KU LEUVEN

FACULTEIT INGENIEURSWETENSCHAPPEN

Solving Systems of Polynomial Equations

Simon Telen

Thesis voorgedragen tot het behalen
van de graad van Master of Science
in de ingenieurswetenschappen:
wiskundige ingenieurstechnieken

Promotor:
Prof. dr. ir. Marc Van Barel

Assessor:
Prof. dr. ir. Daan Huybrechs
Prof. dr. ir. Lieven De Lathauwer

Begeleider:
Prof. dr. ir. Marc Van Barel

Academiejaar 2015 – 2016

Preface

I would like to thank all the people who supported and helped me with writing this thesis. Although the rest of this text is written in English, I prefer to do this in Dutch because I feel like it makes it more personal.

Allereerst wil ik graag mijn promotor, prof. Marc Van Barel, bedanken om zijn ideeën in verband met multivariate veeltermstelsels met mij te delen en een deel van de uitwerking aan mij toe te vertrouwen. Onze wekelijkse vergadering was telkens een moment om naar uit te kijken en telkens een zeer leerrijke ervaring. Bedankt voor de interesse in het project, uw kritische opmerkingen, advies en de tijd die u in dit werk geïnvesteerd heeft.

Ik wil ook Kim Batselier en Philippe Dreesen bedanken voor de interesse, de hulp en de antwoorden op mijn vragen. Laurent Sorber wil ik graag bedanken voor het ter beschikking stellen van zijn software en datasets waarop vele van de numerieke experimenten in deze tekst gebaseerd zijn. Verder wil ik nog de assessoren bedanken om deze tekst door te lezen.

Tot slot ben ik ook mijn ouders en zus dankbaar voor de steun en interesse, voor deze thesis maar ook in het algemeen.

Simon Telen

Contents

Abstract

Multivariate systems of polynomial equations find their applications in various fields of science and engineering. Some examples are filter design, parametric system identification, robotics, computer aided design, chemical engineering, ... Solving such a system is a long studied mathematical problem. An extensive amount of literature in the field of algebraic geometry shows that the emphasis in the research has long been mostly on theoretical aspects. An important algorithmic approach based on symbolic computations is the Buchberger algorithm for constructing a Groebner basis for the system. Groebner bases are used by the solvers of a. o. Maple and Mathematica. Important results that connect the problem of solving polynomial systems to linear algebra date from the late 19th - early 20th century when Sylvester and Macaulay introduced the concept of resultants. These ideas have been picked up only recently because of their limiting computational complexity. Methods have been developed that use different kinds of resultants to find the solutions of a polynomial system using a linear algebra approach. Resultant methods are used for example in Chebfun. Another important and successful numerical solving method is that of homotopy continuation, used by PHCpack and Bertini. In this text, the aim is to propose a new numerical linear algebra based method for solving bivariate 0-dimensional systems. By "solving" we mean finding all solutions, both real and complex, and taking multiplicities into account. We start from a two-parameter eigenvalue approach, similar to the one introduced by Plestenjak and Hochstenbach in 2015. We use the concept of degree extension, which is also used to construct Macaulay resultants, to construct a one-parameter square generalized eigenvalue problem directly from the coefficients of the given polynomials. The degree extension allows us to eliminate one of the variables. This is a typical aspect of resultant methods. The coefficients appear directly, without being manipulated in the pencil, which is constructed in a very intuitive manner. We show that a square generalized eigenvalue problem can be constructed for any 0-dimensional system and that the resulting eigenvalues are equal to those of the Sylvester resultant. After obtaining one of the coordinates in this way, we propose some possible approaches for finding the other coordinate of the solutions. The strong link with the Sylvester resultant allows us to give information about the multiplicity of the solutions. Results are promising. Solutions are obtained with small residuals and the computation time is competitive with other solvers. We show that the method can be generalized to other bases than the classical monomial basis and we propose a generalization for more than two dimensions.

List of Abbreviations and Symbols

Abbreviations

GEP	Generalized Eigenvalue Problem
REP	Rectangular Eigenvalue Problem

Symbols

$\mathbb{R}$	The field of real numbers
$\mathbb{C}$	The field of complex numbers
$\mathbb{C}_0$	$\mathbb{C}\backslash\{0\}$
π	The number pi
$\mathcal{P}_\delta^s$	The ring of polynomials in s variables of degree $\leq \delta$
$\boldsymbol{v}$	A complex vector in $\mathbb{C}^n$, $n > 1$
$\boldsymbol{v}_i$	The i-th entry of $\boldsymbol{v}$
M	A complex matrix
M_{ij}	The entry of the matrix M on the i-th row and the j-th column

Chapter 1

Introduction

In this chapter the problem of solving a system of multivariate polynomial equations is formulated in Section 1.1. It turns out that this problem appears in many fields of science and engineering. Section 1.2 contains several examples of applications. An overview of the existing methods for solving multivariate polynomial systems is given in Section 1.3. Finally, in Section 1.4 and Section 1.5, the goal of this thesis is stated and the outline of this text is briefly discussed.

1.1 Multivariate polynomial systems

The concept of multivariate polynomial systems is familiar to almost every engineer or engineering student. The problem can be stated as follows.

Problem 1 (Multivariate Polynomial System)**.** *Find all vectors* $(x_1, x_2, \ldots, x_s)^\top \in \mathbb{C}^s$ *that satisfy*

$$\begin{cases} p_1(x_1, x_2, \ldots, x_s) = 0 \\ p_2(x_1, x_2, \ldots, x_s) = 0 \\ \qquad\qquad \vdots \\ p_s(x_1, x_2, \ldots, x_s) = 0 \end{cases}$$

where $p_i(x_1, x_2, \ldots, x_s)$, $1 \leq i \leq s$ *are polynomials.*

Problem 1 pops up in various disciplines, be it in its direct form or implicitly, e.g. in the form of a polynomial optimization problem. Solving such a polynomial optimization problem comes down to solving a system of polynomial equations, as will be illustrated by Example 1.1.1.

Problem 2 (Polynomial Optimization Problem)**.** *Find* $(x_1^*, x_2^*, \ldots, x_s^*)^\top \in \mathbb{C}^s$ *such that*

$$\begin{aligned} (x_1^*, x_2^* \ldots, x_s^*) = & \operatorname*{argmin}_{x_1, x_2 \ldots, x_s} O(x_1, x_2, \ldots, x_s) \\ \text{subject to} \quad & p_i(x_1, \ldots, x_s) = 0,\ i = 1, \ldots, m \leq s. \end{aligned} \tag{1.1}$$

where O *is a polynomial objective function and the* $\{p_i\}_{1 \leq i \leq m}$ *are polynomial constraints.*

Example 1.1.1. *Suppose we want to calculate the width x^* and the length y^* of a rectangular piece of cardboard with area 1 that minimize the diagonal length. The problem can be formulated as*

$$(x^*, y^*) = \underset{x,y}{\operatorname{argmin}}\ x^2 + y^2$$
$$\text{subject to} \quad p_1(x, y) = xy - 1 = 0.$$

Using the method of Lagrange multipliers, we find for the Lagrangian $\mathcal{L}(x, y, z) = x^2 + y^2 - z(xy - 1)$. Equating its first partial derivatives to zero we obtain for the optimality conditions

$$\begin{cases} \frac{\partial \mathcal{L}}{\partial x} = 2x - zy = 0 \\ \frac{\partial \mathcal{L}}{\partial y} = 2y - zx = 0 \\ \frac{\partial \mathcal{L}}{\partial z} = xy - 1 = 0 \end{cases}$$

which is an instance of Problem 1 with $s = 3$ that can be solved analytically. The real solutions are given by $(1, 1, 2)$ and $(-1, -1, 2)$, of which only the first one has a physical meaning. The desired width and length are both equal to one.

1.2 Applications

Fields in which polynomial systems are encountered are chemical engineering, civil engineering, signal processing and filter design, system identification, robotics, ... There are some typical situations in which they appear. Some problems naturally lead to polynomial relations. In other cases polynomials furnish a natural tool for modelling phenomena that cannot be adequately described by linear equations. We give three motivating examples of polynomial system applications.

1.2.1 Equilibrium concentrations in chemical reactions

In a chemical reaction, the concentrations at a state of equilibrium of all different substances is governed by conservation equations and by reaction equations [24]. Conservation equations state that the total number of atoms of each element in the reaction must stay constant. They are always linear. For example, consider the reaction

$$\mathrm{H_2O} \rightleftharpoons 2\,\mathrm{H} + \mathrm{O}$$

for which the conservation equations are

$$x_H + 2x_{H_2O} = T_H \tag{1.2}$$
$$x_O + x_{H_2O} = T_O \tag{1.3}$$

where x_i stands for the (unknown) equilibrium concentration of molecule i and T_j denotes the (known) constant concentration of atom j. The reaction equation for this simple system is given by

$$K x_{H_2O} = x_H^2 x_O \tag{1.4}$$

where K is the equilibrium constant that depends on external conditions, such as temperature. Together, (1.2), (1.3) and (1.4) form a polynomial system of three equations in the unknown equilibrium concentrations x_H, x_O and x_{H_2O}.

1.2.2 Daubechies wavelets

Wavelet decompositions have many applications in signal processing, image compressing, image denoising, ... Performing a wavelet decomposition of a signal can be interpreted as sending the signal through a filter bank [27, 6]. The goal of this example is to illustrate how the design of the filters of such a filter bank (which corresponds to the design of the wavelet) can be realized by solving a system of polynomial equations. Consider the space $V_0 \subset L^2$ defined as

$$V_0 = \text{span}\{\phi_{0k}(t)\}_{k\in\mathbb{Z}}$$

where $\phi_{nk}(t) = 2^{\frac{n}{2}}\phi_{00}(2^n t - k)$, $n, k \in \mathbb{Z}$ and $\phi_{00} \triangleq \phi(t)$ satisfies a dilation equation

$$\phi(t) = \sum_{k\in\mathbb{Z}} c_k \phi(2t - k) \tag{1.5}$$

for some set of coefficients $\{c_k\}_{k\in\mathbb{Z}}$. V_0 is spanned by shifted versions of $\phi(t)$. Assume for simplicity that the $\{\phi_{0k}(t)\}_{k\in\mathbb{Z}}$ form an orthonormal basis for V_0 (with respect to the standard inner product in L^2). Now, defining analogously

$$V_{-1} = \text{span}\{\phi_{-1k}(t)\}_{k\in\mathbb{Z}},$$

one obtains a space spanned by the orthonormal set $\{\phi_{-1k}(t) = 2^{-\frac{1}{2}}\phi(\frac{t}{2} - k)\}_{k\in\mathbb{Z}}$, which are translates of a 'stretched out' version of $\phi(t)$. Intuitively, V_0 offers more flexibility for approximating quickly varying functions than V_{-1}. In fact, through (1.5) it can be seen that every basis function $\phi_{-1l}(t)$ of V_{-1} can be written as a linear combination of the basis functions $\{\phi_{0k}(t)\}_{k\in\mathbb{Z}}$ of V_0, which means $V_{-1} \subset V_0$. The orthogonal complement of V_{-1} in V_0 is denoted by W_{-1}:

$$V_0 = V_{-1} \oplus W_{-1}.$$

Suppose an orthonormal basis of W_{-1} is given by the set $\{\psi_{-1k}(t)\}_{k\in\mathbb{Z}}$.

Let some continuous time signal $f(t)$ be approximated by the signal $\tilde{f}(t)$ which is contained in the space V_0:

$$\tilde{f}(t) = \sum_{k\in\mathbb{Z}} v_{0k}\phi_{0k}(t)$$

The decomposition

$$\tilde{f}(t) = \underbrace{\sum_{k\in\mathbb{Z}} v_{-1k}\phi_{-1k}(t)}_{\tilde{f}_\phi} + \underbrace{\sum_{k\in\mathbb{Z}} w_{-1k}\psi_{-1k}(t)}_{\tilde{f}_\psi}$$

is one step in the orthogonal wavelet decomposition of $\tilde{f}$. The function $\tilde{f}_\phi$ is the orthogonal projection of $\tilde{f}$ onto V_{-1} and contains the "low resolution" or low frequency information from $\tilde{f}$. $\tilde{f}_\psi$ can be seen as the error that is made approximating $\tilde{f}$ by $\tilde{f}_\phi$ and contains high frequency information. It is shown that in this orthogonal setting the coefficients $\{v_{-1k}\}_{k\in\mathbb{Z}}$ can be found by applying a low pass filter with transfer function $H_*(z) = \overline{H(\frac{1}{\bar{z}})}$ to the high resolution coefficients $\{v_{0k}\}_{k\in\mathbb{Z}}$ followed by a downsampling. More specifically, $H(z)$ can be found as the z-transform of the scaling coefficients $\{h_k\}_{k\in\mathbb{Z}} \triangleq \left\{\frac{c_k}{\sqrt{2}}\right\}_{k\in\mathbb{Z}}$ with $\{c_k\}_{k\in\mathbb{Z}}$ the coefficients that define $\phi(t)$ in the dilation equation (1.5).

Not any set $\{h_k\}_{k\in\mathbb{Z}}$ defines a scaling function $\phi(t)$ that gives rise to an orthonormal basis of V_0. For computational reasons, it is interesting to look for such a filter with compact support, i.e. only a finite number of filter coefficients h_k is different from zero. This is where polynomial relations come into play. Suppose only the $2p$ coefficients $h_0, \ldots, h_{2p-1}$ are allowed to be different from zero. By requiring

$$\int_{-\infty}^{\infty} \phi(t)dt = \theta$$

with θ some constant different from zero, using (1.5) we obtain

$$\sum_{k=0}^{2p-1} h_k = \sqrt{2}. \tag{1.6}$$

For the system $\{\phi_{0k}\}_{k\in\mathbb{Z}}$ to be orthonormal, it is shown that the following property must be satisfied for all $n \in \mathbb{Z}$:

$$\sum_{k\in\mathbb{Z}} h_k h_{k-2n} = \delta_n = \begin{cases} 1 & n = 0 \\ 0 & n \neq 0 \end{cases} . \tag{1.7}$$

Property (1.7) is called *double shift orthogonality*. In case of our compactly supported filter, (1.7) only gives nontrivial equations for $n = 0, \ldots, p-1$. Equations (1.6) and (1.7) furnish $p+1$ polynomial equations in the $2p$ unknowns. The remaining $p-1$ degrees of freedom can be used to add so called *vanishing moments* to the multiresolution analysis. This comes down to making sure that all polynomials up to degree $p-1$ are contained in the space V_0. The number of vanishing moments is referred to as the order of the multiresolution analysis. The higher the order, the better the convergence properties of the sparse representation of signals in the wavelet basis. For the filter coefficients, this translates into the equations

$$\sum_{k\in\mathbb{Z}} (-1)^k k^n h_k = 0 \qquad n = 1, \ldots, p-1 \tag{1.8}$$

which completes the polynomial system of $2p$ equations in $2p$ unknowns. The real solutions of this system correspond to the so called maxflat wavelets. The solutions that generate a filter $H(z)$ with minimal phase were introduced by Ingrid Daubechies [9] and they are still among the most commonly used wavelets in image compression.

1.2.3 Prediction error methods

Another field in which polynomial systems occur is that of parametric system identification [19, 4]. In this example it will be shown that the optimal parameters for linear time invariant (LTI) systems with a single input $u(t)$ and a single output $y(t)$ (SISO) found by prediction error methods are in fact the solutions of a multivariate polynomial system.

It will be assumed that input and output signals are sampled at N discrete time steps $t = 1, \ldots, N$. The collected dataset is given by Z:

$$Z = \begin{pmatrix} u(1) & u(2) & \cdots & u(N) \\ y(1) & y(2) & \cdots & y(N) \end{pmatrix}.$$

The most general form of a LTI SISO model can be written as

$$A(q)y(t) = \frac{B(q)}{F(q)}u(t) + \frac{C(q)}{D(q)}e(t) \tag{1.9}$$

where $e(t)$ is a white noise sequence and A, B, C, D and F are polynomials in the linear delay operator q:

$$qx(t) = x(t-1).$$

In our example, for ease of notation we will look for a simpler model

$$y(t) = \frac{b_1}{1 + f_1 q + f_2 q^2}u(t) + e(t) = \frac{B(q)}{F(q)}u(t) + e(t) \tag{1.10}$$

which is called an output error model where the number of parameters is chosen to be (only) 3 for simplicity. The parameter set is denoted by $\theta = \{b_1, f_1, f_2\}$. At time t, the prediction error $\epsilon(t)$ is defined as the difference between $y(t)$ and a one-step ahead predictor $\hat{y}(t|\theta)$:

$$\epsilon(t) = y(t) - \hat{y}(t|\theta).$$

The predictor $\hat{y}(t|\theta)$ is defined in such a way that if the parameters θ are the exact parameters of the underlying system, the prediction error is the white noise sequence $e(t)$. In this case

$$\epsilon(t) = y(t) - \frac{B(q)}{F(q)}u(t)$$

or

$$F(q)y(t) - B(q)u(t) - F(q)\epsilon(t) = 0. \tag{1.11}$$

Because $F(q)$ is a polynomial of degree 2, (1.11) can be written down for $t > 2$. The problem that is solved to find a suitable set of parameters θ^* is stated as follows.

Find θ^* such that

$$\begin{aligned} \theta^* &= \operatorname*{argmin}_{\theta} \frac{1}{N}\sum_{k=1}^{N} \frac{\epsilon(k)^2}{2} \\ \text{subject to} \quad & F(q)y(t) - B(q)u(t) - F(q)\epsilon(t) = 0,\ t = 3, \ldots, N. \end{aligned} \tag{1.12}$$

Problem (1.12) is an instance of (1.1) in the variables b_1, f_1, f_2, $\epsilon(1)$, ..., $\epsilon(N)$. The Lagrangian for $N = 5$ is given by

$$\begin{aligned}\mathcal{L}(b_1, f_1, f_2,&\epsilon(1), \epsilon(2), \epsilon(3), \epsilon(4), \epsilon(5), \lambda_1, \lambda_2, \lambda_3) = \\ &\frac{1}{10}(\epsilon(1)^2 + \epsilon(2)^2 + \epsilon(3)^2 + \epsilon(4)^2 + \epsilon(5)^2)+ \\ &\lambda_1(y(3) + f_1 y(2) + f_2 y(1) - b_1 u(3) - \epsilon(3) - f_1\epsilon(2) - f_2\epsilon(1))+ \\ &\lambda_2(y(4) + f_1 y(3) + f_2 y(2) - b_1 u(4) - \epsilon(4) - f_1\epsilon(3) - f_2\epsilon(2))+ \\ &\lambda_3(y(5) + f_1 y(4) + f_2 y(3) - b_1 u(5) - \epsilon(5) - f_1\epsilon(4) - f_2\epsilon(3)).\end{aligned} \tag{1.13}$$

In practice, the value of N is much larger, but the expression for $\mathcal{L}$ would become too long to write out in this text. Note that the variables $\epsilon(1), \ldots, \epsilon(N)$ can easily be eliminated because they appear linearly in the partial derivatives of $\mathcal{L}$ with respect to themselves. For example, the optimality condition

$$\frac{\partial \mathcal{L}}{\partial \epsilon(1)} = \frac{\epsilon(1)}{5} - \lambda_1 f_2 = 0$$

leads to $\epsilon(1) = 5\lambda_1 f_2$.

1.3 State of the art

Polynomial root finding is a long studied discipline with a rich history [21, 11, 33]. It is among the oldest mathematical problems. Greek writings by Diophantus of Alexandria date from the third century and give numerical solutions of univariate polynomial equations with rational coefficients. The notation and graph representation in a cartesian coordinate system for multivariate polynomials as we know it is mostly due to René Descartes. In his book *La Géométrie* (1637) [12, 32], he popularized using superscripts as exponentials, using letters from the beginning of the alphabet as coefficients and using letters from the end of the alphabet to denote variables. In *La Géométrie*, Descartes introduced the description of circles, lines, parabolas, ... as bivariate polynomial equations. Doing so he was among the first to establish the relation between geometrical problems and polynomial algebra. Contemporary scientists that contributed important results in the field are a.o. Isaac Newton and Pierre de Fermat. These insights gave rise to the birth of algebraic geometry, nowadays an important branch of mathematics. From the 18th until the 20th century, names such as Etienne Bézout, Evariste Galois, James Joseph Sylvester, David Hilbert, ... were responsible for an extensive amount of new literature in the field of algebraic geometry. The emphasis in the research of the polynomial root finding problem was mainly on theoretical results and algorithms based on symbolic manipulations. An important notion is that of a Groebner basis.

1.3.1 Groebner bases: an algorithmic approach

A symbolic tool that is definitely worth mentioning is the Buchberger algorithm [5] for transforming a given polynomial system F into a so called *Groebner basis* G [8, 26, 29]. Roughly speaking, a Groebner basis has "nice properties" such that F and G are equivalent in terms of their solution sets and the solutions of G can be computed easily. We will illustrate the concept of Groebner bases by means of a bivariate example. As the emphasis in this text is on the bivariate case, we introduce the notation

$$\begin{cases} p(x,y) = 0 \\ q(x,y) = 0 \end{cases} \tag{1.14}$$

for the version of Problem 1 where $s = 2$, which we will use throughout the text.

Example 1.3.1 (Groebner basis)**.** *Consider the system*

$$\begin{cases} p(x,y) = xy - 2y = 0 \\ q(x,y) = 2y^2 - x^2 = 0 \end{cases}$$

defined by the basis $F = \{xy - 2y, 2y^2 - x^2\}$*. A Groebner basis for* F *is calculated by using the Buchberger algorithm (with a lexicographic ordering that ranks* y *higher than* x*). The result is* $G = \{-2x^2 + x^3, xy - 2y, 2y^2 - x^2\}$*. Therefore, the solutions to the system defined by* F *coincide with the solutions to*

$$\begin{cases} g_1(x,y) = -2x^2 + x^3 = 0 \\ g_2(x,y) = xy - 2y = 0 \\ g_3(x,y) = 2y^2 - x^2 = 0 \end{cases}$$

which is easy to solve since all possible values for x *can be found as the roots of the univariate polynomial* g_1*. The roots of* g_1 *are 0, 0 and 2. Plugging these values into* g_2 *and* g_3*, the solution set* $\{(0,0), (2,-\sqrt{2}), (2,\sqrt{2})\}$ *is obtained. Figure 1.1 illustrates the equivalence of the two systems.*

Groebner bases are still used in computer algebra systems such as Maple and Mathematica. The complexity of calculating Groebner bases increases exponentially with the degree of the system. Moreover, when computing in finite precision, the algorithm suffers from numerical instability, which makes the method useless for high degree systems. Modern day developments in computer algebra are the main reason for the interest in approaches that require numeric computations (as opposed to the classical symbolic approach). A number of such approaches exist and a brief introduction follows in the next subsections.

1.3.2 Polynomial systems and numerical linear algebra

The close relation between polynomial root finding and numerical linear algebra is well known for the univariate case [21, 26]. To each univariate polynomial p a so

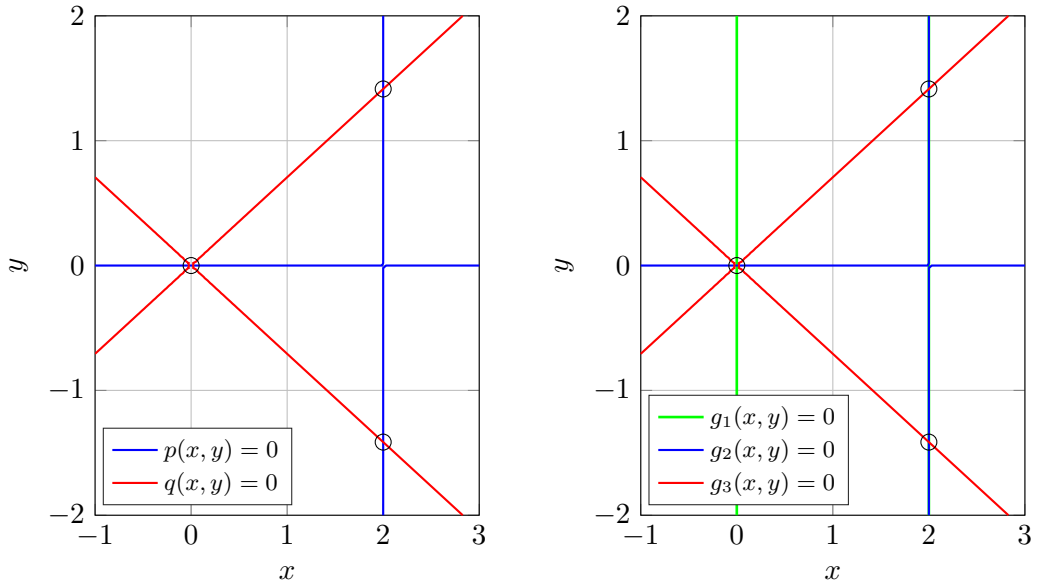


Figure 1.1: Real picture of the zero level sets of the systems defined by F (left) and G (right).

called companion matrix C can be associated of which the characteristic polynomial is equal to p (up to a nonzero constant factor). Assume p is monic and given by

$$p(x) = x^\delta - p_1 x^{\delta-1} - p_2 x^{\delta-2} - \cdots - p_{\delta-1} x - p_\delta,$$

then a possible companion matrix for p is

$$C = \begin{pmatrix} p_1 & p_2 & \cdots & p_{\delta-1} & p_\delta \\ 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & 0 \end{pmatrix}.$$

Several choices for the companion matrix are possible based on straightforward similarity transformations of this one. The assumption that p is monic is not restrictive since $\forall \alpha \in \mathbb{C}_0$, $\alpha p(x) = 0 \Leftrightarrow p(x) = 0$, hence any p can be normalized to be monic. Other companion-type matrices (comrade, confederate, congenial, ...) [1] can be found by representing p in another basis. The roots of p can be found as the eigenvalues of C. An advantage of this approach is that it is guaranteed that all solutions are found at once. Other iterative methods like Newton-Raphson iteration converge (in most cases) quite quickly to one certain root. Which one depends on the choice of the initial guess but it is a nontrivial problem to guarantee that all solutions are found.

An important notion that can be traced back to Sylvester (1853) and Macaulay (1902,1916) is that of resultants. Resultants provide a linear algebra approach to

determine (the existence of) common roots of multivariate polynomials. Because of their limiting computational complexity, the ideas of Sylvester and Macaulay came back to play only in the 1980's, when they were picked up by Lazard and Stetter. Lazard established the relation between the calculation of a Groebner basis and the triangularization of a large matrix, similar to the Macaulay matrix [18]. Stetter linked polynomial system solving to an eigenvalue decomposition a few years later. The method requires some knowledge of algebraic geometry and for a detailed description we refer to [26]. In this book [26, p. 52], Stetter states "with a grain of salt":

> The numerical solution of 0-dimensional systems of polynomial equations is a task of numerical linear algebra.

The work of Stetter and Lazard made the interest in the linear algebra approach to the problem grow. Some of the recently developed linear algebra solving methods are presented below.

Resultant methods

Resultant methods for bivariate polynomial system solving [29, 26, 2, 8] typically select one out of two variables, say x, and consider $p(x, y)$ and $q(x, y)$ as univariate polynomials in y with coefficients that are univariate polynomials in x. They are based on the idea that if $p(x^*, y^*) = q(x^*, y^*) = 0$, then the univariate polynomials $p(x^*, y)$ and $q(x^*, y)$ must have at least one common zero. One can prove that a necessary condition for this to happen is that an associated matrix polynomial[1] $R(x)$, called a *resultant matrix*, is singular for $x = x^*$. This means that all candidate x-coordinates for the solutions of (1.14) are eigenvalues of this matrix polynomial. The determinant of such a resultant matrix is a univariate polynomial called a *resultant*. In other words, a resultant (with respect to x) is a polynomial $\text{res}^{p,q}(x)$ with a set of roots that contains the set [25]

$$\{x \in \mathbb{C} \mid \exists y \in \mathbb{C} : p(x, y) = q(x, y) = 0\}.$$

A standard way to calculate the eigenvalues of such a matrix polynomial, i.e. solving the matrix polynomial eigenvalue problem, is called *linearization* [16, 17]. A linear pencil $A - xB$ is calculated that satisfies $\det(A - xB) = \alpha \det R(x)$ with $\alpha \in \mathbb{C}_0$. By definition, the eigenvalues of $R(x)$ and $A - xB$ coincide. The eigenvalues of $A - xB$ can be calculated in a numerically stable way using the QZ-algorithm.

Different resultants can be obtained by representing p and q in different bases. One resultant, called the *Sylvester resultant* deserves some special attention because

[1] By matrix polynomial we mean a polynomial whose coefficients are matrices, for example

$$R(x) = A_0 + A_1 x + A_2 x^2 + \cdots + A_\delta x^\delta$$

is a matrix polynomial where the coefficients $\{A_i\}_{0 \leq i \leq \delta}$ are matrices.

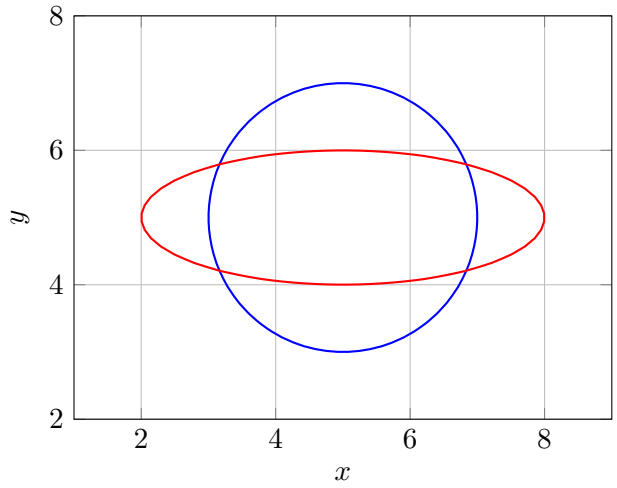


Figure 1.2: Real picture of the zero level curves described by $p(x, y) = 0$ (——) and $q(x, y) = 0$ (——) from the system (1.15).

it is closely related to the root finding method that is described in this text. In section 2.2.2, its construction is discussed and we use it to explain the concept of resultant methods in more detail.

Recently, several resultant methods were developed for solving (1.14). In [25] the aim is to calculate all real solutions to (1.14). Resultants are used to project the solutions onto the real plane associated to the two variables rather than onto the complex plane associated with one variable. In the latest version of Chebfun2, the `roots` command uses a resultant method based on Bézout resultants to calculate all real solutions of a bivariate nonlinear system in a rectangular compact domain of $\mathbb{R}^2$ (polynomial interpolation is used to approximate the given bivariate functions within that domain) [20]. An extension of the concept of the resultant for more than two unknowns ($s > 2$) is due to Macaulay (early 1900s).

Macaulay matrix

In 2013, Ph. Dreesen, K. Batselier et al. [14, 4] have worked out a linear algebra approach to Problem 1. It is based on the so called *Macaulay matrix*. Without formally introducing the Macaulay matrix and its properties, a bivariate example is used to give the reader an idea of how the algorithm works. The method can be generalized to higher dimensional problems. For details, we refer to [14, 4].

Example 1.3.2. *Consider the system*

$$\begin{cases} p(x, y) = 46 - 10x - 10y + x^2 + y^2 = 0 \\ q(x, y) = 241 - 10x - 90y + x^2 + 9y^2 = 0 \end{cases}. \tag{1.15}$$

The level lines of p and q in the real plane are plotted in Figure 1.2 and the solutions are[2] $(3.1629, 4.2094)$, $(3.1629, 5.7906)$, $(6.8371, 4.2094)$ *and* $(6.8371, 5.7906)$. *The system* (1.15) *can be associated to the matrix*

$$M_2 = \left(\begin{array}{c|cc|ccc} 46 & -10 & -10 & 1 & 0 & 1 \\ 241 & -10 & -90 & 1 & 0 & 9 \end{array}\right)$$

which is called the Macaulay matrix of degree 2. The first row of M_2 contains the coefficients of $p(x, y)$, the second row those of $q(x, y)$. The column partitioning of M_2 corresponds to a partitioning of the monomial degrees. E.g. 46 and 241 are the only coefficients that correspond to the monomial 1 (of degree 0). Every column corresponds to a fixed monomial. The ordering of the monomials is not straightforward in the multivariate case. Here, the so called degree negative lexicographic ordering *is used:*

$$1 < x < y < x^2 < xy < y^2 < x^3 < x^2y < xy^2 < y^3 < x^4 < ... \tag{1.16}$$

The Macaulay matrix M_3 of degree 3 can be constructed by multiplying p and q by x and y and extending M_2 with rows for the new equations $xp(x, y) = 0$, $yp(x, y) = 0$, $xq(x, y) = 0$ and $yq(x, y) = 0$ and columns for the third degree monomials x^3, x^2y, xy^2 and y^3. The result is

$$M_3 = \left(\begin{array}{rrr|rrr|cccc} 46 & -10 & -10 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 46 & 0 & -10 & -10 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 46 & 0 & -10 & -10 & 0 & 1 & 0 & 1 \\ \hline 241 & -10 & -90 & 1 & 0 & 9 & 0 & 0 & 0 & 0 \\ 0 & 241 & 0 & -10 & -90 & 0 & 1 & 0 & 9 & 0 \\ 0 & 0 & 241 & 0 & -10 & -90 & 0 & 1 & 0 & 9 \end{array}\right)$$

where the first three rows correspond to the 'shifted' p-equations and the last three rows to the shifted q-equations. It is readily checked that the rank of M_2 is 2 and that of M_3 is 6. The dimensions of the right null spaces of M_2 and M_3 are both equal to 4. We know that the system (1.15) *has 4 solutions and it is clear from the construction of the Macaulay matrices that a couple (x^*, y^*) is a solution to* (1.15) *if and only if the vector $\begin{pmatrix} 1 & x^* & y^* & x^{*2} & x^*y^* & y^{*2} \end{pmatrix}^\top$ belongs to the null space of M_2. This implies that the null space of M_2 is spanned by four such vectors. Analogously, any solution (x^*, y^*) corresponds to a vector*

$$\begin{pmatrix} 1 & x^* & y^* & x^{*2} & x^*y^* & y^{*2} & x^{*3} & x^{*2}y^* & x^*y^{*2} & y^{*3} \end{pmatrix}^\top \tag{1.17}$$

in the null space of M_3. A numerical basis Z for the null space of M_3 can be calculated using the `null` *command in Matlab (Z is a matrix containing the basis vectors in its columns). We know that for every root (x^*, y^*) there is a vector $\boldsymbol{z} = Z\boldsymbol{v}$ in the null space of M_3 that has this multivariate Vandermonde structure* (1.17). *Any vector $\boldsymbol{z}$*

[2] Rounded to 4 decimal digits after the decimal point.

with this structure must satisfy

$$\begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \\ z_6 \end{pmatrix} x^* = \begin{pmatrix} z_2 \\ z_4 \\ z_5 \\ z_7 \\ z_8 \\ z_9 \end{pmatrix}$$

where we used the notation z_i for the i-th element of the vector $\boldsymbol{z}$. In fact, in [14] it is shown that it is sufficient to impose

$$\begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_5 \end{pmatrix} x^* = \begin{pmatrix} z_2 \\ z_4 \\ z_5 \\ z_8 \end{pmatrix}. \tag{1.18}$$

The other equalities are automatically satisfied because of the structure of M_3. Letting S_1 denote a row selection matrix that selects the rows 1, 2, 3 and 5 and letting S_x select the rows 2, 4, 5 and 8, we impose (1.18) *on the vector $Z\boldsymbol{v}$ by writing down the GEP*

$$S_x Z\boldsymbol{v} = xS_1 Z\boldsymbol{v}.$$

The eigenvalues x correspond to the x-coordinates of the four solutions. Using Matlab, we find the eigenvalues $\{6.8371, 3.1629, 3.1629, 6.8371\}$, which are indeed the correct x-values.

Note that the same thing cannot be done for the matrix M_2 because the monomial xy cannot be shifted by x to a monomial of degree ≤ 2. In general, a degree extension up to a certain degree d^* is needed to make the nullity of the Macaulay matrix equal to the number of solutions and to make the construction of S_x possible. An interesting aspect of this method is that the shift factor can be chosen to be any polynomial $g(x, y)$ as long as the degree of the Macaulay matrix is large enough to construct the selection matrices. For example, for every solution (x^*, y^*) the corresponding null vector $\boldsymbol{z}$ must satisfy

$$\begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_5 \end{pmatrix} (x^* + 5y^*) = \begin{pmatrix} z_2 + 5z_3 \\ z_4 + 5z_5 \\ z_5 + 5z_6 \\ z_8 + 5z_9 \end{pmatrix}$$

which corresponds to the generalized eigenvalue problem

$$\underbrace{\begin{pmatrix} 0 & 1 & 5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 5 & 0 \end{pmatrix}}_{S_{g(x,y)}} Z\boldsymbol{v} = g(x, y) \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}}_{S_1} Z\boldsymbol{v}$$

where the eigenvalues are the shift function $g(x, y) = x + 5y$ evaluated at the roots (x^*, y^*). Using Matlab, we find $\{3.5790e1, 2.4210e1, 2.7884e1, 3.2116e1\}$ as the set of eigenvalues. This is interesting for example when one is not so much interested in the minimizers of an optimization problem (1.1), but more in the optimal value of the objective function. The objective function can simply be plugged in as the shift polynomial $g(x, y)$. The Macaulay matrices M_d form an elegant linear algebra tool to provide information about the geometric properties of the solution set of a given system of polynomial equations. For instance, the multiplicity structure of a solution and solutions at infinity are revealed by their null spaces.

Two-parameter eigenvalue approach

In the univariate case, finding the roots of $p(x)$ can be written as an eigenvalue problem because there exists a linear matrix polynomial $C - xI$, where I is the identity matrix, which satisfies

$$\det(C - xI) = \alpha p(x) \quad \text{with} \quad \alpha \in \mathbb{C}_0.$$

Similarly, in the bivariate case, studies have shown [23] that there are square linear matrix polynomials in x and y that satisfy

$$\det(A_p - xB_p - yC_p) = p(x, y) \quad \text{and} \quad \det(A_q - xB_q - yC_q) = q(x, y).$$

Hence the solutions to (1.14) can be found as the eigenvalues of the two-parameter eigenvalue problem

$$\begin{cases} (A_p - xB_p - yC_p)\boldsymbol{u}_p = \mathbf{0} \\ (A_q - xB_q - yC_q)\boldsymbol{u}_q = \mathbf{0} \end{cases}. \tag{1.19}$$

If a pair (x, y) and a pair of nonzero vectors $\boldsymbol{u}_p$ and $\boldsymbol{u}_q$ are found that satisfy (1.19), then (x, y) is said to be an eigenvalue of (1.19) with corresponding eigenvector $\boldsymbol{w} = \boldsymbol{u}_p \otimes \boldsymbol{u}_q$ ($\otimes$ denotes the Kronecker product). Eigenvalues and eigenvectors can be calculated as follows. First, calculate the so called *operator determinants*:

$$\begin{aligned} \Delta_0 &= B_p \otimes C_q - C_p \otimes B_q, \\ \Delta_1 &= C_p \otimes A_q - A_p \otimes C_q, \\ \Delta_2 &= A_p \otimes B_q - B_p \otimes A_q, \end{aligned}$$

and then find the eigenvalues by solving the generalized (one-parameter) eigenvalue problems

$$\begin{aligned} \Delta_1 \boldsymbol{w} &= x\Delta_0 \boldsymbol{w}, \\ \Delta_2 \boldsymbol{w} &= y\Delta_0 \boldsymbol{w}. \end{aligned}$$

Suppose the total degree of (1.14) is equal to δ. The size of the linear pencils from (1.19) typically grows like δ^2. The operator determinants, since they are calculated by taking Kronecker products between the coefficient matrices, grow like δ^4. Solving a generalized eigenvalue problem of size n by using the QZ-algorithm has a complexity

of order $\mathcal{O}(n^3)$. This makes the overall complexity of this method $\mathcal{O}(\delta^{12})$. In [23], a recent two-parameter eigenvalue approach is compared to Mathematica's `Nsolve`, which uses Groebner bases, and PHCpack [31], which uses homotopy continuation. Because of the huge complexity, the method is found to be competitive only for $\delta < 10$.

1.3.3 Other computational methods

The following methods are not based on numerical linear algebra but they are definitely worth mentioning.

Homotopy continuation methods

The idea of homotopy continuation methods [30, 3] is to start from an easy initial polynomial system that can be continuously transformed into the given system (1.14). Denote

$$P(\boldsymbol{x}) = \begin{pmatrix} p(x,y) \\ q(x,y) \end{pmatrix}$$

and let $I(\boldsymbol{x})$ represent the initial system of which the solutions can be calculated easily. Consider the following set of problems

$$H(\boldsymbol{x},t) = (1-t)I(\boldsymbol{x}) + tP(\boldsymbol{x}) = \boldsymbol{0}, \quad t \in [0,1], \tag{1.20}$$

which is called a *homotopy* to $P(\boldsymbol{x})$. Suppose the number of solutions of $I(\boldsymbol{x}) = 0$ is equal to the number of solutions of (1.14) (counting multiplicities and solutions at infinity). In that case, the constructed homotopy using $I(\boldsymbol{x}) = 0$ as initial system is called the *Total Degree Homotopy*. The parameter t parametrizes the solution set of $H(\boldsymbol{x},t)$. For t varying from 0 to 1, this solution set consists out of a finite number of smooth paths that start at the solutions of $H(\boldsymbol{x},0) = \gamma I(\boldsymbol{x}) = 0$ and end at the solutions of $H(\boldsymbol{x},1) = P(\boldsymbol{x}) = 0$. Homotopy continuation methods track these paths numerically to obtain the solutions of (1.14). The reasoning can be generalized to more variables [30, 3]. Software packages that implement these methods to compute the isolated solutions of a system of s equations in s unknowns are PHCpack and Bertini [30, 31, 3]. Both packages are among the most competitive solvers today.

