillars of Scientific Computing. In an analogous fashion, the presently evolving *numerical polynomial algebra* studies the modifications and extensions of classical polynomial algebra necessary to accomodate the presence of inaccurate data and inexact computation. An outstanding token of this new development is the Strategic Alliance of 1998 between the scientific software houses WaterlooMaple and NAg which has shown its effects in the latest release Maple6.

Situations with geometric aspects are a typical source of polynomial systems; e.g., many problems in *robotics* permit such a formulation where some coefficients represent lengths and angles of robotic agents and relative positions of objects to be manipulated. It is obvious that these data have only a limited meaningful accuracy. In other areas of Scientific Computing (e.g. biology or economy), the indetermination of some coefficients may amount to several percent! In most real-life situations, the coefficient 4.865 in a polynomial $p(x,y) := x^3 - 2.913\ x^2y + 4.865\ xy^2 - 0.649\ y^3$ does not signify that precise rational number but rather *any real number* from a sufficiently small neighborhood of 4.865, say from the interval [4.864, 4.866]. The boundaries of that neighborhood are themselves *not sharply defined*: The appropriate interpretation is that there is an indetermination *of order* 10^{-3}.

In this lecture (cf. also [2]), we formalize the concept of *data with an indetermination* and show how the continued use of traditional mathematics is possible under these circumstances:

Valid approximate results are exact results for data within the tolerance neighborhood of the specified data. This principle leads to an *embedding of polynomial algebra into analysis*; such an embedding has become standard in computational linear algebra (linear equations, eigenproblems, least squares problems etc.). We introduce the necessary conceptual and algorithmic framework for this numerial polynomial alebra, and we consider a number of particular problems from the new point of view:

- Zeros of univariate polynomials,
- zeros of systems of multivariate polynomials,
- clusters of polynomial zeros,
- "solutions" of near-consistent overdetermined polynomial systems,
- greatest common divisors,
- factorization of multivariate polynomials.

Finally, we discuss the benefits which may accrue from the implementation of numerical polynomial algebra into current software systems for scientific computing.

2 Empirical Polynomials

In dealing with problems in Technical Scientific Computing, we observe that here are two categories of numerical data:

- *intrinsic data* represent an *exact value* from $\mathbb{R}$ or $\mathbb{C}$ in the sense of classical mathematics;
- *empirical data* have a *specified value* from $\mathbb{R}$ or $\mathbb{C}$ and a *tolerance* $\varepsilon \in \mathbb{R}_+$ which indicate the *range of potential values* for that quantitiy as described below.

A polynomial with at least one empirical coefficient will be called a *polynomial with a tolerance* or an *empirical polynomial*. The polynomial with the specified values of its empirical coefficients is the *specified polynomial*.

Empirical data and tolerances are a part of our everyday experience:

a) The specification of the electric voltage supplied to our homes has a specified value and a tolerance which is commonly interpreted in this way: "Almost always", the deviation of the momentary voltage from the specified one will be less than the tolerance; occasionally, the deviation may exceed this tolerance somewhat. An excess by an order of magnitude is not to be expected.

b) When the level of some larger body of water is reported as "5.16 m" it is clear that the reporting of more decimal digits would be meaningless: We assume that several measurements of the water level, at the same location and within a short time interval, will produce slightly deviating values. If most of these lie between 5.15 and 5.17, with occasional values like 5.14 or 5.19, we would associate a tolerance of .01 with the specified value 5.16.

Formally, with an empirical data quantity with M components, we associate a *family of neighborhoods*, parametrized by a positive real δ in the following way (cf. also [2]): Let

$$\bar{a} \;=\; (\bar{\alpha}_1, \ldots, \bar{\alpha}_M) \;\in\; \mathbb{C}^M \text{ or } \mathbb{R}^M\,, \tag{1}$$

be the *specified value* of the empirical data quantitiy, with the associated *tolerance*

$$e \;:=\; (\varepsilon_1, \ldots, \varepsilon_M)\,, \quad \varepsilon_j > 0\,. \tag{2}$$

Definition 1: The *δ-neighborhood* $N_\delta(\bar{a}, e)$ of the empirical data quantity $(\bar{a}, e)$ contains those values $\tilde{a} = (\tilde{\alpha}_1, \ldots, \tilde{\alpha}_M)$ of the quantity for which

$$\|\tilde{a} - \bar{a}\|_e^* \;:=\; \|(\ldots, \tfrac{|\tilde{\alpha}_j - \bar{\alpha}_j|}{\varepsilon_j}, \ldots)\|^* \;\;\leq\;\; \delta\,. \tag{3}$$

For $\bar{a}$ with *real* components $\bar{\alpha}_j$, it must be specified (or clear from the context) whether the $\tilde{a} \in N_\delta(\bar{p}, e)$ are restricted to *real* values or not. □

The norm $\|..\|^*$ in (3) is a vector norm in the *dual vector space* $(\mathbb{R}^M)^*$ of $\mathbb{R}^M$, i.e. the vector space of the linear functionals on $\mathbb{R}^M$: For an *absolute norm* $\|..\|$ on $\mathbb{R}^M$, the associated *dual norm* or *operator norm* is defined as

$$\|v^T\|^* \;:=\; \sup_{u \neq 0} \frac{|v^T u|}{\|u\|}\,. \tag{4}$$

In a finite-dimensional vector space, the sup in (4) is always attained.

In this lecture, we will exclusively use the combination

$$\|u\| \;:=\; \sum_j |u_j|\,, \qquad \|v^T\|^* \;:=\; \max_j |v_j|\,; \tag{5}$$

for $a \in N_\delta(\bar{a}, e)$, it requires

$$|\tilde{\alpha}_j - \bar{\alpha}_j| \;\leq\; \varepsilon_j\,\delta\,, \quad j = 1(1)M\,. \tag{6}$$

This is the only norm where the requirements on the individual components of $\tilde{a}$ remain *separated*.