Example 1.3.3. *Consider the target system*

$$P(\boldsymbol{x}) = \begin{pmatrix} p(x,y) \\ q(x,y) \end{pmatrix} = \begin{pmatrix} (x-5)^2 + (y-5)^2 - 4 \\ \frac{(x-5)^2}{9} + (y-5)^2 - 1 \end{pmatrix} \tag{1.21}$$

which is equal to the system (1.15) *and consider the initial system of the same degree*

$$I(\boldsymbol{x}) = \begin{pmatrix} i_1(x,y) \\ i_2(x,y) \end{pmatrix} = \begin{pmatrix} x^2 - 1 \\ y^2 - 1 \end{pmatrix}$$

of which the solutions $(-1,-1)$, $(-1,1)$, $(1,-1)$ *and* $(1,1)$ *can be obtained analytically. Constructing the total degree homotopy* H *as in* (1.20)*, the real part of the solution*

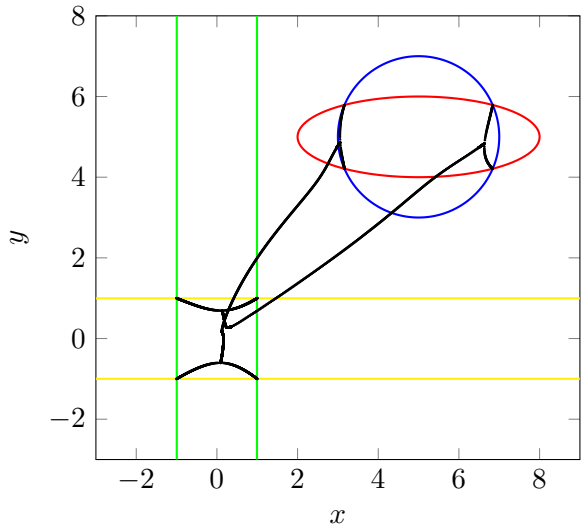


Figure 1.3: The colored lines represent the zero level lines in $\mathbb{R}^2$ of the polynomials that define the initial and the target system: $p(x, y) = 0$ (——), $q(x, y) = 0$ (——), $i_1(x, y) = 0$ (——), $i_2(x, y) = 0$ (——). The black dots represent a discretization of the real part of the solution paths. The paths start at the solutions of the initial system and they converge to the solutions of the target system.

paths is plotted in Figure 1.3. As shown in the figure, the paths converge to the solutions $(3.1629, 4.2094)$*,* $(3.1629, 5.7906)$*,* $(6.8371, 4.2094)$ *and* $(6.8371, 5.7906)$*.*

Contouring algorithms

Contouring algorithms generate the real zero level curves of $p(x, y)$ and $q(x, y)$ numerically, for example by using the marching squares algorithm, and use their intersections as a starting value for Newton-Raphson iteration. This is what was done in an older version of Chebfun. Contouring algorithms for root finding have several drawbacks [20] but for most practical applications it is an adequate approach. Unlike linear algebra methods and homotopy continuation, contouring algorithms can only be used to calculate the real roots of (1.14) on a compact domain of $\mathbb{R}^2$.

1.4 Goal

The aim in this text is to propose a method for solving the bivariate version of Problem 1 using a numerical linear algebra approach. More precisely, it is our aim to calculate all finite solutions to Problem 1 with $s = 2$, counting multiplicities and without limiting ourselves to only real solutions or only a compact domain of $\mathbb{C}^2$. We will assume that the problem is well-posed in a particular sense: the solutions are isolated and finite in number. We will derive the proposed method in a rather intuitive manner and then show its correctness by proving equivalence with the Sylvester resultant approach. We will implement this method and its variants in

Matlab using the results from this text and we will suggest a way to extend the approach to higher dimensional problems.

1.5 Contents

The structure of this text is as follows. In Chapter 2, after briefly introducing the most frequently used concepts and notations, a next section elaborates on the theory that will be needed to obtain our results. This involves not only some useful properties of the zero sets of multivariate polynomial systems, but also a brief introduction to Sylvester resultant theory which will greatly support our proposed method. The emphasis lies on bivariate systems. The second chapter and this introductory chapter can be considered a brief summary of the literature study that is the foundation of this text. The next chapters contain a detailed description of our linear algebra based method. The different steps will be described both formally and through examples. Numerical issues, efficiency, comparison to existing methods and extension to other polynomial bases will be the subjects of the remainder of this thesis. The ideas can be generalized to systems with more variables ($s > 2$). A rigorous treatment of this generalization is beyond the scope of this text. An illustrative example is given in Appendix G.

Chapter 2

Polynomial Systems

In the first section of this chapter the definition of a polynomial system is given and the notation that is used throughout the text is introduced. The second section summarizes some basic properties related to bivariate polynomial systems that will be needed in the remainder of this thesis.

2.1 Definitions and notations

2.1.1 Multivariate polynomials

The polynomial system of Problem 1 consists of s polynomial equations in s variables. The essential building blocks of such equations are defined as follows.

Definition 2.1. *[26, p. 4] A* monomial *in the s variables x_1, x_2, ..., x_s is a power product of the form $x_1^{j_1} x_2^{j_2} \cdots x_s^{j_s}$ where $(j_1, j_2, \ldots, j_s) \in \mathbb{N}^s$. A complex polynomial in s variables is a finite linear combination of monomials in s variables with coefficients from $\mathbb{C}$.*

The set of all monomials in s variables is denoted by $\mathcal{T}^s$. Let $\deg(\cdot)$ be the operator that maps any monomial in $\mathcal{T}^s$ to its total degree:

$$\deg : \mathcal{T}^s \to \mathbb{N}, \quad \deg(x_1^{j_1} x_2^{j_2} \cdots x_s^{j_s}) \triangleq \sum_{i=1}^{s} j_i.$$

Let $\mathcal{P}^s$ denote the ring of all complex polynomials in s variables. By definition, any polynomial $p(x_1, x_2, \ldots, x_s) \in \mathcal{P}^s$ can be written as

$$p(x_1, x_2, \ldots, x_s) = \sum_{i=1}^{N} c_i m_i(x_1, x_2, \ldots, x_s)$$

where $N \geq 1$, $m_i(x_1, x_2, \ldots, x_s) \in \mathcal{T}^s$ and $c_i \in \mathbb{C}$, $1 \leq i \leq N$. Clearly, $\mathcal{T}^s \subset \mathcal{P}^s$ and the domain of $\deg(\cdot)$ can be extended to $\mathcal{P}^s$ by

$$\deg : \mathcal{P}^s \to \mathbb{N} \cup \{-\infty\}, \quad \deg(p) \triangleq \begin{cases} \max_{i \in \mathrm{Supp}} \{\deg(m_i)\}, & p \neq 0 \\ -\infty, & p = 0 \end{cases}$$

where $p = \sum_{i=1}^{N} c_i m_i$, $m_i \in \mathcal{T}^s$, $1 \leq i \leq N$ and Supp $= \{i \mid c_i \neq 0, 1 \leq i \leq N\}$. Obviously, the operator $\deg(\cdot)$ is defined for any $s \in \mathbb{N}_0$.

Definition 2.2. *A* homogeneous polynomial *in s variables is a polynomial whose nonzero terms all have the same degree. In other words, $p(x_1, x_2, \ldots, x_s)$ is a homogeneous polynomial in s variables if*

$$p(x_1, x_2, \ldots, x_s) = \sum_{i=1}^{N} c_i m_i(x_1, x_2, \ldots, x_s)$$

where $\deg(m_1) = \deg(m_2) = \cdots = \deg(m_N)$ and $\{c_i \neq 0\}_{1 \leq i \leq N}$. The zero polynomial is homogeneous by definition. The set of all homogeneous polynomials in s variables will be denoted by $\mathcal{P}^{s,h}$.

Definition 2.3. *The set $\mathcal{P}_\delta^s \subset \mathcal{P}^s$ of all complex polynomials in s variables of total degree at most δ is defined as*

$$\mathcal{P}_\delta^s \triangleq \{p \in \mathcal{P}^s : \deg(p) \leq \delta\}.$$

Analogously, the set of all homogeneous polynomials in s variables of total degree δ is denoted by $\mathcal{P}_\delta^{s,h}$.

Example 2.1.1. *Consider the polynomial $p(x_1, x_2, x_3) = 1 + 5x_1x_3 + 2x_2 + (3 + 2i)x_1x_2^2$. By definition, the following statements hold true: $p \in \mathcal{P}^3$, $\deg(p) = \deg(x_1x_2^2) = 3$, $p \in \mathcal{P}_5^3$, $p \in \mathcal{P}_3^3$.*

Example 2.1.2. *The polynomial $p(x_1, x_2, x_3, x_4) = x_1x_2^3 + x_2x_3^3 + x_3^2x_4^2$ has total degree 4. Therefore $p \in \mathcal{P}_4^4$. Moreover, it is homogeneous: $p \in \mathcal{P}_4^{4,h}$. It is clear that in general $\mathcal{P}_\delta^{s,h} \subset \mathcal{P}_\delta^s$.*

2.1.2 Bivariate polynomials

In the major part of this text, $\mathcal{P}^2$ is the space of interest. Consider the polynomial $p \in \mathcal{P}^2$ that is a finite linear combination of monomials in the variables x and y. Let $\deg(p(x, y)) = \delta$. Throughout this text, we will use several ways of representing a given polynomial $p(x, y)$:

$$p(x, y) \triangleq \sum_{i=0}^{\delta} \sum_{j=0}^{\delta - i} p_{ij} x^j y^i \triangleq \sum_{i=0}^{\delta} p_i^x(x) y^i \triangleq \sum_{i=0}^{\delta} p_i^y(y) x^i \tag{2.1}$$

where $\deg(p_i^x(x)) \leq \delta - i$ and $\deg(p_i^y(y)) \leq \delta - i$ (in particular: $\deg(p_\delta^x(x)) = 0$ and $\deg(p_\delta^y(x)) = 0$). Moreover, this relation implies that

$$p_i^x(x) = \sum_{j=0}^{\delta - i} p_{ij} x^j$$

and

$$p_i^y(y) = \sum_{j=0}^{\delta-i} p_{ji} y^j.$$

Yet another way of representing $p(x,y)$ is by a matrix $P \in \mathbb{C}^{(\delta+1)\times(\delta+1)}$ with the coefficients of p as its entries: $P_{ij} \triangleq p_{i-1,j-1}$, $1 \leq i \leq \delta+1$ and $p_{ij} = 0$ for any couple (i,j) such that $i > \delta$ or $j > \delta$. Conversely, any matrix $P \in \mathbb{C}^{m\times n}$ defines a complex polynomial $p \in \mathcal{P}^2_{m+n-2}$ as $p(x,y) \triangleq \sum_{i=1}^{m}\sum_{j=1}^{n} P_{ij} x^{j-1} y^{i-1}$.

Example 2.1.3. *Consider the following polynomial in $\mathcal{P}^2$:*

$$\begin{aligned} p(x,y) &= -3 - 2x + x^2 + xy + y^2 \\ &= (-3 + y^2)x^0 + (-2 + y)x^1 + 1x^2 \\ &= (-3 - 2x + x^2)y^0 + xy^1 + 1y^2. \end{aligned}$$

Using the notation as explained in this section, it should be clear that

$$p_{00} = -3, \quad p_{01} = -2, \quad p_{02} = 1, \quad p_{10} = 0, \quad p_{11} = 1, \quad p_{20} = 1,$$

$$p_0^y(y) = -3 + y^2, \quad p_1^y(y) = -2 + y, \quad p_2^y(y) = 1,$$

$$p_0^x(x) = -3 - 2x + x^2, \quad p_1^x(x) = x, \quad p_2^x(x) = 1,$$

$$\text{and} \quad P = \begin{pmatrix} -3 & -2 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

are four different ways of representing p.

When dealing with bivariate polynomials, one can consider not only the total degree but also the degree in the two variables x and y separately.

Definition 2.4. *Given any bivariate polynomial $p(x,y)$, its* degree in x *is defined as*

$$\deg_x(p(x,y)) \triangleq \max_{y^* \in \mathbb{C}} \deg(p(x,y^*))$$

where $p(x,y^*) \in \mathcal{P}^1, \forall y^* \in \mathbb{C}$. Similarly, for the degree of $p(x,y)$ in the variable y we define

$$\deg_y(p(x,y)) \triangleq \max_{x^* \in \mathbb{C}} \deg(p(x^*,y)).$$

Example 2.1.4. *Let $p(x,y) = 2x^9y^2 + x^2y^4 + x^{10} + xy + y^4$, then $\deg(p(x,y)) = 11$, $\deg_x(p(x,y)) = 10$ and $\deg_y(p(x,y)) = 4$.*

Corollary 2.1.1. *Denoting $\deg_x(p(x,y)) = \delta_p^x$ and $\deg_y(p(x,y)) = \delta_p^y$, the most compact matrix representation of $p(x,y)$ in the classical monomial basis is $P \in \mathbb{C}^{(\delta_p^y+1)\times(\delta_p^x+1)}$ such that*

$$p(x,y) = \begin{pmatrix} 1 & y & y^2 & \dots & y^{\delta_p^y} \end{pmatrix} P \begin{pmatrix} 1 & x & x^2 & \dots & x^{\delta_p^x} \end{pmatrix}^\top.$$

Definition 2.5. *Any nonzero bivariate polynomial $p(x, y)$ defines a plane affine algebraic curve:*

$$\mathcal{V}_p \triangleq \{(x, y) \mid p(x, y) = 0\} \subset \mathbb{C}^2.$$

The projective completion *of this affine curve is defined as the plane projective curve given by $p_h(x, y, z) = 0$ where*

$$p_h(x, y, z) \triangleq z^{\deg(p)} p\left(\frac{x}{z}, \frac{y}{z}\right)$$

is a homogeneous polynomial in 3 variables: $p_h \in \mathcal{P}^{3,h}_{\deg(p)}$. The projective curve will be denoted by

$$\Pi\mathcal{V}_p \triangleq \{[x, y, z] \mid p_h(x, y, z) = 0\}$$

where the square brackets indicate projective coordinates[1]*. To visualize the algebraic curve $\mathcal{V}_p$ we will plot the elements of $\mathcal{V}_p \cap \mathbb{R}^2$ and call it the* real picture *of $\mathcal{V}_p$.*

Clearly, $p_h(x, y, 1) = 0$ is the equation of the affine curve $\mathcal{V}_p$, which consists out of the points of the projective curve with a third projective coordinate different from zero. The affine curve can be seen as part of its associated projective curve. The points that belong to the projective completion but not to the affine part are given by the projective coordinates

$$\{[x, y, z] \mid p_h(x, y, z) = 0,\ z = 0\}$$

and they are referred to as the points of $\mathcal{V}_p$ at infinity.

Example 2.1.5. *Consider the polynomial $p(x, y) = y^2 - x(x^2 - 1)$. Its zero set $\mathcal{V}_p = \{(x, y) \mid y^2 - x(x^2 - 1) = 0\}$ defines a so called elliptic curve of which the real picture is given by Figure 2.1. The projective completion $\Pi\mathcal{V}_p$ of $\mathcal{V}_p$ is given by the homogeneous equation $y^2 z - x(x^2 - z^2) = 0$. For $z = 1$ we obtain again the definition of $\mathcal{V}_p$. $\mathcal{V}_p$ has one point at infinity which can be found by setting $z = 0$. It is given by the projective coordinates $[0, 1, 0]$.*

Definition 2.6 (Bivariate polynomial system)**.** *Consider two polynomials p, $q \in \mathcal{P}^2$ that are finite linear combinations of monomials in the variables x and y and let them define a polynomial system*

$$\begin{cases} p(x, y) = 0 \\ q(x, y) = 0 \end{cases}. \tag{2.2}$$

Such a system is referred to as a bivariate polynomial system of degree δ *if* $\max\{\deg(p(x, y)), \deg(q(x, y))\} = \delta$*, i.e.* $\delta \triangleq \min_k\{k \mid p, q \in \mathcal{P}^2_k\}$.

[1] The projective plane is defined as the equivalence classes in $\mathbb{C}^3\backslash\{0\}$ with respect to the equivalence relation

$$(x_1, y_1, z_1) \sim (x_2, y_2, z_2) \Leftrightarrow \exists \lambda \in \mathbb{C}_0 :\ (x_1, y_1, z_1) = (\lambda x_2, \lambda y_2, \lambda z_2).$$

The fact that the set $\Pi\mathcal{V}_p$ is well defined can be seen from the fact that if $[x, y, z] \in \Pi\mathcal{V}_p$, then the same holds for an equivalent set of projective coordinates $[\lambda x, \lambda y, \lambda z]$, since $p_h(\lambda x, \lambda y, \lambda z) = \lambda^{\deg(p)} p_h(x, y, z) = 0$, $(\lambda \in \mathbb{C}_0)$.

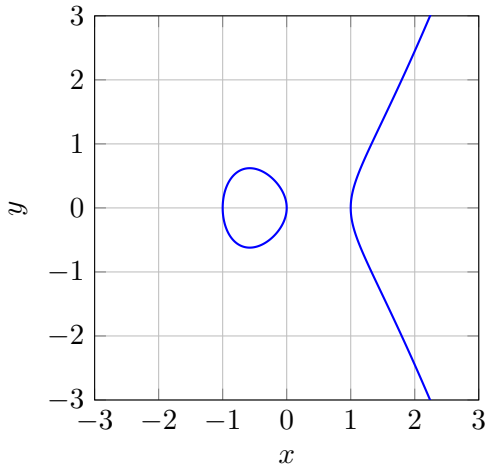


Figure 2.1: Real picture of the elliptic curve defined by $p(x, y) = y^2 - x(x^2 - 1)$.

For the bivariate polynomial system (2.2) the following notations are used:

$$\deg(p(x,y)) \triangleq \delta_p, \quad \deg(q(x,y)) \triangleq \delta_q,$$
$$\deg_x(p(x,y)) \triangleq \delta_p^x, \quad \deg_x(q(x,y)) \triangleq \delta_q^x,$$
$$\deg_y(p(x,y)) \triangleq \delta_p^y, \quad \deg_y(q(x,y)) \triangleq \delta_q^y.$$

2.2 Preliminary theory

2.2.1 Solution sets

Definition 2.7. *The* solution set, zero set or set of roots *of the system* (2.2) *is given by* $\mathcal{V}_{p,q} \triangleq \mathcal{V}_p \cap \mathcal{V}_q$*, i.e. the set of points in* $\mathbb{C}^2$ *where the plane curves* $\mathcal{V}_p$ *and* $\mathcal{V}_q$ *meet. Analogously, the projective solution set is defined as* $\Pi\mathcal{V}_{p,q} \triangleq \Pi\mathcal{V}_p \cap \Pi\mathcal{V}_q$*. The set* $\Pi\mathcal{V}_{p,q}$ *contains the projective coordinates corresponding to the elements of* $\mathcal{V}_{p,q}$ *together with the so called* solutions at infinity*, which are the common points of* $\mathcal{V}_p$ *and* $\mathcal{V}_q$ *at infinity.*

Example 2.2.1. *Consider the second degree polynomial system*

$$\begin{cases} p(x,y) = x^2 + y^2 - 1 = 0 \\ q(x,y) = (x-2)^2 + y^2 - 5 = 0 \end{cases}.$$

The real pictures of the curves $\mathcal{V}_p$ *and* $\mathcal{V}_q$ *are plotted in Figure 2.2. For this system,* $\mathcal{V}_{p,q} = \{(0,1),(0,-1)\}$*. The projective curves are given by:*

$$\Pi\mathcal{V}_p = \{[x,y,z] \mid x^2 + y^2 - z^2 = 0\}$$
$$\Pi\mathcal{V}_q = \{[x,y,z] \mid x^2 - 4xz + 4z^2 + y^2 - 5z^2 = 0\}$$

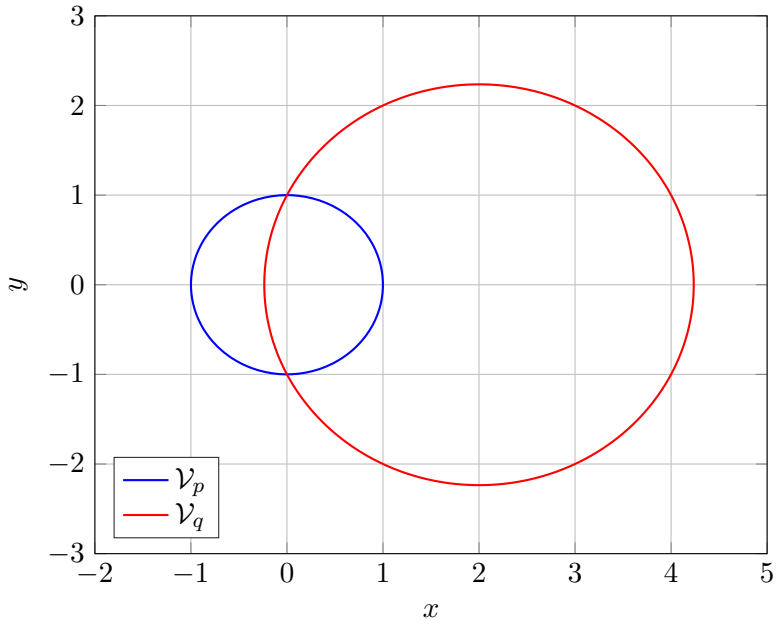


Figure 2.2: Real picture of $\mathcal{V}_p$ and $\mathcal{V}_q$ as defined in Example 2.2.1

.

and thus $\Pi\mathcal{V}_{p,q} = \{\underbrace{[0,1,1],[0,-1,1]}_{\sim\mathcal{V}_{p,q}}, \underbrace{[1,i,0],[1,-i,0]}_{\text{solutions at infinity}}\}$. *The result can be interpreted as follows: the two circles intersect twice in the real plane and twice on the line at infinity* $[x,y,0]$. *Moreover, it is easily verified that any two non identical circles intersect at infinity in the projective points* $[1,i,0]$ *and* $[1,-i,0]$.

Example 2.2.2. *The system*

$$\begin{cases} p(x,y) = xy = 0 \\ q(x,y) = x(xy + x^2 - 1) = 0 \end{cases}$$

has degree $\delta = 3$ *and its solution set consists out of infinitely many points:*

$$\mathcal{V}_{p,q} = \{(1,0),(-1,0)\} \cup \{(x,y) \mid x = 0\}$$

$$\Pi\mathcal{V}_{p,q} = \{[1,0,1],[-1,0,1]\} \cup \{[x,y,1] \mid x = 0\} \cup \{[0,1,0]\}.$$

In this last example, there are infinitely many solutions to the system because the curves $\mathcal{V}_p$ and $\mathcal{V}_q$ have a common component: they are defined by two polynomials p and q that have a greatest common polynomial divisor of degree > 0. From now on, it is always assumed that the polynomials p and q are such that they have a constant greatest common divisor. Such a pair of polynomials is also called *coprime*. The corresponding system has finitely many isolated solutions and is called *0-dimensional*.

Multiplicity of a solution

As for the zeros of a polynomial in one variable, the solutions of a bivariate system can have a multiplicity greater than one. That is, the curves $\mathcal{V}_p$ and $\mathcal{V}_q$ can *intersect multiple times* in the same point. Suppose that the affine curves $\mathcal{V}_p$ and $\mathcal{V}_q$ intersect in the closure of an open set $S \subset \mathbb{C}^2$ in only one point s. Then, suppose the coefficients of p and q (and thus the associated curves) are slightly perturbed. If the perturbations are small enough, the resulting curves, say $\tilde{\mathcal{V}}_p$ and $\tilde{\mathcal{V}}_q$ will intersect in n points in S. The number n coincides with the multiplicity of s. The multiplicity of the intersection s is identified by its intersection number: $M_s(\mathcal{V}_p, \mathcal{V}_q) = n$.

Example 2.2.3. *The real pictures of the affine plane curves corresponding to the system*

$$\begin{cases} p(x,y) = x^2 + y^2 - 1 = 0 \\ q(x,y) = 4x^2 + y^2 + 6x + 2 = 0 \end{cases}$$

are plotted in Figure 2.3. From the figure, it is clear that $s = (-1, 0)$ is a solution of the system. In order to investigate the multiplicity of s one can consider the following perturbation. Let $\tilde{p}$ be a bivariate second degree polynomial with real coefficients in the classical monomial basis that are samples from a normal distribution with mean 0 and standard deviation 1. The system is perturbed in the following way:

$$\begin{cases} p(x,y) + \epsilon\tilde{p}(x,y) = 0 \\ q(x,y) = 0 \end{cases} \tag{2.3}$$

For some small values of ϵ. The system can be interpreted as a parametrized second degree system with parameter ϵ. The multiplicity of s corresponds to the number of branches of the algebraic function $\{x(\epsilon), y(\epsilon)\}$ that meet in s for $\epsilon = 0$. To plot the results, all solutions are calculated for $\epsilon \in \{0, 10^{-5}, 2 \times 10^{-5}, 3 \times 10^{-5}, \ldots, 10^{-3}\}$. All real solutions are plotted in red. For every complex solution, the imaginary part plus the real part is plotted, so that the branches start at $(-1, 0)$ and conjugate pairs can be distinguished. Results are shown for one realization of $\tilde{p}(x, y)$ in Figure 2.3. From the figure, it is clear that $M_s(\mathcal{V}_p, \mathcal{V}_q) = 4$.

A more extended discussion on the multiplicity of a solution requires the notion of polynomial ideals and the associated residue classes. For the interested reader we refer to Appendix A.

Number of solutions

Theorem 2.2.1 (Bézout's theorem)**.** *Let $\Pi\mathcal{V}_p$ and $\Pi\mathcal{V}_q$ be the projective plane curves defined by the polynomials $p(x, y)$ and $q(x, y)$ respectively. Let δ_p and δ_q represent the degree of p and q respectively and assume that the greatest common divisor of p and q is a constant. Then*

$$\sum_{s \in \Pi\mathcal{V}_{p,q}} M_s(\Pi\mathcal{V}_p, \Pi\mathcal{V}_q) = \delta_p \delta_q$$

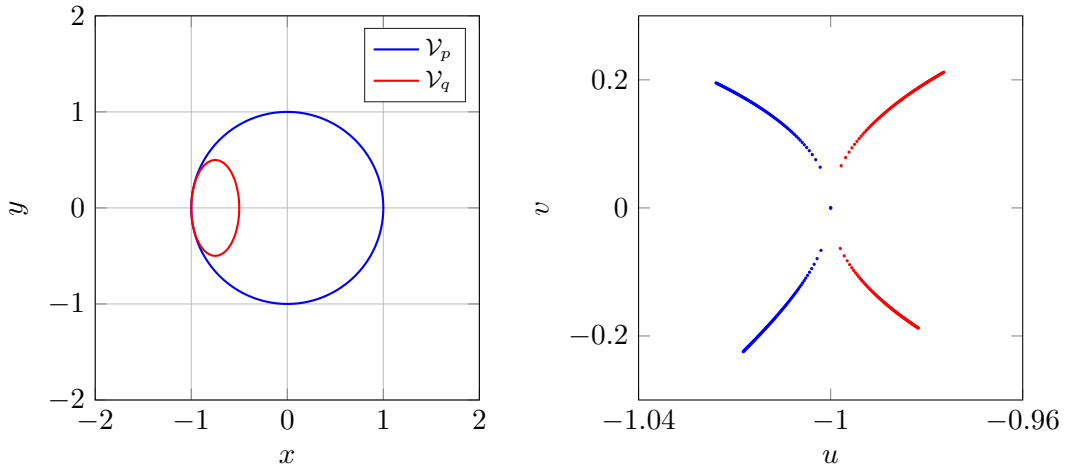


Figure 2.3: Left: real picture of the zero level lines of p and q as defined in Example 2.2.3. Right: all solutions of the perturbed system (2.3) for $\epsilon \in \{0, 10^{-5}, 2\times10^{-5}, 3\times10^{-5}, \ldots, 10^{-3}\}$, plotted in the coordinates $u = \Re(x) + \Im(x)$ and $v = \Re(y) + \Im(y)$. Solutions with $\Im(x) = \Im(y) = 0$ are plotted in red, others in blue.

where $M_s(\Pi\mathcal{V}_p, \Pi\mathcal{V}_q)$ denotes the intersection number of solution s with respect to the projective curves $\Pi\mathcal{V}_p$ and $\Pi\mathcal{V}_q$. In other words: the number of intersections of two projective plane curves, counting multiplicities, is equal to the product of their degrees.

Corollary 2.2.1. *The number of distinct finite solutions of the bivariate polynomial system (2.2), assuming a constant greatest common divisor, is always less than or equal to $\delta_p\delta_q$. Equivalently:*

$$|\mathcal{V}_{p,q}| = |\mathcal{V}_p \cap \mathcal{V}_q| \leq \delta_p\delta_q$$

where $|\cdot|$ denotes the cardinality of a set. The equality only holds when there are no solutions at infinity and $M_s(\mathcal{V}_p, \mathcal{V}_q) = 1, \forall s \in \mathcal{V}_{p,q}$.

Example 2.2.4. *Consider again the problem in Example 2.2.1. The degrees of p and q are given by $\delta_p = 2$ and $\delta_q = 2$. The projective completions $\Pi\mathcal{V}_p$ and $\Pi\mathcal{V}_q$, according to Bézout's theorem, intersect in exactly four points. All the intersections have multiplicity one. The affine plane curves $\mathcal{V}_p$ and $\mathcal{V}_q$ intersect in only two points in $\mathbb{C}^2$, so indeed $|\mathcal{V}_{p,q}| = |\mathcal{V}_p \cap \mathcal{V}_q| = 2 \leq \delta_p\delta_q = 4$.*

A stricter bound on the number of affine roots exists. It takes the sparsity of the polynomials p and q in the monomial basis into account. The bound is associated with the names of Bernstein, Khovanski and Kushnirenko and it is referred to as $\mathrm{BKK}(p, q)$:

$$|\mathcal{V}_p \cap \mathcal{V}_q| \leq \mathrm{BKK}(p, q) \leq \delta_p\delta_q.$$

A description of $\mathrm{BKK}(\mathcal{S})$ for a general system $\mathcal{S}$ of coprime polynomials can be found in Appendix B.

2.2.2 Sylvester matrix formulation [29, 26, 2, 8]

Common zeros of univariate polynomials.

Given two polynomials $f, g \in \mathcal{P}^1$, this subsection deals with the question whether these two polynomials have any zeros (roots) in common. To that end, consider the following definition.

Definition 2.8 (Sylvester matrix)**.** *Given two univariate polynomials $f, g \in \mathcal{P}^1$:*

$$\begin{aligned} f(x) &= f_{\delta_f} x^{\delta_f} + f_{\delta_f - 1} x^{\delta_f - 1} + \cdots + f_1 x + f_0 \\ g(x) &= g_{\delta_g} x^{\delta_g} + g_{\delta_g - 1} x^{\delta_g - 1} + \cdots + g_1 x + g_0 \end{aligned}$$

where $\delta_f = \deg(f)$ and $\delta_g = \deg(g)$, the Sylvester matrix *associated to f and g is given by*

$$S^{f,g} \triangleq \left(\begin{array}{cccccccc} f_0 & f_1 & \cdots & f_{\delta_f} & & & & \\ & f_0 & f_1 & \cdots & f_{\delta_f} & & & \\ & & \ddots & \ddots & & \ddots & & \\ & & & f_0 & f_1 & \cdots & f_{\delta_f} & \\ & & & & f_0 & f_1 & \cdots & f_{\delta_f} \\ \hline g_0 & g_1 & g_2 & \cdots & g_{\delta_g} & & & \\ & g_0 & g_1 & g_2 & \cdots & g_{\delta_g} & & \\ & & \ddots & \ddots & \ddots & & \ddots & \\ & & & g_0 & g_1 & g_2 & \cdots & g_{\delta_g} \end{array}\right), \tag{2.4}$$

where the coefficients of f occur in the first δ_g rows and those of g in the last δ_f rows.

From its definition, it is clear that $S^{f,g}$ is a square matrix of size $\delta_f + \delta_g$. Its significance becomes clear from the following proposition.

Proposition 2.2.1. *If the polynomials f and g have a common root, the Sylvester matrix $S^{f,g}$ is singular.*

Proof. Suppose f and g have a common root for $x = \alpha \in \mathbb{C}$. Let

$$\boldsymbol{v} = \begin{pmatrix} 1 & \alpha & \alpha^2 & \dots & \alpha^{\delta_f + \delta_g - 1} \end{pmatrix}^\top.$$

Now, by construction it holds that $S^{f,g}\boldsymbol{v} = \boldsymbol{0}$. This implies that $S^{f,g}$ has a non-trivial kernel. □

The proof of Proposition 2.2.1 suggests that the common roots of f and g may be determined from the right null space of the Sylvester matrix $S^{f,g}$. The idea is illustrated by a simple example.

Example 2.2.5. *Consider the polynomials*

$$f(x) = x^2 + 1$$

$$g(x) = -3 + x - 3x^2 + x^3$$

of which the associated Sylvester matrix is given by

$$S^{f,g} = \left(\begin{array}{ccccc} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ \hline -3 & 1 & -3 & 1 & 0 \\ 0 & -3 & 1 & -3 & 1 \end{array}\right).$$

Its singular values are[2] *5.6028, 3.5996, 1.2854, 0.0000 and 0.0000. Hence it has a two dimensional right null space. The null space is calculated numerically by using the SVD, it is spanned by the vectors*

$$\begin{pmatrix} -0.5608 & 0.1680 & 0.5608 & -0.1680 & -0.5608 \end{pmatrix}^\top$$

and

$$\begin{pmatrix} -0.1372 & -0.6869 & 0.1372 & 0.6869 & -0.1372 \end{pmatrix}^\top.$$

Now, remark that f and g can be factored as

$$f(x) = (x+i)(x-i),$$

$$g(x) = (x+i)(x-i)(x-3).$$

It is clear that f and g have two common roots: $x = i$ and $x = -i$. The above mentioned orthogonal basis for the kernel of $S^{f,g}$ does not reveal these common roots. However, it is easy to check that

$$\begin{pmatrix} -0.5608 & -0.1372 \\ 0.1680 & -0.6869 \\ 0.5608 & 0.1372 \\ -0.1680 & 0.6869 \\ -0.5608 & -0.1372 \end{pmatrix} \underbrace{\begin{pmatrix} -1.6825 + 0.3359i & -1.6825 - 0.3359i \\ -0.4115 - 1.3737i & -0.4115 + 1.3737i \end{pmatrix}}_{T} = \begin{pmatrix} 1 & 1 \\ i & -i \\ i^2 & (-i)^2 \\ i^3 & (-i)^3 \\ i^4 & (-i)^4 \end{pmatrix}$$

with T non-singular. The rightmost matrix furnishes a basis for the kernel of $S^{f,g}$ that reveals the common roots perfectly.

In Example 2.2.5, the vectors $\begin{pmatrix} 1 & i & \dots & i^4 \end{pmatrix}^\top$ and $\begin{pmatrix} 1 & -i & \dots & (-i)^4 \end{pmatrix}^\top$ reveal the common roots of f and g which are equal to the ratio between their subsequent entries. A vector for which this ratio is constant is called a *univariate Vandermonde vector*. All univariate Vandermonde vectors in the null space of $S^{f,g}$ are of the form

$$\alpha \begin{pmatrix} 1 & x^* & \dots & (x^*)^{\delta_f + \delta_g - 1} \end{pmatrix}^\top$$

[2] The results are given using 4 decimal digits after the decimal point.

where $\alpha \in \mathbb{C}_0$, $\delta_f = \deg(f)$, $\delta_g = \deg(g)$ and x^* is a common zero of f and g. Any vector $\boldsymbol{v}$ that has such a univariate Vandermonde structure must satisfy

$$\begin{pmatrix} -x^* & 1 & & & \\ & -x^* & 1 & & \\ & & \ddots & & \\ & & & -x^* & 1 \end{pmatrix} \boldsymbol{v} \triangleq (S_x - x^* S_1)\boldsymbol{v} = \boldsymbol{0}.$$

Suppose the columns of the matrix Z form a basis for $\operatorname{null}(S^{f,g})$, then any vector in $\operatorname{null}(S^{f,g})$ can be written as $Z\boldsymbol{c}$ for some coefficient vector $\boldsymbol{c}$. We are thus interested in finding all vectors $Z\boldsymbol{c}$ in $\operatorname{null}(S^{f,g})$ that satisfy

$$(S_x - xS_1)Z\boldsymbol{c} = \boldsymbol{0}, \tag{2.5}$$

for some finite x, which is a rectangular eigenvalue problem in x. The eigenvalues of (2.5) are the common zeros of f and g.

Proposition 2.2.1 states that every common zero of f and g generates a vector in $\operatorname{null}(S^{f,g})$. A stronger result about the kernel of $S^{f,g}$ is given by the following proposition [7].

Proposition 2.2.2. *Let $f, g \in \mathcal{P}^1$, then we have* $\dim \operatorname{null}(S^{f,g}) = \deg(\gcd(f,g))$, *where* gcd *denotes the greatest common polynomial divisor.*

Proof. First of all, note that $(S^{f,g})^\top$ can be thought of as the matrix of the linear map

$$\sigma : \mathcal{P}^1_{\delta_g - 1} \times \mathcal{P}^1_{\delta_f - 1} \to \mathcal{P}^1_{\delta_f + \delta_g - 1}$$

defined as

$$\sigma(q_1, q_2) = q_1 f + q_2 g$$

where $q_1 \in \mathcal{P}^1_{\delta_g - 1}$ and $q_2 \in \mathcal{P}^1_{\delta_f - 1}$ are represented in the monomial bases $\{1, x, \ldots, x^{\delta_g - 1}\}$ and $\{1, x, \ldots, x^{\delta_f - 1}\}$ respectively. The image $\sigma(q_1, q_2)$ is represented in the basis $\{1, x, \ldots, x^{\delta_f + \delta_g - 1}\}$. Take for example $f(x) = 1 + x$, $g(x) = -1 + x^2$, $q_1(x) = x$ and $q_2(x) = 1$. Then

$$(S^{f,g})^\top \begin{pmatrix} 0 \\ 1 \\ \hline 1 \end{pmatrix} = \left(\begin{array}{cc|c} 1 & 0 & -1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{array}\right) \begin{pmatrix} 0 \\ 1 \\ \hline 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}$$

and $q_1 f + q_2 g = x(1 + x) - 1 + x^2 = -1 + x + 2x^2$. Now, let $d(x) \triangleq \gcd(f, g)$. It is readily checked that the kernel of σ is the set of pairs

$$\left\{ (q_1, q_2) \mid q_1(x) = r(x)\frac{g(x)}{d(x)}, q_2(x) = -r(x)\frac{f(x)}{d(x)} \text{ for some } r(x) \in \mathcal{P}^1_{\deg(d)-1} \right\},$$

which has dimension $\deg(d)$. Therefore $\dim \operatorname{null}(S^{f,g}) = \dim \operatorname{null}((S^{f,g})^\top) = \deg(d)$. □

Sylvester resultant of bivariate systems

A bivariate system of polynomial equations, defined by $p(x,y), q(x,y) \in \mathcal{P}^2$ can be thought of as a parametrized set of univariate equations. For every value of $\alpha \in \mathbb{C}$, define $f(y|\alpha) \triangleq p(\alpha, y)$ and $g(y|\alpha) \triangleq q(\alpha, y)$. It is clear that $f(y|\alpha)$, $g(y|\alpha) \in \mathcal{P}^1, \forall \alpha \in \mathbb{C}$. Using the notation (2.1), we get

$$f(y|\alpha) = \sum_{i=0}^{\delta_p^y} p_i^x(\alpha) y^i \quad g(y|\alpha) = \sum_{i=0}^{\delta_q^y} q_i^x(\alpha) y^i,$$

so the coefficients of f and g are the coefficient polynomials p_i^x and q_i^x evaluated at $x = \alpha$.[3] Suppose $(\alpha, \beta) \in \mathbb{C}^2$ is an isolated solution to the system defined by p and q, then β must be a common zero of the univariate polynomials $f(y|\alpha)$ and $g(y|\alpha)$.

Proposition 2.2.3. *The existence of a common root of the polynomials $f(y|\alpha)$ and $g(y|\alpha)$ is a necessary condition for the existence of a pair $(\alpha, \beta) \in \mathbb{C}^2$ that satisfies $p(\alpha, \beta) = q(\alpha, \beta) = 0$. Moreover, for all such pairs (α, β) the possible values of β are given by the distinct common roots of $f(y|\alpha)$ and $g(y|\alpha)$.*

As explained above, the Sylvester matrix is a tool for checking the existence of a common root of two univariate polynomials. Since $f(y|\alpha)$ and $g(y|\alpha)$ depend on the choice of the parameter α (i.e. the value that is assigned to the variable x to construct f from p and g from q), so does the associated Sylvester matrix:

$$S^{p,q}(x) = \left(\begin{array}{cccccc}
p_0^x(x) & p_1^x(x) & \dots & p_{\delta_p^y}^x(x) & & \\
 & p_0^x(x) & p_1^x(x) & \dots & p\delta_p^{yx}(x) & \\
 & & \ddots & \ddots & & \ddots \\
 & & & p_0^x(x) & p_1^x(x) & \dots & p_{\delta_p^y}^x(x) \\
\hline
q_0^x(x) & q_1^x(x) & \dots & q_{\delta_q^y}^x(x) & & \\
 & q_0^x(x) & q_1^x(x) & \dots & q_{\delta_q^y}^x(x) & \\
 & & \ddots & \ddots & & \ddots \\
 & & & q_0^x(x) & q_1^x(x) & \dots & q_{\delta_q^y}^x(x)
\end{array}\right) \tag{2.6}$$

which is a matrix polynomial of degree $\max\{\delta_p^x, \delta_q^x\}$. There are δ_q^y p-rows and δ_p^y q-rows. From proposition (2.2.1), for some choice of α, a necessary condition for $f(y|\alpha)$ and $g(y|\alpha)$ to have a common root is that $S^{p,q}(\alpha)$ is singular.

Definition 2.9 (Sylvester resultant)**.** *For two bivariate polynomials $p, q \in \mathcal{P}^2$, the* Sylvester resultant *with respect to the variable y is defined as*

$$\mathrm{res}^{p,q}(x) \triangleq \det S^{p,q}(x)$$

with $S^{p,q}(x)$ as defined in (2.6)*. An analogous definition holds for the resultant with respect to the variable x.*

[3] Of course, one could follow the same procedure using the variable y as a parameter, obtaining parametrized univariate polynomials in x. The rest of the reasoning in this section would be completely analogous. It is chosen to proceed using the value of x as a parameter.

Theorem 2.2.2. *[29, 26, 2, 8] If p and $q \in \mathcal{P}^2$ do not have a nontrivial greatest common divisor (p and q are coprime), the distinct roots of $\mathrm{res}^{p,q}(x)$ are the x-coordinates in $\mathbb{C}$ of the isolated roots of the bivariate system $p(x,y) = q(x,y) = 0$ and the common roots of the leading coefficient polynomials $p^x_{\delta^y_p}(x)$ and $q^x_{\delta^y_q}(x)$. In other words*

$$\{x \in \mathbb{C} \mid \mathrm{res}^{p,q}(x) = 0\} = \mathcal{V}^{(x)}_{p,q} \cup \{x \in \mathbb{C} \mid p^x_{\delta^y_p}(x) = q^x_{\delta^y_q}(x) = 0\}$$

where $\mathcal{V}^{(x)}_{p,q} \triangleq \{x \in \mathbb{C} \mid \exists y \in \mathbb{C} : (x,y) \in \mathcal{V}_{p,q}\}$. The multiplicity of a zero x^* of $\mathrm{res}^{p,q}(x)$ is equal to the sum of the multiplicities of all roots of the form (x^*, y)[4].

Theorem 2.2.2 states that the x-values for which $\mathrm{res}^{p,q}(x)$ vanishes are the x-values in $\mathbb{C}$ corresponding to the affine solutions of the system (2.2) along with some spurious values of x in case $p^x_{\delta^y_p}(x)$ and $q^x_{\delta^y_q}(x)$ have common zeros.

Example 2.2.6. *Consider the system*

$$\begin{cases} p(x,y) = -1 + (1+x)y^2 = 0 \\ q(x,y) = x + (x^2-1)y = 0 \end{cases}$$

which has the finite solutions $(2.2470, -0.5550)$, $(0.5550, 0.8019)$ and $(-0.8019, -2.2470)$. The real picture of the affine curves $\mathcal{V}_p$ and $\mathcal{V}_q$ is plotted in Figure 2.4. The associated Sylvester matrix is given by

$$S^{p,q}(x) = \begin{pmatrix} -1 & 0 & 1+x \\ x & x^2-1 & \\ & x & x^2-1 \end{pmatrix}.$$

The resultant is calculated as the determinant of this matrix: $\mathrm{res}^{p,q}(x) = -(x^2-1)^2 + (1+x)x^2$. From Figure 2.4, it can be seen that the resultant vanishes for $x \in \{-0.8019, 0.5550, 2.2470\}$, which are the zeros that correspond to the finite solutions of the system. It also vanishes for $x = -1$, which corresponds to the common root of the leading coefficient polynomials $p^x_2(x) = 1 + x$ and $q^x_1(x) = x^2 - 1$. Note that such a common root corresponds to a solution "at infinity" with a finite x-coordinate. In this case, p and q intersect at $(-1, \pm\infty)$, as can be seen from Figure 2.4.

[4] Solutions of the form $(x, \pm\infty)$ have to be taken into account. These solutions correspond to values of x that are in $\{x \in \mathbb{C} \mid p^x_{\delta^y_p}(x) = q^x_{\delta^y_q}(x) = 0\}$.

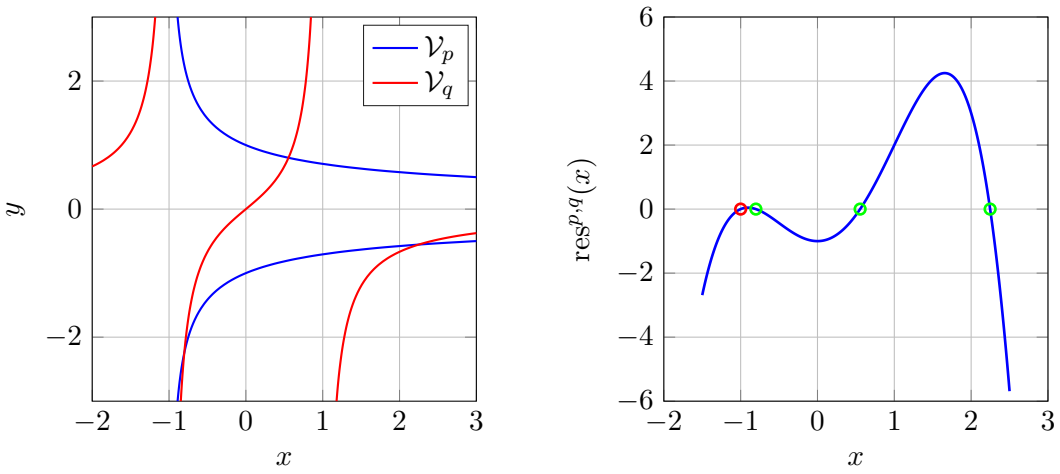


Figure 2.4: Left: real picture of the zero level sets of p and q as defined in Example 2.2.6. Right: Sylvester resultant of the system. Roots that correspond to affine solutions of the bivariate system are indicated by a green mark, the spurious root is marked red.

Chapter 3

Linearization of Bivariate Polynomial Systems

In this chapter, the first section shows how solving a bivariate polynomial system can be interpreted as solving an associated two-parameter eigenvalue problem. The second section describes what is meant by degree extension and how it is used to construct an extended eigenvalue problem. The resulting problem will be separable, which allows for the (otherwise computationally expensive) two-parameter problem to be solved in an efficient way [23]. The last section describes how the extended pencil can be reduced in size by deleting redundant columns and rows.

3.1 A two-parameter eigenvalue formulation

3.1.1 The coefficient matrix Φ

Consider a bivariate system of polynomial equations (2.2) of degree δ defined by the polynomials p, $q \in \mathcal{P}_\delta^2$:

$$\begin{cases} p(x,y) = 0 \\ q(x,y) = 0 \end{cases}. \tag{3.1}$$

Equivalently, the system can be written as

$$\begin{cases} \sum_{i=0}^{\delta} \sum_{j=0}^{\delta-i} p_{ij} x^j y^i = 0 \\ \sum_{i=0}^{\delta} \sum_{j=0}^{\delta-i} q_{ij} x^j y^i = 0 \end{cases}. \tag{3.2}$$

A first step in translating the problem into an eigenvalue problem is thinking of p and q as the result of a left multiplication of a vector $\boldsymbol{v}(x,y)$ of monomials by a coefficient matrix Φ_{pq}. From (3.2), it can be seen that

$$\begin{pmatrix} p(x,y) \\ q(x,y) \end{pmatrix} = \Phi_{pq} \boldsymbol{v}(x,y)$$

where

$$\Phi_{pq} \triangleq \left(\begin{array}{ccccc|cccc|c|cc|c} p_{00} & p_{01} & p_{02} & \cdots & p_{0\delta} & p_{10} & p_{11} & \cdots & p_{1,\delta-1} & \cdots & p_{\delta-1,0} & p_{\delta-1,1} & p_{\delta 0} \\ q_{00} & q_{01} & q_{02} & \cdots & q_{0\delta} & q_{10} & q_{11} & \cdots & q_{1,\delta-1} & \cdots & q_{\delta-1,0} & q_{\delta-1,1} & q_{\delta 0} \end{array}\right)$$

and

$$\boldsymbol{v}(x,y) \triangleq \left(\begin{array}{ccccc|cccc|c|cc|c} 1 & x & x^2 & \cdots & x^\delta & y & xy & \cdots & x^{\delta-1}y & \cdots & y^{\delta-1} & xy^{\delta-1} & y^\delta \end{array}\right)^\top.$$

The vector $\boldsymbol{v}(x,y)$ represents the classical monomial basis of the space $\mathcal{P}_\delta^2$ (thought of as a vector space). The fact that the entries of $\boldsymbol{v}(x,y)$ construct a basis of $\mathcal{P}_\delta^2$ makes sure that any system of degree δ can be written in this way[1]. The matrix Φ_{pq} is determined by the polynomials p and q: it contains the coefficients of p and q in the classical monomial basis. From now on, the subscript $\cdot_{pq}$ is dropped to simplify the notation. It is chosen to order the bivariate monomials of degree $\leq \delta$ first in groups of increasing degree in y and then, within each group, by increasing degree in x. This type of ordering will simplify some of the notation in the next sections and chapters and it is found to be the most suitable choice for describing our results. Note that the matrix Φ is partitioned into blocks that correspond to monomials of increasing degree in y. We are thus interested in finding all pairs $(x^*, y^*) \in \mathbb{C}^2$ for which

$$\Phi \boldsymbol{v}(x^*, y^*) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{3.3}$$

It is clear that every solution (x^*, y^*) generates a vector in the right null space of the matrix Φ. However, the converse is not true.

Example 3.1.1. *Consider the following system of degree 2:*

$$\begin{cases} x^2 + y^2 = 1 \\ y = 0 \end{cases}$$

which is known to have the solutions $(-1, 0)$ *and* $(1, 0)$*. The corresponding matrix* Φ *is:*

$$\Phi = \left(\begin{array}{ccc|cc|c} -1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{array}\right)$$

The monomial vector of degree 2 is

$$\boldsymbol{v}(x,y) = \left(\begin{array}{ccc|cc|c} 1 & x & x^2 & y & xy & y^2 \end{array}\right)^\top$$

It is easy to check that indeed $\boldsymbol{v}(-1,0) = \begin{pmatrix} 1 & -1 & 1 & 0 & 0 & 0 \end{pmatrix}^\top$ *and* $\boldsymbol{v}(1,0) = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \end{pmatrix}^\top$ *are two independent vectors that are contained in the right null space of* Φ*. Also, the vector* $\boldsymbol{w} = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}^\top$ *satisfies* $\Phi \boldsymbol{w} = \boldsymbol{0}$*, but it does not correspond to a solution of the system. Moreover, in this case, the right null space of* Φ *has dimension 4 while there are only 2 solutions to the system.*

[1] This choice of basis has been made to introduce the basic concepts of the proposed method to solve the problem (3.1). One could come to analogous results making use of a different basis. This issue is discussed later on in this text.

3.1.2 The Vandermonde structure

The question is how to select the "useful" right null vectors of Φ. To that end, consider the following definition.

Definition 3.1. *A vector* $\boldsymbol{w} \in \mathbb{C}^{\frac{(\delta+2)(\delta+1)}{2}}$ *is said to have a* Vandermonde structure *in the classical monomial basis of* $\mathcal{P}^2_\delta$ *if* $\exists (x, y) \in \mathbb{C}^2, c \in \mathbb{C}_0$ *s.t.*

$$\boldsymbol{w} = c \begin{pmatrix} 1 & x & x^2 & \dots & x^\delta \mid y & xy & \dots & x^{\delta-1}y \mid \dots \mid y^{\delta-1} & xy^{\delta-1} \mid y^\delta \end{pmatrix}^\top.$$

It is clear that the monomial vector $\boldsymbol{v}(x, y)$ has a Vandermonde structure $\forall (x, y) \in \mathbb{C}^2$. The null vectors of Φ that are directly related to a solution (x^*, y^*) of (3.1) are those with a Vandermonde structure. A straightforward thing to do is to add equations to $\Phi \boldsymbol{v} = \boldsymbol{0}$ that impose the Vandermonde structure on the vector $\boldsymbol{v}$. For example, the requirement that the second entry of $\boldsymbol{v}$ is equal to x times the first entry leads to the equation

$$\begin{pmatrix} -x & 1 & 0 & \dots & 0 \mid 0 & 0 & \dots & 0 \mid \dots \mid 0 & 0 \mid 0 \end{pmatrix} \boldsymbol{v} = 0.$$

Analogous equations need to be added for the other entries of $\boldsymbol{v}$. These equations are not unique. For example, to impose that $\boldsymbol{v}_{\delta+1} = x^\delta \boldsymbol{v}_1$, the equation

$$\begin{pmatrix} 0 & -x^{\delta-1} & 0 & \dots & 1 \mid 0 & 0 & \dots & 0 \mid \dots \mid 0 & 0 \mid 0 \end{pmatrix} \boldsymbol{v} = 0$$

can be added, but also

$$\begin{pmatrix} 0 & 0 & \dots & -x & 1 \mid 0 & 0 & \dots & 0 \mid \dots \mid 0 & 0 \mid 0 \end{pmatrix} \boldsymbol{v} = 0$$

will do (supposing that $\boldsymbol{v}_\delta = x^{\delta-1}\boldsymbol{v}_1$ is guaranteed by one of the other equations). Also, the following two equations will make sure $\boldsymbol{v}_{\delta+3} = xy\boldsymbol{v}_1$

$$\begin{pmatrix} 0 & -y & 0 & \dots & 0 \mid 0 & 1 & \dots & 0 \mid \dots \mid 0 & 0 \mid 0 \end{pmatrix} \boldsymbol{v} = 0,$$

$$\begin{pmatrix} 0 & 0 & 0 & \dots & 0 \mid -x & 1 & \dots & 0 \mid \dots \mid 0 & 0 \mid 0 \end{pmatrix} \boldsymbol{v} = 0.$$

In order to obtain a linear eigenvalue problem, x and y should appear only linearly in the equations. Also, for reasons that will become clear in the next sections, it is chosen to avoid using y as much as possible. These choices lead to a unique set of equations that imposes the Vandermonde structure on a nonzero vector $\boldsymbol{v}$. Define

$$\underline{I}_i \triangleq \begin{pmatrix} I_i & \boldsymbol{0}_i \end{pmatrix}, \overline{I}_i \triangleq \begin{pmatrix} \boldsymbol{0}_i & I_i \end{pmatrix}, \tag{3.4}$$

where I_i stands for the identity matrix of size $i \times i$ and $\boldsymbol{0}_i$ stands for a column vector of length i filled with zeros. $\underline{I}_i$ and $\overline{I}_i$ are row truncated identity matrices of size

$i \times (i+1)$. Let $\boldsymbol{e}_i$ be the first column of I_i and let

$$
\begin{aligned}
\mathcal{B}_x &= \begin{pmatrix} \overline{I}_\delta & & & & & \\ & \overline{I}_{\delta-1} & & & & \\ & & \overline{I}_{\delta-2} & & & \\ & & & \ddots & & \\ & & & & \overline{I}_1 & 0 \end{pmatrix}, \\
\mathcal{C}_x &= \begin{pmatrix} \underline{I}_\delta & & & & & \\ & \underline{I}_{\delta-1} & & & & \\ & & \underline{I}_{\delta-2} & & & \\ & & & \ddots & & \\ & & & & \underline{I}_1 & 0 \end{pmatrix}, \\
\mathcal{B}_y &= \begin{pmatrix} \mathbf{0}_{\delta+1}^\top & \boldsymbol{e}_\delta^\top & & & & \\ \mathbf{0}_{\delta+1}^\top & & \boldsymbol{e}_{\delta-1}^\top & & & \\ \mathbf{0}_{\delta+1}^\top & & & \boldsymbol{e}_{\delta-2}^\top & & \\ \vdots & & & & \ddots & \\ \mathbf{0}_{\delta+1}^\top & & & & & \boldsymbol{e}_1^\top \end{pmatrix}, \\
\mathcal{C}_y &= \begin{pmatrix} \boldsymbol{e}_{\delta+1}^\top & & & & 0 \\ & \boldsymbol{e}_\delta^\top & & & 0 \\ & & \boldsymbol{e}_{\delta-1}^\top & & 0 \\ & & & \ddots & \vdots \\ & & & \boldsymbol{e}_2^\top & 0 \end{pmatrix}.
\end{aligned}
\tag{3.5}
$$

The dimensions of these matrices are given by $\mathcal{B}_x, \mathcal{C}_x \in \mathbb{C}^{\frac{\delta(\delta+1)}{2} \times \frac{(\delta+1)(\delta+2)}{2}}$ and $\mathcal{B}_y, \mathcal{C}_y \in \mathbb{C}^{\delta \times \frac{(\delta+1)(\delta+2)}{2}}$ (blank spaces indicate zero entries). A vector $\boldsymbol{v}$ has the desired Vandermonde structure if it satisfies the following linear equations:

$$
\begin{pmatrix} \mathcal{B}_x \\ \mathcal{B}_y \end{pmatrix} \boldsymbol{v} = x \begin{pmatrix} \mathcal{C}_x \\ \end{pmatrix} \boldsymbol{v} + y \begin{pmatrix} \\ \mathcal{C}_y \end{pmatrix} \boldsymbol{v}. \tag{3.6}
$$

3.1.3 A two-parameter eigenvalue problem

Adding the equations (3.6) to $\Phi \boldsymbol{v} = \mathbf{0}$ we find that the solutions to the problem (3.1) are all eigenvalues (x^*, y^*) that satisfy

$$
\begin{pmatrix} \Phi \\ \mathcal{B}_x \\ \mathcal{B}_y \end{pmatrix} \boldsymbol{v} = x^* \begin{pmatrix} \\ \mathcal{C}_x \\ \end{pmatrix} \boldsymbol{v} + y^* \begin{pmatrix} \\ \\ \mathcal{C}_y \end{pmatrix} \boldsymbol{v}. \tag{3.7}
$$

The size of the problem is $(2 + \frac{\delta(\delta+1)}{2} + \delta) \times (\frac{(\delta+1)(\delta+2)}{2})$, which means the problem is "nearly square": there is one more row than there are columns. This means that,

defining Φ_p as the first row of Φ and Φ_q as the second row of Φ, the problem

$$\begin{cases} \underbrace{\begin{pmatrix} \Phi_p \\ \mathcal{B}_x \\ \mathcal{B}_y \end{pmatrix}}_{A_p} \boldsymbol{v} = x \underbrace{\begin{pmatrix} \\ \mathcal{C}_x \\ \\ \end{pmatrix}}_{C_x} \boldsymbol{v} + y \underbrace{\begin{pmatrix} \\ \\ \mathcal{C}_y \end{pmatrix}}_{C_y} \boldsymbol{v} \\ \underbrace{\begin{pmatrix} \Phi_q \\ \mathcal{B}_x \\ \mathcal{B}_y \end{pmatrix}}_{A_q} \boldsymbol{v} = x \begin{pmatrix} \\ \mathcal{C}_x \\ \\ \end{pmatrix} \boldsymbol{v} + y \begin{pmatrix} \\ \\ \mathcal{C}_y \end{pmatrix} \boldsymbol{v} \end{cases} \leftrightarrow \begin{cases} (A_p + xC_x + yC_y)\boldsymbol{v} = 0 \\ (A_q + xC_x + yC_y)\boldsymbol{v} = 0 \end{cases} \tag{3.8}$$

is a square two-parameter eigenvalue problem as introduced in Section 1.3.2. An important disadvantage of this type of problem is that, in order to solve it, current algorithms translate it into a classical generalized eigenvalue problem with a dimension equal to the dimension of the two-parameter problem squared [23]. We will propose a way to translate the problem to a problem in only x or only y that avoids this dramatic "blow up" of the dimension.

Example 3.1.2. *Let us consider the problem*

$$\begin{cases} p(x,y) = x^2 + y^2 - 4 \\ q(x,y) = -3 - 2x + x^2 + xy + y^2 \end{cases}.$$

In this case, $\delta = 2$ and we expect, from Bézout's theorem, $\delta_p \delta_q = 4$ (projective) solutions. Now, one can easily verify that for this problem, the matrices $\Phi, \mathcal{B}_x, \mathcal{B}_y, \mathcal{C}_x$ and $\mathcal{C}_y$ are given by

$$\Phi = \left(\begin{array}{ccc|cc|c} -4 & 0 & 1 & 0 & 0 & 1 \\ -3 & -2 & 1 & 0 & 1 & 1 \end{array}\right),$$

$$\mathcal{B}_x = \left(\begin{array}{ccc|cc|c} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right), \quad \mathcal{C}_x = \left(\begin{array}{ccc|cc|c} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 \end{array}\right),$$

$$\mathcal{B}_y = \left(\begin{array}{ccc|cc|c} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right), \quad \mathcal{C}_y = \left(\begin{array}{ccc|cc|c} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{array}\right).$$

The (6×6) *multi-parameter eigenvalue problem* (3.8) *is given by*

$$\begin{cases} \left(\begin{array}{ccc|ccc} -4 & 0 & 1 & 0 & 0 & 1 \\ \hline 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right) \boldsymbol{v} = x \left(\begin{array}{ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \boldsymbol{v} + y \left(\begin{array}{ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{array}\right) \boldsymbol{v} \\ \left(\begin{array}{ccc|ccc} -3 & -2 & 1 & 0 & 1 & 1 \\ \hline 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right) \boldsymbol{v} = x \left(\begin{array}{ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \boldsymbol{v} + y \left(\begin{array}{ccc|ccc} 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{array}\right) \boldsymbol{v} \end{cases}$$

and it is solved using the `twopareig` *function in the MultiParEig toolbox written in Matlab [22]. All solutions turn out to be finite and they are given by* $(-0.8722 - 1.3810i, 2.3269 - 0.5176i)$*,* $(-0.8722 + 1.3810i, 2.3269 + 0.5176i)$*,* $(1.4934, 1.3303)$ *and* $(0.2510, -1.9842)$*.*

3.1.4 A rectangular linear pencil in x and y.

We have formulated (3.7) as a square two-parameter eigenvalue problem. Alternatively, we can write

$$L(x, y)\boldsymbol{v} = \mathbf{0} \tag{3.9}$$

where $L(x, y)$ is defined as the linear pencil

$$L(x, y) \triangleq \left(\begin{array}{c} \Phi \\ \hline \mathcal{B}_x - x\mathcal{C}_x \\ \hline \mathcal{B}_y - y\mathcal{C}_y \end{array}\right) \triangleq \left(\begin{array}{c} \Pi_x(x) \\ \hline \Pi_y(y) \end{array}\right) \in \mathbb{C}^{(2+\frac{\delta(\delta+1)}{2}+\delta)\times(\frac{(\delta+1)(\delta+2)}{2})}$$

in the variables x and y. We denote

$$\begin{cases} p(x, y) = 0 \\ q(x, y) = 0 \end{cases} \quad \xrightarrow[C]{} \quad L(x, y)$$

where C stands for construction of the pencil as explained in the previous subsections. $L(x, y)$ can be subdivided into a coefficient block row (Φ) and two block rows that define the monomial basis, one using only x $(\mathcal{B}_x - x\mathcal{C}_x)$ and one using only y $(\mathcal{B}_y - y\mathcal{C}_y)$. Constructing $L(x, y)$ in the above mentioned way, the following theorem holds.

Theorem 3.1.1. *Let*

$$\begin{cases} p(x, y) = 0 \\ q(x, y) = 0 \end{cases} \quad \xrightarrow[C]{} \quad L(x, y).$$

A pair $(x^*, y^*) \in \mathbb{C}^2$ *is a solution to problem* (3.1) *if and only if* $\exists \boldsymbol{v} \neq \mathbf{0}$ *s.t.* $L(x^*, y^*)\boldsymbol{v} = \mathbf{0}$. *The solutions to* (3.1) *are the couples* (x^*, y^*) *for which* $L(x, y)$ *loses full column rank.*

Proof. If $\exists \boldsymbol{v} \neq \mathbf{0}$ that satisfies (3.9), $\boldsymbol{v}$ has a Vandermonde structure because it satisfies

$$\begin{pmatrix} \mathcal{B}_x - x^*\mathcal{C}_x \\ \mathcal{B}_y - y^*\mathcal{C}_y \end{pmatrix} \boldsymbol{v} = \mathbf{0}.$$

Therefore

$$\Phi \boldsymbol{v} = \begin{pmatrix} cp(x^*, y^*) \\ cq(x^*, y^*) \end{pmatrix} = \mathbf{0}$$

for some $c \in \mathbb{C}_0$, which proves the if direction. For the only if direction, let $\boldsymbol{v}$ be the monomial vector $\boldsymbol{v}(x, y)$ of degree δ evaluated at $(x, y) = (x^*, y^*)$, where (x^*, y^*) is a solution to (3.1). □

3.2 Degree extension

Consider Example 3.1.2. The resulting system in the form (3.7) consists out of seven equations: two equations that originate from the coefficients of p and q in the classical monomial basis and five equations that somehow "define" the basis. From these last five equations, only two contain the variable y. Leaving out these two equations would thus lead to a 5×6 rectangular generalized eigenvalue problem (see Appendix E) in one parameter x. Such a "flat" eigenvalue problem has infinitely many eigenvalues: the matrix

$$\Pi_x(x) = \begin{pmatrix} \Phi \\ \mathcal{B}_x - x\mathcal{C}_x \end{pmatrix} \in \mathbb{C}^{5\times 6}$$

of Example 3.1.2 has a non trivial rigt null space for all values of $x \in \mathbb{C}$. Let us have another look at the dimensions of $\mathcal{B}_x$ and $\mathcal{B}_y$ (of which the number of rows represents the number of equations in x and in y respectively):

$$\mathcal{B}_x \in \mathbb{C}^{\frac{\delta(\delta+1)}{2} \times \frac{(\delta+1)(\delta+2)}{2}},$$

$$\mathcal{B}_y \in \mathbb{C}^{\delta \times \frac{(\delta+1)(\delta+2)}{2}}.$$

The number of equations in x grows like $\frac{\delta^2}{2}$ while the number of equations in y only grows like δ. This is due to the construction of the matrices $\mathcal{B}_x, \mathcal{B}_y, \mathcal{C}_x$ and $\mathcal{C}_y$, where the use of y is avoided as much as possible. For degree δ, $\mathcal{B}_x - x\mathcal{C}_x$ provides us with $\frac{1}{2}(\delta^2 + \delta)$ equations in x. The number of monomials of degree δ (the number of columns) is equal to $\frac{1}{2}(\delta^2 + 3\delta + 2)$. This means that in general there is a shortage of $\delta + 1$ equations in x for $\mathcal{B}_x - x\mathcal{C}_x$ to be square. The block row Φ can only make

up for two of these missing equations, so in the end we need to add $\delta - 1$ equations to $\Pi_x(x)$ to obtain a square pencil. Indeed, for $\delta = 1$ we have

$$\begin{cases} \alpha_1 x + \beta_1 y + \gamma_1 = 0 \\ \alpha_2 x + \beta_2 y + \gamma_2 = 0 \end{cases} \xrightarrow[C]{} \left(\begin{array}{ccc} \gamma_1 & \alpha_1 & \beta_1 \\ \gamma_2 & \alpha_2 & \beta_2 \\ -x & 1 & \\ \hline -y & & 1 \end{array}\right)$$

and $\Pi_x(x)$ is a square pencil. For $\delta > 1$, we use a procedure that will be referred to as *degree extension* to add equations to $\Pi_x(x)$. Suppose we want to *extend* the degree δ by 1. The monomial basis is extended by all monomials of degree $\delta + 1$ and the (extended) matrices $\hat{\Phi}$, $\hat{\mathcal{B}}_x$, $\hat{\mathcal{B}}_y$, $\hat{\mathcal{C}}_x$ and $\hat{\mathcal{C}}_y$ are constructed as follows. The $\hat{\mathcal{B}}$ and $\hat{\mathcal{C}}$ matrices can be defined as in (3.5) replacing δ by $\delta + 1$. For the construction of $\hat{\Phi}$, p and q can be thought of as polynomials of total degree $\delta + 1$ with zero coefficients corresponding to monomials of degree $\delta + 1$. To find the missing equations for $\Pi_x(x)$ we make use of the following property.

Property 3.2.1. *The following equivalence holds, for any pair of polynomials $p(x, y), q(x, y) \in \mathcal{P}^2$ and for any $\Delta\delta_p \geq 0$, $\Delta\delta_q \geq 0$, $(x, y) \in \mathbb{C}^2$:*

$$\begin{cases} p(x,y) = 0 \\ q(x,y) = 0 \end{cases} \Leftrightarrow \begin{cases} p(x,y) = 0 \\ q(x,y) = 0 \\ yp(x,y) = 0 \\ \quad \vdots \\ y^{\Delta\delta_p} p(x,y) = 0 \\ yq(x,y) = 0 \\ \quad \vdots \\ y^{\Delta\delta_q} q(x,y) = 0 \end{cases} . \tag{3.10}$$

Proof. The $\Leftarrow$ implication follows from the first two equations from the extended system. The $\Rightarrow$ implication is also straightforward: if p and q vanish for some couple (x, y), so does every left hand side expression from the equations of the extended system[2]. □

Definition 3.2. *The system at the right side of the '$\Leftrightarrow$' sign in* (3.10) *will be referred to as the* extended system *and its degree will be called the* extended degree*, which will be denoted by $\hat{\delta} = \max(\hat{\delta}_p, \hat{\delta}_q)$ where $\hat{\delta}_p = \delta_p + \Delta\delta_p$ and $\hat{\delta}_q = \delta_q + \Delta\delta_q$. The difference between δ and $\hat{\delta}$ will be referred to as the* shift degree *and it is denoted by $\Delta\delta = \hat{\delta} - \delta$.*

[2] This can also be seen by the fact that $\langle p(x, y), q(x, y)\rangle$ and $\langle p, q, yp, \ldots, y^{\Delta\delta_p} p, yq, \ldots, y^{\Delta\delta_q} q\rangle$ generate the same polynomial ideal. Indeed, we can keep adding polyomial combinations of p and q to the extended system without altering this ideal and thus the solution set. More information on polynomial ideals and polynomial combinations can be found in Appendix A.

Now, we can add another block row Ψ to $\Pi_x(x)$ containing two rows corresponding to the coefficients of $yp(x,y)$ and $yq(x,y)$ in the extended monomial basis. Dropping the y-rows, the resulting equation looks like this:

$$\begin{pmatrix} \hat{\Phi} \\ \Psi \\ \hat{\mathcal{B}}_x - x\hat{\mathcal{C}}_x \end{pmatrix} \boldsymbol{v} = \mathbf{0}. \tag{3.11}$$

Example 3.2.1. *Consider again p and q from Example 3.1.2:*

$$p(x,y) = x^2 + y^2 - 4$$

$$q(x,y) = -3 - 2x + x^2 + xy + y^2$$

Increasing the degree by 1 ($\Delta\delta_p = \Delta\delta_q = \Delta\delta = 1$) results in an extended degree of $\hat{\delta} = \delta + \Delta\delta = 2 + 1 = 3$. The monomial basis vector $\boldsymbol{v}(x,y)$ is

$$\begin{pmatrix} 1 & x & x^2 & x^3 \mid y & xy & x^2y \mid y^2 & xy^2 \mid y^3 \end{pmatrix}^\top$$

and the matrices from (3.11) *can be found as*

$$\hat{\Phi} = \left(\begin{array}{cccc|ccc|cc|c} -4 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ -3 & -2 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \end{array}\right),$$

$$\Psi = \left(\begin{array}{cccc|ccc|cc|c} 0 & 0 & 0 & 0 & -4 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -3 & -2 & 1 & 0 & 1 & 1 \end{array}\right),$$

$$\hat{\mathcal{B}}_x = \left(\begin{array}{cccc|ccc|cc|c} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{array}\right),\ \hat{\mathcal{C}}_x = \left(\begin{array}{cccc|ccc|cc|c} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{array}\right).$$

The dropped y-rows in (3.11) *are given by $\hat{\mathcal{B}}_y - y\hat{\mathcal{C}}_y$ where*

$$\hat{\mathcal{B}}_y = \left(\begin{array}{cccc|ccc|cc|c} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right),\ \hat{\mathcal{C}}_y = \left(\begin{array}{cccc|ccc|cc|c} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{array}\right).$$

In general, the dimensions of the matrices involved in the x-pencil for an extended system of degree $\hat{\delta} \geq \delta$ are given by:

$$\hat{\Phi} \in \mathbb{C}^{2\times\alpha}, \quad \Psi \in \mathbb{C}^{(\Delta\delta_p + \Delta\delta_q)\times\alpha} \text{ and } \quad (\hat{\mathcal{B}}_x - x\hat{\mathcal{C}}_x) \in \mathbb{C}^{(\alpha - \hat{\delta} - 1)\times\alpha}, \forall x \in \mathbb{C}$$

where $\alpha = \frac{(\hat{\delta}+1)(\hat{\delta}+2)}{2}$ is the number of monomials in 2 variables of degree $\leq \hat{\delta}$. $\hat{\Phi}$ accounts for the original equations $p(x,y) = 0$ and $q(x,y) = 0$, Ψ accounts for all the shifted equations in the extended system and $(\hat{\mathcal{B}}_x - x\hat{\mathcal{C}}_x)$ imposes a partial

Vandermonde structure on the eigenvector $\boldsymbol{v}$. The left out equations in y are given by

$$(\hat{\mathcal{B}}_y - y\hat{\mathcal{C}}_y)\boldsymbol{v} = \mathbf{0}$$

where $(\hat{\mathcal{B}}_y - y\hat{\mathcal{C}}_y)\boldsymbol{v} \in \mathbb{C}^{\hat{\delta}\times\alpha}$. Together, $(\hat{\mathcal{B}}_x - x\hat{\mathcal{C}}_x)$ and $(\hat{\mathcal{B}}_y - y\hat{\mathcal{C}}_y)$ impose a complete Vandermonde structure of degree $\hat{\delta}$ on $\boldsymbol{v}$.

Using some shift degrees $\Delta\delta_p \geq 0$ and $\Delta\delta_q \geq 0$, we define the resulting extended pencil $\hat{L}(x, y)$ related to the original pencil $L(x, y)$ used in (3.9) as

$$\hat{L}(x,y) \triangleq \left(\begin{array}{c} \hat{\Phi} \\ \Psi \\ (\hat{\mathcal{B}}_x - x\hat{\mathcal{C}}_x) \\ \hline (\hat{\mathcal{B}}_y - y\hat{\mathcal{C}}_y) \end{array}\right) \triangleq \left(\begin{array}{c} \hat{\Pi}_x(x) \\ \\ \hline \hat{\Pi}_y(y) \end{array}\right). \tag{3.12}$$

We will use the notation

$$\begin{cases} p(x,y) = 0 \\ q(x,y) = 0 \end{cases} \quad \xrightarrow[C]{} \quad L(x,y) \quad \xrightarrow[E]{} \quad \hat{L}(x,y)$$

where E stands for extension. Note from Example 3.2.1 that $\hat{L}(x, y)$ has $\hat{\delta} + 1$ block columns consisting of a decreasing number of columns: the first block column has $\hat{\delta} + 1$ columns, the last one only 1. In analogy with Theorem (3.1.1), we have the following more general result.

Theorem 3.2.1. *Let*

$$\begin{cases} p(x,y) = 0 \\ q(x,y) = 0 \end{cases} \quad \xrightarrow[C]{} \quad L(x,y) \quad \xrightarrow[E]{} \quad \hat{L}(x,y).$$

For any $\Delta\delta_p \geq 0$, any $\Delta\delta_q \geq 0$, a pair $(x^*, y^*) \in \mathbb{C}^2$ is a solution to problem (3.1) *if and only if $\exists \boldsymbol{v} \neq \mathbf{0}$ s.t. $\hat{L}(x^*, y^*)\boldsymbol{v} = \mathbf{0}$. The solutions to* (3.1) *are the couples (x^*, y^*) for which $\hat{L}(x, y)$ loses full column rank.*

Proof. The proof is completely analogous to that of Theorem (3.1.1) using a monomial vector of degree $\hat{\delta}$ instead of δ and using (3.10). □

3.3 Reducing the pencil size

It is not necessary to keep a column in $\hat{L}(x, y)$ for every monomial of degree $\hat{\delta}$ in case $\hat{\delta} > \delta$. It might happen that neither p and q, nor their shifted versions in the extended system have a nonzero coefficient corresponding to some monomial $x^i y^j$ of degree $\leq \hat{\delta}$. For example, consider the system

$$\begin{cases} p(x,y) = -1 + x^2 = 0 \\ q(x,y) = -1 + y^2 = 0 \end{cases} \quad \xrightarrow[C]{} \quad L(x,y) \quad \xrightarrow[E]{} \quad \hat{L}(x,y)$$

where we use $\Delta\delta_p = \Delta\delta_q = 1$ for the extension step. It can be verified that

$$\hat{L}(x,y) = \left(\begin{array}{ccc|ccc|cc|c} -1 & 0 & 1 & 0 & 0 & & 0 & & \\ -1 & 0 & 0 & 0 & 0 & & 1 & & \\ \hline & & & -1 & 0 & 1 & 0 & 0 & 0 \\ & & & -1 & 0 & 0 & 0 & 0 & 1 \\ \hline -x & 1 & & & & & & & \\ & -x & 1 & & & & & & \\ & & -x & 1 & & & & & \\ \hline & & & -x & 1 & & & & \\ & & & & -x & 1 & & & \\ \hline & & & & & & -x & 1 & \\ \hline -y & & & 1 & & & & & \\ & & & -y & & & 1 & & \\ & & & & & & -y & & 1 \end{array}\right).$$

Now, it is clear that the monomial x^3 (fourth column) does not belong to the support of p and q. It does not belong to the support of yp and yq either. Therefore, we might think of dropping the fourth column of $\hat{L}(x,y)$. However, this would interfere with our basis definition. Consider the first column block of $(\hat{\mathcal{B}}_x - x\hat{\mathcal{C}}_x)$, it defines the recurrence

$$1 \cdot x = x,\ x \cdot x = x^2,\ x^2 \cdot x = x^3.$$

Leaving out the fourth column of $\hat{L}(x,y)$, we are left with the recursion

$$1 \cdot x = x,\ x \cdot x = x^2,\ x^2 \cdot x = 0,$$

which is incorrect. We can, however, leave out the third row of $(\hat{\mathcal{B}}_x - x\hat{\mathcal{C}}_x)$ as well. This does not interfere with any of the other recurrence relations and the rows that remain in $(\hat{\mathcal{B}}_x - x\hat{\mathcal{C}}_x)$ and $(\hat{\mathcal{B}}_y - y\hat{\mathcal{C}}_y)$ define the Vandermonde structure of the *reduced* monomial basis (that is, leaving out x^3) perfectly. The same thing can be done for the ninth column. Leaving it out of the pencil forces us to drop the last row of $(\hat{\mathcal{B}}_x - x\hat{\mathcal{C}}_x)$ along with it. We cannot apply this procedure for every monomial that does not appear in the extended system. Consider for example the sixth column of $\hat{L}(x,y)$. Leaving it out would force us to drop both the fourth and the fifth row of $(\hat{\mathcal{B}}_x - x\hat{\mathcal{C}}_x)$. Moreover, to restore the recurrence, it would force us to add a quadratic equation in x:

$$y \cdot x^2 = x^2 y$$

which would destroy the linearity of the pencil. In general, the block columns of $(\hat{\mathcal{B}}_x - x\hat{\mathcal{C}}_x)$ can be thought of as recurrence chains that we can shorten from the right to reduce the pencil size if the supports of p and q allow us to. The chains cannot be shortened from the left, since this would interfere with the recurrences in $\hat{\Pi}_y(y)$, nor can they be interrupted somewhere in the middle. The y-recurrence chain in the $\hat{\Pi}_y(y)$ block row can be shortened from the right too if there are entire column blocks of $\hat{\Phi}$ and Ψ that are filled with zeros. It cannot be shortened from the left if we assume p and q to be coprime. Indeed, if the leftmost block column of $\hat{\Phi}$ is filled

with zeros, p and q share the common factor y. An algorithmic approach to perform the reduction can be summarized as follows.