For an *empirical polynomial* $(\bar{p}, e)$ in $s \geq 1$ variables, with the specified polynomial $\bar{p} \in \mathbb{C}[x_1, \ldots, x_s]$,

$$\bar{p} \;=\; \sum_{j \in J} \bar{\alpha}_j\, x^j\,, \quad J \subset \mathbb{N}^s, \quad \bar{\alpha}_j \in \mathbb{C} \text{ or } \mathbb{R}\,, \tag{7}$$

we let

$$\emptyset \;\neq\; \tilde{J} \;:=\; \{j \in J \,:\, \alpha_j \text{ is an empirical coefficient of } (\bar{p}, e)\} \;\subset\; J\,;$$

be the *empirical support* of $(\bar{p}, e)$, with $|\tilde{J}| = M$. Thus, the components ε_j of the *tolerance* e of the empirical polynomial $(\bar{p}, e)$ refer only to subscripts in the empirical support $\tilde{J}$:

$$e \;:=\; (\varepsilon_j > 0\,, \quad j \in \tilde{J})\,. \tag{8}$$

Definition 2: The *δ-neighborhood* $N_\delta(\bar{p}, e)$ contains those polynomials $\tilde{p} \in \mathbb{C}[x_1, \ldots, x_s]$, $\tilde{p}(x) = \sum_{j \in J} \tilde{\alpha}_j x^j$, with

$$\begin{gathered} \|\tilde{p} - \bar{p}\|_e^* \;:=\; \|(\ldots, \tfrac{|\tilde{\alpha}_j - \bar{\alpha}_j|}{\varepsilon_j}, \ldots,\; j \in \tilde{J})\|^* \;\;\leq\;\; \delta\,, \\ \tilde{\alpha}_j \;=\; \bar{\alpha}_j\,, \qquad j \in J \setminus \tilde{J}\,. \end{gathered} \tag{9}$$

I.e., $\tilde{p} \in \mathbb{P}^s$ is a *δ-neighbor* of $\bar{p}$ with tolerance e iff it has the *same support* as $\bar{p}$ and if its coefficients satisfy (9). For $\bar{p}$ with *real* coefficients $\bar{\alpha}_j$, it must be specified (or clear from the context) whether $N_\delta(\bar{p}, e)$ is restricted to *real* polynomials or not. □

In accordance with (6) and our intuitive introduction of the term "tolerance", we will say that

$$\left.\begin{array}{l}\tilde{a}\in N_1(\bar{a},e)\\ \tilde{p}\in N_1(\bar{p},e)\end{array}\right\} \quad\Longleftrightarrow\quad \left.\begin{array}{l}\tilde{a}\\ \tilde{p}\end{array}\right\} \text{ is } \textit{indistinguishable} \text{ from } \left\{\begin{array}{l}\bar{a}\\ \bar{p}\end{array}\right. . \tag{10}$$

Here, the term "indistinguishable" in (10) is to mean that all data in $N_1(\bar{a},e)$ furnish an *equally valid model* for the situation under consideration. In particular, it implies that there are *no better or worse data* in $N_1(\bar{a},e)$ for that purpose!

But (6) with $\delta = 1$ cannot be considered as a *strict bound* for valid values of the empirical quantity $(\bar{a},e)$. Such an assumption would simply shift the discontinuity between the statements $\tilde{a}=\bar{a}$ and $\tilde{a}\neq\bar{a}$ to another point. Such discontinuities are meaningless for data in $\mathbb{C}$ or $\mathbb{R}$ and also destroy the possibility of approximate (e.g. floating-point) computation; therefore, we have introduced the parametrized *family of neighborhoods* N_δ. Here, $\delta = 1$ is not a sharp bound for the validity of a data quantity $\tilde{a}$ but only a mark on the *continuous* δ-scale: The validity of $\tilde{a}$ decreases with an increase of the value of δ necessary to achieve $\tilde{a}\in N_\delta(\bar{a},e)$. The same interpretation applies for empirical polynomials.

A more precise adjustment of a validity scale to particular values of δ will vary with the application and must be left to the experts who have modelled a particular situation. But quite generally, one may conceive of an interpretation of δ like

δ	≤ 1	$\ldots$	3	$\ldots$	10	$\ldots$	30	$\ldots$	100
	valid		probably valid		possibly valid		probably invalid		invalid

(11)

Example 1: Consider the empirical polynomial $(\bar{p},e)$ with

$$\bar{p}(x,y) \;=\; x^3 + 4.865\,xy^2 - y^3 + 2.902\,x^2 + 0.0\,xy - 8.389\,x + 2\,y - 17.54\;,$$

with

$$\begin{aligned} J \;&=\; \{(3,0),(1,2),(0,3);(2,0),(1,1);(1,0),(0,1);(0,0)\}\;,\\ \tilde{J} \;&=\; \{(1,2);(2,0),(1,1);(1,0);(0,0)\}\;.\\ e \;&=\; (10^{-3};\; 5\cdot 10^{-4}, 10^{-4};\; 10^{-3};\; 5\cdot 10^{-3})\;, \end{aligned}$$

With our norm definiton (5), the following polynomials are indistinguishable from $\bar{p}$ and hence valid representations of the empirical polynomial $(\bar{p},e)$:

$$\tilde{p}_1(x,y) \;=\; x^3 + 4.8642\,xy^2 - y^3 + 2.9025\,x^2 - 8.3888\,x + 2\,y - 17.541\;,$$

$$\tilde{p}_2(x,y) \;=\; x^3 + 4.865\,xy^2 - y^3 + 2.9018\,x^2 - 0.0001\,xy - 8.3896\,x + 2\,y - 17.536\;.$$

Note that the xy-term has an *empiric* coefficient 0 in $(\bar{p},e)$; therefore, it must appear in the supports J and $\tilde{J}$ and it may appear with a sufficiently small coefficient in a valid representation of $(\bar{p},e)$, cf. $\tilde{p}_2$. $\square$

3 Valid Approximate Results

It is our goal to solve algebraic problems with polynomial data of limited accuracy in a meaningful way. Since we have a dense infinite set $\tilde{A}_1 \subset \mathcal{A}$ of *equally valid* input data for such a problem, we must generally expect a dense infinite set of result values which are *equally valid* as results of the given problem. We will now formalize this situation; cf. also [2].

For an empirical algebraic problem, we consider the mapping from the space $\mathcal{A}$ of the empirical data $(\bar{a}, e)$ to a result space $\mathcal{Z}$ which assigns to a particular input value $\tilde{a} \in N_{\bar{\delta}}(\bar{a}, e)$ the value $\tilde{z} \in \mathcal{Z}$ of the *exact* result of the algebraic problem for $\tilde{a}$. Here, $\tilde{z}$ may be a real or a complex number or a set of such numbers. When there are several independent result quantities, we may restrict our mapping to one particular result quantity. (Integer results are not considered in this lecture.) $\bar{\delta} > 1$ denotes a value beyond which we do not wish to consider the indetermination of our input data; cf. (3) and (11). We consider the intrinsic data as a fixed part of the specification of our algebraic problem and also of the mapping introduced above.

Definition 3: For an empirical algebraic problem, with empirical data $(\bar{a}, e)$, the mapping

$$F : \quad \tilde{A}_{\bar{\delta}} \subset \mathcal{A} \quad \longrightarrow \quad \mathcal{Z} \tag{12}$$

which assigns to each data value $\tilde{a} = (\tilde{\alpha_j})$ in $\tilde{A}_{\bar{\delta}}$ the *exact* result $z \in \mathcal{Z}$ of the algebraic problem with these data, is called the *data→result mapping* for that problem. □

Definition 4: For a data→result mapping (12), the sets

$$Z_\delta \;:=\; \{z := F(a)\,, \quad a \in N_\delta(\bar{a}, e)\}\,, \quad \delta \le \bar{\delta}\,, \tag{13}$$

are called (δ-)*pseudoresult sets* of (12). In the case of *real* specified data $\bar{a}$, it must be clear whether the data neighborhoods $N_\delta(\bar{a}, e)$ are restricted to the real domain or not. □

Obviously, each result value $\tilde{z} \in Z_\delta$ is the *exact result* of the algebraic problem for some input data $\tilde{a} \in N_\delta(\bar{a}, e)$. In particular, each result value in Z_1 is the exact result for some input data which are *indistinguishable* from the specified data $\bar{a}$ (cf. (10)) and thus constitute a valid representation of the model under discussion. Therefore, *each value* in the 1-pseudoresult set Z_1 is a valid result of the algebraic problem for that model, as valid as any other value in Z_1. Moreover, in the relaxed sense of (11), this is also true of values in Z_δ with $\delta = \mathrm{O}(1)$.

Definition 5: In the situation described by a data→result mapping (12), the values in the δ-pseudoresult sets Z_δ with $\delta = \mathrm{O}(1)$ are *valid approximate results* of the empirical algebraic problem. □

The word "approximate" in Definition 5 serves to emphasize that any δ-pseudoresult – while being the exact result for some $\tilde{a} \in N_\delta(\bar{a}, e)$ in the classical sense – can only represent an *approximate result* for the problem at hand.

Example 2: Let the univariate monic polynomial

$$\bar{p}(x) \;:=\; x^4 - 2.83088\,x^3 + 0.00347\,x^2 + 5.66176\,x - 4.00694 \tag{14}$$

be the specified polynomial of an empirical polynomial $(\bar{p}, e)$ with the tolerance vector $e = (\varepsilon_j = 10^{-5},\ j = 0(1)3)$. This means that all coefficients except the leading one are known only to within one unit of their trailing digit; we assume that they are restricted to the real domain.

We wish to determine the zeros of the empirical polynomial $(\bar{p}, e)$. As $\bar{p}$ is a perturbed version of $(x-\sqrt{2})^3(x+\sqrt{2})$ we expect one real zero in the left halfplane (near $-\sqrt{2}$) and three zeros in the right halfplane (near $\sqrt{2}$).

Let us first look at the zero in the left halfplane: With a considerate shifting of the coefficients to appropriate corners of the domain N_1 and with the help of a zerofinding code, we may convince ourselves that the negative zero can take values in the interval $[-1.4142168, -1.4142104]$ (rounded to 7 digits), for $a \in N_1$. Thus, e.g., any of the 6-digit values $-1.414216, \ldots, -1.414211$ is a valid approximate zero $\tilde{z}$ of $(\bar{p}, e)$.

This changes dramatically when we consider the three zeros in the right halfplane which form a cluster. For the specified polynomial $\bar{p}$, a zerofinder yields the three close real zeros (rounded to 5 digits) 1.41421, 1.41481, 1.41607. However, the equally valid polynomial

$$\tilde{p}(x) \;:=\; x^4 - 2.83087\,x^3 + 0.00348\,x^2 + 5.66177\,x - 4.00693 \;\in\; N_1(\bar{p}, e)$$

yields the exact zeros (rounded to 5 digits) 1.38583 and 1.42963 ± 0.02578 i in the right halfplane!

One can easily convince oneself that it is not possible to attribute individual pseudozero sets Z_1 to each of the three zeros in the right halfplane but that we must treat the three zeros as *one* result which may as well consist of three real zeros as of one real zero and a complex conjugate pair. Quantitatively, it appears that this pseudozero set Z_1 for the zero triple in the right halfplane has a diameter of nearly $3 \cdot 10^{-2}$! Each $\tilde{z} \in Z_1$ is an exact zero of some $\tilde{p} \in N_1(\bar{p}, e)$ and there are exactly three zeros (counting potential multiplicities) in Z_1 for each $\tilde{p} \in N_1(\bar{p}, e)$. □

So far, we have tacitly assumed some basic *regularity* of the algebraic problem: We expect that the domain $\tilde{A}_{\bar{\delta}}$ of the data→result mapping F is *open* in the data space $\mathcal{A}$ and that it contains the neighborhood $N_{\bar{\delta}}(\bar{a}, e)$ so that results exist for *all* $\tilde{a}$ in a neighborhood $N_\delta(\bar{a}, e)$, $\delta < \bar{\delta}$. We also expect that F is surjective so that the family $\{Z_\delta\}$ of pseudoresult sets consists of *compact connected* sets in the natural topology of the result space $\mathcal{Z}$ and that a full neighborhood of an approximate result value $\tilde{z} \in Z_\delta$ also consists of approximate results, perhaps for a slightly larger δ. These are natural assumptions for a problem whose results are to be determined by approximate computations.

Proposition 1: For an algebraic problem with empirical data, assume that

$$M \;:=\; \dim \mathcal{A} \;\geq\; \dim \mathcal{Z} \;=:\; m\,, \tag{15}$$

that the data→result mapping F is *continuous* on the open domain $\tilde{A}_{\bar{\delta}} \supset N_{\bar{\delta}}(\bar{a}, e)$, $\bar{\delta} > 1$, and that the image of $\tilde{A}_{\bar{\delta}}$ by F has dimension m in $\mathcal{Z}$. Then, the δ-pseudoresult sets Z_δ are compact connected sets in $\mathcal{Z}$.

Proof: By (3) and our assumptions, the neighborhoods $N_\delta(\bar{a}, e)$ are compact connected sets of dimension M in $\mathcal{A}$. A continuous surjective mapping maps a compact connected set onto a compact connected set in the image space. □

Definition 6: An empirical algebraic problem whose data→result mapping satisfies the hypotheses of Proposition 1 is called *well-posed*, otherwise it is called *ill-posed*. □

There are various non-trivial ways in which an algebraic problem may be ill-posed:

1) A (proper) result of the algebraic problem is only defined for data on a manifold $\mathcal{S}$ of a dimension $< M$ in $\mathcal{A}$:

In this case, if the intersection of $\mathcal{S}$ with $N_{\bar{\delta}}(\bar{a}, e)$ is empty, the family of pseudoresult sets is empty and no valid representation of the algebraic problem has an exact solution. On the other hand, if the domain of F on the manifold $\mathcal{S}$ has a component of dimension $d \geq m$ (the result space dimension) which has a nonempty intersection with the neighborhoods $N_\delta(\bar{a}, e)$ for all $\delta \geq \delta_0$, $\delta_0 \leq 1$, we may restrict F to this component. Moreover, if this intersection is homeomorphic to a compact part of a linear space of dimension d, we have – in a sense – a well-posed empirical problem. Note that the exact problem with the *specified* data $\bar{a}$ has no solution if $\delta_0 > 0$!