Algorithm 1 (Pencil reduction)**.** *We divide the algorithm in two parts.*

1. y-reduction. *First, we remove all the block columns corresponding to monomials of a degree in y that is higher than the maximum degree in y attained by p, q and their shifted versions. That is, we shorten the y-recurrence chain as much as possible. Consider the last block column of $\hat{L}(x, y)$.*

   while *$\hat{\Phi}$ and Ψ are completely filled with zeros in the considered block column* ***do***
   Remove this block column.
   Remove the last row of $\hat{L}(x, y)$ (the last y-equation).
   if *This is not the last block column* ***then***
   Remove the last block row of $\hat{\mathcal{B}}_x - x\hat{\mathcal{C}}_x$ in $\hat{L}(x, y)$.
   end if
   Consider the next block column (the one to the left of this one).
   end while

2. x-reduction. *Next, within each block column, we shorten the chain of x-recurrences as much as possible.*

   for *all remaining block columns of $\hat{L}(x, y)$* ***do***
   Start by the rightmost column of this block column.
   while *$\hat{\Phi}$ and Ψ are completely filled with zeros in this column* ***do***
   Remove this column.
   Remove the row of $\hat{L}(x, y)$ in which this column of $\hat{\mathcal{B}}_x$ has a nonzero entry.
   Consider the column to the left of the deleted column.
   end while
   end for

We will refer to the reduced pencil as $\hat{L}_r(x, y)$ and we will denote

$$\begin{cases} p(x,y) = 0 \\ q(x,y) = 0 \end{cases} \xrightarrow[C]{} L(x,y) \xrightarrow[E]{} \hat{L}(x,y) \xrightarrow[R]{} \hat{L}_r(x,y)$$

where R stands for reduction. For the corresponding block rows, we denote

$$\begin{cases} p(x,y) = 0 \\ q(x,y) = 0 \end{cases} \xrightarrow[C]{} \begin{pmatrix} \Pi_x(x) \\ \hline \Pi_y(y) \end{pmatrix} \xrightarrow[E]{} \begin{pmatrix} \hat{\Pi}_x(x) \\ \hline \hat{\Pi}_y(y) \end{pmatrix} \xrightarrow[R]{} \begin{pmatrix} \hat{\Phi}_r \\ \Psi_r \\ \hat{\mathcal{B}}_{x,r} - x\hat{\mathcal{C}}_{x,r} \\ \hline \hat{\mathcal{B}}_{y,r} - y\hat{\mathcal{C}}_{y,r} \end{pmatrix} \triangleq \begin{pmatrix} \hat{\Pi}_{x,r}(x) \\ \hline \hat{\Pi}_{y,r}(y) \end{pmatrix}.$$

For the previous example, we have

$$\hat{L}(x,y) \quad \underset{R}{\longrightarrow} \quad \left(\begin{array}{ccc|ccc|cc} -1 & 0 & 1 & 0 & 0 & & 0 & \\ -1 & 0 & 0 & 0 & 0 & & 1 & \\ \hline & & & -1 & 0 & 1 & 0 & 0 \\ & & & -1 & 0 & 0 & 0 & 1 \\ \hline -x & 1 & & & & & & \\ & -x & 1 & & & & & \\ \hline & & & -x & 1 & & & \\ & & & & -x & 1 & & \\ \hline -y & & & 1 & & & & \\ & & & -y & & & 1 & \\ & & & & & & -y & 1 \end{array}\right). \tag{3.13}$$

Theorem 3.3.1. *Let*

$$\begin{cases} p(x,y)=0 \\ q(x,y)=0 \end{cases} \quad \underset{C}{\longrightarrow} \quad L(x,y) \quad \underset{E}{\longrightarrow} \quad \hat{L}(x,y) \quad \underset{R}{\longrightarrow} \quad \hat{L}_r(x,y).$$

For any $\Delta\delta_p \geq 0$, any $\Delta\delta_q \geq 0$, a pair $(x^*,y^*) \in \mathbb{C}^2$ is a solution of (3.1) *if and only if $\exists \boldsymbol{v} \neq \mathbf{0}$ s.t. $\hat{L}_r(x^*,y^*)\boldsymbol{v} = \mathbf{0}$. The solutions of* (3.1) *are the couples (x^*,y^*) for which $\hat{L}_r(x,y)$ loses full column rank.*

Proof. The proof is completely analogous to that of Theorem (3.1.1) using the appropriate reduced monomial vector of degree $\hat{\delta}$ and using (3.10). □

Definition 3.3. *A nonzero vector $\boldsymbol{v}$ that satisfies*

$$\begin{pmatrix} \hat{\mathcal{B}}_{x,r} - x\hat{\mathcal{C}}_{x,r} \\ \hat{\mathcal{B}}_{y,r} - y\hat{\mathcal{C}}_{y,r} \end{pmatrix} \boldsymbol{v} = \mathbf{0}$$

for some couple $(x,y) \in \mathbb{C}^2$ is said to have a Vandermonde structure *in the reduced monomial basis. A vector $\boldsymbol{v}$ that satisfies*

$$(\hat{\mathcal{B}}_{x,r} - x\hat{\mathcal{C}}_{x,r})\boldsymbol{v} = \mathbf{0}$$

for some $x \in \mathbb{C}$ is said to have a blockwise Vandermonde structure *in the reduced monomial basis.*

Example 3.3.1. *For the reduced pencil* (3.13)*, the vector*

$$\boldsymbol{v}_1 = \begin{pmatrix} 1 & -1 & 1 \mid 1 & -1 & 1 \mid 1 \mid 1 \end{pmatrix}^\top$$

has a Vandermonde structure: it is in the right null space of $\begin{pmatrix} \hat{\mathcal{B}}_{x,r} + \hat{\mathcal{C}}_{x,r} \\ \hat{\mathcal{B}}_{y,r} - \hat{\mathcal{C}}_{y,r} \end{pmatrix}$. The vector $\boldsymbol{v}_2 = \begin{pmatrix} 1 & -1 & 1 \mid 0 & 0 & 0 \mid 1 \mid 0 \end{pmatrix}^\top$ has a blockwise Vandermonde structure: $(\hat{\mathcal{B}}_{x,r} + \hat{\mathcal{C}}_{x,r})\boldsymbol{v}_2 = \mathbf{0}$. The reduced monomial basis is given by

$$\begin{pmatrix} 1 & x & x^2 \mid y & xy & x^2y \mid y^2 \mid y^3 \end{pmatrix}^\top.$$

Chapter 4

A Square One-Parameter GEP

The first aim of this chapter is to show that for appropriate choices of $\Delta\delta_p$ and $\Delta\delta_q$ in the extension step (E), a square $\hat{\Pi}_{x,r}(x)$ can always be obtained. Next, we will show that the eigenvalues of the square pencil $\hat{\Pi}_{x,r}(x)$ for the right $\Delta\delta_p$ and $\Delta\delta_q$ contain the x-values of the solutions to (3.1). In other words: $\det \hat{\Pi}_{x,r}(x)$ is a resultant. In fact, it turns out to be equivalent to that of Sylvester [29, 26, 2, 8, 7]. In the last section, we extend the results to a more general class of tensor product bases.

4.1 The right shift degrees

Theorem 4.1.1. *Constructing $\hat{L}_r(x,y)$ as explained in Chapter 3, the pencil $\hat{\Pi}_{x,r}(x)$ is square if in the extension step the shift degrees $\Delta\delta_p = \deg_y(q(x,y)) - 1 = \delta_q^y - 1$ and $\Delta\delta_q = \deg_y(p(x,y)) - 1 = \delta_p^y - 1$ are used.*

Proof. In case $\Delta\delta_p = \delta_q^y - 1$ and $\Delta\delta_q = \delta_p^y - 1$, we have

$$\begin{aligned} \hat{\delta} &= \max(\delta_p + \Delta\delta_p, \delta_q + \Delta\delta_q) \\ &= \max(\delta_p + \delta_q^y - 1, \delta_q + \delta_p^y - 1). \end{aligned}$$

We will assume (without loss of generality) that $\delta_p + \delta_q^y - 1 \geq \delta_q + \delta_p^y - 1$, so $\hat{\delta} = \delta_p + \delta_q^y - 1$. Consider the extended (but not reduced) pencil $\hat{L}(x,y)$ first. Denote the number of rows and columns of $\hat{\Pi}_x(x)$ by $\hat{m}$ and $\hat{n}$ respectively and let α be the number of monomials in two variables of degree $\leq \hat{\delta}$. We have by construction

$$\hat{n} = \alpha = \sum_{i=1}^{\hat{\delta}+1} i = \frac{(\hat{\delta}+1)(\hat{\delta}+2)}{2},$$

$$\hat{m} = \underbrace{2}_{\hat{\Phi}} + \underbrace{\Delta\delta_p + \Delta\delta_q}_{\Psi} + \underbrace{\alpha - (\hat{\delta}+1)}_{\hat{\mathcal{B}}_x - x\hat{\mathcal{C}}_x}.$$

For $\Delta\delta_p = \delta_q^y - 1$ and $\Delta\delta_q = \delta_p^y - 1$, using $\hat{\delta} = \delta_p + \delta_q^y - 1$ we find

$$\hat{m} = \delta_p^y + \alpha - \delta_p.$$

During the reduction step some of the rows and columns of $\hat{\Pi}_x(x)$ will be deleted. We denote

$$\hat{m}, \hat{n} \xrightarrow[R]{} \hat{m}_r, \hat{n}_r.$$

In the y-reduction of Algorithm 1, the rightmost zero block columns of $\hat{\Phi}$ and Ψ are removed from $\hat{\Pi}_x(x)$ along with the corresponding rows. After the extension using $\Delta\delta_p = \delta_q^y - 1$ and $\Delta\delta_q = \delta_p^y - 1$, the highest degree in y that is reached by the shifts of p and q is equal to $\delta_p^y + \delta_q^y - 1$. Therefore, the number of block columns that can be dropped in the y-reduction step is equal to $\hat{\delta} - \delta_p^y - \delta_q^y + 1 = \delta_p - \delta_p^y$. The total number of removed columns will be denoted by γ_n. It is found as

$$\gamma_n = \sum_{i=1}^{\delta_p - \delta_p^y} i = \frac{(\delta_p - \delta_p^y)(\delta_p - \delta_p^y + 1)}{2}, \quad \delta_p \geq \delta_p^y.$$

Performing the y-reduction, we only start removing block rows of $\hat{\mathcal{B}}_x - x\hat{\mathcal{C}}_x$ from the second deleted block column on. The total number of deleted rows in the y-part of the pencil reduction is denoted by γ_m and it is found as

$$\gamma_m = \sum_{i=1}^{\delta_p - \delta_p^y - 1} i = \frac{(\delta_p - \delta_p^y - 1)(\delta_p - \delta_p^y)}{2} = \gamma_n - (\delta_p - \delta_p^y), \quad \delta_p \geq \delta_p^y.$$

In case $\delta_p = \delta_p^y$ no columns or rows are removed in the first part of the reduction algorithm and $\hat{m} = \hat{n} = \alpha$. If $\delta_p = \delta_p^y + 1$, we find $(\hat{m} - \gamma_m) - (\hat{n} - \gamma_n) = (\alpha - 1) - (\alpha - 1) = 0$. When $\delta_p > \delta_p^y + 1$, we have

$$(\hat{m} - \gamma_m) - (\hat{n} - \gamma_n) = (\delta_p^y + \alpha - \delta_p - (\gamma_n - (\delta_p - \delta_p^y))) - (\alpha - \gamma_n) = 0$$

This means that in every possible case, after the first part of Algorithm 1, the number of rows in the x-pencil is equal to the number of columns. In the x-part of the algorithm, the x-reduction, each time a column is removed a corresponding row is removed along with it. Therefore, some number s of columns and rows is removed in addition to γ_n and γ_m and we have

$$\hat{m}_r - \hat{n}_r = (\hat{m} - \gamma_m - s) - (\hat{n} - \gamma_n - s) = 0,$$

which proves the theorem. □

Corollary 4.1.1. *The size of the resulting square pencil $\hat{\Pi}_{x,r}(x)$ is bounded by $2\delta^2 + \delta$.*

Proof. The size of the pencil is given by $\alpha - \gamma_n - s$. For the extended degree $\hat{\delta}$ we have

$$\hat{\delta} = \max(\delta_p + \delta_q^y - 1, \delta_q + \delta_p^y - 1) \leq \delta_p + \delta_q - 1.$$

Therefore

$$\alpha \leq \frac{(\delta_p + \delta_q)(\delta_p + \delta_q + 1)}{2}.$$

Because $\gamma_n \geq 0$ and $s \geq 0$ we find

$$\alpha - \gamma_n - s \leq \alpha \leq \frac{(\delta_p + \delta_q)(\delta_p + \delta_q + 1)}{2} \leq 2\delta^2 + \delta.$$

□

In fact, Corollary 4.1.1 is a very pessimistic bound in many cases. For the interested reader, a more realistic bound is derived in Appendix C.

Example 4.1.1. *In Example 3.2.1, it can be seen that for the considered system of degree $\delta = 2$, using $\Delta\delta_p = \delta_q^y - 1 = 1$ and $\Delta\delta_q = \delta_p^y - 1 = 1$ we find that*

$$\hat{\Pi}_{x,r}(x) = \left(\begin{array}{ccc|ccc|cc|c} -4 & 0 & 1 & 0 & 0 & & 1 & & \\ -3 & -2 & 1 & 0 & 1 & & 1 & & \\ \hline & & & -4 & 0 & 1 & 0 & 0 & 1 \\ & & & -3 & -2 & 1 & 0 & 1 & 1 \\ \hline -x & 1 & & & & & & & \\ & -x & 1 & & & & & & \\ \hline & & & -x & 1 & & & & \\ & & & & -x & 1 & & & \\ \hline & & & & & & -x & 1 & \end{array}\right)$$

is square and of size 9. The upper bound from Corollary 4.1.1 is equal to 10.

4.2 The eigenvalues of $\hat{\Pi}_{x,r}(x)$

Theorem 4.2.1. *Let $\hat{\Pi}_{x,r}(x)$ be the x-pencil associated to the system*

$$\begin{cases} p(x,y) = 0 \\ q(x,y) = 0 \end{cases} \xrightarrow[C]{} L(x,y) \xrightarrow[E]{} \hat{L}(x,y) \xrightarrow[R]{} \hat{L}_r(x,y)$$

with $p, q \in \mathcal{P}_\delta^2$ and where $\Delta\delta_p = \delta_q^y - 1$ and $\Delta\delta_q = \delta_p^y - 1$ is used in the degree extension step E, then $\hat{\Pi}_{x,r}(x)$ is square and

$$\det \hat{\Pi}_{x,r}(x) = \gamma \mathrm{res}^{p,q}(x)$$

where $\gamma \in \{-1, 1\}$.

Proof. We know from Theorem 4.1.1 that $\hat{\Pi}_{x,r}(x)$ is square. Let $\boldsymbol{m}_\psi^x$ be defined as the vector of univariate x-monomials of increasing degree up to $x^{\psi-1}$:

$$\boldsymbol{m}_\psi^x \triangleq \begin{pmatrix} 1 & x & x^2 & \dots & x^{\psi-1} \end{pmatrix}^\top.$$

Let the matrix $\mathcal{M}$ be constructed as follows:

$$\mathcal{M} \triangleq \left(\begin{array}{ccc|ccc} \boldsymbol{m}_{\psi_1}^x & & & \overline{I}_{\psi_1-1}^\top & & \\ & \boldsymbol{m}_{\psi_2}^x & & & \overline{I}_{\psi_2-1}^\top & \\ & & \ddots & & & \ddots \\ & & \boldsymbol{m}_{\psi_r}^x & & & \overline{I}_{\psi_r-1}^\top \end{array}\right) \triangleq \begin{pmatrix} M_x \mid M_1 \end{pmatrix} \quad (4.1)$$

with $r \triangleq \delta_p^y + \delta_q^y$, $\overline{I}_i$ as defined in (3.4) ($\overline{I}_0$ is an empty matrix) and the numbers ψ_i represent the number of columns of the i-th block column of $\hat{\Pi}_{x,r}(x)$. In words: for the i-th block column of $\hat{\Pi}_{x,r}(x)$ we add a column to M_x containing the univariate monomial vector in the i-th block row and we add a block column to M_1 consisting of $\psi_i - 1$ columns of the identity matrix. For instance, for $\hat{\Pi}_{x,r}(x)$ from Example 4.1.1, $\mathcal{M}$ is given by

$$\mathcal{M} = \left(\begin{array}{ccc|ccccc} 1 & & & 0 & & & & \\ x & & & 1 & & & & \\ x^2 & & & & 1 & & & \\ & 1 & & & & 0 & & \\ & x & & & & 1 & & \\ & x^2 & & & & & 1 & \\ & & 1 & & & & & 0 \\ & & x & & & & & 1 \\ & & & 1 & & & & 0 \end{array}\right).$$

Now, consider the matrix product

$$\hat{\Pi}_{x,r}(x)\mathcal{M} = \left(\begin{array}{c|c} \begin{pmatrix}\hat{\Phi}\\ \Psi\end{pmatrix} M_x & \begin{pmatrix}\hat{\Phi}\\ \Psi\end{pmatrix} M_1 \\ \hline (\hat{\mathcal{B}}_{x,r} - x\hat{\mathcal{C}}_{x,r})M_x & (\hat{\mathcal{B}}_{x,r} - x\hat{\mathcal{C}}_{x,r})M_1 \end{array}\right). \tag{4.2}$$

Note that every column of M_x has the blockwise Vandermonde structure imposed by $(\hat{\mathcal{B}}_{x,r} - x\hat{\mathcal{C}}_{x,r})$, so each column of M_x lies in the null space of $(\hat{\mathcal{B}}_{x,r} - x\hat{\mathcal{C}}_{x,r})$ for all values of x and the lower left block of $\hat{\Pi}_{x,r}(x)\mathcal{M}$ is filled with zeros. The matrix M_1 can be thought of as a column selector. It selects columns from $\hat{\Pi}_{x,r}(x)\mathcal{M}$ that correspond to monomials $m_i(x,y)$ with $\deg_x(m_i(x,y)) > 0$ (it skips the first column of every block column of $\hat{\Pi}_{x,r}(x)$). By construction, $\hat{\mathcal{B}}_{x,r}M_1$ is the identity matrix and

$$\hat{\mathcal{C}}_{x,r}M_1 = \begin{pmatrix} \Delta_{\psi_1-1} & & & \\ & \Delta_{\psi_2-1} & & \\ & & \ddots & \\ & & & \Delta_{\psi_r-1} \end{pmatrix}$$

where Δ_ψ is an $\psi \times \psi$ matrix with ones on the first subdiagonal ($\Delta_1 = 0$ and Δ_0 is the empty matrix). Therefore $(\hat{\mathcal{B}}_{x,r} - x\hat{\mathcal{C}}_{x,r})M_1$ is a lower triangular square matrix with a diagonal full of ones $\forall x \in \mathbb{C}$, which implies $\det(\hat{\mathcal{B}}_{x,r} - x\hat{\mathcal{C}}_{x,r})M_1 = 1$. As an illustration, consider again the pencil in Example 4.1.1, where we have

$$\hat{\mathcal{C}}_{x,r}M_1 = \begin{pmatrix} 0 & & & & \\ 1 & 0 & & & \\ & & 0 & & \\ & & 1 & 0 & \\ & & & & 0 \end{pmatrix}, \qquad (\hat{\mathcal{B}}_{x,r} - x\hat{\mathcal{C}}_{x,r})M_1 = \begin{pmatrix} 1 & & & & \\ -x & 1 & & & \\ & & 1 & & \\ & & -x & 1 & \\ & & & & 1 \end{pmatrix}.$$

The upper left block of $\hat{\Pi}_{x,r}(x)\mathcal{M}$ is given by

$$\begin{pmatrix}\hat{\Phi}\\ \Psi\end{pmatrix} M_x = \begin{pmatrix}
p_0^x(x) & p_1^x(x) & \dots & p_{\delta_p^y}^x(x) & & & & \\
q_0^x(x) & q_1^x(x) & \dots & q_{\delta_p^y}^x(x) & \dots & q_{\delta_q^y}^x(x) & & \\
 & p_0^x(x) & p_1^x(x) & \dots & p_{\delta_p^y}^x(x) & & & \\
 & & p_0^x(x) & p_1^x(x) & \dots & p_{\delta_p^y}^x(x) & & \\
 & & & p_0^x(x) & p_1^x(x) & \dots & p_{\delta_p^y}^x(x) & \\
 & & & & \ddots & & \ddots & \\
 & & & & p_0^x(x) & p_1^x(x) & \dots & p_{\delta_p^y}^x(x) \\
 & q_0^x(x) & q_1^x(x) & \dots & q_{\delta_p^y}^x(x) & \dots & q_{\delta_q^y}^x(x) & \\
 & & & \ddots & & & \ddots & \\
 & & q_0^x(x) & q_1^x(x) & \dots & q_{\delta_p^y}^x(x) & \dots & q_{\delta_q^y}^x(x)
\end{pmatrix}$$

where we have assumed (merely for the visualization of the matrix) that $\delta_q^y > \delta_p^y$. Applying the appropriate row permutation R_p we get

$$R_p \begin{pmatrix}\hat{\Phi}\\ \Psi\end{pmatrix} M_x = \left(\begin{array}{cccccccc}
p_0^x(x) & p_1^x(x) & \dots & p_{\delta_p^y}^x(x) & & & & \\
 & p_0^x(x) & p_1^x(x) & \dots & p_{\delta_p^y}^x(x) & & & \\
 & & p_0^x(x) & p_1^x(x) & \dots & p_{\delta_p^y}^x(x) & & \\
 & & & p_0^x(x) & p_1^x(x) & \dots & p_{\delta_p^y}^x(x) & \\
 & & & & \ddots & & \ddots & \\
 & & & & & p_0^x(x) & p_1^x(x) & \dots & p_{\delta_p^y}^x(x) \\
\hline
q_0^x(x) & q_1^x(x) & \dots & q_{\delta_p^y}^x(x) & \dots & q_{\delta_q^y}^x(x) & & & \\
 & q_0^x(x) & q_1^x(x) & \dots & q_{\delta_p^y}^x(x) & \dots & q_{\delta_q^y}^x(x) & & \\
 & & \ddots & & & & & \ddots & \\
 & & & q_0^x(x) & q_1^x(x) & \dots & q_{\delta_p^y}^x(x) & \dots & q_{\delta_q^y}^x(x)
\end{array}\right).$$

From these considerations and from (4.2) we have

$$\det \hat{\Pi}_{x,r}(x)\mathcal{M} = \det\left(\begin{pmatrix}\hat{\Phi}\\ \Psi\end{pmatrix} M_x\right) \det((\hat{\mathcal{B}}_{x,r} - x\hat{\mathcal{C}}_{x,r})M_1).$$

Note that the columns of $\mathcal{M}$ can be permuted, say by C_p, to obtain a lower triangular matrix with a diagonal full of ones:

$$\det \mathcal{M}C_p = 1 \quad \forall x \in \mathbb{C}.$$

Since $\det C_p$ can only take on the values 1 and -1, we have that $\det \hat{\Pi}_{x,r}(x)\mathcal{M} = \det \hat{\Pi}_{x,r}(x) \det \mathcal{M} = \det C_p \det \hat{\Pi}_{x,r}(x)$, so $\det \hat{\Pi}_{x,r}(x)\mathcal{M}$ is equal to $\det \hat{\Pi}_{x,r}(x)$ up

to a sign. As we have shown, $\det(\hat{\mathcal{B}}_{x,r} - x\hat{\mathcal{C}}_{x,r})M_1 = 1$ and using (2.6) we get

$$\begin{aligned}\det\left(\begin{pmatrix}\hat{\Phi}\\ \Psi\end{pmatrix}M_x\right) &= \det\left(R_p^{-1}R_p\begin{pmatrix}\hat{\Phi}\\ \Psi\end{pmatrix}M_x\right)\\ &= \det R_p \det\left(R_p\begin{pmatrix}\hat{\Phi}\\ \Psi\end{pmatrix}M_x\right)\\ &= \det R_p \det S^{p,q}(x).\end{aligned}$$

Therefore

$$\det\hat{\Pi}_{x,r}(x) = \frac{\det R_p}{\det C_p}\mathrm{res}^{p,q}(x).$$

□

Corollary 4.2.1. *From Theorem 4.2.1 and Theorem 2.2.2, it follows that if p and q from* (3.1) *are coprime, then $\hat{\Pi}_{x,r}(x)$ is a regular pencil and its finite eigenvalues contain the x-coordinates of the solutions to* (3.1)*. Moreover, each eigenvalue x^* appears a number of times equal to the sum of the multiplicities of all roots of* (3.1) *that have x^* as x-component. More specifically:*

$$\{x \in \mathbb{C} \mid \det\hat{\Pi}_{x,r}(x) = 0\} = \mathcal{V}_{p,q}^{(x)} \cup \{x \in \mathbb{C} \mid p_{\delta_p^y}^x(x) = q_{\delta_q^y}^x(x) = 0\}.$$

4.3 A change of basis

So far we have always used the classical monomial basis to make the construction process of $\hat{L}_r(x,y)$ clear. We have started by constructing the coefficient matrix Φ that contains the coordinates of p and q in the classical monomial basis. The aim of this section is to provide a more general framework for obtaining a linear pencil in which we can choose another basis to represent p and q in. First, we introduce the concept of a tensor product basis. In the next subsections we show how the construction, extension and reduction step can be generalized to a general class of tensor product bases. Finally, we show that the obtained x-pencil is equivalent to $\hat{\Pi}_{x,r}(x)$, in the sense that it has the same eigenvalues.

4.3.1 A tensor product basis

Consider the basis $B_x \triangleq \{b_0^x(x), b_1^x(x), \ldots, b_\delta^x(x)\}$ for $\mathcal{P}_\delta^1$ where $\deg(b_i^x(x)) = i, \forall i$. We will take $b_0^x(x) = 1$. Any such basis satisfies a recurrence relation:

$$b_k^x(x) = \alpha_k^x x b_{k-1}^x(x) + \sum_{i=0}^{k-1}\beta_{k,i}^x b_i^x(x), \qquad 1 \le k \le \delta \tag{4.3}$$

where $\beta_{k,i}^x \in \mathbb{C}, 0 \le i \le k \le \delta$ and $\alpha_k^x \in \mathbb{C}_0, 1 \le k \le \delta$. Analogously, we consider another basis $B_y \triangleq \{b_0^y(y), b_1^y(y), \ldots, b_\delta^y(y)\}$, with $\deg(b_i^y(y)) = i, \forall i$ and

$$b_k^y(y) = \alpha_k^y y b_{k-1}^y(y) + \sum_{i=0}^{k-1}\beta_{k,i}^y b_i^y(y), \qquad 1 \le k \le \delta \tag{4.4}$$

where $\beta_{k,i}^y \in \mathbb{C}, 0 \leq i \leq k \leq \delta$ and $\alpha_k^y \in \mathbb{C}_0, 1 \leq k \leq \delta$. Now, consider the tensor product basis

$$B \triangleq B_x \otimes B_y \triangleq \{b_{ij}(x,y)\}_{0 \leq i,j \leq \delta}$$

for $\mathcal{P}_\delta^2$ where

$$b_{ij}(x,y) \triangleq b_j^x(x) b_i^y(y).$$

We will use the notation $\{\cdot\}_B$ to indicate that we are working in a basis B different from the classical monomial basis.

4.3.2 Coefficient matrix and basis definition

Suppose two polynomials $p, q \in \mathcal{P}_\delta^2$ are given with their coordinates in B:

$$p(x,y) = \sum_{i=0}^{\delta} \sum_{j=0}^{\delta-i} p_{ij} b_j^x(x) b_i^y(y), \qquad q(x,y) = \sum_{i=0}^{\delta} \sum_{j=0}^{\delta-i} q_{ij} b_j^x(x) b_i^y(y).$$

For this basis, we define the Vandermonde vector $\{\boldsymbol{v}\}_B(x,y)$ as

$$\Big(b_{00}(x,y) \quad b_{01}(x,y) \quad \dots \quad b_{0\delta}(x,y) \;\big|\; \dots \;\big|\; b_{\delta-1,0}(x,y) \quad b_{\delta-1,1}(x,y) \;\big|\; b_{\delta 0}(x,y)\Big)^\top.$$

Again we have for every solution x^*, y^* of

$$\begin{cases} p(x,y) = 0 \\ q(x,y) = 0 \end{cases}$$

that $\{\Phi\}_B \{\boldsymbol{v}\}_B(x^*, y^*) = \mathbf{0}$ with

$$\{\Phi\}_B \triangleq \left(\begin{array}{ccccc|cccc|c|cc|c} p_{00} & p_{01} & p_{02} & \cdots & p_{0\delta} & p_{10} & p_{11} & \cdots & p_{1,\delta-1} & \cdots & p_{\delta-1,0} & p_{\delta-1,1} & p_{\delta 0} \\ q_{00} & q_{01} & q_{02} & \cdots & q_{0\delta} & q_{10} & q_{11} & \cdots & q_{1,\delta-1} & \cdots & q_{\delta-1,0} & q_{\delta-1,1} & q_{\delta 0} \end{array}\right).$$

For the basis definition blocks $\{\mathcal{B}_x - x\mathcal{C}_x\}_B$ and $\{\mathcal{B}_y - y\mathcal{C}_y\}_B$ we make use of the recurrence relations (4.3) and (4.4). We denote

$$\begin{cases} p(x,y) = 0 \\ q(x,y) = 0 \end{cases} \xrightarrow[C]{B} \{L(x,y)\}_B = \begin{pmatrix} \{\Phi\}_B \\ \{\mathcal{B}_x - x\mathcal{C}_x\}_B \\ \{\mathcal{B}_y - y\mathcal{C}_y\}_B \end{pmatrix}.$$

Let $\gamma_i^x \triangleq \alpha_i^x x + \beta_{i,i-1}^x$ and $\gamma_i^y \triangleq \alpha_i^y y + \beta_{i,i-1}^y$. The matrix

$$\begin{pmatrix} \{\mathcal{B}_x - x\mathcal{C}_x\}_B \\ \{\mathcal{B}_y - y\mathcal{C}_y\}_B \end{pmatrix}$$

is given by

$$\left(\begin{array}{cccc|ccc|c|c|c}
-\gamma_1^x & 1 & & & & & & & & \\
-\beta_{20}^x & -\gamma_2^x & 1 & & & & & & & \\
\vdots & & \ddots & & & & & & & \\
-\beta_{\delta 0}^x & -\beta_{\delta 1}^x & \dots & -\gamma_\delta^x \; 1 & & & & & & \\
\hline
 & & & & -\gamma_1^x & 1 & & & & \\
 & & & & \vdots & \ddots & & & & \\
 & & & & -\beta_{\delta-1,0}^x & \dots & -\gamma_{\delta-1}^x \; 1 & & & \\
\hline
 & & & & & & & \ddots & & \\
\hline
 & & & & & & & & -\gamma_1^x \; 1 & \\
\hline
-\gamma_1^y & & & & 1 & & & & & \\
-\beta_{20}^y & & & & -\gamma_2^y & & & & & \\
 & & & & & & & \ddots & & \\
-\beta_{\delta 0}^y & & & & -\beta_{\delta 1}^y & & & \dots & -\gamma_\delta^y & 1
\end{array}\right).$$

4.3.3 Degree extension

For the degree extension we extend the bases B_x and B_y to $\hat{B}_x$ and $\hat{B}_y$ respectively. $\hat{B}_x$ and $\hat{B}_y$ are bases for $\mathcal{P}^1_{\hat{\delta}}$ and we assume that they satisfy a recurrence relation similar to (4.3) and (4.4). we cannot just shift the coefficients in Φ to the next block row like in the classical monomial basis. Indeed, in general $b_k^y(y) \neq b_1^y(y) b_{k-1}^y(y)$ for $k \neq 1$. We have to calculate the coordinates of every shifted polynomial in the basis $\hat{B} \triangleq \hat{B}_x \otimes \hat{B}_y$ before we can plug the corresponding row into the extended pencil. To obtain more equations in x, we extend the system in the following way:

$$\begin{cases} p(x,y) = 0 \\ q(x,y) = 0 \end{cases} \rightarrow \begin{cases} p(x,y) = 0 \\ q(x,y) = 0 \\ s_1^p(y) p(x,y) = 0 \\ \quad \vdots \\ s_{\Delta\delta_p}^p(y) p(x,y) = 0 \\ s_1^q(y) q(x,y) = 0 \\ \quad \vdots \\ s_{\Delta\delta_q}^q(y) q(x,y) = 0 \end{cases}$$

with $\Delta\delta_p = \delta_q^y - 1$, $\Delta\delta_q = \delta_p^y - 1$ and $\deg(s_i^p) = \deg(s_i^q) = i$. The functions s_i^p and s_i^q will be referred to as the *shift functions*. We construct $\{\hat{L}(x,y)\}_{\hat{B}}$ by extending the tensor product basis and adding a block $\{\Psi\}_{\hat{B}}$ to $\{L(x,y)\}_B$, like we did for the classical monomial basis. The pencil reduction algorithm is completely analogous to Algorithm 1. We will denote the x-pencil obtained by constructing $\{\hat{L}_r(x,y)\}_{\hat{B}}$ in the basis $\hat{B}$ by $\{\hat{\Pi}_{x,r}(x)\}_{\hat{B}}$.

4.3.4 The eigenvalues of $\{\hat{\Pi}_{x,r}(x)\}_{\hat{B}}$

In the following we will use the subscript $\cdot_{\tilde{r}}$ to denote the matrices of the reduced linear pencil $\hat{L}_r(x,y)$ before the x-reduction has been performed. The number of columns in the k-th block column of $\{\hat{L}_{\tilde{r}}(x,y)\}_{\hat{B}}$ is equal to $\hat{\delta}+2-k$ for any appropriate basis $\hat{B}$ and we have that

$$\det\{\hat{\Pi}_{x,\tilde{r}}(x)\}_{\hat{B}} = \gamma \det\{\hat{\Pi}_{x,r}(x)\}_{\hat{B}} \tag{4.5}$$

with $\gamma \in \{-1,1\}$. This can easily be seen by developing $\det\{\hat{\Pi}_{x,\tilde{r}}(x)\}_{\hat{B}}$ with respect to any column that is removed during the x-reduction.

Theorem 4.3.1. *Let B_x and B_y be two bases for $\mathcal{P}^1_\delta$ that satisfy* (4.3) *and* (4.4). *Constructing $\{\hat{\Pi}_{x,r}(x)\}_{\hat{B}}$ as explained above, we have that*

$$\det\{\hat{\Pi}_{x,r}(x)\}_{\hat{B}} = C \det \hat{\Pi}_{x,r}(x)$$

where $\hat{\Pi}_{x,r}(x)$ is the pencil that is obtained by constructing $\hat{L}_r(x,y)$ in the classical monomial basis and C is a nonzero constant.

Proof. Consider the isomorphism π that maps a row vector to the corresponding polynomial in the reduced monomial basis:

$$\pi : \mathbb{C}^{|\mathcal{T}^2_{\hat{\delta},\tilde{r}}|} \to \operatorname{span}(\mathcal{T}^2_{\hat{\delta},\tilde{r}}) \subset \mathcal{P}^2_{\hat{\delta}}$$

where $\mathcal{T}^2_{\hat{\delta},\tilde{r}}$ represents the set of monomials that is contained in the reduced monomial vector after the y-reduction step and $|\cdot|$ denotes the cardinality of a set. Analogously, we define $\pi_{\hat{B}}$ as the map of a row vector to the corresponding polynomial in the basis $\hat{B}$. For example, consider the reduced monomial basis vector

$$\begin{pmatrix} 1 & x & x^2 & x^3 \mid y & xy & x^2y \mid y^2 & xy^2 \end{pmatrix}$$

and the basis $\hat{B} = \{1, 1+x, (1+x)^2, (1+x)^3, y, (1+x)y, (1+x)^2y, y^2, (1+x)y^2\}$. Then we have for

$$\boldsymbol{v} = \begin{pmatrix} -1 & 0 & 1 & 0 \mid 0 & 1 & 0 \mid 1 & 0 \end{pmatrix}$$

that $\pi(\boldsymbol{v}) = -1 + x^2 + xy + y^2$ and $\pi_{\hat{B}}(\boldsymbol{v}) = -1 + (1+x)^2 + (1+x)y + y^2$. Then, let T be the invertible matrix of the linear map that realizes the change of basis from the classical monomial basis to $\hat{B}$:

$$\pi(\boldsymbol{v}T) = \pi_{\hat{B}}(\boldsymbol{v}). \tag{4.6}$$

We extend the domain of π and $\pi_{\hat{B}}$ to matrices by defining the image of a matrix as the vector of images of the rows of that matrix:

$$\pi(M) = \pi\left(\begin{pmatrix} \boldsymbol{m}_1^\top \\ \vdots \\ \boldsymbol{m}_m^\top \end{pmatrix}\right) \triangleq \begin{pmatrix} \pi(\boldsymbol{m}_1^\top) \\ \vdots \\ \pi(\boldsymbol{m}_m^\top) \end{pmatrix}$$

where $M \in \mathbb{C}^{m\times n}$ and $\boldsymbol{m}_i^\top$ denotes the i-th row of M. It is clear that

$$\pi_{\hat{B}}(\{\hat{\Phi}_{\tilde{r}}\}_{\hat{B}}) = \begin{pmatrix} p(x,y) \\ q(x,y) \end{pmatrix} = \pi(\hat{\Phi}_{\tilde{r}})$$

and

$$\begin{aligned} \pi_{\hat{B}}\left(\begin{pmatrix} \{\hat{\Phi}_{\tilde{r}}\}_{\hat{B}} \\ \{\Psi_{\tilde{r}}\}_{\hat{B}} \end{pmatrix}\right) = \begin{pmatrix} p(x,y) \\ q(x,y) \\ s_1^p(y)p(x,y) \\ \vdots \\ s^p_{\Delta\delta_p}(y)p(x,y) \\ s_1^q(y)q(x,y) \\ \vdots \\ s^q_{\Delta\delta_q}(y)q(x,y) \end{pmatrix} &= L_{\phi,\psi} \begin{pmatrix} p(x,y) \\ q(x,y) \\ yp(x,y) \\ \vdots \\ y^{\Delta\delta_p}p(x,y) \\ yq(x,y) \\ \vdots \\ y^{\Delta\delta_q}q(x,y) \end{pmatrix} \\ &= L_{\phi,\psi}\pi\left(\begin{pmatrix} \hat{\Phi}_{\tilde{r}} \\ \Psi_{\tilde{r}} \end{pmatrix}\right) \end{aligned}$$

with $L_{\phi,\psi}$ a regular lower triangular matrix. Now, let $\{\boldsymbol{r}_{ki}\}_{\hat{B}}$ be the i-th row of the $(k+1)$-st block row of $\{\mathcal{B}_{x,\tilde{r}} - x^*\mathcal{C}_{x,\tilde{r}}\}_{\hat{B}}$ for some $x^* \in \mathbb{C}$ (the block row corresponding to degree k in y). Then

$$\pi_{\hat{B}}(\{\boldsymbol{r}_{ki}\}_{\hat{B}}) = b_k^y(y)\left(-\sum_{j=0}^{i-1}\beta_{ij}^x b_j^x(x) - \alpha_i^x x^* b_{i-1}^x(x) + b_i^x(x)\right)$$

and using (4.3) we find that

$$\pi_{\hat{B}}(\{\boldsymbol{r}_{ki}\}_{\hat{B}}) = \alpha_i^x(x-x^*)b_{i-1}^x(x)b_k^y(y), \qquad 1 \le i \le \hat{\delta} - k, 0 \le k < \delta_p^y + \delta_q^y.$$

Therefore

$$\pi_{\hat{B}}(\{\boldsymbol{r}_{ki}\}_{\hat{B}}) \in \operatorname{span}(\{(x-x^*)x^j y^l\}_{0\le j<i, 0\le l\le k}), \qquad 1 \le i \le \hat{\delta} - k, 0 \le k < \delta_p^y + \delta_q^y.$$

Defining $\boldsymbol{r}_{ki}$ as the i-th row of the $(k+1)$-st block row of $\hat{\mathcal{B}}_{x,\tilde{r}} - x^*\hat{\mathcal{C}}_{x,\tilde{r}}$ we have

$$\pi(\boldsymbol{r}_{ki}) = (x-x^*)x^{i-1}y^k, \qquad 1 \le i \le \hat{\delta} - k, 0 \le k < \delta_p^y + \delta_q^y.$$

We observe that

$$\operatorname{span}(\{(x-x^*)x^j y^l\}_{0\le j<i, 0\le l\le k}) = \operatorname{span}(\{\pi(\boldsymbol{r}_{lj})\}_{1\le j\le i, 0\le l\le k}),$$

for $1 \le i \le \hat{\delta} - k, 0 \le k < \delta_p^y + \delta_q^y$. This proves that any polynomial in $\pi_{\hat{B}}(\{\mathcal{B}_{x,\tilde{r}} - x^*\mathcal{C}_{x,\tilde{r}}\}_{\hat{B}})$ can be written as a linear combination of the corresponding polynomial in $\pi(\mathcal{B}_{x,\tilde{r}} - x^*\mathcal{C}_{x,\tilde{r}})$ (the one with the same position in the vector) and the previous polynomials in $\pi(\mathcal{B}_{x,\tilde{r}} - x^*\mathcal{C}_{x,\tilde{r}})$. Therefore

$$\pi_{\hat{B}}(\{\mathcal{B}_{x,\tilde{r}} - x^*\mathcal{C}_{x,\tilde{r}}\}_{\hat{B}}) = L_b\pi(\mathcal{B}_{x,\tilde{r}} - x^*\mathcal{C}_{x,\tilde{r}})$$

with L_b a regular lower triangular matrix that does not depend on x^*. Combining the results, we have that

$$\pi_{\hat{B}}(\{\hat{\Pi}_{x,\tilde{r}}(x^*)\}_{\hat{B}}) = \underbrace{\begin{pmatrix} L_{\phi,\psi} & \\ & L_b \end{pmatrix}}_{L} \pi(\hat{\Pi}_{x,\tilde{r}}(x^*)), \forall x^* \in \mathbb{C}$$

and thus, using (4.6), the linearity of π and applying π^{-1} to both sides:

$$\{\hat{\Pi}_{x,\tilde{r}}(x^*)\}_{\hat{B}} T = L\hat{\Pi}_{x,\tilde{r}}(x^*), \forall x^* \in \mathbb{C}.$$

Finally, this results in

$$\det\{\hat{\Pi}_{x,\tilde{r}}(x)\}_{\hat{B}} = \tilde{C} \det \hat{\Pi}_{x,\tilde{r}}(x)$$

where $\tilde{C} = \frac{\det L}{\det T} \neq 0$. Together with (4.5) this proves the theorem. □

For the interested reader, an example in the Chebyshev tensor product basis is worked out in Appendix D.