Example 3: Consider two real univariate empirical polynomials $(\bar{p}_i, e_i), i = 1, 2$, of degrees > 1 and assume that the $\bar{p}_i$ have disjoint simple zeros near some $\zeta_0 \neq 0$ while their remaining zeros are sufficiently distinct so that $\gcd(\bar{p}_1, \bar{p}_2) = 1$. But we assume that $\zeta_0 \in Z_1(\bar{p}_i, e_i)$, $i = 1, 2$. In the data space $\mathcal{A}$ of all the empirical coefficients of the two polynomials, let $\mathcal{S}$ be that component of the coefficient manifold where $\gcd(\tilde{p}_1, \tilde{p}_2)$ has positive degree which intersects with $N_1((\bar{p}_1, \bar{p}_2), (e_1, e_2))$. With this restriction, the problem of determining a nontrivial gcd of the two empirical polynomials is regular and has valid approximate solutions. □

2) The dimension of the image $F(\tilde{A}_{\bar{\delta}})$ in the result space $\mathcal{Z}$ is lower than m. This implies that, in each arbitrarily small neighborhood in $\mathcal{Z}$ of a valid approximate result $\tilde{z}$, there are infinitely many values z which cannot be interpreted as exact results of the algebraic problem for whatever values of the empirical data. Thus, the perturbations induced by numerical computation will generally prevent a computed result to be an element of a pseudoresult set although it may be very close to a valid approximate result.

This phenomenon may have two reasons: Either the dimension condition (15) of Proposition 1 is violated, e.g. when there are fewer empirical data items than there are components in a result of the problem. Or the data→result mapping is not surjective because the algebraic problem restricts its results to some lower-dimensional manifold in the result space. In both cases, the problem is not well-posed by our Definition 5.

Example 4: Consider a monic univariate polynomial $\bar{p}$ with *one* empiric coefficient; we wish to determine numerically an approximate quadratic polynomial divisor $\tilde{s}$. Here, the empirical data space has dimension $M = 1$ while the result space has dimension $m = 2$ so that condition (15) is violated. Nevertheless, we may wish to call $\tilde{s}$ a valid approximate divisor if it is sufficiently close to an exact divisor of some $\tilde{p} \in N_1(\bar{p}, e)$ in a sense to be specified. □

Ill-posed problems abound in polynomial algebra; thus, their numerical treatment for the case of empirical data is of particular interest. Later in this lecture, we will show how this can be achieved in particular instances; cf. sections 7 and 8. At the moment, we assume that we are dealing with well-posed problems: Then, the formal framework introduced in the previous two sections poses three fundamental tasks:

(i) Given a value $\tilde{z} \in \mathcal{Z}$ from whatever source, determine whether this is a valid approximate result of the empirical problem.

(ii) If $\tilde{z}$ is *invalid* by a moderate margin only, find a valid approximate result by some further computation.

(iii) For the numerical specification of a valid approximate result $\tilde{z}$, how many digits are meaningful ?

The next three sections will be devoted to the solution of these three tasks.

4 Backward Error of Approximate Results

The pseudoresult sets of empirical algebraic problems are an important conceptual tool; but their explicit determination, even to a low degree of relative accuracy, requires a very high computational effort in all but trivial situations. If the results have several complex components, even a representation of some Z_δ appears infeasible. The more is it important that we are able to *check and verify* whether some $\tilde{z} \in \mathcal{Z}$ is a valid approximate solution.

According to Definition 3, $\tilde{z}$ is a valid approximate result if there exist data $\tilde{a}$ in some neighborhood $N_\delta(\bar{a}, e)$ with $\delta = \mathrm{O}(1)$ such that $\tilde{z}$ is the *exact* result of our problem with data $\tilde{a}$. In the verification of this condition, the set of *all* data $\tilde{a} \in \mathcal{A}$ for which this condition holds plays an important role.

Definition 7: For an empirical problem with data→result function $F : \mathcal{A} \to \mathcal{Z}$, and for a *given* result value $\tilde{z} \in \mathcal{Z}$, the *equivalent-data set* is defined by

$$\mathcal{M}(\tilde{z}) \;:=\; \{\tilde{a} \in \mathcal{A} \,:\, F(\tilde{a}) = \tilde{z}\}\,, \tag{16}$$

For algebraic problems, the equivalent-data set is generally an algebraic manifold in the space $\mathcal{A}$ of the empirical data. □

Example 5: Consider an empirical polynomial $(\bar{p}, e)$, with $\bar{p}(x) = \sum_{j\in J} \bar{\alpha}_j x^j$ and the empirical support $\tilde{J} \subset J$; cf. (7). In order that a specified $\tilde{z}$ is a zero of $\tilde{p}(x) = \sum_{j\in J} \tilde{\alpha}_j x^j$, the coefficients $\tilde{\alpha}$ must satisfy

$$\tilde{p}(\tilde{z}) \;=\; \sum_{j\in J} \tilde{\alpha}_j \tilde{z}^j \;=\; \sum_{j\in J} (\tilde{\alpha}_j - \bar{\alpha}_j)\, \tilde{z}^j + \bar{p}(\tilde{z}) \;=\; 0\,;$$

Since $\tilde{\alpha}_j = \bar{\alpha}_j$ for $j \in J \setminus \tilde{J}$,

$$\mathcal{M}(\tilde{z}) \;:=\; \{a \in \mathcal{A} : \sum_{j\in \tilde{J}} (\tilde{\alpha}_j - \bar{\alpha}_j)\, \tilde{z}^j + \bar{p}(\tilde{z}) = 0\}\,. \tag{17}$$

Thus, the equivalent-data set is a *linear manifold* in the space $\mathcal{A}$ of the empirical coefficients; its representation requires merely the computation of the *residual* $\bar{p}(\tilde{z})$. □

As in this example, explicit or implicit representations for the equivalent-data manifold can generally be obtained for computational algebraic problems. Moreover, $\mathcal{M}(z)$ is often a *linear manifold* in $\mathcal{A}$; this is the case if the empirical data are coefficients of *polynomials* which occur in the problem in a linear fashion: polynomials are *linear in their coefficients*!

The verification task

$$\text{Given } \tilde{z} \in \mathcal{Z} \;:\quad \exists\, \tilde{a} \in N_\delta(\bar{a}, e) \;:\; F(\tilde{a}) = \tilde{z}\;? \tag{18}$$

may now be reduced to the following two steps:

(i) Determine the equivalent-data manifold $\mathcal{M}(\tilde{z})$;

(ii) Check whether $\mathcal{M}(\tilde{z})$ has a non-empty intersection with $N_\delta(\bar{a}, e)$.

Since the neighborhood $N_\delta(\bar{a}, e)$ is defined by the metric (3) step (ii) is equivalent to:

(ii)' Find the shortest distance $\delta(\tilde{z})$ of $\mathcal{M}(\tilde{z})$ from $\bar{a}$ in the metric (3).

Definition 8: In the situation just described,

$$\delta(\tilde{z}) \;:=\; \min_{a\in\mathcal{M}(\tilde{z})} \|a - \bar{a}\|_e^* \tag{19}$$

is the *backward error*[1] of the approximate result $\tilde{z}$; cf. (3) - (6) for the definition of $\|..\|_e^*$. The shortest distance from $\mathcal{M}(\tilde{z})$ to $\bar{a}$ is well-defined because $\mathcal{M}(\tilde{z}) \cap N_{\bar{\delta}}(\bar{a}, e)$ is a compact set. □

To simplify the notation and without loss of generality, we shift the origin of the data space $\mathcal{A}$ to $\bar{a}$; intuitively, this means that we use the *deviations* $\Delta\alpha_j := \alpha_j - \bar{\alpha}_j$ as variables in $\mathcal{A}$. For a *linear manifold*, the task of finding the shortest norm distance from the origin is classical. Here, we consider only the norm and dual norm (5) and the associated norms $\|..\|_e^*$ from (3) and $\|u\|_e := \sum_j \varepsilon_j \, |u_j|$; for other norms, see, e.g., [2].

For a linear manifold of codimension 1 (a "hyperplane"), there are explicit expressions for the shortest norm distance from the origin as well as for the point where it is attained:

Proposition 2: Consider the linear manifold $\mathcal{M}(c)$ of codimension 1 in $\mathbb{C}^M$ specified by the linear equation $\sum_{j=1}^M \gamma_j \, \alpha_j = \gamma_0$. The shortest max-norm distance of $\mathcal{M}(c)$ from the origin is

$$\min_{a \in \mathcal{M}(c)} \max_j |\alpha_j| \;=\; \frac{|\gamma_0|}{\sum_{j=1}^M |\gamma_j|} \,, \tag{20}$$

which is attained for (here $..^*$ denotes the complex-conjugate value)

$$a_{\min} \;=\; \frac{\gamma_0}{\sum_{j=1}^M |\gamma_j|} \cdot \left(\ldots \; \frac{\gamma_j^*}{|\gamma_j|} \; \ldots \right) . \tag{21}$$

Proof: From $|\sum \gamma_j \, \alpha_j| \leq \sum |\gamma_j| \cdot \max |\alpha_j|$, we have

$$\max |\alpha_j| \;\geq\; \frac{|\sum \gamma_j \, \alpha_j|}{\sum |\gamma_j|} \;=\; \frac{|\gamma_0|}{\sum |\gamma_j|} \,.$$

It is easily confirmed that equality is attained for $a_{\min}$ from (21). □

Example 5, continued: For an approximate zero $\tilde{z} \in \mathbb{C}^s$, $s \geq 1$, of an empirical polynomial $(\bar{p}, e)$ with empirical support $\tilde{J}$, $|\tilde{J}| = M$, we had obtained the (shifted) equivalent-data manifold

$$\mathcal{M}(\tilde{z}) \;=\; \{\Delta a \in \mathcal{A} \,:\, \sum_{j \in \tilde{J}} \Delta\alpha_j \, \tilde{z}^j + \bar{p}(\tilde{z}) = 0\} \,. \tag{22}$$

Thus, by (20), the backward error of $\tilde{z}$ with the weighted norm (3) is

$$\delta(\tilde{z}) \;:=\; \min_{\Delta a \in \mathcal{M}(\tilde{z})} \|\Delta a\|_e^* \;=\; \frac{|\bar{p}(\tilde{z})|}{\|\, (\tilde{z}^j) \,\|_e} \;=\; |\bar{p}(\tilde{z})| / \sum_{j \in \tilde{J}} \varepsilon_j |\tilde{z}|^j \,. \qquad \square \tag{23}$$

If the set $\mathcal{M}(\tilde{z})$ is a linear manifold of a codimension > 1 or a nonlinear algebraic manifold, the determination of the backward error is not so simple; but there exist well-known numerical algorithms for solving the associated minimization problems (19). Also, we may generally assume that the equivalent-data manifold of a reasonably computed approximate result passes close by the origin and restrict our search to a domain about the origin. Finally, if we have

[1]The term "backward error" has been introduced by J. Wilkinson in his classical text of 1963. His idea to interpret the deviation (or "error" in numerical analysis terminology) of an approximate result $\tilde{z}$ as the effect of a deviation in the data of the original problem has become a central tool in numerical analysis and, more generally, in applied mathematics.

located some point on $\mathcal{M}(\tilde{z})$ with a norm distance ≤ 1 from the origin, we are finished: then $\tilde{z}$ is a valid approximate result.

Example 6: For the empirical polynomial (14), $\tilde{s}(x) = (x - \sqrt{2})^3$ is an approximate divisor. To determine the *backward error* $\delta(\tilde{s})$ of $\tilde{s}$ as a divisor of $(\bar{p}, e)$, we note that the equivalent-data manifold $\mathcal{M}(\tilde{s}) \subset \mathcal{A} = \mathbb{C}^4$ is the 1-dimensional linear manifold of the coefficients of $\tilde{p}(x) = (x - \beta)\,\tilde{s}(x) =: p(x, \bar{a} + \Delta a(\beta))$, i.e. (to 5 decimal digits) $\Delta\alpha_0 = 4.00694 - 2.82843\,\beta$, $\Delta\alpha_1 = -8.49019 + 6\,\beta$, $\Delta\alpha_2 = 5.99653 - 4.24264\,\beta$, $\Delta\alpha_3 = -1.41176 + \beta$. Its minimal distance from the origin is found as

$$\delta(\tilde{s}) \;=\; \min_{\beta} \|\Delta a(\beta)\|_e^* \;\approx\; 555\,;$$

hence $\tilde{s}$ is not a valid approximate divisor of $(\bar{p}, e)$. $\square$

5 Iterative Refinement of Approximate Results

We consider task (ii) of the end of section 3: For a well-posed problem, we have an approximate result $\tilde{z} \in \mathcal{Z}$ with a backward error moderately larger than O(1); how may we correct it into a valid approximate result? The most flexible scheme for that purpose is the method of *local linearization* or *Newton's method* which has been well-known and widely used for a long time, particularly for the refinement of a solution of one equation in one variable. We address the more general case of m equations in m results, with empirical data from a data space $\mathcal{A}$.