Chapter 5

Solving Bivariate Polynomial Systems

So far, we have proposed a way to find the x-coordinates (or y-coordinates) of the solutions of the bivariate system

$$\begin{cases} p(x,y) = 0 \\ q(x,y) = 0 \end{cases} . \tag{5.1}$$

Namely, they can be found among the roots of $\det \hat{\Pi}_{x,r}(x)$ (see Corollary 4.2.1). In this chapter, we will discuss the problem of finding the corresponding y-coordinates. This can be done in several ways. For each approach, we will discuss the theoretical results as well as the numerical aspects. We will use the following notation throughout this chapter. $\mathcal{X}$ represents the set of x-solutions, counting multiplicities, whereas X represents the set of distinct x-solutions: $X = \mathcal{V}_{p,q}^{(x)}$. Their numerical approximations will be denoted by $\tilde{\mathcal{X}}$ and $\tilde{X}$ respectively. We will use a similar notation for the y-solutions: $\mathcal{Y} \approx \tilde{\mathcal{Y}}$ for the y-values counting multiplicities and $Y \approx \tilde{Y}$ for the distinct y-values. The solution set of (5.1) counting multiplicities is denoted by $\mathcal{S}$, the set of distinct solutions is $S = \mathcal{V}_{p,q}$. The numerically found solutions will be denoted by $\tilde{\mathcal{S}}$ and $\tilde{S}$ respectively. Our aim is to find $\tilde{\mathcal{S}}$ rather than just $\tilde{S}$.
Suppose that we use the pencil $\det \hat{\Pi}_{x,r}(x)$ as a tool for finding $\tilde{\mathcal{X}}$. We do not calculate $\tilde{\mathcal{Y}}$ explicitly. This leads to a first class of versions of our method where our goal will be to associate one (and only one) y-value to each value in $\tilde{\mathcal{X}}$ so that the multiplicities of all couples are correct. Some ways to do this are explained in the first section of this chapter. Another approach is to calculate both $\tilde{\mathcal{X}}$ and $\tilde{\mathcal{Y}}$. The goal is then to find an appropriate coupling between $\tilde{\mathcal{X}}$ and $\tilde{\mathcal{Y}}$ to construct $\tilde{\mathcal{S}}$. In both cases, each solution in $\mathcal{S}$ has a numerical representative in $\tilde{\mathcal{S}}$. In the second section we propose two ways of realizing such a coupling. Finally, in a third section we propose a variable precision method.

Example 5.0.1. *Consider again the problem of Example 2.2.3 where*

$$\begin{cases} p(x,y) = x^2 + y^2 - 1 = 0 \\ q(x,y) = 4x^2 + y^2 + 6x + 2 = 0 \end{cases}$$

with a 4-fold zero at $(-1, 0)$*. For this problem*

$$\mathcal{X} = \{-1, -1, -1, -1\}, \quad X = \{-1\}, \quad \mathcal{Y} = \{0, 0, 0, 0\}, \quad Y = \{0\}.$$

Calculating the roots of $\det \hat{\Pi}_{x,r}(x)$ *using Matlab, we find*

$$\begin{aligned}\tilde{\mathcal{X}} &= \{-9.999999999999998 \times 10^{-1} + 1.484190404154479 \times 10^{-8} i,\\ &- 9.999999999999998 \times 10^{-1} - 1.484190404154479 \times 10^{-8} i,\\ &- 1.000000000000000 \times 10^{0} + 0.000000000000000 \times 10^{0} i,\\ &- 1.000000000000000 \times 10^{0} + 0.000000000000000 \times 10^{0} i\}.\end{aligned}$$

A possible (and in this case very good) numerical approximation $\tilde{X}$ *for* X *could be the mean of all values in* $\tilde{\mathcal{X}}$*:*

$$\tilde{X} = \{-9.999999999999999 \times 10^{-1} - 8.271806125530277 \times 10^{-25} i\}.$$

To evaluate the different versions of our method, we will make use of the *residual* of the numerically found solutions. The residual will be defined in the second section. We will assume that spurious eigenvalues of $\hat{\Pi}_{x,r}(x)$ due to common zeros of the highest degree coefficient polynomials can be filtered out easily because they do not correspond to a finite y-value that generates a small residual for (5.1). We must take care when these spurious eigenvalues correspond to both finite and infinite solutions. For such an eigenvalue x^*, the multiplicity is equal to the sum of the multiplicities of all solutions of the form (x^*, y), including (x^*, ∞). We will propose a solution to this problem at the end of the first section.

5.1 Finding $\tilde{\mathcal{S}}$ without solving the GEP in y

Clearly, the set $\tilde{\mathcal{Y}}$ can be found as the eigenvalues of the pencil $\hat{\Pi}_{x,r}(x)$ constructed for the system

$$\begin{cases}\tilde{p}(x, y) = 0\\ \tilde{q}(x, y) = 0\end{cases}$$

where $\tilde{p}(x, y) = p(y, x)$ and $\tilde{q}(x, y) = q(y, x)$. For the approaches in this section we avoid the construction of a second square GEP. We will find $\tilde{\mathcal{Y}}$ by finding the possible y-values for each value in $\tilde{\mathcal{X}}$. This can be done by making use of the eigenspace of each $x^* \in \tilde{\mathcal{X}}$, by thinking of $p(x^*, y)$ and $q(x^*, y)$ as two univariate polynomials or by using the linear pencil $L(x^*, y)$ in y. To discuss the theory behind each approach we will assume infinite precision: we will show that the correct set $\mathcal{Y}$ is found in exact arithmetic. We will also highlight some of the numerical issues that are encountered by each approach in practice (finite precision).

5.1.1 The eigenvectors of $\hat{\Pi}_{x,r}(x)$

During the construction of $\hat{\Pi}_{x,r}(x)$ we have imposed a blockwise Vandermonde structure on its eigenvectors. For an x-value that occurs only once in $\mathcal{X}$, the

associated eigenspace is one-dimensional. The following proposition is a direct consequence.

Proposition 5.1.1. *Let $x^* \in \mathcal{X}$ be such that there is no $x \in \mathcal{X}$, apart from x^* itself, such that $x^* = x$. Also, we assume that x^* is such that $\exists y \in \mathbb{C}$ such that $p(x^*, y) = q(x^*, y) = 0$. Then the corresponding eigenvector $\boldsymbol{w}$:*

$$\hat{\Pi}_{x,r}(x^*)\boldsymbol{w} = \mathbf{0}$$

has the Vandermonde structure in the reduced monomial basis.

Proof. Under these assumptions the eigenvector $\boldsymbol{w}$ and its scalar multiples are the only vectors $\boldsymbol{u}$ that satisfy

$$\hat{\Pi}_{x,r}(x^*)\boldsymbol{u} = \mathbf{0}.$$

Also, there is only one solution to (5.1), say (x^*, y^*), with x-coordinate x^*. Indeed, if there were others, x^* would appear multiple times in $\mathcal{X}$. The Vandermonde vector corresponding to this solution $\boldsymbol{v}(x^*, y^*)$ must lie in the right null space of $\hat{L}_r(x^*, y^*)$ and thus in that of $\hat{\Pi}_{x,r}(x^*)$. Therefore $\boldsymbol{w} = C\boldsymbol{v}(x^*, y^*)$, $C \in \mathbb{C}_0$. □

Corollary 5.1.1. *For a simple root of* (5.1) *that is such that no other root has the same x-coordinate, we can find the corresponding y-value by dividing the $(\psi_1 + 1)$-st entry of the corresponding eigenvector $\boldsymbol{w}$ by the first entry:*

$$y^* = \frac{\boldsymbol{w}_{\psi_1+1}}{\boldsymbol{w}_1}.$$

For $x^* \in \mathcal{X}$ that does not satisfy the assumptions of Proposition 5.1.1 things are slightly more complicated. If x^* appears c_{x^*} times in $\mathcal{X}$, the null space of $\hat{\Pi}_{x,r}(x^*)$ can have a dimension up to c_{x^*}. Moreover, this null space may contain vectors that do not have the Vandermonde structure. We illustrate this with an example.

Example 5.1.1. *Consider again the system from Example 2.2.3*

$$\begin{cases} p(x, y) = x^2 + y^2 - 1 = 0 \\ q(x, y) = 4x^2 + y^2 + 6x + 2 = 0 \end{cases}$$

with a 4-fold zero at $(-1, 0)$. We have

$$\begin{cases} p(x, y) = 0 \\ q(x, y) = 0 \end{cases} \xrightarrow[C]{} L(x, y) \xrightarrow[E]{} \hat{L}(x, y) \xrightarrow[R]{} \left(\begin{array}{ccc|ccc|cc} -1 & 0 & 1 & 0 & 0 & 0 & 1 & \\ 2 & 6 & 4 & 0 & 0 & 0 & 1 & \\ \hline & & & -1 & 0 & 1 & 0 & 1 \\ & & & 2 & 6 & 4 & 0 & 1 \\ \hline -x & 1 & & & & & & \\ & -x & 1 & & & & & \\ \hline & & & -x & 1 & & & \\ & & & & -x & 1 & & \\ \hline -y & & & 1 & & & & \\ & & & -y & & & 1 & \\ & & & & & & -y & 1 \end{array}\right).$$

The reduced monomial vector is

$$\boldsymbol{v}(x,y) = \begin{pmatrix} 1 & x & x^2 \mid y & xy & x^2y \mid y^2 \mid y^3 \end{pmatrix}^\top$$

and the finite eigenvalues of $\hat{\Pi}_{x,r}(x)$ *are* $\mathcal{X} = \{-1,-1,-1,-1\}$*. The null space of* $\hat{\Pi}_{x,r}(-1)$ *has dimension 2 and it is spanned by*

$$\boldsymbol{w}_1 = \begin{pmatrix} 1 & -1 & 1 \mid 0 & 0 & 0 \mid 0 \mid 0 \end{pmatrix}^\top$$

and

$$\boldsymbol{w}_2 = \begin{pmatrix} 0 & 0 & 0 \mid 1 & -1 & 1 \mid 0 \mid 0 \end{pmatrix}^\top.$$

Note that $\boldsymbol{w}_1$ *has the Vandermonde structure, whereas* $\boldsymbol{w}_2$ *does not:*

$$\hat{L}_r(-1,0)\boldsymbol{w}_2 \neq \boldsymbol{0}.$$

Of course, both $\boldsymbol{w}_1$ *and* $\boldsymbol{w}_2$ *have a blockwise Vandermonde structure. We can understand the presence of* $\boldsymbol{w}_2$ *in* $\operatorname{null}(\hat{\Pi}_{x,r}(-1))$ *better by investigating the multiplicity structure of* $\boldsymbol{z} = (-1,0)$ *(see Appendix A). It is easy to check that*

$$\frac{\partial}{\partial y}\Big|_{\boldsymbol{z}} \begin{pmatrix} p \\ q \\ yp \\ yq \end{pmatrix} = \begin{pmatrix} \hat{\Phi}_r \\ \Psi_r \end{pmatrix} \frac{\partial \boldsymbol{v}(x,y)}{\partial y}\Big|_{\boldsymbol{z}} = \boldsymbol{0}$$

where

$$\frac{\partial \boldsymbol{v}(x,y)}{\partial y} = \begin{pmatrix} 0 & 0 & 0 \mid 1 & x & x^2 \mid 2y \mid 3y^2 \end{pmatrix}^\top$$

has a blockwise Vandermonde structure. Therefore $\boldsymbol{w}_2 = \frac{\partial \boldsymbol{v}(x,y)}{\partial y}\Big|_{\boldsymbol{z}}$ *is contained in the null space of* $\hat{\Pi}_{x,r}(-1)$*.*

In general, we can find the possible values of y corresponding to an x-solution x^* by looking for Vandermonde structured vectors in the right null space of $\hat{\Pi}_{x,r}(x^*)$, which is the eigenspace corresponding to the eigenvalue x^*. Suppose the columns of V_{x^*} form a basis for this eigenspace. A null vector of $\hat{\Pi}_{x,r}(x^*)$ has the Vandermonde structure if it is in the null space of $\hat{\mathcal{B}}_{y,r} - y\hat{\mathcal{C}}_{y,r}$ for some $y \in \mathbb{C}$. Therefore, we are looking for a vector $V_{x^*}\boldsymbol{c} \in \operatorname{null}(\hat{\Pi}_{x,r}(x^*))$ that satisfies

$$(\hat{\mathcal{B}}_{y,r} - y\hat{\mathcal{C}}_{y,r})V_{x^*}\boldsymbol{c} = \boldsymbol{0}. \tag{5.2}$$

This is a rectangular eigenvalue problem (see Appendix E) of size $(\delta_p^y + \delta_q^y - 1) \times \dim\operatorname{null}(\hat{\Pi}_{x,r}(x^*))$. The finite eigenvalues are the possible y-coordinates for solutions of the form (x^*, y). This way we can find S by adding all couples (x^*, y^*) that are not already contained in S. It is never guaranteed, however, that we can find $\mathcal{S}$. Suppose we have a 3-fold eigenvalue x^* that corresponds to the two solutions (x^*, y_1) (multiplicity 1) and (x^*, y_2) (multiplicity 2), $y_1 \neq y_2$. Then we would find for each

time x^* appears in $\mathcal{X}$ that (x^*, y_1) and (x^*, y_2) are possible couples. We would end up with three times (x^*, y_1) and three times (x^*, y_2) and we loose the multiplicity information. We can avoid this by applying a "generic" transformation of variables T:

$$\begin{cases} p(x,y) = 0 \\ q(x,y) = 0 \end{cases} \longrightarrow \begin{cases} p(T(x,y)) \triangleq p_t(x,y) = 0 \\ q(T(x,y)) \triangleq q_t(x,y) = 0 \end{cases} \tag{5.3}$$

so that we are "almost" sure that for the transformed problem there are no solutions with the same x-coordinates and different y-coordinates. If we take T to be linear we can think of it as a matrix and we can find every solution to the original problem (x^*, y^*) from the solutions to the transformed system (x_T^*, y_T^*) as

$$\begin{pmatrix} x^* \\ y^* \end{pmatrix} = T \begin{pmatrix} x_T^* \\ y_T^* \end{pmatrix}.$$

Note that this is also a trick for avoiding solutions at infinity with the same x-coordinate as another finite solution.

In finite precision, this method suffers from some drawbacks. In practice, we need to start from the set $\tilde{\mathcal{X}}$. The matrix V_{x^*} must contain a (numerical) basis for the eigenspace of x^*. Therefore we must take into account all the numerical representatives of x^*, which can only be found by accepting two x-values x_1 and x_2 to be "sufficiently equal" if the difference $|x_1 - x_2|$ is smaller than some threshold value. A possible strategy is to sort the eigenvalues in $\tilde{\mathcal{X}}$ by their distance to x^* and to determine how many of the first eigenvectors there are needed to construct V_{x^*}. This requires a number of heuristics, which is not desirable. Another way to proceed is to calculate a numerical basis for the right null space of $\hat{\Pi}_{x,r}(x^*)$ for each $x^* \in \tilde{\mathcal{X}}$ and using the basis vectors as V_{x^*}. Although this requires more computational effort, it is numerically more reliable and the method is implemented in this way. Also, a transformation of variables is, conceptually, an interesting thing to do to perserve the multiplicity information. However, it turns out to be less interesting from a numerical point of view. After a linear transformation of variables, the maximal ratio between the absolute values of the coefficients of a bivariate polynomial p can increase drastically. This leads to an unbalanced eigenvalue problem of which the eigenvalues are calculated less accurately.

5.1.2 Common roots of univariate polynomials

All possible y-values that correspond to a solution (x^*, y) must satisfy

$$\begin{cases} p(x^*, y) = 0 \\ q(x^*, y) = 0 \end{cases}$$

where $p(x^*, y) \triangleq f(y)$ and $q(x^*, y) \triangleq g(y)$ are univariate polynomials in y. We have shown in Chapter 2 that the common roots of f and g can be found as the eigenvalues

of the rectangular eigenvalue problem

$$(S_y - yS_1)Z\boldsymbol{v} = 0, \tag{5.4}$$

where the columns of Z form a basis for the null space of $S^{f,g}$ and

$$S_y - yS_1 = \begin{pmatrix} -y & 1 & & \\ & -y & 1 & \\ & & \ddots & \\ & & & -y \quad 1 \end{pmatrix}$$

is a $(\deg(f) + \deg(g) - 1) \times (\deg(f) + \deg(g))$ matrix. This leads to another strategy for finding the set Y or $\mathcal{Y}$. For each $x^* \in \mathcal{X}$, calculate a basis Z for the null space of $S^{f,g}$ and solve the eigenvalue problem (5.4). Then, add all the distinct finite eigenvalues to the set Y. In case we want to find $\mathcal{Y}$, we can apply a generic transformation of variables T as in (5.3) so that every x^* generates only one distinct eigenvalue y^*. The size of the eigenvalue problem that has to be solved for every $x^* \in \mathcal{X}$ is $(\deg(f) + \deg(g) - 1) \times \dim \operatorname{null}(S^{f,g})$.

Note that for this approach, no numerical threshold is needed to construct the rectangular eigenvalue problem. Of course, the eigenvalue problem still has to be solved numerically (see Appendix E).

5.1.3 The eigenvalues of $L(x^*, y)$

According to Theorem 3.1.1, all possible y-values corresponding to $x^* \in \mathcal{X}$ are the values of y for which $L(x^*, y)$ loses full column rank. Equivalently, we can say that all possible y-values corresponding to $x^* \in \mathcal{X}$ are the eigenvalues of the rectangular eigenvalue problem

$$L(x^*, y)\boldsymbol{v} = \mathbf{0}.$$

Again, Y and $\mathcal{Y}$ can be found as in the previous approaches. The size of the eigenvalue problem can be reduced by applying Algorithm 1 to $L(x, y)$ rather than $\hat{L}(x, y)$.

The eigenvalue problem is constructed without using any numerical threshold and it is solved as explained in Appendix E.

5.2 Finding $\tilde{\mathcal{S}}$ by coupling $\tilde{\mathcal{X}}$ and $\tilde{\mathcal{Y}}$

Another way to proceed is to calculate $\tilde{\mathcal{Y}}$ by constructing the generalized eigenvalue problem like we did for $\tilde{\mathcal{X}}$ and then find an appropriate coupling between the two sets. In order to find such a coupling, we need a tool to determine whether a certain couple is feasible or not.

Definition 5.1 (residual)**.** *For every couple* $(x^*, y^*) \in \mathbb{C}^2$ *the* residual *with respect to* (5.1) *is defined as*

$$r(x^*, y^*) = \frac{|p(x^*, y^*)|}{|p|(|x^*|, |y^*|) + 1} + \frac{|q(x^*, y^*)|}{|q|(|x^*|, |y^*|) + 1}$$

where $|p|(x, y) \triangleq \sum_{i=0}^{\delta} \sum_{j=0}^{\delta-i} |p_{ij}| x^j y^i$ *and* $|q|(x, y) \triangleq \sum_{i=0}^{\delta} \sum_{j=0}^{\delta-i} |q_{ij}| x^j y^i$.

Note that in Definition 5.1, the denominators contain a constant term +1 so the residual is determined based on a mixed (both relative and absolute) criterion. This way the residual is also defined for $|p|(|x^*|, |y^*|) = 0$ or $|q|(|x^*|, |y^*|) = 0$ and it does not become too pessimistic for small values of $|p|(|x^*|, |y^*|)$ or $|q|(|x^*|, |y^*|)$. Suppose we have the sets $\mathcal{X}$ and $\mathcal{Y}$ and we have indexed the elements in some way. Denote the i-th element in $\mathcal{X}$ as $x^{(i)}$. Analogously, for $\mathcal{Y}$ we use the notation $y^{(i)}$. For the numerically obtained sets $\tilde{\mathcal{X}}$ and $\tilde{\mathcal{Y}}$ we use $\tilde{x}^{(i)}$ and $\tilde{y}^{(i)}$ respectively. The *residual matrix* for the sets $\mathcal{X}$ and $\mathcal{Y}$ is defined as

$$R(\mathcal{X}, \mathcal{Y})_{ij} = r(x^{(j)}, y^{(i)}).$$

The residual matrix $R(\mathcal{X}, \mathcal{Y})$ is square ($\mathcal{X}$ and $\mathcal{Y}$ contain the same number of elements). We have experimented with two coupling procedures: applying a "minimax" principle to the diagonal of the residual matrix and finding an appropriate coupling based on a so called *connection diagram*.

A minimax principle

Suppose we have found $\tilde{\mathcal{X}}$ and $\tilde{\mathcal{Y}}$ and we have calculated the residual matrix $R(\tilde{\mathcal{X}}, \tilde{\mathcal{Y}})$. An intuitive way of finding a coupling between $\tilde{\mathcal{X}}$ and $\tilde{\mathcal{Y}}$ is to look for a row permutation R_p for $R(\tilde{\mathcal{X}}, \tilde{\mathcal{Y}})$ that minimizes the maximal element on its diagonal. After permuting the elements of $\tilde{\mathcal{Y}}$ in the same way, we can propose $\tilde{\mathcal{S}} = \{(\tilde{x}^{(i)}, \tilde{y}^{(i)})\}_{1 \leq i \leq n}$ as a numerical solution set where n is the number of elements in $\tilde{\mathcal{X}}$ and $\tilde{\mathcal{Y}}$. The problem can be formulated as follows. Find R_p such that

$$R_p = \operatorname{argmin}_{R_p \in \mathcal{R}_p} \max \operatorname{diag}(R_p R(\tilde{\mathcal{X}}, \tilde{\mathcal{Y}}))$$

where $\mathcal{R}_p$ is the set of all row permutation matrices. There are algorithms that provide a suboptimal solution to this optimization problem. One such algorithm that is based on Jacobi-like sweeps is implemented. Residuals are small but the resulting multiplicities are not guaranteed to be correct. Also, some solutions might not be found. Figure 5.1 illustrates the result of a suboptimal permutation to minimize the diagonal elements. In this case, the permutation results in a correct coupling. Another possibility is to approach the problem as a minimum weight matching in a bipartite graph. We do not elaborate on this approach. The coupling in the next section is found to be more effective.

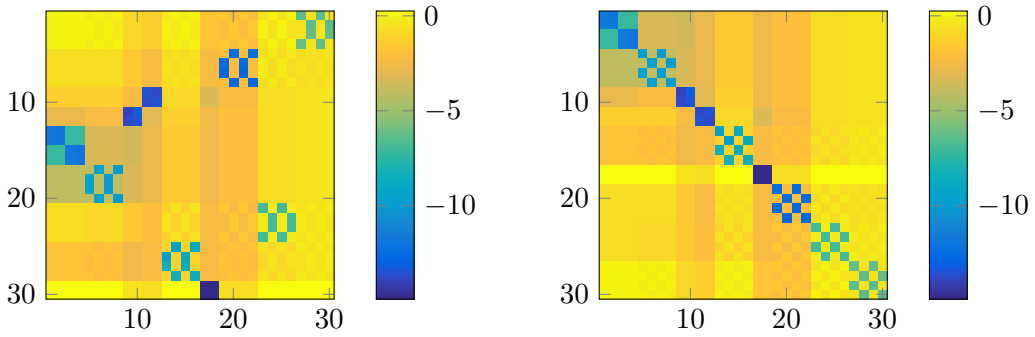


Figure 5.1: Residual matrix $R(\tilde{\mathcal{X}}, \tilde{\mathcal{Y}})$ before (left) and after (right) applying a row permutation to minimize the diagonal elements. The color bars represent the amplitude of the elements in $R(\tilde{\mathcal{X}}, \tilde{\mathcal{Y}})$ on a $\log_{10}$-scale.

Connection diagrams

To explain the ideas in this paragraph, we will use the following toy solution set:

$$S = \{(x_1, y_1), (x_1, y_2), (x_2, y_2), (x_2, y_3), (x_3, y_1), (x_3, y_4)\} \tag{5.5}$$

with $x_i \in \mathbb{C}, i = 1, \ldots, 3$, $y_j \in \mathbb{C}, j = 1, \ldots, 4$, $x_i \neq x_j, j \neq i$ and $y_i \neq y_j, j \neq i$. We will denote the multiplicities by $M_{(x_i, y_j)}(\mathcal{V}_p, \mathcal{V}_q) = \mu_{ij}$ and they are given by

$$\mu_{11} = 3, \quad \mu_{12} = \mu_{23} = 2, \quad \mu_{22} = \mu_{31} = \mu_{34} = 1. \tag{5.6}$$

We use the following notation to distinguish between exact and numerical results:

$$\begin{aligned} \mathcal{X} &= \{x_1, x_1, x_1, x_1, x_1, x_2, x_2, x_2, x_3, x_3\}, \\ \tilde{\mathcal{X}} &= \{\tilde{x}_{11}, \tilde{x}_{12}, \tilde{x}_{13}, \tilde{x}_{14}, \tilde{x}_{15}, \tilde{x}_{21}, \tilde{x}_{22}, \tilde{x}_{23}, \tilde{x}_{31}, \tilde{x}_{32}\} \end{aligned}$$

where we assume that $\mathcal{X}$ and $\tilde{\mathcal{X}}$ are ordered so that $\tilde{x}^{(i)}$ is a numerical representative for $x^{(i)}$: $\tilde{x}_{ij} = x_i + \epsilon_{ij}$ with ϵ_{ij} some (small) complex number. The set $\mathcal{X}$ is ordered by increasing distance to x_1. Note that the number of appearances c_i^x of x_i in $\mathcal{X}$ follows directly from (5.6): $c_i^x = \sum_j \mu_{ij}$. We will refer to the number c_i^x as the *projected multiplicity* of x_i. Analogously, for the y-coordinates we denote

$$\begin{aligned} \mathcal{Y} &= \{y_1, y_1, y_1, y_1, y_2, y_2, y_2, y_3, y_3, y_4\}, \\ \tilde{\mathcal{Y}} &= \{\tilde{y}_{11}, \tilde{y}_{12}, \tilde{y}_{13}, \tilde{y}_{14}, \tilde{y}_{21}, \tilde{y}_{22}, \tilde{y}_{23}, \tilde{y}_{31}, \tilde{y}_{32}, \tilde{y}_{41}\}. \end{aligned}$$

Note that obtaining $\mathcal{X}$ and $\mathcal{Y}$ by solving the eigenvalue problem described in Chapter 4 we find the projected multiplicities c_i^x and c_i^y (in theory) but not the multiplicities μ_{ij}[1]. Therefore in the following, the projected multiplicties will be considered as known, whereas the multiplicities of the solutions are the unknowns. Let us first consider the sets $\mathcal{X}$ and $\mathcal{Y}$ which are the finite eigenvalues of the associated pencils, calculated in infinite precision. The question is how we can construct $\mathcal{S}$ based on these two sets.

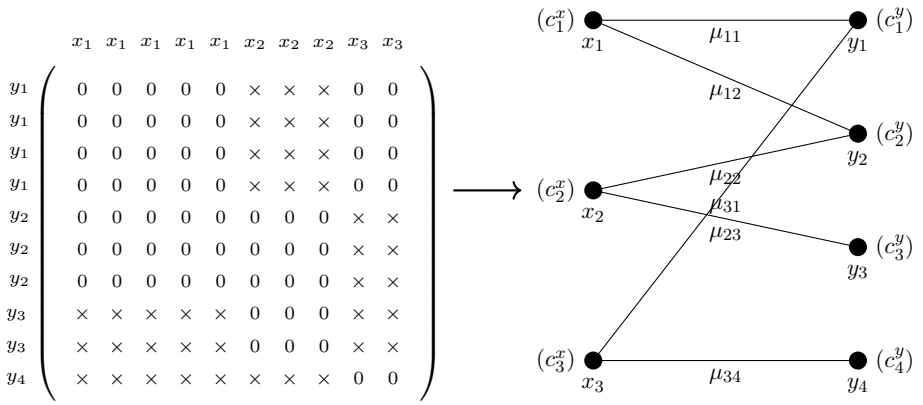


Figure 5.2: Illustration of the construction of the connection diagram from $R(\mathcal{X}, \mathcal{Y})$. A little cross (×) represents any positive nonzero number.

First, we calculate the residual matrix $R(\mathcal{X}, \mathcal{Y})$. Its structure is shown in Figure 5.2. Note that the columns of $R(\mathcal{X}, \mathcal{Y})$ corresponding to the same x-value (for example the first 5 columns) are identical. The same holds for the rows corresponding to the same y-value. To check which y-values form a solution with x_1, we can simply check which entries in the first column of the residual matrix are zero. This is true for all rows corresponding to y_1 and y_2. Therefore, we know that S contains the solutions (x_1, y_1) and (x_1, y_2). In $\mathcal{S}$, we will find these solutions with a certain multiplicity. The unknown multiplicities of (x_1, y_1) and (x_1, y_2) are μ_{11} and μ_{12} respectively. For the second column, no new solutions are found and no new unknown multiplicities are introduced. The sixth column is the first one that corresponds to a new x-value, and we find that also (x_2, y_2) and (x_2, y_3) must be added to S and they lead to two new unknowns μ_{22} and μ_{23}. Proceeding like this for the remaining columns we can construct a *connection diagram* for the sets $\mathcal{X}$ and $\mathcal{Y}$. The residual matrix and connection diagram are presented in Figure 5.2 for our example sets. From the diagram in Figure 5.2 we can see that for a feasible solution set $\mathcal{S}$, it must hold that $\mu_{11} + \mu_{12} = c_1^x = 5$. Indeed, the number of solutions with an x-coordinate x_1 must be equal to the multiplicity of the eigenvalue x_1 of the pencil $\hat{\Pi}_{x,r}(x)$, which is equal to the projected multiplicity of x_1. This way we obtain an equation for each node in the connection diagram. The resulting system is

$$\begin{cases} \mu_{11} + \mu_{12} = c_1^x \\ \mu_{22} + \mu_{23} = c_2^x \\ \mu_{31} + \mu_{34} = c_3^x \\ \mu_{11} + \mu_{31} = c_1^y \\ \mu_{12} + \mu_{22} = c_2^y \\ \mu_{23} = c_3^y \\ \mu_{34} = c_4^y \end{cases} \rightarrow \begin{pmatrix} 1 & 1 & & & & \\ & & 1 & 1 & & \\ & & & & 1 & 1 \\ 1 & & & & 1 & \\ & 1 & 1 & & & \\ & & & 1 & & \\ & & & & & 1 \end{pmatrix} \begin{pmatrix} \mu_{11} \\ \mu_{12} \\ \mu_{22} \\ \mu_{23} \\ \mu_{31} \\ \mu_{34} \end{pmatrix} = \begin{pmatrix} c_1^x \\ c_2^x \\ c_3^x \\ c_1^y \\ c_2^y \\ c_3^y \\ c_4^y \end{pmatrix} \rightarrow M\boldsymbol{\mu} = \boldsymbol{c} \qquad (5.7)$$

[1] In practice, we can find the projected multiplicities by clustering the numerically found sets $\tilde{\mathcal{X}}$ and $\tilde{\mathcal{Y}}$, more about this later.

where $\boldsymbol{c}$ is a known vector containing the projected multiplicities of the elements of $\mathcal{X}$ and $\mathcal{Y}$. The unique solution of this system is

$$\mu_{11} = 3, \quad \mu_{12} = \mu_{23} = 2, \quad \mu_{22} = \mu_{31} = \mu_{34} = 1. \tag{5.8}$$

These are indeed the multiplicities (5.6) that we were looking for.

We now step away from the assumption of infinite precision. A first numerical difficulty is determining the projected multiplicities. The projected multiplicity of x_1 is equal to the number of numerical representatives of x_1 in $\tilde{\mathcal{X}}$. Therefore, we have to find an appropriate clustering for the values in $\tilde{\mathcal{X}}$. That is, we have to find the clusters $\tilde{x}_1$, $\tilde{x}_2$ and $\tilde{x}_3$ of numerical approximations for x_1, x_2 and x_3 respectively (and the same for the y-values). The projected multiplicity of x_i is $|\tilde{x}_i|$ with $|\cdot|$ the cardinality of a cluster. Another issue in finite precision is that the residuals for the solutions in $\tilde{\mathcal{S}}$ are not exactly zero. In order to construct the connection diagram, we will choose a threshold value $\epsilon > 0$ to decide whether we accept a residual as small enough for the corresponding couple to be considered a solution. We will propose a way to cluster the values in $\tilde{\mathcal{X}}$ and $\tilde{\mathcal{Y}}$ using the residual matrix $R(\tilde{\mathcal{X}}, \tilde{\mathcal{Y}})$. The clustering is based on the assumption that x-values of the same cluster generate a small ($\leq \epsilon$) residual with the same y-values in $\tilde{\mathcal{Y}}$. Based on that assumption, the residual matrix looks like this:

$$R(\tilde{\mathcal{X}}, \tilde{\mathcal{Y}}) = \begin{array}{c} \\ \tilde{y}_{11} \\ \tilde{y}_{12} \\ \tilde{y}_{13} \\ \tilde{y}_{14} \\ \tilde{y}_{21} \\ \tilde{y}_{22} \\ \tilde{y}_{23} \\ \tilde{y}_{31} \\ \tilde{y}_{32} \\ \tilde{y}_{41} \end{array} \begin{array}{c} \begin{array}{cccccccccc} \tilde{x}_{11} & \tilde{x}_{12} & \tilde{x}_{13} & \tilde{x}_{14} & \tilde{x}_{15} & \tilde{x}_{21} & \tilde{x}_{22} & \tilde{x}_{23} & \tilde{x}_{31} & \tilde{x}_{32} \end{array} \\ \left(\begin{array}{cccccccccc} \tilde{0} & \tilde{0} & \tilde{0} & \tilde{0} & \tilde{0} & \times & \times & \times & \tilde{0} & \tilde{0} \\ \tilde{0} & \tilde{0} & \tilde{0} & \tilde{0} & \tilde{0} & \times & \times & \times & \tilde{0} & \tilde{0} \\ \tilde{0} & \tilde{0} & \tilde{0} & \tilde{0} & \tilde{0} & \times & \times & \times & \tilde{0} & \tilde{0} \\ \tilde{0} & \tilde{0} & \tilde{0} & \tilde{0} & \tilde{0} & \times & \times & \times & \tilde{0} & \tilde{0} \\ \tilde{0} & \tilde{0} & \tilde{0} & \tilde{0} & \tilde{0} & \tilde{0} & \tilde{0} & \tilde{0} & \times & \times \\ \tilde{0} & \tilde{0} & \tilde{0} & \tilde{0} & \tilde{0} & \tilde{0} & \tilde{0} & \tilde{0} & \times & \times \\ \tilde{0} & \tilde{0} & \tilde{0} & \tilde{0} & \tilde{0} & \tilde{0} & \tilde{0} & \tilde{0} & \times & \times \\ \times & \times & \times & \times & \times & \tilde{0} & \tilde{0} & \tilde{0} & \times & \times \\ \times & \times & \times & \times & \times & \tilde{0} & \tilde{0} & \tilde{0} & \times & \times \\ \times & \times & \times & \times & \times & \times & \times & \times & \tilde{0} & \tilde{0} \end{array} \right) \end{array}.$$

The symbol $\tilde{0}$ represents any nonnegative number $\leq \epsilon$ and the symbol $\times$ stands for any positive number $> \epsilon$. We see that $\tilde{x}_{11}$, $\tilde{x}_{12}$, $\tilde{x}_{13}$, $\tilde{x}_{14}$ and $\tilde{x}_{15}$ generate a small residual when coupled to the first seven y-values in $\tilde{\mathcal{Y}}$. For the last three y-values, they do not. We conclude that these five x-values belong to the same cluster: $\tilde{x}_1 = \{\tilde{x}_{11}, \tilde{x}_{12}, \tilde{x}_{13}, \tilde{x}_{14}, \tilde{x}_{15}\}$. For the y-values $\tilde{y}_{31}$ and $\tilde{y}_{32}$, a coupling generates a small residual for the same three x-values. Therefore they belong to the same cluster in $\tilde{\mathcal{Y}}$. The obtained connection diagram and clustering is shown in the left part

of Figure 5.3. Note that it is a "numerical equivalent" for the connection diagram of Figure 5.2. The projected multiplicities of the clusters are given by their cardinality.

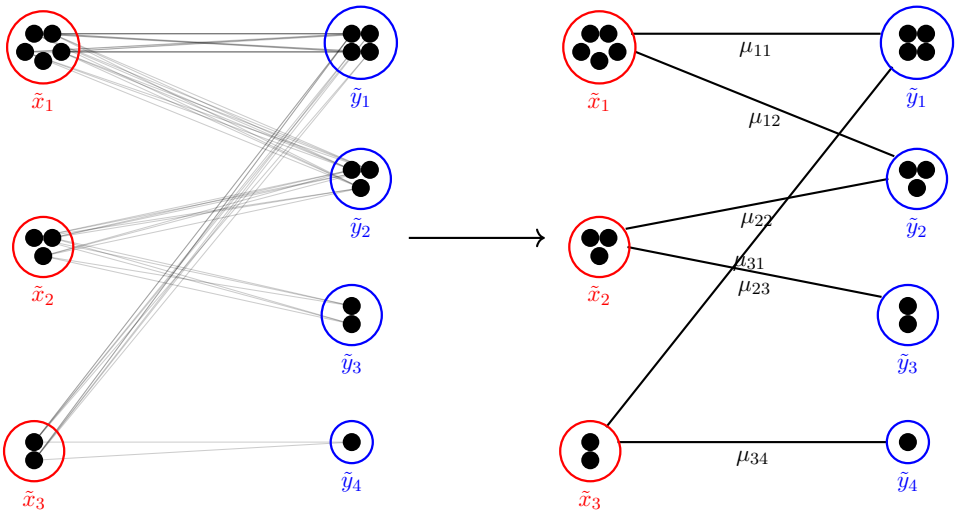


Figure 5.3: Left: the numerically obtained connection diagram. Every connection represents a $\tilde{0}$ element in the residual matrix. Black dots represent the elements of $\tilde{\mathcal{X}}$ (left) and $\tilde{\mathcal{Y}}$ (right). The red circles indicate the x-clusters, the blue circles represent the y-clusters. Right: the corresponding clustered connection diagram.

A next step is to cluster all the individual connections in the connection diagram in the left part of Figure 5.3. Rather than considering connections between elements of $\tilde{\mathcal{X}}$ and $\tilde{\mathcal{Y}}$, we consider connections *between clusters*. All connections that connect the same x-cluster with the same y-cluster represent one cluster connection with an unknown multiplicity. The resulting diagram is called the *clustered connection diagram* and it is shown in the right part of Figure 5.3. To determine these multiplicities, a system equal to (5.7) is obtained from the clustered connection diagram. Finally, the solution set $\tilde{\mathcal{S}}$ can be constructed by adding μ_{ij} couples of x-values in $\tilde{x}_i$ and y-values in $\tilde{y}_j$ for each cluster connection. For example, one can choose to select μ_{ij} times the combination that generates the smallest residual or μ_{ij} times the combination of the mean value of $\tilde{x}_i$ with the mean value of $\tilde{y}_j$.