Assume that our empirical problem may be formulated in terms of m *equations*:

$$G(x;\, a) \;=\; 0\,, \tag{24}$$

with a mapping $G : \mathbb{C}^m \times \mathbb{C}^M \to \mathbb{C}^m$ so that the data→result mapping F is defined by

$$G(F(a);\, a) \;=\; 0 \tag{25}$$

for all a in a sufficiently large neighborhood of the specified data $\bar{a}$. Assume that G is differentiable w.r.t. to both arguments in that neighborhood. We want to determine Δz such that

$$0 \;=\; G(\tilde{z} + \Delta z, \bar{a}) \;=\; G(\tilde{z}, \bar{a}) + \frac{\partial G}{\partial x}(\tilde{z}, \bar{a})\,\Delta z + \mathrm{O}(\|\Delta z\|^2)\,. \tag{26}$$

After the omission of the quadratic term in (26), we have a *linear problem* of the same nature as our given problem. Its solution will not solve the left-hand equality in (26) exactly; but often the resulting corrected result $\tilde{z} + \Delta z$ will be a valid approximate result of (24) which is all we want.

For a well-posed polynomial system of m equations in m variables, $\frac{\partial G}{\partial x}$ is simply the *Jacobian* matrix of the system and the *Newton correction step* (26) is also well-known. We illustrate the formation of a Newton correction for a typical algebraic problem:

Example 7: Assume that we have found a poor approximate factor $\tilde{s}(x)$ (monic of degree m) for a given univariate empirical polynomial $p(x)$ (monic of degree n) and we wish to correct it into a valid approximate factor $\tilde{s}(x) + \Delta s(x)$. The associated correction step

$$p(x) \;=\; \tilde{s}(x)\,q(x) + r(x) \;=\; (\tilde{s}(x) + \Delta s(x))\,(q(x) + \Delta q(x)) \tag{27}$$

represents a system of n bilinear equations for the m coefficients of $\Delta s(x)$ and the $n-m$ coefficients of $\Delta q(x)$. The linearization of (27) is straightforward and generates the *linear system*

$$\Delta s(x)\, q(x) + \tilde{s}(x)\, \Delta q(x) \;=\; r(x)\,. \tag{28}$$

Since we are only interested in Δs, we can further simplify (28) by taking remainders modulo $\tilde{s}$; this leaves Δs and r (of degree $m-1$) unchanged while we may have to form the remainder of q mod $\tilde{s}$. We obtain

$$\mathrm{rem}(\Delta s(x) \,\mathrm{rem}\, q(x)) \;=\; r(x)$$

which is a linear system for the m coefficients of Δs only. $\square$

6 Condition Estimates

The remaining task (iii) of section 3 asks for an estimate of the extension of a pseudoresult set Z_δ, $\delta = \mathrm{O}(1)$: If equally valid pseudoresults $\tilde{z}$ can differ by $\mathrm{O}(10^{-r})$, it is meaningless or even misleading to specify more than r decimal digits of a $\tilde{z}$. Since the variation of the values in Z_δ is the consequence of the variation of the data in N_δ, this is a classical case of a *condition estimate*. We apply the following well-known result to the data→result mapping F of our empirical problem:

Proposition 3: Let $F : \mathcal{A} \to \mathcal{Z}$ have a Lipschitz continuous Frechet derivative, with Lipschitz constant L', in some convex domain $A \subset \mathcal{A}$; then, for $a, a+\Delta a \in A$,

$$F(a+\Delta a) - F(a) \;=\; F'(a)\,\Delta a + r(a,\Delta a) \quad \text{with } \|r(a,\Delta a)\| \;\le\; \frac{L'}{2}\,(\|\Delta a\|^*)^2\,. \tag{29}$$

Thus, in a reasonably regular situation with a moderate L', the linear mapping $F'(a)$ describes the effect of *small perturbations* of the data well. In the following, we will always neglect quadratic terms in $\|\Delta a\|^*$ and assume

$$Z_\delta(\bar{a}, e) \;=\; \{\tilde{z} \in \mathcal{Z} \,:\, \tilde{z} = F(a) + F'(a)\,(\tilde{a}-\bar{a})\,,\ \tilde{a} \in N_\delta(\bar{a}, e)\}\,, \tag{30}$$

which implies

$$\mathrm{diam}\ Z_\delta(\bar{a}, e) \;=\; 2\,\delta\,|F'(\bar{a})\,e| \;\le\; 2\,\delta\,\|F'(\bar{a})\|\,\|e\|^*\,. \tag{31}$$

A large diameter of $Z_\delta(\bar{a}, e)$ may be due to only one stretched direction in the result space $\mathcal{Z}$. Therefore, it may be interesting also to look at the *singular values* of $F'(\bar{a})$ which represent the half axes of the ellipsoidal image of a Euclidean ball about $\bar{a}$ under the linear map $F'(\bar{a})$.

Generally, the Jacobian F' of the data→result mapping F is not explicitly available because F is implicitly defined by a system of equations (24). Again we assume that G is differentiable w.r.t. to both arguments in a suitable domain; then differentiation of (25) w.r.t. a yields, with $\bar{z} = F(\bar{a})$,

$$\frac{\partial G}{\partial x}(\bar{z}, \bar{a}) \cdot F'(\bar{a}) + \frac{\partial G}{\partial a}(\bar{z}, \bar{a}) \;=\; 0\,, \tag{32}$$

or

$$F'(\bar{a}) \;=\; -\left[\frac{\partial G}{\partial x}(\bar{z}, \bar{a})\right]^{-1} \frac{\partial G}{\partial a}(\bar{z}, \bar{a})\,. \tag{33}$$

Example 8: For an isolated zero z of a univariate polynomial p we have $G(x;\alpha_0,\ldots,\alpha_n) = \sum_{\nu=0}^{n}\alpha_\nu x^\nu$ and

$$F'(\bar{a}) \;=\; \frac{1}{p'(\bar{z})}\;(1,\,\bar{z},\,\ldots,\,\bar{z}^n)\;;$$

by (31),

$$\text{diam}\;Z(\bar{a},e) \;=\; 2\,\delta\;|F'(\bar{a})\,e| \;=\; \frac{2\,\delta}{|p'(\bar{z})|}\;\sum_{\nu=0}^{n}\varepsilon_\nu|\bar{z}|^\nu \;=\; \frac{2\,\delta}{|p'(\bar{z})|}\;\|\begin{pmatrix}1\\ \bar{z}\\ \vdots\\ \bar{z}^n\end{pmatrix}\|_e \qquad (34)$$

Due to the occurrence of p' in the denominator, this estimate is not applicable for "zero clusters", i.e. pseudozero domains which contain more than one zero of each $\tilde{p}\in N_\delta(\bar{p},e)$. □

7 Zero Clusters

In Example 2, we had observed that, for an empirical polynomial, the zeros in an m-cluster have *no individuality* in the following sense: If we mark a certain simple zero of $p(x,\bar{a})$ in $Z_\delta(\bar{a},e)$ and then vary the coefficients of $p(x,\tilde{a})$ continuously over $N_\delta(\bar{a},e)$, we cannot follow that zero in a unique fashion: There occurs at least one confluence of zeros in this process. It may even happen that there exists an $\tilde{a}\in N_\delta(\bar{a},e)$ where $p(x,\tilde{a})$ has an m-fold zero in $Z_\delta(\bar{a},e)$. When we consider a cluster of m zeros as one m-dimensional result quantity, then we find that not each m-tuple of points in $Z_\delta(\bar{a},e)$ can appear as an m-tuple of exact zeros of some $p(x,\tilde{a})$. Thus, by Definition 6, our problem is *ill-posed*.

The appropriate tool for the description of an m-cluster is the monic polynomial $s(x,c)$ of degree m which vanishes at the zeros of the cluster; cf. [1]. The *coefficients of* s depend *Lipschitz continuously* and smoothly on the coefficients of p, due to the following result (cf.[4], Theorem 7):

Proposition 4: Let

$$p(x,a) \;=\; q(x,b)\cdot s(x,c)\,, \qquad (35)$$

with monic p,q,s of degrees n, $n-m$, m resp., with s and q coprime. Then there exists an $n\times m$-matrix C such that $p(x,a+\Delta a) = q(x,b+\Delta b)\cdot s(x,c+\Delta c)$ implies

$$(\;\ldots\Delta c^T\ldots\;) = (\;\ldots\Delta a^T\ldots\;)\cdot C + \;\mathrm{O}(\|\Delta a\|^2)\,, \qquad \text{i.e.}\quad C \;=\; \left(\frac{dc}{da}\right)_a\,. \qquad (36)$$

Remark: The coprimality of s and q requires that s contains all its zeros with their full multiplicity. C has a moderate norm if the zeros represented by s are well separated from the remaining zeros of p. □

Proposition 4 implies that the computation of the coefficients of the *"cluster polynomial"* s is a *well-posed problem* so that we can apply all concepts and algorithms of the previous sections to it. It also has the important practical consequence that the *moments* of the zeros in a cluster depend Lipschitz continuously and smoothly on the coefficients of p because, by Vieta's Theorem, the coefficients of s are the fundamental symmetric functions of the cluster zeros, and the moments of a set of m points $z_\mu\in\mathbb{C}$ may be expressed as polynomials in their fundamental symmetric functions. For the *arithmetic mean* of the cluster zeros, this result

has been known for some time. Our approach shows that *all* moments of the m cluster zeros vary only linearly with a small change of the coefficients of p. This suggests that – for an empirical polynomial – the *moments* of the zeros of a cluster should be computed and specified rather than the locations of the individual zeros. These moments lie in pseudoresult sets whose approximate diameters can be estimated from (36).

Example 9: In Example 6, we have found that $\tilde{s}(x) = (x - \sqrt{2})^3$ is not a valid approximate cluster polynomial for the 3-cluster of the empirical polynomial (14) about $\sqrt{2}$; therefore we perform one Newton correction step as described in Example 7 which yields, from a 3×3 linear system, $\Delta s(x) \approx -.00491 + .00694\,x - .00246\,x^2$. The refined cluster polynomial

$$\tilde{s}_{\rm new}(x) \;=\; \tilde{s}(x) + \Delta s(x) \;\approx\; x^3 - 4.24510\,x^2 + 6.00694\,x - 2.83334\,,$$

with $\delta(\tilde{s}_{\rm new}) < .5$, is a valid approximate description of the 3-cluster of our empirical polynomial. On the other hand, from the columns of C (cf. (36)), one finds that the pseudoresult sets for the coefficients $\gamma_0, \gamma_1, \gamma_2$ of s have diameters of $\approx (2.3, 3.6, 2.1) \times 10^{-5}$. This displays the well-conditioned character of the determination of the cluster polynomial and shows that it is meaningful to specify it to 5 digits, with our tolerance level of p. This well-conditioning carries over to the determination of the moments of the cluster zeros from the c_μ. E.g., the *arithmetic mean* of the 3 cluster zeros (=the first moment) varies only by $\approx 7 \times 10^{-6}$ within the tolerance neighborhood of p while the zeros can vary by up to 6×10^{-2} ! $\Box$

8 Overdetermined Problems

In the previous section, we have been able to deal with an ill-posed empirical problem by a change of the requested results. Now we consider ill-posed problems of the kind described as type 1) after Definition 6. These are *overdetermined problems* which can have (non-trivial) solutions only for data satisfying certain relations, i.e. data which lie on a manifold $\mathcal{S}$ in the data space $\mathcal{A}$. Typical polynomial problems of this kind are greatest common divisors (cf. Example 3) and the *factorization of multivariate polynomials*. In this section, we use this last problem to exemplify our approach.

Consider an s-variate polynomial (7) of total degree k. A factorization of p into two factors u of total degree l and v of total degree $k - l \geq l$ can exist only if the *overdetermined system of equations* for the coefficients β_j of u and γ_j of v resulting from a comparison of coefficients in

$$p(x;a) \;:=\; \sum_{j \in J} \alpha_j\, x^j \;=\; \Big(\sum_{|j| \leq l} \beta_j\, x^j\Big)\,\Big(\sum_{|j| \leq k-l} \gamma_j\, x^j\Big) \tag{37}$$

has at least one solution. This is a very restrictive requirement because the number of coefficients grows exponentially with the total degree for $s > 1$. (E.g., for $s = 3$, $k = 5$, $l = 2$, there are 56 equations for the 30 coefficients of u and v.) However, if the real-world phenomenon which has been modelled by the empirical polynomial $(\bar{p}, e)$ is decomposible, there should exist some factorizable polynomials $\tilde{p} \in N_\delta(\bar{p}, e)$, with $\delta = {\rm O}(1)$, and we should be able to find *pseudofactors* of $(\bar{p}, e)$, i.e. exact factors of such a $\tilde{p}$.