The clustering procedure for finding the x- and y-clusters has to be completed by a second clustering step in some special cases. For example, consider the solution set $\{(1,0),(-1,0)\}$ and the numerically found sets $\tilde{\mathcal{X}} = \{-1+\epsilon_1, 1+\epsilon_2\}$ and $\tilde{\mathcal{Y}} = \{\epsilon_3, \epsilon_4\}$. It is clear that both x-values will generate a small residual when coupled to both y-values, yet they should not be in the same cluster. Therefore, each intermediate cluster after the first clustering step based on the residual matrix might consist out of several clusters. We can split the clusters for example by detecting groups within a cluster of which the mean values are more than some threshold distance apart.

An advantage of this coupling strategy is that it allows us to find $\mathcal{S}$ even if there are different solutions that have one coordinate in common. Our toy solution set is an example of this. Therefore, a transformation of variables is only required in case there are infinite solutions that have one finite coordinate that coincides with that of a finite solution. If, however, a transformation of variables is performed, there is no need for the second clustering step.

A drawback of this approach is that the system $M\boldsymbol{\mu} = \boldsymbol{c}$ might be underdetermined. To find the solution we must take into account that all entries of $\boldsymbol{\mu}$ should be positive and integer. This is an integer programming problem, which is NP-hard. For the dataset that is used to test our method, the system is overdetermined in more than 90% of the cases. In the cases where it is underdetermined, the smallest norm least squares solution

$$\boldsymbol{\mu} = M^{+}\boldsymbol{c}$$

where M^{+} denotes the Moore-Penrose pseudo-inverse of M, is the desired positive integer solution. The reason for this is at this point unclear to the author and requires further research. If the system is underdetermined, the x- and y-values are strongly interconnected: there are more connections than the total number of elements in $\tilde{\mathcal{X}}$ and $\tilde{\mathcal{Y}}$. This interconnection can be avoided by applying a transformation of variables.

5.3 Variable precision

The idea in this section is the following. Denote the machine epsilon in double precision by ϵ_d and in quadruple precision by ϵ_q. To solve the problem in double precision, we manipulate the coefficients of p and q by a generic perturbation of order $\epsilon_d \times 10^{-2}$ and perform all calculations in quadruple precision. The perturbation is of the following particular form. Let P and Q be the matrix representations of p and q where $P \in \mathcal{C}^{(\delta_p+1)\times(\delta_p+1)}$ and $Q \in \mathcal{C}^{(\delta_q+1)\times(\delta_q+1)}$ (these representations may be redundant). The perturbed polynomials will have a matrix representation given by

$$\begin{aligned} \tilde{P} &= P + P_\epsilon, \\ \tilde{Q} &= Q + Q_\epsilon, \end{aligned} \tag{5.9}$$

where $\Re(P_\epsilon)$, $\Im(P_\epsilon)$, $\Re(Q_\epsilon)$ and $\Im(Q_\epsilon)$ are matrices with random normally distributed entries with mean zero and standard deviation $\epsilon_d \times 10^{-2}$ on and above the antidiagonal. Equivalently, we add a polynomial p_ϵ to p with normally distributed coefficients in the monomial basis for $\mathcal{P}^2_{\delta_p}$. We do the same for q. This way the solutions to (5.1) are slightly perturbed and they will (generically) all $\delta_p\delta_q$ be simple and finite in quadruple precision. Solutions corresponding to infinite solutions of the original problem are easily detected because they correspond to very large solutions of the perturbed problem. All multiple precision calculations are done using the Multiprecision Computing Toolbox for Matlab (`http://www.advanpix.com/`).

5.3.1 Another form of the linear pencil

For this version of our method we will need yet another form of the linear pencil. We start from the extended version $\hat{L}(x,y)$:

$$\begin{cases} p(x,y)=0 \\ q(x,y)=0 \end{cases} \quad \xrightarrow[C]{} \quad L(x,y) \xrightarrow[E]{} \quad \hat{L}(x,y).$$

Recall that in $\hat{L}(x,y)$ we have used as few as possible equations in y to define the extended monomial basis. What we will now do is add all of the other linear y-recurrences to $(\hat{\mathcal{B}}_y - y\hat{\mathcal{C}}_y)$, so that $(\hat{\mathcal{B}}_y - y\hat{\mathcal{C}}_y)$ is extended to a matrix of the same size as $(\hat{\mathcal{B}}_x - x\hat{\mathcal{C}}_x)$. We will denote this operation by

$$\hat{L}(x,y) \quad \xrightarrow[Y]{} \quad \hat{L}_y(x,y) = \begin{pmatrix} \hat{\Pi}_x(x) \\ \hat{\mathcal{B}}_{y,e} - y\hat{\mathcal{C}}_{y,e} \end{pmatrix}.$$

After that, we apply a y-reduction to $\hat{L}_y(x,y)$ so that the x-pencil becomes square. We remove all rows in $(\hat{\mathcal{B}}_{y,e} - y\hat{\mathcal{C}}_{y,e})$ that can no longer be used in the reduced monomial basis. We denote

$$\hat{L}(x,y) \quad \xrightarrow[Y]{} \quad \hat{L}_y(x,y) \quad \xrightarrow[y\text{-red}]{} \quad \begin{pmatrix} \hat{\Pi}_{x,\tilde{r}}(x) \\ \hat{\Pi}_{y,e,\tilde{r}}(y) \end{pmatrix} = \begin{pmatrix} \hat{\Pi}_{x,\tilde{r}}(x) \\ \hat{\mathcal{B}}_{y,e,\tilde{r}} - y\hat{\mathcal{C}}_{y,e,\tilde{r}} \end{pmatrix}$$

Example 5.3.1. *Consider the system*

$$\begin{cases} p(x,y) = -1 + x^2 + y = 0 \\ q(x,y) = -1 + y = 0 \end{cases} \quad \xrightarrow[C]{} \quad L(x,y) \quad \xrightarrow[E]{} \quad \hat{L}(x,y).$$

We have that

$$\hat{L}_y(x,y) = \left(\begin{array}{ccc|cc|c} -1 & 0 & 1 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 \\ \hline -x & 1 & & & & \\ & -x & 1 & & & \\ & & & -x & 1 & \\ \hline -y & & & 1 & & \\ & -y & & & 1 & \\ & & & -y & & 1 \end{array}\right) \xrightarrow[y\text{-red}]{} \left(\begin{array}{ccc|cc} -1 & 0 & 1 & 1 & 0 \\ -1 & 0 & 0 & 1 & 0 \\ \hline -x & 1 & & & \\ & -x & 1 & & \\ & & & -x & 1 \\ \hline -y & & & 1 & \\ & -y & & & 1 \end{array}\right) = \begin{pmatrix} \hat{\Pi}_{x,\tilde{r}}(x) \\ \hat{\Pi}_{y,e,\tilde{r}}(y) \end{pmatrix}$$

where the green entries indicate the row that is introduced in the $\xrightarrow[Y]{}$ *step.*

It is clear that we still have a one to one relation between the null vectors of the resulting pencil and the solutions of (5.1). Indeed, every nonzero vector $\boldsymbol{w}$ that satisfies

$$\begin{pmatrix} \hat{\Pi}_{x,\tilde{r}}(x) \\ \hat{\Pi}_{y,e,\tilde{r}}(y) \end{pmatrix} \boldsymbol{w} = \boldsymbol{0}$$

has a Vandermonde structure in the reduced monomial basis. Such a vector $\boldsymbol{w}$ is a scalar multiple of the monomial vector $\boldsymbol{v}(x,y)$ evaluated at a solution of (5.1).

5.3.2 A blockwise triangularization

Suppose Q_1 and Z are the unitary matrices from the generalized Schur decomposition of $\hat{\Pi}_{x,\tilde{r}}(x)$ so that

$$Q_1 \hat{\Pi}_{x,\tilde{r}}(x) Z = U_1 - xU_2$$

where U_1 and U_2 are upper triangular. We assume that the factorization is ordered so that all finite eigenvalues of $\hat{\Pi}_{x,\tilde{r}}(x)$ are found as the roots of the first n diagonal elements of $U_1 - xU_2$: for the i-th eigenvalue we find

$$x^{(i)} = \frac{(U_1)_{ii}}{(U_2)_{ii}}, \quad 1 \leq i \leq n.$$

We have

$$\begin{pmatrix} Q_1 & \\ & I \end{pmatrix} \begin{pmatrix} \hat{\Pi}_{x,\tilde{r}}(x) \\ \hat{\Pi}_{y,e,\tilde{r}}(y) \end{pmatrix} Z = \begin{pmatrix} U_1 - xU_2 \\ (\hat{\mathcal{B}}_{y,e,\tilde{r}} - y\hat{\mathcal{C}}_{y,e,\tilde{r}})Z \end{pmatrix},$$

where I is the identity matrix of appropriate size. Next, we take the QR-factorization of $\hat{\mathcal{C}}_{y,e,\tilde{r}} Z$:

$$\hat{\mathcal{C}}_{y,e,\tilde{r}} Z = Q_2 R.$$

This leads to

$$\begin{pmatrix} I & \\ & Q_2^* \end{pmatrix} \begin{pmatrix} U_1 - xU_2 \\ (\hat{\mathcal{B}}_{y,e,\tilde{r}} - y\hat{\mathcal{C}}_{y,e,\tilde{r}})Z \end{pmatrix} = \begin{pmatrix} U_1 - xU_2 \\ Q_2^* \hat{\mathcal{B}}_{y,e,\tilde{r}} Z - yR \end{pmatrix}.$$

We will use the following proposition to explain the rest of the reasoning.

Proposition 5.3.1. *For the generic system defined by* (5.9) *the number of rows of* $(\hat{\mathcal{B}}_{y,e,\tilde{r}} - y\hat{\mathcal{C}}_{y,e,\tilde{r}})$ *is greater than or equal to the number of finite solutions of* (5.1).

Proof. The matrix $(\hat{\mathcal{B}}_{y,e,\tilde{r}} - y\hat{\mathcal{C}}_{y,e,\tilde{r}})$ contains all of the possible linear y-recurrences needed to determine the reduced monomial basis corresponding to the columns of $\hat{L}_y(x,y)$. The total number of y-rows is equal to

$$n_y \triangleq \sum_{i=1}^{\hat{\delta}} i - \sum_{i=1}^{\Delta\delta_{y,\max}} i$$

where $\Delta\delta_{y,\max} \triangleq \max(\delta_p - \delta_p^y, \delta_q - \delta_q^y)$. The number of finite solutions is equal to the degree of the x-resultant, which is bounded by the number of rows in $\hat{\mathcal{B}}_{x,r} - x\hat{\mathcal{C}}_{x,r}$. This leads, using the results of Chapter 4 and Appendix C, to

$$n \leq \alpha - (\hat{\delta} + 1) - \gamma_m - s = \sum_{i=1}^{\hat{\delta}} i - \sum_{i=1}^{\Delta\delta_{y,\max}-1} i - \sum_{i=1}^{\delta_{y,\min}-1} i,$$

where $\delta_{y,\min} \triangleq \min(\delta_p^y, \delta_q^y)$. Therefore, if $\Delta\delta_{y,\max} - \sum_{i=1}^{\delta_{y,\min}-1} i \leq 0$ we have

$$n \leq \sum_{i=1}^{\hat{\delta}} i - \sum_{i=1}^{\Delta\delta_{y,\max}-1} i - \sum_{i=1}^{\delta_{y,\min}-1} i \leq \sum_{i=1}^{\hat{\delta}} i - \sum_{i=1}^{\Delta\delta_{y,\max}} i = n_y.$$

The condition $\Delta\delta_{y,\max} - \sum_{i=1}^{\delta_{y,\min}-1} i \leq 0$ is satisfied for a generic system (5.9) because generically $\Delta\delta_{y,\max} = 0$. □

Proposition 5.3.2. *For a generic system defined by* (5.9), $W \triangleq Q_2^* \hat{\mathcal{B}}_{y,e,\tilde{r}} Z$ *is upper triangular in its first* n *columns.*

Proof. For each $x \in X$, there must exist a value of $y \in \mathbb{C}$ such that

$$\begin{pmatrix} \hat{\Pi}_{x,\tilde{r}}(x) \\ \hat{\Pi}_{y,e,\tilde{r}}(y) \end{pmatrix}$$

loses full column rank. Note that from the $(n+1)$-st column on, there can be no linear dependencies. Indeed, for $k > n$ we have $(U_2)_{kk} = 0$ and $(U_1)_{kk} \neq 0$ because the k-th diagonal element corresponds to an infinite eigenvalue and the pencil is regular. We will consider only the first n columns and use a Matlab notation $(\cdot)_{:,1:n}$ to indicate this. Let $x = x^{(1)}$, then

$$\begin{pmatrix} U_1 - x^{(1)}U_2 \\ W - yR \end{pmatrix}_{:,1:n} = \left(\begin{array}{cccc} 0 & \times & \dots & \times \\ & \otimes & \dots & \times \\ & & \ddots & \vdots \\ & & & \otimes \\ & & & 0 \\ & & & \vdots \\ & & & 0 \\ \hline (W_{11} - yR_{11}) & \times & \dots & \times \\ \bullet & (W_{22} - yR_{22}) & \dots & \times \\ \bullet & \bullet & \ddots & \vdots \\ \bullet & \bullet & \bullet & (W_{nn} - yR_{nn}) \\ \bullet & \bullet & \bullet & \bullet \\ \vdots & \vdots & \vdots & \vdots \\ \bullet & \bullet & \bullet & \bullet \end{array}\right)$$

where we used the $\times$ sign to indicate any complex number, $\otimes$ to indicate nonzero complex numbers and $\bullet$ to indicate elements below the diagonal of W. Note that we used Proposition 5.3.1 for the visualization of the matrix. The diagonal elements of the upper block, except the first one, are all nonzero because of the assumption that, generically, there are no multiple eigenvalues. Because of the structure of the upper block, it is clear that the only possibility for this matrix to loose full column rank is when the first column is a zero column. This happens when all dots in the first column of the lower block are zero and $y = \frac{W_{11}}{R_{11}}$. Note that $R_{ii} \neq 0, 1 \leq i \leq n$

because $\hat{\mathcal{C}}_{y,e,\tilde{r}}$ is of full row rank by construction. Now, let $x = x^{(2)}$, we get

$$\begin{pmatrix} U_1 - x^{(2)}U_2 \\ W - yR \end{pmatrix}_{:,1:n} = \left(\begin{array}{cccc} \otimes & \times & \dots & \times \\ & 0 & \dots & \times \\ & & \ddots & \vdots \\ & & & \otimes \\ & & & 0 \\ & & & \vdots \\ & & & 0 \\ \hline (W_{11} - yR_{11}) & \times & \dots & \times \\ \bullet & (W_{22} - yR_{22}) & \dots & \times \\ \bullet & \bullet & \ddots & \vdots \\ \bullet & \bullet & \bullet & (W_{nn} - yR_{nn}) \\ \bullet & \bullet & \bullet & \bullet \\ \vdots & \vdots & \vdots & \vdots \\ \bullet & \bullet & \bullet & \bullet \end{array}\right).$$

A necessary condition for this matrix to be column rank deficient is again that all dots in the second column of the lower block are zero and $y = \frac{W_{22}}{R_{22}}$. We can repeat this reasoning for $x = x^{(i)}, 1 \leq i \leq n$, which proves the proposition. □

From the proof of Theorem 5.3.2 it is clear that in case $\hat{\Pi}_{x,\tilde{r}}(x)$ has only simple eigenvalues, the n finite solutions to (5.1) are found as

$$\left(\frac{(U_1)_{ii}}{(U_2)_{ii}}, \frac{W_{ii}}{R_{ii}}\right), 1 \leq i \leq n.$$

Applying the perturbation (5.9), the eigenvalues of $\hat{\Pi}_{x,\tilde{r}}(x)$ are generically all simple (in quadruple precision) and we can apply this result. The method gives small residuals but it requires more computation time because the calculations for the eigenvalue problem and the factorizations have to be performed in quadruple precision.

Chapter 6

Numerical Results

In this chapter we discuss the numerical results that are obtained by the method that is proposed in this text. We will also compare different versions of our method with each other and with other existing solvers. We will use *performance profiles* [13] to visualize the overall performance of one method with respect to another. Such a profile is constructed as follows. Suppose we are comparing all solvers in a set S on a set of problems P. The *performance curve* for a solver $s \in S$ is given by

$$\rho_s(\tau) = \frac{|\{p \in P \mid t_{p,s} \leq 2^\tau (\min_{s \in S} t_{p,s})\}|}{|P|} \tag{6.1}$$

where $|\cdot|$ denotes the cardinality of a set and $t_{p,s}$ denotes the time it took for solver s to "successfully" solve problem p. If solver s did not solve s successfully, $t_{p,s} = \infty$. What is meant by solving a problem "successfully" will be explained for the different comparisons. Note that the curve $\rho_s(\tau)$ is nondecreasing. $\rho_s(0)$ is the percentage of the problems in P that was solved by solver s the fastest. As τ increases, $\rho_s(\tau)$ converges to the percentage of problems that was solved successfully by solver s. In the first section we compare the different versions of our method (Chapter 5). In the next section we will compare our method to PHClab [31, 30], Bertini [3] and the PNLA package [4, 14]. These packages are chosen because they also intend to find all solutions with the correct multiplicities. All methods are tested on a set of 60 lower degree problems that range from simple problems to problems with challenging multiplicity structures and on a set of random problems of total degree $\delta = 1, \ldots, 40$. By "random" we mean they have normally distributed random coefficients with mean 0 and standard deviation 1 for all monomials of degree $\leq \delta$. All numerical results are reported in the tables in Appendix F. Finally, in a last section, we will give some interesting numerical examples. All of our software is implemented in Matlab R2015b. Some examples that illustrate how to use the programs are given in Appendix H. PNLA is implemented in Matlab but the QR-factorization of sparse matrices is done by a routine from SuiteSparse [10], which is implemented in C++. PHCpack is implemented in C and the Matlab interface PHClab is used for the experiments. Bertini is implemented in C++. All experiments are done on a server with an Intel Core i7-5500U CPU and 16 GB RAM.

6.1 Testing the approaches of Chapter 5

The different versions of our method are presented in Chapter 5. To evaluate their performance, each version is tested on a set of 60 problems of which the essential information is given in Table F.2 in Appendix F. The reference solution set $\mathcal{S}_{\text{ref}}$ is calculated using Bertini with adaptive precision, which is a very reliable solver but, due to the variable precision, not an efficient one.

For the methods described in 5.1.1 (using the eigenvectors of $\hat{\Pi}_{x,r}(x)$), 5.1.2 (using the Sylvester matrix) and 5.1.3 (using $L(x,y)$) we cannot expect to find the solutions with the correct multiplicities. Therefore, the criterion for "success" in this case is chosen to be the following. Let $\tilde{S}$ be the set of numerically found solutions. Recall that in Definition 5.1, the residual of a solution (x^*, y^*) is defined as

$$r(x^*, y^*) = \frac{|p(x^*, y^*)|}{|p|(|x^*|, |y^*|) + 1} + \frac{|q(x^*, y^*)|}{|q|(|x^*|, |y^*|) + 1}$$

where $|p|(x,y) \triangleq \sum_{i=0}^{\delta} \sum_{j=0}^{\delta-i} |p_{ij}| x^j y^i$ and $|q|(x,y) \triangleq \sum_{i=0}^{\delta} \sum_{j=0}^{\delta-i} |q_{ij}| x^j y^i$. A problem is successfully solved if all residuals are smaller than 10^{-6} (backward error) and for every reference solution $\boldsymbol{s} \in \mathcal{S}_{\text{ref}}$ there is a solution $\tilde{\boldsymbol{s}} \in \tilde{S}$ such that

$$\|\boldsymbol{s} - \tilde{\boldsymbol{s}}\|_2 \leq 10^{-2}(1 + \|\boldsymbol{s}\|_2)$$

(forward error). This criterion leads to the performance curves in Figure 6.1. The figure shows that the method using the Sylvester matrix (——) solves more than 50% of the problems in the fastest way. There is, however, one problem that is not successfully solved by this version. Relaxing the criterion for success to a 2% allowable forward error, all performance curves reach the value 1. Details can be found in Table F.3 and Table F.4 in Appendix F. Note that all three methods in many cases find too many solutions. This is due to the fact that several y-values might be coupled to the same x-value and the solutions are not clustered or filtered. The success criterion of 99% forward accuracy is for most problems very mild. In general, the forward error is much smaller than 1%. The forward accuracy depends strongly on the multiplicity properties of the problem.

For the versions of our method that intend to find the multiplicities of all solutions, we must define the notion of "successfully" solving a problem in another way. We will use the following criterion. A problem is successfully solved, taking multiplicities into account if the number of solutions is equal to the number of reference solutions and the solutions pass the following test. There must be a bijective map $b : \mathcal{S}_{\text{ref}} \to \tilde{\mathcal{S}}$ such that for each solution $\boldsymbol{s}$ in $\mathcal{S}_{\text{ref}}$ we have

$$\|\boldsymbol{s} - \tilde{\boldsymbol{s}}\|_2 \leq 10^{-2}(1 + \|\boldsymbol{s}\|_2),$$

where $\tilde{\boldsymbol{s}} \triangleq b(\boldsymbol{s})$. This test is implemented using the `bipartite_matching` function from the gaimc Matlab software for graph algorithms [15]. This criterion is used

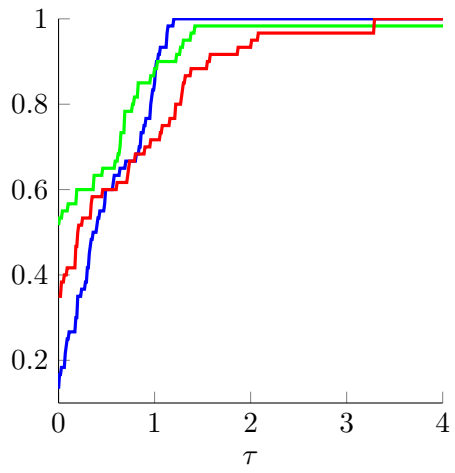


Figure 6.1: Performance profile for the comparison of the methods described in 5.1.1 (——), 5.1.2 (——) and 5.1.3 (——).

for the coupling method based on connection diagrams (5.2), the variable precision method (5.3) and the Sylvester matrix method (5.1.2) with a linear transformation of variables of the following form. Let θ be a random number with a uniform distribution between $3\pi/8$ and $5\pi/8$. The transformed system is

$$\begin{cases} p_t(x,y) = p(T_{11}x + T_{12}y, T_{21}x + T_{22}y) = 0 \\ q_t(x,y) = q(T_{11}x + T_{12}y, T_{21}x + T_{22}y) = 0 \end{cases}$$

where

$$T = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

This transformation is chosen for now because it is the one that gives the best results among the transformations that are experimented with. As mentioned before, this transformation of variables has bad numerical consequences. The results are shown on the left part of Figure 6.2. It can be seen that the transformation of variables causes the method to fail for about 20% of the problems. Nearly all problems are solved successfully by the variable precision method, but it requires more computational effort. The blue performance curve shows that the algorithm based on the connection diagrams solves all problems accurately and quickly. Details can be found in Table F.5 and F.6 in Appendix F.

6.2 Comparison with other solvers

In this section we compare the version of our method that finds a coupling between the x- and y-values based on connection diagrams to the solvers Bertini (in double precision) [3], PHClab [31, 30] and PNLA [14, 4]. Bertini and PHClab use homotopy continuation, whereas PNLA is based on Macaulay resultants. The two-parameter

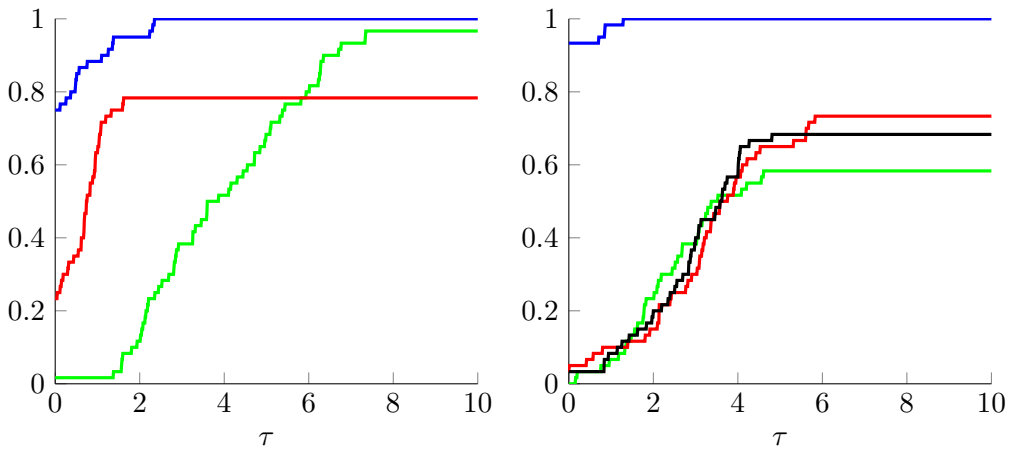


Figure 6.2: Left: performance profile for the comparison of the methods described in 5.2 (——) (connection diagrams), 5.3 (——) and 5.1.2 with a transformation of variables (——). Right: performance profile for the comparison of the coupling method using connection diagrams 5.2 (——), PHClab (——), Bertini (——) and PNLA (——).

eigenvalue approach [23] also intends to find all solutions with their corresponding multiplicities. We have not used this method for the comparison with other solvers because, as described in Section 1.3.2, the size of the resulting generalized eigenvalue problem is too large. In [23], Plestenjak and Hochstenbach propose two determinantal representations of the bivariate rootfinding problem that reduce the resulting pencil size. Figure 6.3 illustrates that for degrees higher than 4 the two-parameter eigenvalue approach results in a larger generalized eigenvalue problem than the one that is solved using our method. The performance profile for our solver, PHClab, Bertini and PNLA is given on the right part of Figure 6.2. We used standard settings for all methods, no refinement options or variable precision options are used[1]. It can be seen from the figure that the other solvers solve less than 80% of the problems in the dataset successfully[2]. Moreover, the coupling method is faster than all other solvers for more than 90% of the problems. We stress the fact that the results for the other solvers could be improved by changing the settings. Recent versions of PHCpack offer the possibility to perform calculations in higher precision. This would lead to much more accurate results but it would take more computation time. PNLA offers the possibility to calculate the so called "radical system" associated to a given problem which is such that the roots remain the same but they all become simple. This leads to much higher accuracy.

To give an idea about the performance of the solvers for higher degree problems

[1]For Bertini we used `MPTYPE: 0` for double precision, since the default setting (`MPTYPE: 1`) uses adaptive precision. For PNLA, the `sparf` function is used.

[2]The success criterion is still the one that was used to compare the coupling method to the variable precision method.

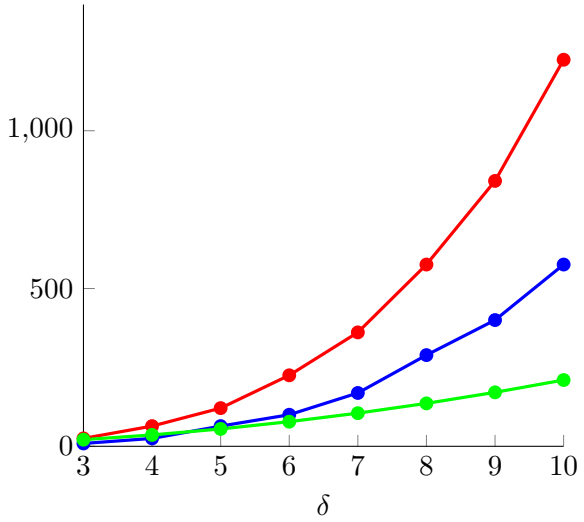


Figure 6.3: Size of the generalized eigenvalue problem constructed using the first (—●—) and the second (—●—) linearization proposed in [23] and the (pessimistic) upper bound $2\delta^2 + \delta$ for the size of $\hat{\Pi}_{x,r}(x)$ (—●—), with respect to the degree δ.

we have generated a set of "generic" problems of degree $\delta = 1, \ldots, 40$ in the following way. The polynomials p and q that define the generic problem of degree δ have normally distributed coefficients with mean 0 and standard deviation 1 corresponding to all monomials of degree $\leq \delta$. The number of solutions to such a generic problem is δ^2 (from Bézout's theorem) and all solutions are generically simple and finite. The criterion for success for these problems is taken to be the following. A generic problem of degree δ is solved successfully if a number of solutions between $0.99\delta^2$ and δ^2 is found, all with a residual $< 10^{-6}$. The results are shown in figure 6.4 and details can be found in Table F.9 and Table F.10 in Appendix F. The performance curve for PNLA is not plotted because the tests have not been completed for this solver. It took PNLA about 25 hours to solve the generic problem of degree 25, so the tests were too time consuming. The method `qdsparf` from PNLA solves the problems much faster than `sparf` but the accuracy deteriorates. Results for `qdsparf` are not reported. As can be seen from Figure 6.4, our solver has passed the test for success for all 40 problems. Bertini has for more than 90% of the problems and PHClab for 70% of them. It must be noted that the higher degree problems are solved faster by Bertini and PHClab and the obtained residuals are very small. However, the timing results in Figure 6.4 (right part) show that our solver is competitive at least up to degree 40. It may take longer, but all δ^2 solutions have been found for each problem (Table F.10). The residual for our coupling method increases with the degree of the problem. Then again, the obtained solutions are good starting values for Newton-Raphson refinement, which has not yet been implemented. When it comes to computation time, PHClab is the absolute winner for higher degree problems. It finds consistently more than 98%, but not 99% of all solutions.

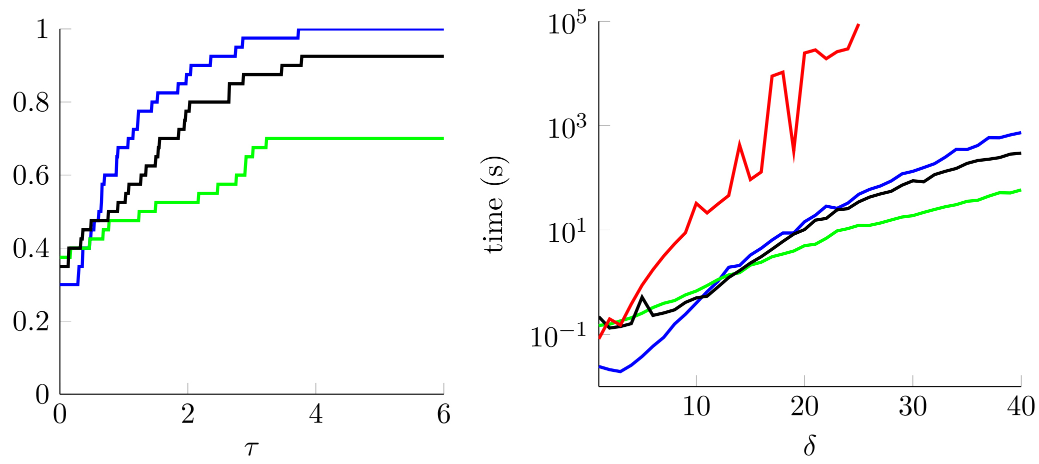


Figure 6.4: Left: performance profile for the solver described in this text, based on a coupling using connection diagrams (—), PHClab (—) and Bertini (—) for random test problems of degree 1 up to 40. Right: timing results for all four solvers, the computation time for PNLA is indicated in red (—).

6.3 Some interesting examples

In this section we will give a few examples of bivariate problems and the solutions found using the version of our method described in Section 5.2 based on connection diagrams.

- Consider the system given by

$$\begin{cases} p(x,y) = -4 + 5x - 3x^2 + x^3 + 5y - 2xy - 3y^2 + y^3 = 0 \\ q(x,y) = -4 + x - 2x^2 + 2x^3 + 9y + 2xy - 4x^2y - 8y^2 + 3xy^2 + y^3 = 0 \end{cases}, \tag{6.2}$$

which has only real finite solutions, among which $(2,2)$ is a 3-fold and $(1,1)$ a 5-fold zero. The real picture of the zero sets of p and q is given in Figure 6.5. The numerical solutions found by our solver are given in Table 6.1. Note that even the 5-fold zero is found up to machine precision. Residuals are plotted in the right part of Figure 6.5.

$\Re(x)$	$\Im(x)$	$\Re(y)$	$\Im(y)$
$2.00000000000000 \cdot 10^{0}$	$4.59237660016485 \cdot 10^{-15}$	$2.00000000000002 \cdot 10^{0}$	$6.48002114568442 \cdot 10^{-15}$
$2.00000000000000 \cdot 10^{0}$	$4.59237660016485 \cdot 10^{-15}$	$2.00000000000002 \cdot 10^{0}$	$6.48002114568442 \cdot 10^{-15}$
$2.00000000000000 \cdot 10^{0}$	$4.59237660016485 \cdot 10^{-15}$	$2.00000000000002 \cdot 10^{0}$	$6.48002114568442 \cdot 10^{-15}$
$1.57142857142857 \cdot 10^{0}$	$1.20727320460183 \cdot 10^{-15}$	$-1.42857142857144 \cdot 10^{-1}$	$-1.52306944438596 \cdot 10^{-16}$
$1.00000000000000 \cdot 10^{0}$	$-5.98962175057873 \cdot 10^{-16}$	$1.00000000000000 \cdot 10^{0}$	$9.60530618801872 \cdot 10^{-16}$
$1.00000000000000 \cdot 10^{0}$	$-5.98962175057873 \cdot 10^{-16}$	$1.00000000000000 \cdot 10^{0}$	$9.60530618801872 \cdot 10^{-16}$
$1.00000000000000 \cdot 10^{0}$	$-5.98962175057873 \cdot 10^{-16}$	$1.00000000000000 \cdot 10^{0}$	$9.60530618801872 \cdot 10^{-16}$
$1.00000000000000 \cdot 10^{0}$	$-5.98962175057873 \cdot 10^{-16}$	$1.00000000000000 \cdot 10^{0}$	$9.60530618801872 \cdot 10^{-16}$
$1.00000000000000 \cdot 10^{0}$	$-5.98962175057873 \cdot 10^{-16}$	$1.00000000000000 \cdot 10^{0}$	$9.60530618801872 \cdot 10^{-16}$

Table 6.1: Numerical solutions to (6.2). Every row represents one solution.

- Figure 6.6 shows the real part of the zero level sets of Problem 7 and Problem 51 from the test problem set. Some information about these problems can be

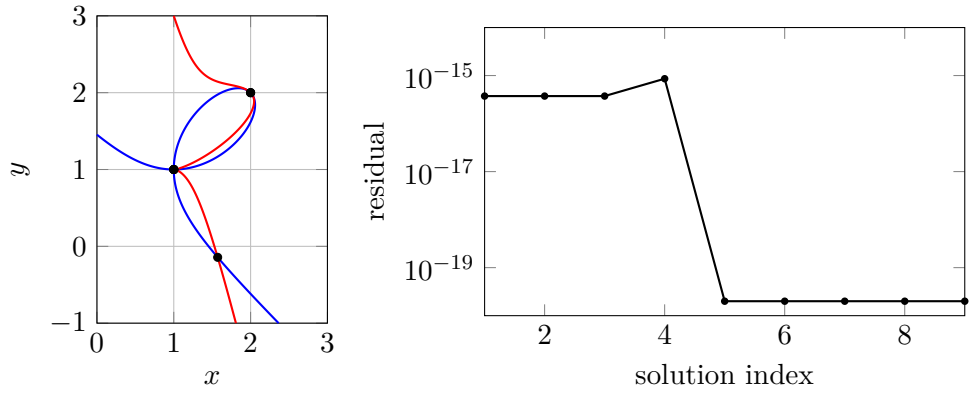


Figure 6.5: Left: real zero level lines of p (—) and q (—) from (6.2) and the real part of the numerical solutions (•). Right: residual for all 9 numerical solutions.

found in Table F.2 in Appendix F. The results[3] are shown in Figure 6.7 and in Figure 6.8. For Problem 7 all residuals are very small, the 18-fold zero is found up to machine precision. For problem 51, all solutions are simple. Residuals are small and all reference solutions are retrieved with a forward error of order 10^{-12} or smaller.

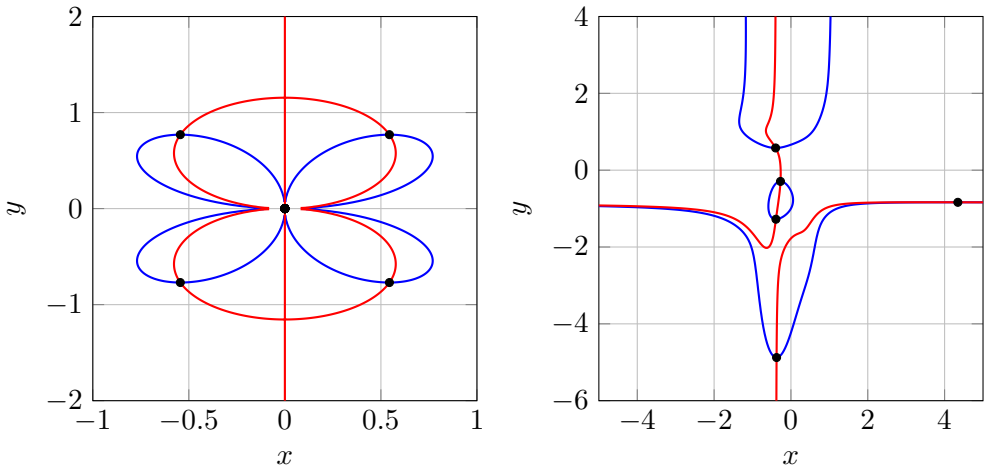


Figure 6.6: Real picture of the zero level sets of Problem 7 (left) and Problem 51 (right) from the test problem set. Black dots (•) indicate the numerical solutions.

[3]For solutions with a very small residual ($< 10^{-20}$) the value 10^{-20} is plotted to avoid trouble with taking the logarithm of zero.

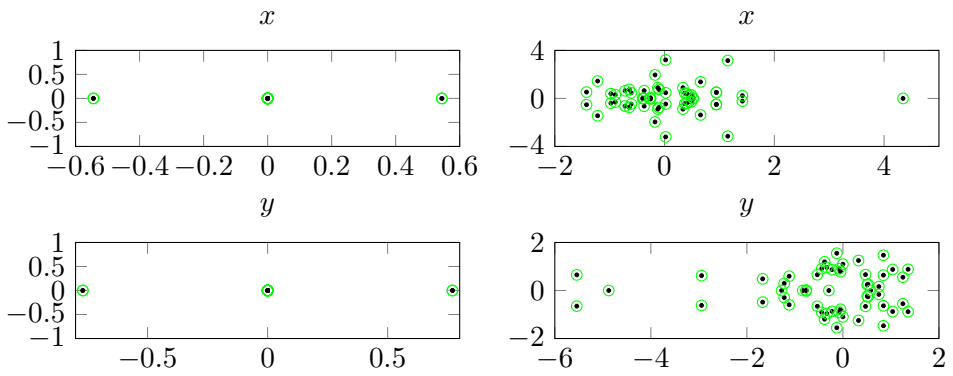


Figure 6.7: Left: x- and y-coordinates of the reference solutions ($\cdot$) and the found solutions ($\circ$) in the complex plane for Problem 7. Right: the same picture for Problem 51.

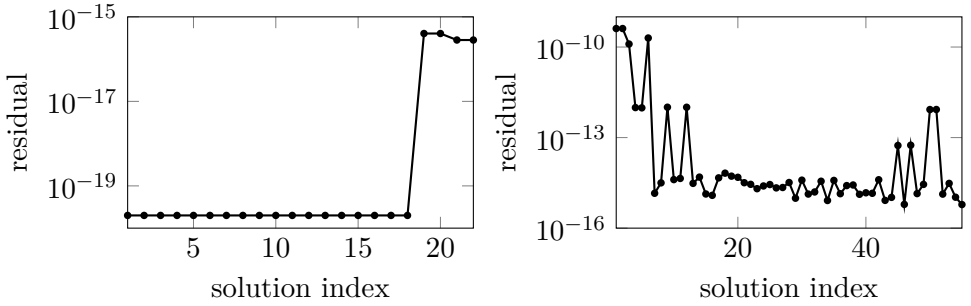


Figure 6.8: Left: residual for all 22 solutions of Problem 7. Right: residual for all 55 solutions of Problem 51.

- Figure 6.9 shows the real part of the zero level sets of two other examples that are not in the test problem set. The real numerical solutions are indicated with black dots. For the first problem (left part of Figure 6.9) p has degree 10 and q has degree 9. There are 90 solutions and they are all found with small residuals. For the second problem p has degree 16 and q has degree 15. All solutions have a multiplicity > 1 and there are 192 solutions in total (counting multiplicities). For example, the point $(0, 0)$ is a 16-fold solution. Results are shown in Figure 6.10 and Figure 6.11. The residuals are higher for the second problem because of the higher degree and the multiplicity structure of this problem. For both problems, the numerical solution set $\tilde{\mathcal{S}}$ satisfies the success criterion as formulated for the solvers that intend to find the multiplicities for the test set of 60 problems.

- Let us now consider a random problem (as specified previously) of degree 20. The real zero level sets are given in Figure 6.12. All 400 solutions are found with a residual of order 10^{-12} or smaller. Doing some timing tests, we conclude that about 90% of the time that our solver needs for solving this problem is

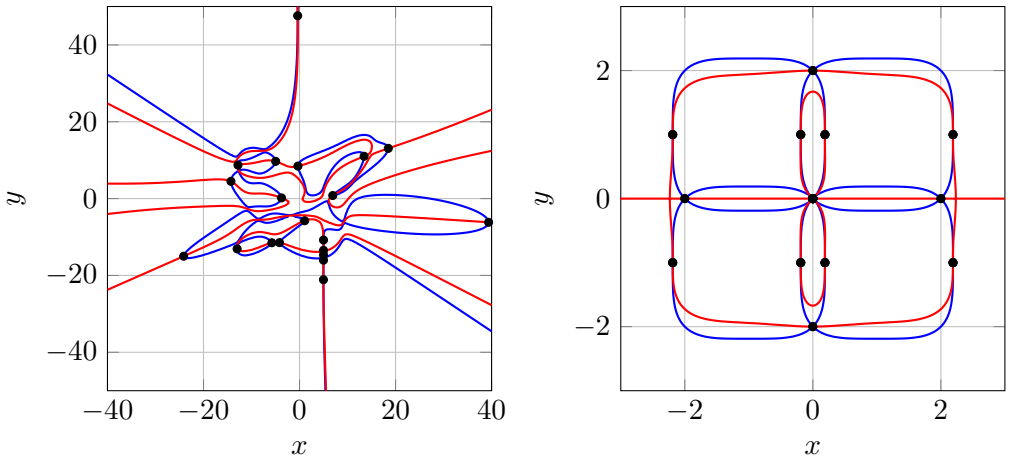


Figure 6.9: Real picture of the zero level sets of two example problems. Black dots (•) indicate the numerical solutions.

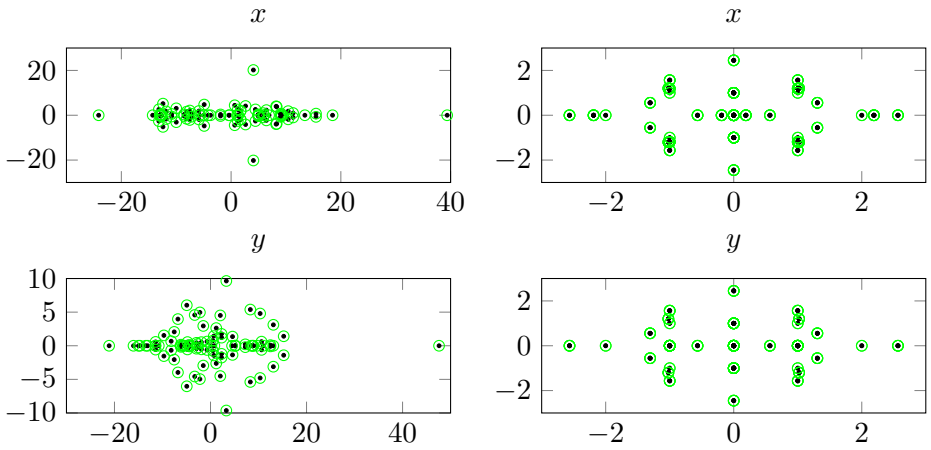


Figure 6.10: Left: x- and y-coordinates of the reference solutions (·) and the found solutions (○) in the complex plane for the problem on the left side of Figure 6.9. Right: the same picture for the right problem of Figure 6.9.

used for constructing and solving the two generalized eigenvalue problems for finding all values of x and y. The time that is needed to calculate the residual matrix is only a third of a percent of the total time. The remaining ±10% of the time is needed to construct and to solve the system for finding the couples and their multiplicities. This is a general result: timings for other random problems confirm this partitioning of the total computation time. The same is true for problems of other degrees. Of the first ±90% of the time, a negligible part (about 2% for $\delta = 20$) is spent on constructing the generalized eigenvalue problems. The majority of the computation time is spent on applying the QZ-algorithm.

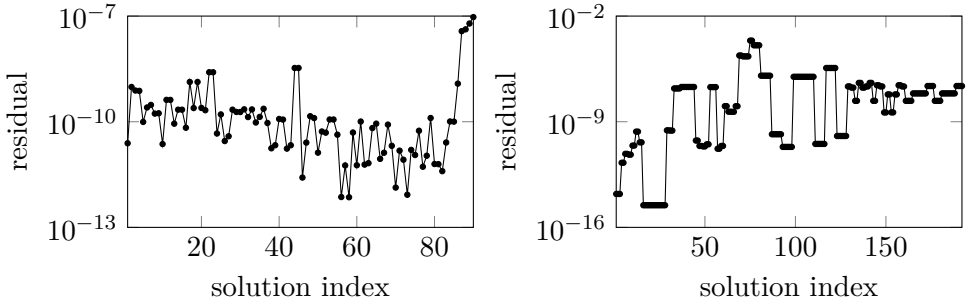


Figure 6.11: Left: residual for all 90 solutions of the first problem. Right: residual for all 192 solutions of the second problem.

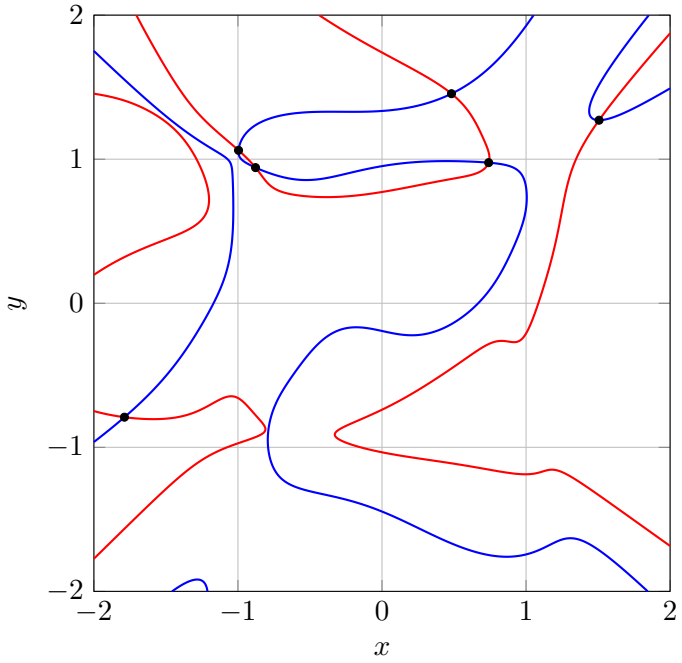


Figure 6.12: Real picture of the zero level sets of a generic problem of degree 20. Black dots (•) indicate the numerical real solutions.

Chapter 7

Conclusion and Future Work

In this final chapter, we summarize the most important results and we present some suggestions for future work.

7.1 Conclusions

In this text, we have presented a numerical linear algebra based solver for 0-dimensional bivariate systems of polynomial equations. A square generalized eigenvalue problem is constructed in an intuitive way. We have shown that the obtained eigenvalues coincide with the projections of the solutions onto the complex plane associated with one out of the two variables (except for some spurious eigenvalues that are easy to detect). The other coordinates can be found in various ways. Due to the strong connection between the GEP that is solved in our approach and the Sylvester resultant, the solver is capable of retrieving the correct multiplicity information about the solutions. The pencil is such that the coefficients of the given polynomials can be plugged in directly, without any manipulations. This is advantageous for the accuracy. From our numerical experiments we learn that the solver manages to find all solutions with the right multiplicities for moderate degree problems ($\delta \leq 20$). Solutions are found fast and with high accuracy, even though no Newton-Raphson refinement is used. All solutions are found for generic problems of degree at least 40. For higher degree problems, homotopy continuation based solvers are faster but they fail to retrieve all solutions for $\delta > 35$ (using standard double precision settings).

7.2 Future work

We will now list some suggestions for future work.

- The pencil $\hat{\Pi}_{x,r}(x)$ whose eigenvalues are the x-coordinates of the solutions has many infinite eigenvalues. Numerically, however, these eigenvalues might be finite but large. If they are very large, such spurious "infinite" eigenvalues are easy to detect. However, it turns out that even for moderate degrees these eigenvalues might be of magnitude $\mathcal{O}(10^3)$ or even $\mathcal{O}(10^2)$. Such eigenvalues

lead to large solutions that still have a small residual. It is therefore needed to find a way to distinguish such spurious large solutions from actual solutions that have a large norm. This is essentially a scaling problem.

- It can be seen from Table F.3 that the methods presented in Section 5.1 find too many solutions in many cases. These solutions could be clustered or filtered in order to obtain only one good representative for each isolated solution of the bivariate problem.

- Although the title of this thesis sounds more general, we have mainly focused on bivariate problems. The ideas can, however, be extended to higher dimensions. An example in $\mathbb{C}^3$ is given in Appendix G. The method is based on the same principles (multi-parameter eigenvalue problems, degree extension) as our bivariate solver. Of course, the connection with the Sylvester resultant is lost and there is no guarantee (yet) of finding a square pencil in one variable. A detailed description and analysis of the generalization to higher dimensions needs further research.

- In order to avoid solutions with the same x- or y-coordinates we have suggested to apply a generic transformation of variables. Out of all transformations that are experimented with, a generic rotation over an angle between $3\pi/8$ and $5\pi/8$ radians gives the best results. As shown in Chapter 6, the accuracy deteriorates significantly. Finding a generic transformation that does not have these unwanted effects could enable us to find the correct multiplicities using the methods presented in Section 5.1 and it could enhance the results for systems that have solutions at infinity with one finite coordinate.

- As stated before, for the coupling method based on connection diagrams roughly 90% of the computation time is devoted to solving square generalized eigenvalue problems. The pencil $\hat{\Pi}_{x,r}(x)$ is sparse and it is very structured. Efficiency might be enhanced by further exploiting this structure. Also, if one is only (or mainly) interested in finding solutions in a certain region of $\mathbb{C}^2$, iterative methods could be used to speed up the calculations.

- In Chapter 4 we have shown that the choice of basis is not restricted to the classical monomial basis. It is an open question how much improvement can be made by representing p and q in other product tensor bases. For example, the Chebyshev basis is expected to give better results for real solutions in the square $[-1, 1] \times [-1, 1]$.

- As mentioned in Section 5.2, the linear system that is solved for finding the right multiplicities might be underdetermined. It is arguable that, instead of using the minimal norm least squares solution, the system should be solved as an integer programming problem in the underdetermined case. Again, this can be avoided by performing a transformation of variables.

Appendices

Appendix A

Polynomial Ideals, their Quotient Rings and Dual Spaces

In this appendix, we give a brief introduction to polynomial ideals and their significance for the understanding of the structure of solution sets of multivariate polynomial systems [26]. The emphasis will be on notions that are useful to gain insight in the multiplicity structure of a 0-dimensional solution set.

A.1 Polynomial ideals

Consider a set of n polynomials $\mathcal{S} = \{p_i\}_{1 \leq i \leq n}$ in s variables: $p_i \in \mathcal{P}^s, \forall i$. A linear combination of the elements of $\mathcal{S}$ is any polynomial $p \in \mathcal{P}^s$ that can be written as

$$p = \sum_{i=1}^{n} c_i p_i$$

for some set of complex coefficients $\{c_i\}_{1 \leq i \leq n}$. One can extend this notion of a linear combination to *polynomial combinations* in the following way.

Definition A.1 (polynomial combination)**.** *Any polynomial p that can be written as*

$$p = \sum_{i=1}^{n} c_i p_i$$

for some set $\{c_i\}_{1 \leq i \leq n}$ and $c_i \in \mathcal{P}^s, \forall i$ is called a polynomial combination *of the elements of $\mathcal{S}$.*

Now, just as a subspace of an s-dimensional Euclidian vectorspace, say $\mathbb{R}^s$, is defined as a set which is closed under linear combination, we introduce the concept of a *polynomial ideal* through the following definition.

Definition A.2 (polynomial ideal)**.** *A* polynomial ideal *in $\mathcal{P}^s$ is a subset of $\mathcal{P}^s$ which is closed under polynomial combination. The ideal that contains all polynomial combinations of the polynomials in $\mathcal{S}$ is said to be* generated *by $\mathcal{S}$ and is denoted by*

$$\mathcal{I}_{\mathcal{S}} = \langle \mathcal{S} \rangle = \langle p_1, p_2, \ldots, p_n \rangle.$$

The polynomials p_i in $\mathcal{S}$ are called a basis *for $\langle \mathcal{S} \rangle$.*

Given this definition of a polynomial ideal, consider a set Z of points in $\mathbb{C}^s$. A polynomial $p \in \mathcal{P}^s$ is said to "vanish at Z" if it satisfies $p(\boldsymbol{z}) = 0, \forall \boldsymbol{z} \in Z$. One can verify that the set $\mathcal{P}^s_Z$ of all polynomials in $\mathcal{P}^s$ that vanish at Z is an ideal in $\mathcal{P}^s$. Indeed, any polynomial that is a polynomial combination of elements of $\mathcal{P}^s_Z$ must vanish at Z, so $\mathcal{P}^s_Z$ is closed under polynomial combinations. Conversely, let us start from a set $\mathcal{S}$ of polynomials. Suppose all polynomials in $\mathcal{S}$ vanish at some point $z \in \mathbb{C}^s$, then it must hold by definition that every polynomial in $\langle \mathcal{S} \rangle$ vanishes at z. This illustrates the close connection between polynomial ideals and zero sets of systems of polynomials.

Definition A.3 (zero set of an ideal)**.** *Given an ideal $\mathcal{I}$, the zero set of $\mathcal{I}$ is the set*

$$Z(\mathcal{I}) = \{\boldsymbol{z} \in \mathbb{C}^s \mid p(\boldsymbol{z}) = 0, \forall p \in \mathcal{I}\}$$

of all points at which all of the polynomials in $\mathcal{I}$ vanish.

In particular, suppose $\mathcal{S}$ is a set of s polynomials that defines a 0-dimensional system like (1). All points $\boldsymbol{z} \in \mathbb{C}^s$ of the solution set of the associated polynomial system are zeros of the ideal. Also, all zeros of the ideal $\langle \mathcal{S} \rangle$ are solutions to the system defined by $\mathcal{S}$. Calculating the solutions to (1) defined by the polynomials $\{p_i\}_{1 \leq i \leq s}$ in $\mathcal{S}$ is equivalent to determining the zeros of $\langle \mathcal{S} \rangle$.

A.2 The quotient ring of $\langle \mathcal{S} \rangle$

Let $\mathcal{S} = \{p_i\}_{1 \leq i \leq n}$. For ease of notation, denote $\langle \mathcal{S} \rangle$ by $\mathcal{I}$ and suppose $\mathcal{Z}(\mathcal{I})$ is 0-dimensional. Let $\boldsymbol{z} \in Z(\mathcal{I})$ be a point in the zero set of $\mathcal{I}$ and consider a polynomial p that satisfies $p(\boldsymbol{z}) = \alpha$. From the previous discussion it is clear that if $\alpha \neq 0$, $p \notin \mathcal{I}$. We can write p as the sum of a polynomial in $\mathcal{I}$ and some remainder term r:

$$p = \sum_{i=1}^{n} c_i p_i + r$$

where the first term represents a polynomial combination of the polynomials in $\mathcal{S}$. It follows directly from $p(\boldsymbol{z}) = \alpha$ that (independent from the choice of the coefficients c_i), $r(\boldsymbol{z}) = \alpha$. More generally, the remainder term r of any polynomial p takes on the same values $\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$ in the respective points $Z(\mathcal{I}) = \{\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_m\}$ as p. If $p \in \mathcal{I}$, it vanishes at $Z(\mathcal{I})$ and so does r. The converse is not necessarily true.

Example A.2.1. *Consider the set $\mathcal{S} = \{x^2 + y^2 - 1, x^2 + 4y^2 - 1\}$ and the corresponding ideal in $\mathcal{P}^2$: $\mathcal{I} = \langle x^2 + y^2 - 1, x^2 + 4y^2 - 1 \rangle$. The zero set of this ideal is given by*

$$Z(\mathcal{I}) = \{(-1, 0), (1, 0)\}.$$

The polynomial $p = y \in \mathcal{P}^2$ vanishes at $Z(\mathcal{I})$, yet it is not a member of $\mathcal{I}$. Indeed, there is no polynomial combination of the members of $\mathcal{S}$ that results into p.

However, one can show that if p vanishes at $Z(\mathcal{I})$ and all of the points in $Z(\mathcal{I})$ are *simple* (have multiplicity 1), then $p \in \mathcal{I}$. In other words, in the case of an ideal with all simple isolated zeros, the question whether a polynomial p is a member of $\mathcal{I}$ can be answered by simply evaluating p at $Z(\mathcal{I})$ and checking whether it vanishes or not. The ideal is completely characterized by the evaluations at $Z(\mathcal{I})$. In this case we could subdivide $\mathcal{P}^s$ into equivalence classes that contain polynomials that take on the same values at $Z(\mathcal{I})$. One such equivalence class, the one that vanishes at $Z(\mathcal{I})$, coincides with $\mathcal{I}$.

Definition A.4. *Given two polynomials p and q in $\mathcal{P}^s$. We say that p and q are equivalent with respect to $\sim_{\mathcal{I}}$ if the difference between p and q is in $\mathcal{I}$. We write*

$$p \sim_{\mathcal{I}} q \Leftrightarrow p - q \in \mathcal{I}.$$

Note that in the simple case (all simple zeros), $\sim_{\mathcal{I}}$ realizes exactly the subdivision of $\mathcal{P}^s$ we described earlier. Indeed, in that case

$$p \sim_{\mathcal{I}} q \Leftrightarrow p(Z(\mathcal{I})) = q(Z(\mathcal{I})). \tag{A.1}$$

If not all zeros are simple, it is still a necessary (but not sufficient) condition for the equivalence:

$$p \sim_{\mathcal{I}} q \Rightarrow p(Z(\mathcal{I})) = q(Z(\mathcal{I})). \tag{A.2}$$

Any polynomial $p \in \mathcal{P}^s$ can be associated to its so called *residue class* $\bmod \mathcal{I}$.

Definition A.5. *The set*

$$[p]_{\mathcal{I}} \triangleq \{r \in \mathcal{P}^S : p - r \in \mathcal{I}\}$$

is called the residue class of $p \mod \mathcal{I}$. *It is the set of all remainders of p modulo the ideal $\mathcal{I}$.*

Note that this notation uses p as a representative for its residue class. That is, for any $q \in \mathcal{P}^s$ that satisfies $p \sim_{\mathcal{I}} q$ it holds that $[p]_{\mathcal{I}} = [q]_{\mathcal{I}}$. In particular, for $p \in \mathcal{I}$ we have that $[p]_{\mathcal{I}} = [0]_{\mathcal{I}} = \mathcal{I}$. The set of all distinct residue classes is denoted by $\mathcal{R}[\mathcal{I}]$ and it is called the *quotient ring* of the polynomial ideal $\mathcal{I}$. In $\mathcal{R}[\mathcal{I}]$, define the addition and scalar multiple operations as follows. Let $p, q \in \mathcal{P}^s$ and $\alpha \in \mathbb{C}$, then

$$\alpha \cdot [p]_{\mathcal{I}} \triangleq [\alpha p]_{\mathcal{I}}$$

$$[p]_{\mathcal{I}} + [q]_{\mathcal{I}} \triangleq [p + q]_{\mathcal{I}}$$

and it can be shown that $[\alpha p]_{\mathcal{I}}$ and $[p + q]_{\mathcal{I}}$ are independent of the choice of representatives p and q for the considered residue classes. This means $\mathcal{R}[\mathcal{I}]$ is a linear vector space[1]. When all zeros of $\mathcal{I}$ are simple, we have shown that any residue class is completely defined by the values it takes on the set of zeros $Z(\mathcal{I})$. Therefore, we have

$$\dim(\mathcal{R}[\mathcal{I}]) = |Z(\mathcal{I})| = m.$$

[1] In fact, $\mathcal{R}[\mathcal{I}]$ is called a quotient *ring* because also the multiplication

$$[p]_{\mathcal{I}} \cdot [q]_{\mathcal{I}} = [pq]_{\mathcal{I}}$$

is well-defined and commutative. So $\mathcal{R}[\mathcal{I}]$ is a *commutative ring*.

A.3 Dual spaces of polynomial ideals

Consider again the quotient ring $\mathcal{R}[\mathcal{I}]$ where we assume $\mathcal{I}$ to be an ideal with a finite set of simple zeros $Z(\mathcal{I}) = \{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m\}$. We have shown that $\mathcal{R}[\mathcal{I}]$ is a vector space of dimension m. To any vector space, a *dual space* is associated.

Definition A.6. *The dual space V^* of a vector space V is the vector space of all linear functionals $l : V \to \mathbb{C}$, where any such linear functionals $l, l_1, l_2 \in V^*$ satisfy, for any $v \in V$, any $\alpha \in \mathbb{C}$,*

$$(\alpha l)(v) \triangleq \alpha l(v) \quad \textit{and} \quad (l_1 + l_2)(v) \triangleq l_1(v) + l_2(v).$$

It can be shown that for finite dimensional vector spaces, the associated dual spaces have the same dimension. We will denote

$$\mathcal{R}[\mathcal{I}]^* \triangleq \mathcal{D}[\mathcal{I}]$$

and we have $\dim(\mathcal{D}[\mathcal{I}]) = \dim(\mathcal{R}[\mathcal{I}]) = m$.

Definition A.7. *For $\boldsymbol{j} = \begin{pmatrix} j_1 & j_2 & \ldots & j_s \end{pmatrix}^\top \in \mathbb{N}^s$ and for $\boldsymbol{z} \in \mathbb{C}^s$, the* differential functional $\partial_{\boldsymbol{j}}[\boldsymbol{z}] \in (\mathcal{P}^s)^*$ *is defined by*

$$\partial_{\boldsymbol{j}}[\boldsymbol{z}](p) \triangleq \frac{1}{j_1! j_2! \ldots j_s!} \left(\frac{\partial^{\sum_{i=1}^s j_i}}{\partial x_1^{j_1} \partial x_2^{j_2} \ldots \partial x_s^{j_s}} p \right) (\boldsymbol{z})$$

By (A.1) we have that for $p, q \in \mathcal{P}^s$

$$p \sim_{\mathcal{I}} q \Rightarrow \partial_{\boldsymbol{0}}[\boldsymbol{z}_k]p = \partial_{\boldsymbol{0}}[\boldsymbol{z}_k]q$$

for $1 \leq k \leq m$. So the functionals $\{\partial_{\boldsymbol{0}}[\boldsymbol{z}_k]\}_{1 \leq k \leq m}$ can be interpreted as linear functionals on $\mathcal{R}[\mathcal{I}]$:

$$\partial_{\boldsymbol{0}}[\boldsymbol{z}_k] \in \mathcal{D}[\mathcal{I}], \quad 1 \leq k \leq m.$$

Proposition A.3.1. *For a polynomial ideal $\mathcal{I}$ with zeros $Z(\mathcal{I})$, finite in number ($|Z(\mathcal{I})| = m$) and all simple, the set $\{\partial_{\boldsymbol{0}}[\boldsymbol{z}_k]\}_{1 \leq k \leq m}$ is a basis for $\mathcal{D}[\mathcal{I}]$.*

Proof. We know that $\dim \mathcal{D}[\mathcal{I}] = m$ and every functional in $\{\partial_{\boldsymbol{0}}[\boldsymbol{z}_k]\}_{1 \leq k \leq m}$ is contained in $\mathcal{D}[\mathcal{I}]$, so it is sufficient to show that the set $\{\partial_{\boldsymbol{0}}[\boldsymbol{z}_k]\}_{1 \leq k \leq m}$ is linearly independent. Suppose that there is a linear combination

$$l = \sum_{k=1}^{m} c_k \partial_{\boldsymbol{0}}[\boldsymbol{z}_k]$$

such that $c_i \neq 0$ for some i and such that l is equal to the zero functional. Consider the residue class $[p]_{\mathcal{I}}$ that vanishes at $Z(I) \backslash \{\boldsymbol{z}_i\}$ and takes on the value 1 at $\boldsymbol{z}_i$. Then

$$l[p]_{\mathcal{I}} = c_i \partial_{\boldsymbol{0}}[\boldsymbol{z}_i][p]_{\mathcal{I}} = c_i \neq 0,$$

which is a contradiction. □

In the simple case where all zeros of $\mathcal{I}$ are simple, the zero set of the polynomial system that generates $\mathcal{I}$ are those points at which all polynomials in $\mathcal{I}$ vanish and the number of distinct zeros is equal to the dimension of the dual space $\mathcal{D}[\mathcal{I}]$. Let us now step away from the assumption that all zeros have multiplicity 1. In the univariate case ($s = 1$) it is well known that if $p(x) \in \mathcal{P}^1$ has a zero with multiplicity 2 in $x = z_1$, then $p(x)$ can be written as

$$p(x) = \tilde{p}(x)(x - z_1)^2, \quad \tilde{p}(x) \in \mathcal{P}^1.$$

The ideal generated by $p(x)$ is given by

$$\mathcal{I} = \langle p(x) \rangle = \{q \in \mathcal{P}^1 \mid q(x) = r(x)p(x) \text{ for some } r(x) \in \mathcal{P}^1\}$$

so that every $q \in \mathcal{I}$ can be written as

$$q(x) = \tilde{q}(x)(x - z_1)^2, \quad \tilde{q}(x) \in \mathcal{P}^1.$$

By the previous discussion, it is clear that

$$\partial_0[z_1]q = q(z_1) = 0, \quad \forall q \in \mathcal{I}$$

but also

$$\partial_1[z_1]q = \frac{\partial q(x)}{\partial x}\Big|_{x=z_1} = \left(\frac{\partial \tilde{q}(x)}{\partial x}(x - z_1)^2 + 2\tilde{q}(x)(x - z_1)\right)\Big|_{x=z_1} = 0, \quad \forall q \in \mathcal{I}.$$

In this case, for two polynomials q_1 and q_2 to be in the same equivalence class, the difference $q_1 - q_2$ must be contained in $\mathcal{I}$. Therefore not only the function value of $q_1 - q_2$ but also its first derivative must vanish at z_1 and therefore $\partial_1[z_1]q_1 = \partial_1[z_1]q_2$. Suppose p has m zeros $\{z_i\}_{1 \le i \le m}$ of which only z_1 has multiplicity 2 and the other zeros are simple, then the equivalence classes in $\mathcal{P}^1$ induced by $\mathcal{I}$ are completely determined by their m function values at the zeros and their first derivative at $x = z_1$. The dimension of $\mathcal{R}[\mathcal{I}]$ is equal to $m + 1$ and so is $\dim(\mathcal{D}[\mathcal{I}])$. $\mathcal{D}[\mathcal{I}]$ is spanned by the functionals

$$\{\partial_0[z_i]\}_{1 \le i \le m} \cup \partial_1[z_1].$$

The multiplicity structure of the zeros of p is completely determined by the structure of the dual space $\mathcal{D}[\mathcal{I}]$. This holds for the multivariate case too.

Example A.3.1. *Consider again the ideal from Example A.2.1:*

$$\mathcal{I} = \langle x^2 + y^2 - 1, x^2 + 4y^2 - 1 \rangle \triangleq \langle q_1, q_2 \rangle.$$

Let us denote the zeros by $\boldsymbol{z}_1 = (-1, 0)$ *and* $\boldsymbol{z}_2 = (1, 0)$. *We have shown that the polynomial* $p(x, y) = y$ *is not a member of* $\mathcal{I}$. *That is, it is not in the same residue class as* $q_1(x, y) = x^2 + y^2 - 1$ *although it takes on the same values at the zero set* $Z(\mathcal{I})$. *Every polynomial* $q \in \mathcal{I}$ *can be written as*

$$q = p_1 q_1 + p_2 q_2$$

with $p_1, p_2 \in \mathcal{P}^2$. It can be verified that

$$\partial_{00}[\boldsymbol{z}_1]q = q(\boldsymbol{z}_1) = 0, \quad \forall q \in \mathcal{I},$$

$$\partial_{01}[\boldsymbol{z}_1]q = \frac{\partial q}{\partial y}\Big|_{z_1} = \left(\frac{\partial p_1}{\partial y}q_1 + p_1\frac{\partial q_1}{\partial y} + \frac{\partial p_2}{\partial y}q_2 + p_2\frac{\partial q_2}{\partial y}\right)\Big|_{z_1} = 0, \quad \forall q \in \mathcal{I}$$

Therefore, for two polynomials to belong to the same residue class modulo $\mathcal{I}$, they have to take on the same value at $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ but they also have to have the same partial derivative with respect to y at $\boldsymbol{z}_1$. Analogously, it can be shown that

$$\partial_{00}[\boldsymbol{z}_2]q = q(\boldsymbol{z}_2) = 0, \quad \forall q \in \mathcal{I},$$

$$\partial_{01}[\boldsymbol{z}_2]q = \frac{\partial q}{\partial y}\Big|_{z_2} = 0, \quad \forall q \in \mathcal{I},$$

so that the equivalence classes with respect to $\sim_{\mathcal{I}}$ are defined by the values they take on at $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ and the values of their partial derivatives with respect to y at $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$. This means $\dim(\mathcal{R}[\mathcal{I}]) = \dim(\mathcal{D}[\mathcal{I}]) = 4$ and

$$\mathcal{D}[\mathcal{I}] = span\{\partial_{00}[\boldsymbol{z}_1], \partial_{01}[\boldsymbol{z}_1], \partial_{00}[\boldsymbol{z}_2], \partial_{01}[\boldsymbol{z}_2]\}. \tag{A.3}$$

Before stating the formal definition of a μ-fold zero, we need to introduce the notion of a *closed set of functionals.*

Definition A.8 (closed set of functionals)**.** *A set of functionals $\{c_i\}_{1\leq i\leq m} \subset (\mathcal{P}^s)^*$ is said to be closed if*

$$c_i(p) = 0, \forall i \in \{1, \ldots, m\} \text{ for some } p \in \mathcal{P}^s$$

implies that $c_i(qp) = 0, \forall i \in \{1, \ldots, m\}$, $\forall q \in \mathcal{P}^s$.

Definition A.9 (μ-fold zero)**.** *A zero $\boldsymbol{z}_i$ of a 0-dimensional ideal $\mathcal{I} \subset \mathcal{P}^s$ is a μ-fold zero of $\mathcal{I}$ if there exists a closed set of μ linearly independent differentiation functionals $\{c_{ik} = \sum_{\boldsymbol{j}\in J_{ik}} \alpha_{ik\boldsymbol{j}}\partial_{\boldsymbol{j}}[\boldsymbol{z}_i]\}_{1\leq k\leq \mu}$ in the dual space $\mathcal{D}[\mathcal{I}]$ and no such set exists containing more than μ functionals.*

Example A.3.2. *The sets $\{\partial_{00}[\boldsymbol{z}_1], \partial_{01}[\boldsymbol{z}_1]\}$ and $\{\partial_{00}[\boldsymbol{z}_2], \partial_{01}[\boldsymbol{z}_2]\}$ from* (A.3) *are closed sets of functionals. Indeed, if $p(\boldsymbol{z}_1) = \frac{\partial p}{\partial y}\Big|_{\boldsymbol{z}_1} = 0$ then for any $q \in \mathcal{P}^1$*

$$(pq)(\boldsymbol{z}_1) = \frac{\partial(pq)}{\partial y}\Big|_{z_1} = \left(\frac{\partial p}{\partial y}q + p\frac{\partial q}{\partial y}\right)\Big|_{z_1} = 0.$$

The zero $\boldsymbol{z}_1$ is a 2-fold zero because it contributes two linearly independent functionals to this basis of $\mathcal{D}[\mathcal{I}]$. Analogously, the zero $\boldsymbol{z}_2$ has multiplicity 2.

Appendix B

Polynomial Systems and Newton Polytopes: the BKK-Bound

Let $\mathcal{S}$ denote a set of polynomials $\{p_i\}_{1\leq i\leq s}$ in $\mathcal{P}^s$ that define a 0-dimensional polynomial system. Let p_i be given by $p_i = \sum_{k=1}^{N_i} c_{ik} m_k$, $m_k \in \mathcal{T}^s$, $1 \leq k \leq N_i$ and $\text{Supp}_i = \{k \mid c_{ik} \neq 0, 1 \leq k \leq N_i\}$. Now, define the map $\phi : \mathcal{T}^s \to \mathbb{N}^s$ as follows. Let $\boldsymbol{j} = \begin{pmatrix} j_1 & j_2 & \dots & j_s \end{pmatrix}^\top \in \mathbb{N}^s$, then

$$\phi(m) = \boldsymbol{j} \Leftrightarrow m = x_1^{j_1} x_2^{j_2} \dots x_s^{j_s}.$$

We will refer to $\boldsymbol{j}$ as the coordinates of the monomial m in $\mathbb{R}^s$. Next, we map all the monomials in the support of p_i to their coordinates in $\mathbb{R}^s$, using the map ϕ:

$$\mathcal{J}_i \triangleq \{\boldsymbol{j} \in \mathbb{N}^s \mid \exists k \in \text{Supp}_i : \phi(m_k) = \boldsymbol{j}\}, \quad i = 1, \dots, s.$$

In this appendix, we will refer to $\mathcal{J}_i$ as the *support* of p_i. For fixed supports $\{\mathcal{J}_i\}_{1\leq i\leq s}$ it can be shown that for almost all possible sets of complex coefficients $\{c_{ik}\}_{1\leq i\leq s, 1\leq k\leq N_i}$, the system defined by $\mathcal{S}$ has the same number of affine roots (counting multiplicities). This number is defined as BKK($\mathcal{S}$) [28, 26].

B.1 Newton polytopes

The convex hull of some finite set of points $\mathcal{J} \subset \mathbb{R}^s$ is a convex polytope. To any nonzero polynomial $p \in \mathcal{P}^s$, such a convex polytope is associated.

Definition B.1 (Newton polytope)**.** *Given a nonzero polynomial $p \in \mathcal{P}^s$ with support $\mathcal{J}$, the associated* Newton polytope *is given by*

$$\text{New}(p) \triangleq \textit{conv}(\mathcal{J})$$

where conv$(\cdot)$ is the convex hull of a set of points in $\mathbb{R}^s$.

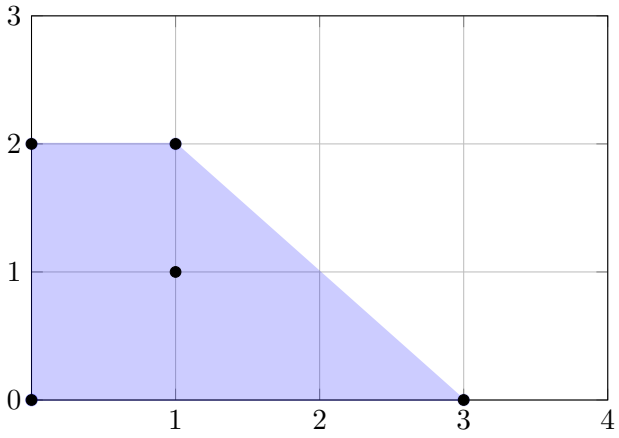


Figure B.1: New(p) (in blue) and $\mathcal{J}$ (•) in $\mathbb{R}^2$ from Example B.1.1.

We will switch to the case where $s = 2$ for simplicity, but all of the results can be readily extended to the general case. A polytope in two dimensions is called a *polygon*.

Example B.1.1. *Consider the polynomial* $p(x, y) = -1 + x^3 + xy + y^2 + xy^2$. *The set* $\mathcal{J}$ *is given by*

$$\mathcal{J} = \{(0, 0), (3, 0), (1, 1), (0, 2), (1, 2)\}$$

and the polygon New(p) *is shown in Figure B.1.*

B.2 Minkowski sum and mixed area

Definition B.2 (Minkowski sum)**.** *Let* P_1 *and* P_2 *be two polygons in* $\mathbb{R}^2$, *their* Minkowski sum $[P_1 + P_2]$ *is defined as*

$$[P_1 + P_2] = \{p_1 + p_2 \in \mathbb{R}^2 \mid p_1 \in P_1, p_2 \in P_2\}.$$

Property B.2.1. *For every* $q_1, q_2 \in \mathcal{P}^2$, *it holds that* $\text{New}(q_1 q_2) = [\text{New}(q_1) + \text{New}(q_2)]$. *More generally, the Minkowski sum of two convex polytopes in* $\mathbb{R}^2$ *is again a convex polytope in* $\mathbb{R}^2$.

Example B.2.1. *consider the polynomial*

$$\begin{aligned} p(x, y) &= 1 + x + y + xy + x^2y + xy^2 + x^3y + 2x^2y^2 + xy^3 + x^3y^2 + x^2y^3 \\ &= \underbrace{(1 + x + y + xy)}_{q_1}\underbrace{(1 + x^2y + xy^2)}_{q_2} \end{aligned}$$

with support $\mathcal{J}$. *Figure B.2 illustrates Property B.2.1.*

We need one more definition in order to state the central theorem of this appendix.

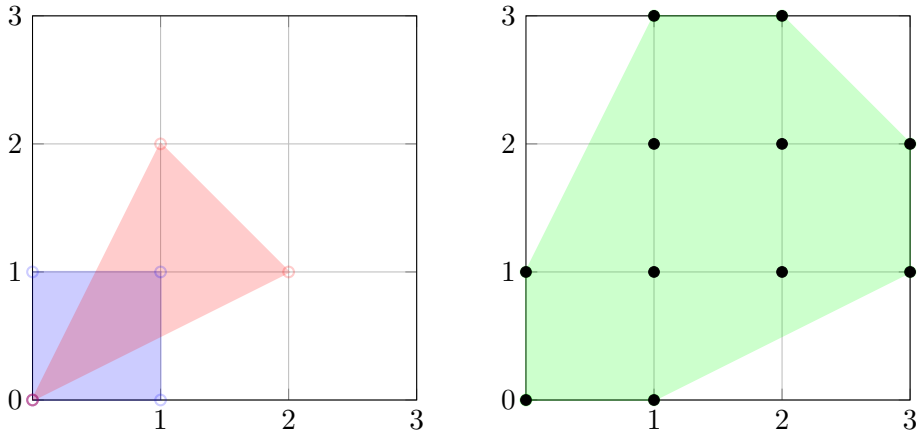


Figure B.2: Left: Newton polytopes for q_1 (blue) and q_2 (red) from example B.2.1. Right: Minkowski sum [New(q_1) + New(q_2)] (green) and the support of p (•) as defined in Example B.2.1.

Definition B.3 (mixed area)**.** *Given two polygons P_1 and P_2 in $\mathbb{R}^2$. Their* mixed area *is defined as*

$$\mathcal{M}(P_1, P_2) \triangleq \mathcal{A}([P_1 + P_2]) - \mathcal{A}(P_1) - \mathcal{A}(P_2)$$

where $\mathcal{A}(\cdot)$ maps a polygon to its area.

B.3 The BKK-bound

Theorem B.3.1. *Let p and q be two bivariate polynomials and let $(\mathbb{C}_0)^2$ be the two-dimensional agebraic torus:*

$$(\mathbb{C}_0)^2 = \{(x, y) \in \mathbb{C}^2 \mid x \neq 0 \text{ and } y \neq 0\}.$$

The number of solutions to $p(x, y) = q(x, y) = 0$ in $(\mathbb{C}_0)^2$, counting multiplicities is bounded by the mixed area $\mathcal{M}(\text{New}(p), \text{New}(q)) \triangleq BKK(p, q)$.

Theorem B.3.1 provides us with a stricter bound than that of Bézout and the result is even stronger than it may seem from the formulation of the theorem. For almost all systems with the same set of supports $\{\mathcal{J}_p, \mathcal{J}_q\}$, BKK$(p, q)$ gives exactly the number of zeros in $(\mathbb{C}_0)^2$ rather than just an upper bound. Deviations from this property are discussed in the next section.

Example B.3.1. *Consider the system defined by q_1 and q_2 from Example B.2.1. The Bézout number for this system is $\delta_{q_1}\delta_{q_2} = 6$. The mixed area $\mathcal{M}(\text{New}(q_1), \text{New}(q_2))$ equals 4. Although Bézout's theorem predicts 6 solutions, by the BKK-bound we know that at least two of those solutions must lie at infinity. Indeed, the problem has four affine solutions, all real and plotted in Figure B.3.*

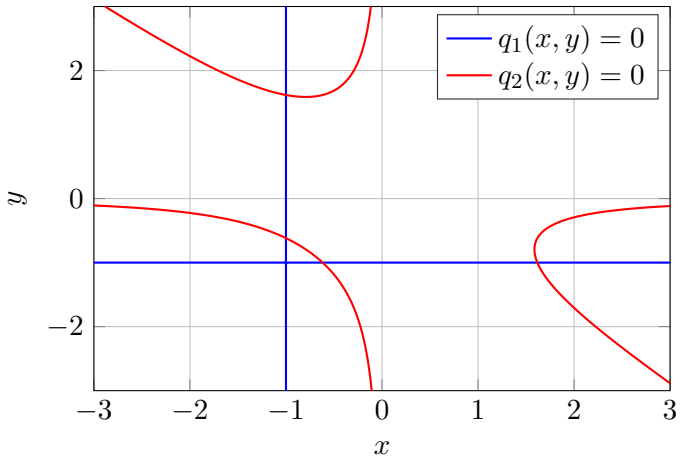


Figure B.3: Real picture of the zero level curves defined by the system $q_1 = q_2 = 0$ as defined in Example B.2.1.