Obviously, we may *normalize* one coefficient in each of the 3 polynomials in (37). If p possesses a nonvanishing constant term α_0, it appears convenient to require

$$\alpha_0 \;=\; \beta_0 \;=\; \gamma_0 \;=\; 1\,; \tag{38}$$

this is also impartial with respect to the variables. If $|\alpha_0| = 0$ or very small, (38) cannot be used; but we will not consider the necessary adaptations in this lecture. As in section 2, we do not require that *all* coefficients $(\bar{\alpha}_j, \varepsilon_j)$ of $(\bar{p}, e)$ have a *positive* tolerance ε_j.

The task of finding a pseudofactorization of a multivariate polynomial has two parts (cf. [3]):

(i) the determination of approximate candidate factors (global analysis);

(ii) the checking and refinement of such factors (local improvement).

In phase (i), one may project the overdetermined system (37) onto a subspace of fewer variables: If it cannot be (nearly) satisfied there, no (pseudo)solution for (37) exists; on the other hand, if there exists such a solution, one may be able to extend the subspace solution into a candidate solution for the full system (37). Therefore, we select two variables (call them x, y here) and form the univariate (approximate) factorization

$$p(x, 0, 0..0; \bar{a}) \;\approx\; \prod_{\kappa=1}^{k} (1 + \beta_{10,\kappa} x)\,, \quad \beta_{10,\kappa} \in \mathbb{C}\,, \tag{39}$$

with $\beta_{10,\kappa} = -1/\xi_\kappa$, where the $\xi_\kappa \neq 0$ (cf. (38)) are the k approximate zeros of $p(x, 0, 0..0; \bar{a})$. For each factor combination in (39), we compute a unique extension of the univariate "germ" into a bivariate polynomial compatible with p. When we set $x = 0$ there, the remaining univariate germ in y should be a pseudofactor of $p(0, y, 0..0; \bar{a})$. (For this extension procedure, we must refer to [3].)

Iff one of the germs from (39) passes this test, we extend it further into two full s-variate factors u, v. For this process, we select relations in (37) which permit the computation of further coefficients for u and v from nonsingular linear equations; all the remaining equations in the overdetermined system (37) are simply disregarded. This brutal approach is reasonable: If the polynomial $(\bar{p}, e)$ is pseudo-factorizable and if we have chosen a suitable germ towards this factorization, then the coefficients which satisfy a subset of the equations in (37) must automatically satisfy the remaining equations approximately! If this hypothesis does not hold, it will be discovered in phase (ii). Again, we must refer to [3] for the technical details.

In phase (ii), we compute the backward errors of our candidate pair u, v with respect to those equations in (37) which have been disregarded in phase (i): If they are all of O(1), we have found a *valid approximate factorization* of the empirical polynomial $(\bar{p}, e)$. Otherwise, we turn to the refinement stage – except if some backward errors are so large that we choose to discard the candidate pair.

For the *correction* of a promising candidate pair with coefficients b_0, c_0, we generalize our Newton approach of section 5: We linearize the complete system (37) about b_0, c_0 which yields a *linear overdetermined system* in the correction coefficients $\Delta b, \Delta c$. Since each equation of that system represents the deviation from an individual coefficient of the specified polynomial $\bar{p}$ to be factored and since we measure these deviations in the *weighted max-norm* $\|..\|_e^*$ of (9), we determine the *minimal residual norm* solution of the linearized system with respect to this norm rather than a least-squares solution; this amounts to solving a standard linear program. The occurrence of tolerances $\varepsilon_j = 0$ for intrinsic coefficients of p formally leads to equality conditions. Instead, we rather introduce suitably chosen *tiny* $\varepsilon_j > 0$ for these j; we have found that this makes the solution algorithm for the minimization problem more stable. Due to the fact that we have *linearized* our problem for the correction step, we have to reevaluate the residuals for the *nonlinear* system (37) and repeat the refinement if necessary.

9 Conclusions

It is a severe shortcoming of present computer algebra software systems that they cannot deal properly with data of limited accuracy in algebraic computations; such data occur in almost all scientific computations arising from real-life problems. In this lecture, we have sketched a conceptual framework for an appropriate and consistent treatment of empirical algebraic problems; some of these concepts have reached wide acceptance in the computational algebra community while others are only beginning to become known. The whole approach is in agreement with standard concepts and procedures in Numerical Analysis; it attempts to close the gap between Computer Algebra and Numerical Analysis which still exists.

These concepts permit the design of algorithms which solve many empirical algebraic problems correctly and efficiently. The set of these problems not only encompasses standard problems, like systems of polynomial equations, but also extends to problems which have appeared intractable for data with limited accuracy, like multivariate factorization and many others. It is the accessibility of such problems which poses a major mathematical challenge for numerical polynomial algebra.

It is evident that a user-friendly implementation of algorithms for numerical polynomial algebra in wide-scope software systems for Scientific Computing – like, e.g., Maple, Mathematica, Matlab – will enhance the use of these systems for a large number of applied scientists from the most diverse areas. It will also pave the way towards a more extended use of nonlinearity in the modelling of real-life phenomena and thus towards more reliability in the prediction and control of these phenomena.

References

[1] H.J. Stetter: Analysis of Zero Clusters in Multivariate Polynomial Systems, in: Proceed. Int.Symp.Symbol.Algebr.Comput. 96 (Ed.: R. Caviness), acm-press, 1996, 127-136

[2] H.J. Stetter: Polynomials with Coefficients of Limited Accuracy, in: Computer Algebra in Scientific Computing (Eds. V.G.Ganzha, E.W.Mayr, E.V.Vorozhtsov), Springer, 1999, 409-430

[3] Y. Huang, H.J. Stetter, W. Wu, L. Zhi: Pseudofactors of Multivariate Polynomials, in: Proceed. Int.Symp.Symbol.Algebr.Comput. 2000 (Ed.: C. Traverso), acm-press, 2000, 161-168

[4] H.J. Stetter: Condition Analysis of Overdetermined Polynomial Systems, in: Computer Algebra in Scientific Computing – CASC 2000 (Eds. V.G.Ganzha, E.W.Mayr, E.V.Vorozhtsov), Springer, 2000, 345-366