As stated in Theorem B.3.1, the result does not account for solutions with zero components ($x = 0$ or $y = 0$). This is due to the fact that the mixed area $\mathcal{M}(\text{New}(p), \text{New}(q))$ does not change by multiplying one of the polynomials by some monomial. Such a manipulation might, however, introduce solutions that are not in $(\mathbb{C}_0)^2$. Example B.3.2 illustrates this phenomenon.

Example B.3.2. *Consider the system*

$$\begin{cases} p(x,y) = xy - 1 \\ q(x,y) = x - y + 1 \end{cases}$$

There are two affine solutions that can be calculated analytically and they are given by (a_1^{-1}, a_1) *and* (a_2^{-1}, a_2) *where*

$$a_1 = \frac{1+\sqrt{5}}{2} \quad \textit{and} \quad a_2 = \frac{1-\sqrt{5}}{2}.$$

Multiplying p *by* y*, we obtain the system*

$$\begin{cases} \tilde{p}(x,y) = xy^2 - y \\ \tilde{q}(x,y) = x - y + 1 \end{cases}$$

which has the solutions (a_1^{-1}, a_1)*,* (a_2^{-1}, a_2) *and* $(-1, 0)$*. The Newton polytopes involved are depicted in Figure B.4. The mixed areas* $\mathcal{M}(\text{New}(\tilde{p}), \text{New}(\tilde{q}))$ *and* $\mathcal{M}(\text{New}(\tilde{p}), \text{New}(\tilde{q}))$ *are identical and equal to 2. The "new" solution* $(-1, 0)$ *is not detected by the BKK-bound.*

Fortunately, there is a simple trick for avoiding this deficiency, based on the continuity of the roots with respect to the coefficients $\{c_{ik}\}_{1 \le i \le s, 1 \le k \le N_i}$.

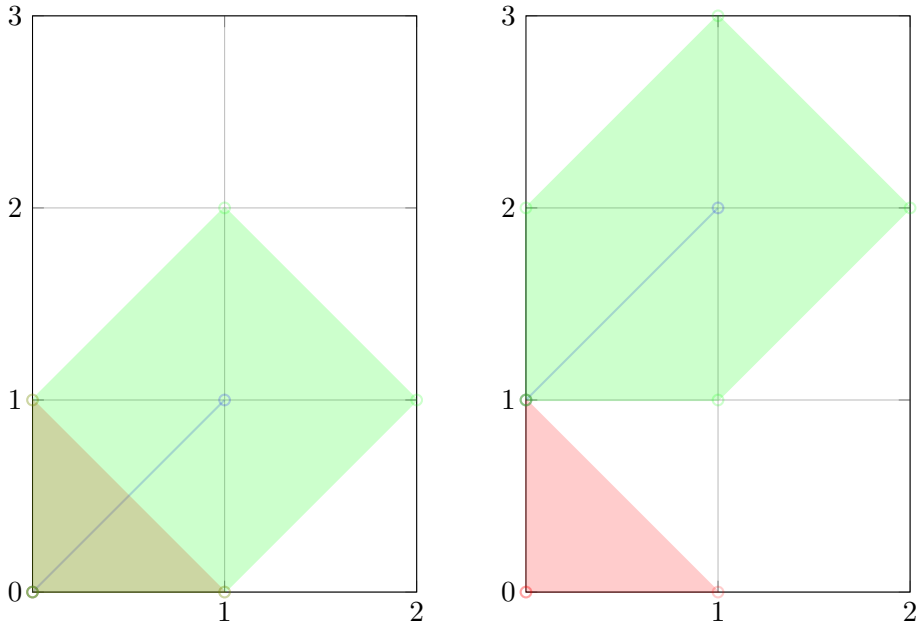


Figure B.4: Left: New(p) (blue), New(q) (red) and [New(p)+New(q)] (green). Right: New($\tilde{p}$) (blue), New($\tilde{q}$) (red) and [New($\tilde{p}$) + New($\tilde{q}$)] (green).

Proposition B.3.1. *The BKK-bound gives the correct number of zeros in all of* $\mathbb{C}^2$ *when we add the point* $(0,0)$ *to the support of both* p *and* q *(or the point* $\mathbf{0}_s$ *to the support of all* p_i *in the general case) for the construction of the Newton polytopes.*

Example B.3.3. *Applying the trick of Proposition B.3.1 to the system* $\tilde{p} = \tilde{q} = 0$ *as defined in Example B.3.2, the BKK-bound becomes equal to 3, which is the right number of solutions. The polytopes are given in Figure B.5.*

B.4 Disappearing roots

As mentioned in the previous section, the BKK-bound almost always gives the exact amount of affine roots of the system. It might happen that for some set of coefficients $\mathcal{C}^*$, some of the BKK(p,q) roots lie at infinity. If we let the set of coefficients 'move' towards $\mathcal{C}^*$ in a continuous manner, we can see these solutions 'move' towards infinity. Such solutions are called *diverging solutions*. Since the "disappearance" of solutions only happens for a lower dimensional manifold in the coefficients space, any arbitrarily small perturbation that results in a system with the same supports makes the solutions "reappear" as solutions with a very large norm.

Example B.4.1. *Consider the system*

$$\begin{cases} p(x,y) = (x-a_1)(y-b_1) - 1 \\ q(x,y) = (x-a_2)(y-b_2) + 1 \end{cases}.$$

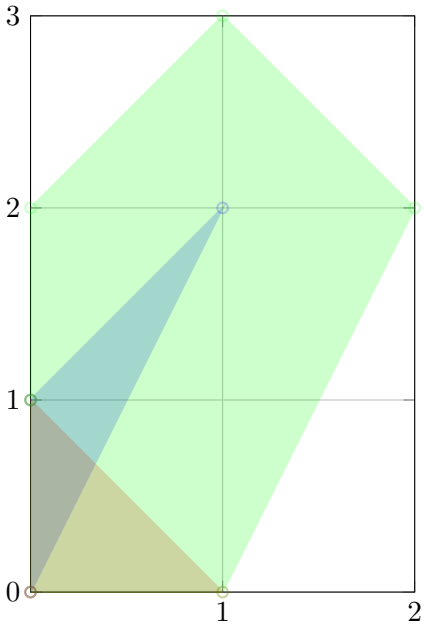


Figure B.5: New($\tilde{p}$) (blue), New($\tilde{q}$) (red) and [New($\tilde{p}$)+New($\tilde{q}$)] (green) after applying proposition B.3.1.

The Bézout number is 4. The BKK-bound is given by BKK(p, q)=2 and for generic coefficients $\{a_1, a_2, b_1, b_2\}$ *there are two affine zeros. For* $a_2 \to a_1$*, one of these zeros disappears (it diverges to infinity). When* $b_2 \to b_1$*, the other one follows and there are no finite solutions left.*

Appendix C

A Stricter Bound on the Pencil Size

In Chapter 4 the upper bound $2\delta^2 + \delta$ was deduced for the size of the square pencil $\hat{\Pi}_{x,r}(x)$. We will show in this appendix that this is a very pessimistic bound. We warn the reader that the derivations are rather technical and they are not essential for the understanding of the rest of this text. For an illustration, the reader can skip the calculations and have a look at the examples in this appendix.

C.1 Calculations

The number s represents the number of columns and rows that are removed during the x-reduction step. Define the numbers ϕ_i as

$$\phi_i \triangleq \begin{cases} \max(\deg(p_{i-1}^x(x)), \deg(q_{i-1}^x(x))) + 1, & i \in \{1, \ldots, \max(\delta_p^y, \delta_q^y) + 1\} \\ 0 & i > \max(\delta_p^y, \delta_q^y) + 1 \end{cases}.$$

Note that ϕ_i represents the number of columns in the i-th nonzero coefficient block in $\hat{\Phi}_r$. Clearly, this number must be bounded by

$$\phi_i \leq \delta + 2 - i = \max(\delta_p, \delta_q) + 2 - i, \forall i \in \{1, \ldots, \max(\delta_p^y, \delta_q^y) + 1\}.$$

The number of columns in the i-th block column of $\begin{pmatrix} \hat{\Phi}_r \\ \Psi_r \end{pmatrix}$ is denoted by ψ_i. Since the block columns of Ψ_r contain shifted versions of the blocks in $\hat{\Phi}_r$, ψ_i is bounded by

$$\psi_i \leq \max_j \phi_j \leq \delta + 1, \forall i \in \{1, \ldots, \delta_p^y + \delta_q^y\}.$$

This means that at the right side of the first block column of $\hat{L}(x, y)$, which corresponds to the $\hat{\delta} + 1$ monomials of degree 0 in y, we are guaranteed to find at least $\hat{\delta} + 1 - (\delta + 1)$ zero columns. In the second block column, there are at least $\hat{\delta} - (\delta + 1)$

Table C.1: All possible values for $\hat{\delta} - \delta$ depending on the values of δ_p, δ_q, δ_p^y and δ_q^y.

$\hat{\delta} - \delta$	$\delta_p > \delta_q$	$\delta_p \leq \delta_q$
$\delta_p + \delta_q^y > \delta_q + \delta_p^y$	$\delta_q^y - 1$	$\delta_p + \delta_q^y - \delta_q - 1$
$\delta_p + \delta_q^y \leq \delta_q + \delta_p^y$	$\delta_q + \delta_p^y - \delta_p - 1$	$\delta_p^y - 1$

zero columns. In general, at the right side of the j-th block column of $\hat{L}(x, y)$ we find at least $\hat{\delta} + 2 - j - (\delta + 1)$ zero columns. This statement holds as long as

$$\begin{aligned} \hat{\delta} + 2 - j - (\delta + 1) &\geq 0 \\ j &\leq \hat{\delta} - \delta + 1, \end{aligned}$$

which means we are assured to find zero columns in the first $\hat{\delta} - \delta$ block columns if they are not removed during the y-reduction. It is, however, not difficult to see that these block columns are always left untouched by the first step of Algorithm 1. Indeed, $\hat{\delta} - \delta$ never exceeds $\max(\delta_p^y, \delta_q^y) + 1$, which is the number of nonzero block columns in $\hat{\Phi}$. This is illustrated by Table C.1. The number s is therefore bounded from below by

$$s \geq \sum_{i=1}^{\hat{\delta}-\delta} i = \frac{(\hat{\delta} - \delta)(\hat{\delta} - \delta + 1)}{2}. \tag{C.1}$$

One can observe from Table C.1 that $\hat{\delta} - \delta \geq \min(\delta_p^y, \delta_q^y) - 1$. For example, in the case $\delta_p + \delta_q^y > \delta_q + \delta_p^y$ and $\delta_p \leq \delta_q$ we have that $\delta_q^y - \delta_q > \delta_p^y - \delta_p$ and thus $\delta_p + \delta_q^y - \delta_q - 1 > \delta_p^y - 1$. Denoting $\min(\delta_p^y, \delta_q^y) \triangleq \delta_{y,\min}$, we get

$$s \geq \frac{(\delta_{y,\min} - 1)\delta_{y,\min}}{2}.$$

For γ_n we have

$$\gamma_n = \frac{\Delta\delta_{y,\max}(\Delta\delta_{y,\max} + 1)}{2}$$

where $\Delta\delta_{y,\max} \triangleq \max(\delta_p - \delta_p^y, \delta_q - \delta_q^y)$. Using these results we get for the pencil size

$$\begin{aligned} \alpha - \gamma_n - s &\leq \frac{(\delta_p + \delta_q)(\delta_p + \delta_q + 1)}{2} - \frac{\Delta\delta_{y,\max}(\Delta\delta_{y,\max} + 1)}{2} - \frac{(\delta_{y,\min} - 1)\delta_{y,\min}}{2} \\ &\leq 2\delta^2 + \delta - \frac{\Delta\delta_{y,\max}(\Delta\delta_{y,\max} + 1)}{2} - \frac{(\delta_{y,\min} - 1)\delta_{y,\min}}{2}. \end{aligned} \tag{C.2}$$

Also, one can observe that

$$\hat{\delta} = \max(\delta_p + \delta_q^y - 1, \delta_q + \delta_p^y - 1) = \delta_p + \delta_q - 1 - \Delta\delta_{y,\min}$$

with $\Delta\delta_{y,\min} \triangleq \min(\delta_p - \delta_p^y, \delta_q - \delta_q^y)$. This results in the exact value of α:

$$\alpha = \frac{(\delta_p + \delta_q - \Delta\delta_{y,\min})(\delta_p + \delta_q - \Delta\delta_{y,\min} + 1)}{2}. \tag{C.3}$$

For the exact value of s we need to take into account the degrees of the coefficient polynomials $p_i^x(x)$ and $q_i^x(x)$. This would lead us to far. We conclude this section by emphasizing the fact that $\hat{\Pi}_{x,r}(x)$ grows quadratically with δ but the support structure of the polynomials p and q can influence the pencil size strongly because of the reduction step (R).

C.2 Examples

We illustrate the ideas with two examples.

Example C.2.1. *In Example 3.2.1, it can be seen that for the considered system of degree $\delta = 2$, using $\Delta\delta_p = \delta_q^y - 1 = 1$ and $\Delta\delta_q = \delta_p^y - 1 = 1$ we find that*

$$\hat{\Pi}_{x,r}(x) = \left(\begin{array}{ccc|ccc|cc|c} -4 & 0 & 1 & 0 & 0 & & 1 & & \\ -3 & -2 & 1 & 0 & 1 & & 1 & & \\ \hline & & & -4 & 0 & 1 & 0 & 0 & 1 \\ & & & -3 & -2 & 1 & 0 & 1 & 1 \\ \hline -x & 1 & & & & & & & \\ & -x & 1 & & & & & & \\ \hline & & & -x & 1 & & & & \\ & & & & -x & 1 & & & \\ \hline & & & & & & -x & 1 & \end{array}\right)$$

is square and of size 9. The numbers ϕ_i and ψ_i are given by

$$\phi_1 = \psi_1 = \psi_2 = 3, \quad \phi_2 = \psi_3 = 2, \quad \phi_3 = \psi_4 = 1 \quad \text{and} \quad \phi_4 = 0.$$

The upper bound from Corollary 4.1.1 is equal to 10. The stricter bound (C.2) *gives an upper bound equal to 9 ($\delta_{y,\min} = 2$ and $\Delta\delta_{y,\max} = 0$).*

Example C.2.2. *Consider a bivariate system where $p(x, y)$ and $q(x, y)$ are represented by two different random matrices of size 8×8 of which the entries are normally distributed with mean 0 and standard deviation 1. For this problem, we have $\delta_p = \delta_q = 14$, $\delta_p^y = \delta_p^x = \delta_q^y = \delta_q^x = 7$. The nonzero structures of the associated pencils $\hat{\Pi}_x(x)$ and $\hat{\Pi}_{x,r}(x)$ are shown in Figure C.1. The size of $\hat{\Pi}_x(x)$ is 224×231. That of $\hat{\Pi}_{x,r}(x)$ is (only) 112×112, while $2\delta^2 + \delta = 406$. Using* (C.3) *instead of the upper bound $2\delta^2 + \delta$ for α in* (C.2) *we get $\alpha - \gamma_n - s \leq 182$.*

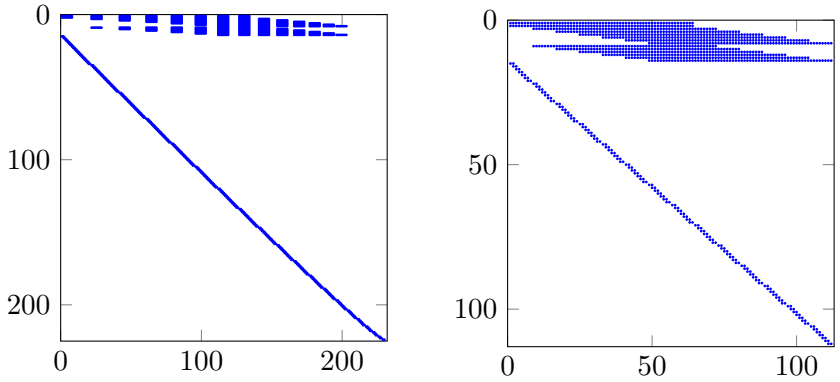


Figure C.1: Nonzero structure of $\hat{\Pi}_x(x)$ (left) and $\hat{\Pi}_{x,r}(x)$ (right) of Example C.2.2.

Appendix D

An Example in the Chebyshev Basis

We work out an example in the Chebyshev basis. The Chebyshev polynomials of the first kind are defined by the 3-term recurrence relation

$$\begin{aligned} T_0(x) &= 1, \\ T_1(x) &= x, \\ T_k(x) &= 2xT_{k-1}(x) - T_{k-2}(x), \qquad k > 1. \end{aligned}$$

We will work with the bases

$$B_x = \{T_0(x), T_1(x), \ldots, T_\delta(x)\} \quad \text{and} \quad B_y = \{T_0(y), T_1(y), \ldots, T_\delta(y)\}$$

and the corresponding tensor product basis

$$B \triangleq B_x \otimes B_y = \{b_{ij}(x,y)\}_{0 \le i,j \le \delta}$$

where $b_{ij}(x,y) \triangleq T_j(x)T_i(y)$. Consider the problem

$$\begin{cases} p(x,y) = x^2 + y^2 - 1 = 0 \\ q(x,y) = xy - \frac{1}{2} = 0 \end{cases}$$

which has two different finite solutions $(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$ and $(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2})$. Both solutions have multiplicity 2. The real zero level lines of p and q are shown in Figure D.1. For the coordinates of p and q in B we write

$$\begin{aligned} p(x,y) &= \frac{1}{2}(2x^2 - 1) + \frac{1}{2}(2y^2 - 1) \\ &= \frac{1}{2}b_{02}(x,y) + \frac{1}{2}b_{20}(x,y) \\ q(x,y) &= xy - \frac{1}{2} \\ &= b_{11}(x,y) - \frac{1}{2}b_{00}(x,y) \end{aligned}$$

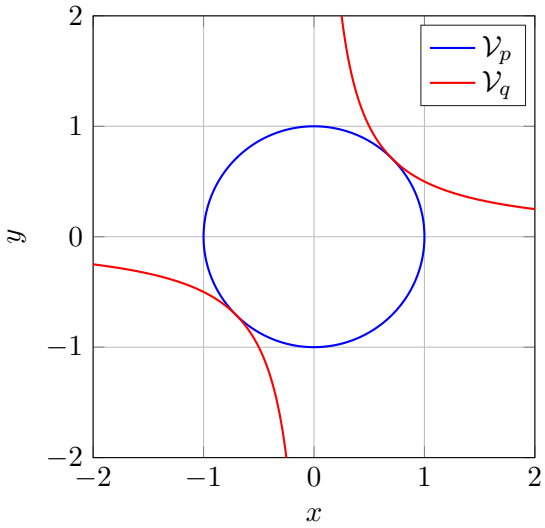


Figure D.1: Real picture of $\mathcal{V}_p$ and $\mathcal{V}_q$ for p and q from the considered example.

and we find

$$\begin{cases} p(x,y)=0 \\ q(x,y)=0 \end{cases} \xrightarrow[C]{B} \{L(x,y)\}_B = \left(\begin{array}{ccc|cc|c} 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ -\frac{1}{2} & 0 & 0 & 0 & 1 & 0 \\ \hline -x & 1 & & & & \\ 1 & -2x & 1 & & & \\ \hline & & & -x & 1 & \\ \hline -y & & & 1 & & \\ 1 & & & -2y & & 1 \end{array}\right).$$

For the degree extension, we use $\Delta\delta_p = \delta_q^y - 1 = 0$, $\Delta\delta_q = \delta_p^y - 1 = 1$ and $s_1^q(y) = y$:

$$yq(x,y) = xy^2 - \frac{1}{2}y = \frac{1}{2}b_{21}(x,y) + \frac{1}{2}b_{01}(x,y) - \frac{1}{2}b_{10}(x,y).$$

We get for $\{\hat{L}(x,y)\}_{\hat{B}}$:

$$\{L(x,y)\}_B \xrightarrow[E]{B} \left(\begin{array}{ccc|ccc|cc|c} 0 & 0 & \frac{1}{2} & 0 & 0 & & \frac{1}{2} & & \\ -\frac{1}{2} & 0 & 0 & 0 & 1 & & 0 & & \\ \hline & \frac{1}{2} & & -\frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 \\ \hline -x & 1 & & & & & & & \\ 1 & -2x & 1 & & & & & & \\ 0 & 1 & -2x & 1 & & & & & \\ \hline & & & -x & 1 & & & & \\ & & & 1 & -2x & 1 & & & \\ \hline & & & & & & -x & 1 & \\ \hline -y & & & 1 & & & & & \\ 1 & & & -2y & & & 1 & & \\ & & & 1 & & & -2y & & 1 \end{array}\right).$$

Applying Algorithm 1 to $\{\hat{L}(x,y)\}_{\hat{B}}$ we obtain for $\{\hat{L}_r(x,y)\}_{\hat{B}}$

$$\{L(x,y)\}_B \xrightarrow[E]{B} \{\hat{L}(x,y)\}_{\hat{B}} \xrightarrow[R]{B} \left(\begin{array}{ccc|cc|cc} 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} & \\ -\frac{1}{2} & 0 & 0 & 0 & 1 & 0 & \\ \hline & \frac{1}{2} & & -\frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \hline -x & 1 & & & & & \\ 1 & -2x & 1 & & & & \\ \hline & & & -x & 1 & & \\ \hline & & & & & -x & 1 \\ \hline -y & & & 1 & & & \\ 1 & & & -2y & & 1 & \end{array}\right).$$

We define $\{\hat{\Pi}_{x,r}(x)\}_{\hat{B}}$ as the rows of $\{\hat{L}_r(x,y)\}_{\hat{B}}$ that do not contain y. The finite eigenvalues of $\{\hat{\Pi}_{x,r}(x)\}_B$ are found using Matlab, they are equal to 0.7071, 0.7071, -0.7071, -0.7071. The eigenvalues are indeed (numerical approximations for) the x-coordinates of the solutions of the considered problem, taking their multiplicities into account. For this example, the matrix T from the proof of Theorem 4.3.1 is found as

$$\left(\begin{array}{c} 1 \\ x \\ 2x^2-1 \\ 4x^3-3x \\ \hline y \\ xy \\ 2x^2y-y \\ \hline 2y^2-1 \\ 2xy^2-x \end{array}\right) = T \left(\begin{array}{c} 1 \\ x \\ x^2 \\ x^3 \\ \hline y \\ xy \\ x^2y \\ \hline y^2 \\ xy^2 \end{array}\right) = \left(\begin{array}{ccccccccc} 1 & & & & & & & & \\ & 1 & & & & & & & \\ -1 & & 2 & & & & & & \\ & -3 & & 4 & & & & & \\ & & & & 1 & & & & \\ & & & & & 1 & & & \\ & & & & -1 & & 2 & & \\ -1 & & & & & & & 2 & \\ & -1 & & & & & & & 2 \end{array}\right) \left(\begin{array}{c} 1 \\ x \\ x^2 \\ x^3 \\ \hline y \\ xy \\ x^2y \\ \hline y^2 \\ xy^2 \end{array}\right).$$

For the matrix L we get

$$\begin{aligned} \{\hat{\Pi}_{x,\tilde{r}}(x)\}_{\hat{B}} T &= L\hat{\Pi}_{x,\tilde{r}}(x), \qquad \forall x \in \mathbb{C}, \\ L &= \{\hat{\Pi}_{x,\tilde{r}}(x)\}_{\hat{B}} T (\hat{\Pi}_{x,\tilde{r}}(x))^{-1}, \qquad \forall x \in \mathbb{C}. \end{aligned}$$

It is found using Matlab as

$$L = \left(\begin{array}{ccc|ccccc} 1 & & & & & & & & \\ & 1 & & & & & & & \\ & & 1 & & & & & & \\ \hline & & & 1 & & & & & \\ & & & & 2 & & & & \\ & & & -2 & & 4 & & & \\ & & & & & & 1 & & \\ & & & & & & & 2 & \\ & & & -1 & & & & & 2 \end{array}\right) = \begin{pmatrix} L_{\phi,\psi} & \\ & L_b \end{pmatrix}.$$

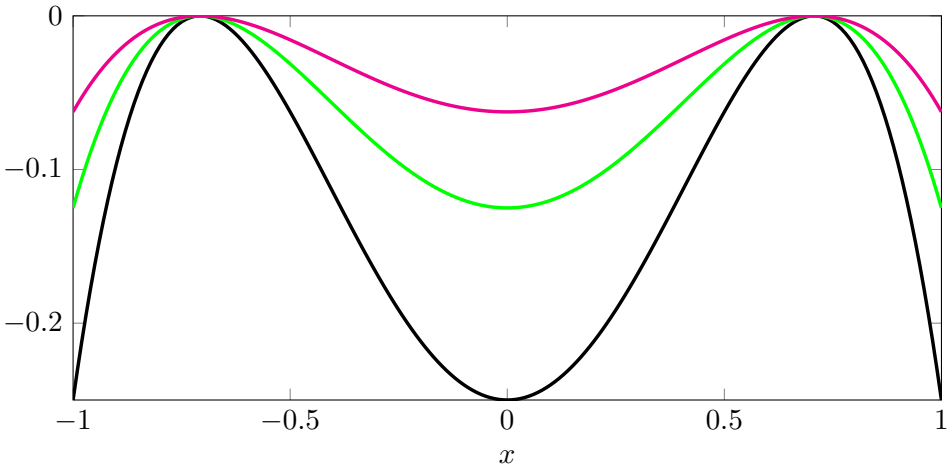


Figure D.2: The resultants $\det\{\hat{\Pi}_{x,r}(x)\}_{\hat{B}}$ (——) and $\det \hat{\Pi}_{x,r}(x)$ (——) for the problem that is considered in this appendix. The constant C from Theorem 4.3.1 is equal to $\frac{1}{2}$. The purple resultant (——) is obtained using the shift function $s_1^q(y) = \frac{1}{2}y + 1$ and the Chebyshev basis, the matrix L changes and $C = \frac{1}{4}$.

It is indeed both regular and lower triangular. The constant $\tilde{C}$ is found as

$$\tilde{C} = \frac{\det L}{\det T} = \frac{1}{2}.$$

Figure D.2 illustrates Theorem 4.3.1. Apparently $\gamma = 1$ and $C = \tilde{C}$.

Appendix E

Rectangular Eigenvalue Problems

In order to find a numerical approximation of the solution set of a bivariate polynomial system, some versions of the proposed method in this text result in solving several rectangular eigenvalue problems. The problem also plays an important role in the generalization to higher dimensions (Appendix G). In this appendix we discuss the problem and we propose a way to solve it in a numerically reliable way.

E.1 The rectangular eigenvalue problem (REP)

Definition E.1 (REP)**.** *The problem of finding all values of $x \in \mathbb{C}$ such that there exists a nonzero vector $\boldsymbol{v}$ that satisfies*

$$A\boldsymbol{v} = xB\boldsymbol{v} \tag{E.1}$$

where $A, B \in \mathbb{C}^{m\times n}$ and $\boldsymbol{v} \in \mathbb{C}^n$ is called a rectangular eigenvalue problem*. A solution x and the corresponding vector(s) $\boldsymbol{v}$ are called an* eigenvalue *and the corresponding* eigenvector(s) *respectively.*

Definition E.1 can be interpreted as follows. Solving the REP, we are looking for all finite values of x such that some rectangular linear pencil $A - xB$ is column rank deficient. It is clear that if $m < n$, for any value of $x \in \mathbb{C}$ there exists a vector $\boldsymbol{v}$ such that (E.1) is satisfied. Such a "flat" REP has infinitely many eigenvalues. We will assume from now on that $m \geq n$. A "tall" pencil $A - xB$ is called singular if it is column rank deficient for every $x \in \mathbb{C}$. A pencil that is not singular is called regular. Generically, a tall REP has no eigenvalues in $\mathbb{C}$.

E.2 Solving a REP

A necessary condition for (E.1) to hold is that

$$A_{1:n,:}\boldsymbol{v} = xB_{1:n,:}\boldsymbol{v} \tag{E.2}$$

where we used the Matlab notation $(\cdot)_{1:n,:}$ to denote the first n rows and all columns of A. This is a square GEP that can be solved using the QZ-algorithm for example. For the resulting eigenvalues x^* the upper square part of the pencil $A - x^*B$ is column rank deficient. It is, however, not guaranteed that the entire pencil is too. Therefore, we must check for each eigenvalue x^* of (E.2) whether $A - x^*B$ has at least one singular value $= 0$. This would work in exact arithmetic. In practice, we test for every numerically obtained eigenvalue x^* of (E.2) if $\boldsymbol{\sigma}_n < \epsilon(\boldsymbol{\sigma}_1 + 1)$ where ϵ is some small threshold value and $\boldsymbol{\sigma}_i$ is the i-th singular value of $A - x^*B$[1]. From a numerical point of view, it is more interesting to use a QR-factorization with optimal column pivoting before splitting the pencil to obtain a GEP. The procedure that is implemented is given by the following algorithm.

Algorithm 2. *Let $A, B \in \mathbb{C}^{m\times n}$ be given matrices with $m > n$ and let $\epsilon > 0$ be a given threshold,*

Let $M = \begin{pmatrix} A & B \end{pmatrix}$ and compute the QR-factorization of M with optimal column pivoting so that $M = QRP^*$ with P some column permutation matrix.
Define $RP^* = \begin{pmatrix} R_1 & R_2 \end{pmatrix}$ with $R_1, R_2 \in \mathbb{C}^{m\times n}$ and calculate the eigenvalues of the GEP

$$(R_1 - xR_2)_{1:n,:}\boldsymbol{v} = 0$$

by using the QZ-algorithm. Denote the set of finite eigenvalues by $\tilde{X}$.
Initialize the set of eigenvalues of the REP as an empty set X_{rep}.
for** every $x^* \in \tilde{X}$ **do
    Let $A - x^*B = U\Sigma V^\top$ be the singular value decomposition of $A - x^*B$.
    Let $\boldsymbol{\sigma}$ be the vector containing the singular values: $\boldsymbol{\sigma} = \operatorname{diag}(\Sigma)$.
    if** $\boldsymbol{\sigma}_n < \epsilon(\boldsymbol{\sigma}_1 + 1)$ **then
        Add x^* to the set X_{rep}.
    end if
end for
Output the set X_{rep}.

[1]The singular values are assumed to be ordered from large to small, as they usually are.

Appendix F

Detailed Numerical Results

In this appendix we report the numerical results in tables. We will use the notations in Table F.1. All tests are performed on a set of 60 problems of which some properties can be found in Table F.2 and on a set of random problems of degree $\delta = 1, \ldots, 40$ as specified in Chapter 6.

p	Problem index, $1 \leq p \leq 60$.
# refsol	The number of reference solutions.
time (s)	Computation time in seconds.
# sol	The number of solutions found by a solver.
$r_{\max}$	Maximal residual of all the solutions found by a solver.
$\epsilon_{\max}$	Maximal (mixed absolute and relative) forward error.
s	Boolean, indicates whether the problem is successfully solved by a solver. Criteria for success are given in Chapter 6.

Table F.1: Notations.

The maximal residual $r_{\max}$ is calculated as in Definition 5.1. The maximal forward error $\epsilon_{\max}$ is calculated as

$$\epsilon_{\max} \triangleq \max_{\boldsymbol{s} \in \mathcal{S}_{\text{ref}}} \left\{ \frac{\min_{\tilde{\boldsymbol{s}} \in \tilde{\mathcal{S}}} \|\boldsymbol{s} - \tilde{\boldsymbol{s}}\|_2}{\|\boldsymbol{s}\|_2 + 1} \right\}$$

where $\tilde{\mathcal{S}}$ indicates the set of numerically found solutions. Note that the criterion for success for solvers that intend to take multiplicities into account is not automatically satisfied if $\epsilon_{\max} < 0.01$. The criterion is more restrictive. The criterion is described in Chapter 6. The reference solution set $\mathcal{S}_{\text{ref}}$ is calculated using Bertini in adaptive precision, which is a very reliable solver. For the random problems, the value of $\epsilon_{\max}$ is not reported because it is not used in the criterion for success. The tests on the random problems for PNLA are only executed up to degree 25 because they were too time consuming.

p	δ_p	δ_q	# refsol	$\mu_{\max}$	p	δ_p	δ_q	# refsol	$\mu_{\max}$
1	8	7	32	13	31	8	7	32	16
2	8	7	32	14	32	4	5	16	4
3	4	3	12	3	33	2	2	4	1
4	2	1	2	1	34	3	3	6	2
5	3	2	4	2	35	4	2	8	4
6	4	3	12	8	36	9	10	90	9
7	6	5	22	18	37	3	4	4	1
8	9	8	72	8	38	3	1	3	1
9	3	2	2	1	39	3	1	2	2
10	3	2	4	2	40	4	2	3	2
11	3	2	6	3	41	3	3	6	2
12	4	3	4	1	42	4	2	8	2
13	3	2	4	2	43	6	3	18	8
14	4	3	8	2	44	6	4	24	6
15	8	6	48	2	45	8	7	49	1
16	8	6	48	2	46	8	6	48	2
17	4	3	10	4	47	5	4	4	1
18	4	3	8	2	48	4	3	10	4
19	6	5	22	2	49	4	3	8	2
20	11	10	55	1	50	6	5	22	2
21	8	6	48	2	51	11	10	55	1
22	8	7	52	12	52	8	7	49	1
23	6	5	30	1	53	8	7	52	12
24	8	1	8	2	54	6	5	30	1
25	9	8	52	1	55	8	7	32	18
26	4	3	12	6	56	9	8	56	6
27	3	2	6	3	57	8	1	8	2
28	7	6	38	10	58	9	8	52	1
29	4	3	11	6	59	4	3	12	6
30	8	7	32	16	60	3	2	6	3

Table F.2: Information about the problem set, $\mu_{\max}$ denotes the highest multiplicity of the solutions of the problem. It is determined by the solution vector $\boldsymbol{\mu}$ of (5.7).

		EV					Syl		
p	# refsol	$r_{\max}$	$\epsilon_{\max}$	s	time (s)	# sol	$r_{\max}$	$\epsilon_{\max}$	s
1	32	$5.42 \cdot 10^{-12}$	$1.02 \cdot 10^{-4}$	1	0.21	122	$6.38 \cdot 10^{-11}$	$9.59 \cdot 10^{-5}$	1
2	32	$4.61 \cdot 10^{-9}$	$1.03 \cdot 10^{-3}$	1	0.23	94	$2.57 \cdot 10^{-7}$	$1.07 \cdot 10^{-3}$	1
3	12	$2.52 \cdot 10^{-16}$	$7.89 \cdot 10^{-14}$	1	$2.23 \cdot 10^{-2}$	48	$2.52 \cdot 10^{-16}$	$7.89 \cdot 10^{-14}$	1
4	2	$1.49 \cdot 10^{-14}$	$3.33 \cdot 10^{-14}$	1	$9.87 \cdot 10^{-3}$	2	$1.35 \cdot 10^{-15}$	$3.32 \cdot 10^{-15}$	1
5	4	$5.88 \cdot 10^{-17}$	$5.65 \cdot 10^{-14}$	1	$6.41 \cdot 10^{-3}$	8	$5.88 \cdot 10^{-17}$	$5.65 \cdot 10^{-14}$	1
6	12	$1.08 \cdot 10^{-7}$	$6.04 \cdot 10^{-9}$	1	$1.68 \cdot 10^{-2}$	42	$1.37 \cdot 10^{-16}$	$1.74 \cdot 10^{-12}$	1
7	22	$1.1 \cdot 10^{-14}$	$1.03 \cdot 10^{-3}$	1	$4.04 \cdot 10^{-2}$	116	$4.44 \cdot 10^{-15}$	$2.35 \cdot 10^{-12}$	1
8	72	$4.43 \cdot 10^{-15}$	$3.27 \cdot 10^{-12}$	1	$4.14 \cdot 10^{-2}$	648	$4.43 \cdot 10^{-15}$	$3.27 \cdot 10^{-12}$	1
9	2	$8.88 \cdot 10^{-16}$	$2.01 \cdot 10^{-15}$	1	$6.31 \cdot 10^{-3}$	2	$3.1 \cdot 10^{-16}$	$1.99 \cdot 10^{-15}$	1
10	4	$1.08 \cdot 10^{-16}$	$4.89 \cdot 10^{-9}$	1	$1.05 \cdot 10^{-2}$	8	$1.05 \cdot 10^{-16}$	$4.89 \cdot 10^{-9}$	1
11	6	$3.26 \cdot 10^{-16}$	$6.04 \cdot 10^{-6}$	1	$6.8 \cdot 10^{-3}$	6	$1.35 \cdot 10^{-15}$	$6.04 \cdot 10^{-6}$	1
12	4	$2.87 \cdot 10^{-15}$	$7.02 \cdot 10^{-16}$	1	$8.63 \cdot 10^{-3}$	4	$1.35 \cdot 10^{-15}$	$6.28 \cdot 10^{-16}$	1
13	4	$5.88 \cdot 10^{-17}$	$3.08 \cdot 10^{-15}$	1	$5.76 \cdot 10^{-3}$	8	$5.88 \cdot 10^{-17}$	$3.08 \cdot 10^{-15}$	1
14	8	$7.42 \cdot 10^{-15}$	$1.63 \cdot 10^{-8}$	1	$1.77 \cdot 10^{-2}$	24	$6.91 \cdot 10^{-15}$	$7.86 \cdot 10^{-9}$	1
15	48	$1.4 \cdot 10^{-11}$	$2.38 \cdot 10^{-6}$	1	0.17	48	$2.58 \cdot 10^{-12}$	$2.38 \cdot 10^{-6}$	1
16	48	$1.4 \cdot 10^{-11}$	$2.38 \cdot 10^{-6}$	1	0.17	48	$2.58 \cdot 10^{-12}$	$2.38 \cdot 10^{-6}$	1
17	10	$1.95 \cdot 10^{-15}$	$1.29 \cdot 10^{-12}$	1	$1.11 \cdot 10^{-2}$	20	$3.27 \cdot 10^{-15}$	$1.29 \cdot 10^{-12}$	1
18	8	$1.2 \cdot 10^{-15}$	$1.35 \cdot 10^{-13}$	1	$1.17 \cdot 10^{-2}$	24	$1.51 \cdot 10^{-15}$	$1.35 \cdot 10^{-13}$	1
19	22	$6.65 \cdot 10^{-15}$	$5.53 \cdot 10^{-15}$	1	$4.49 \cdot 10^{-2}$	26	$6.27 \cdot 10^{-15}$	$5.67 \cdot 10^{-15}$	1
20	55	$3.18 \cdot 10^{-13}$	$1.03 \cdot 10^{-13}$	1	0.17	55	$1.99 \cdot 10^{-12}$	$3.26 \cdot 10^{-13}$	1
21	48	$1.4 \cdot 10^{-11}$	$2.38 \cdot 10^{-6}$	1	0.18	48	$2.58 \cdot 10^{-12}$	$2.38 \cdot 10^{-6}$	1
22	52	$3.08 \cdot 10^{-11}$	$7.92 \cdot 10^{-5}$	1	0.21	156	$1.67 \cdot 10^{-10}$	$2.65 \cdot 10^{-5}$	1
23	30	$2.55 \cdot 10^{-12}$	$3.62 \cdot 10^{-8}$	1	$5.36 \cdot 10^{-2}$	60	$6.44 \cdot 10^{-13}$	$3.62 \cdot 10^{-8}$	1
24	8	$5.29 \cdot 10^{-16}$	$2.81 \cdot 10^{-14}$	1	$1.59 \cdot 10^{-2}$	8	$7.18 \cdot 10^{-16}$	$2.81 \cdot 10^{-14}$	1
25	52	$5.4 \cdot 10^{-11}$	$1.02 \cdot 10^{-10}$	1	0.15	52	$1.79 \cdot 10^{-13}$	$9.09 \cdot 10^{-14}$	1
26	12	$4.25 \cdot 10^{-8}$	$1.08 \cdot 10^{-8}$	1	$1.08 \cdot 10^{-2}$	36	$7.43 \cdot 10^{-16}$	$3.07 \cdot 10^{-13}$	1
27	6	$8.96 \cdot 10^{-16}$	$9.89 \cdot 10^{-9}$	1	$8.77 \cdot 10^{-3}$	7	$2.71 \cdot 10^{-12}$	$9.89 \cdot 10^{-9}$	1
28	38	$5.78 \cdot 10^{-9}$	$1.23 \cdot 10^{-7}$	1	0.11	57	$9.22 \cdot 10^{-9}$	$8.57 \cdot 10^{-8}$	1
29	11	$2.86 \cdot 10^{-15}$	$1.91 \cdot 10^{-11}$	1	$1.49 \cdot 10^{-2}$	27	$3.05 \cdot 10^{-15}$	$9.74 \cdot 10^{-14}$	1
30	32	$3.66 \cdot 10^{-10}$	$3.23 \cdot 10^{-6}$	1	0.17	78	$3.68 \cdot 10^{-10}$	$1.45 \cdot 10^{-6}$	1
31	32	$9.25 \cdot 10^{-10}$	$9.16 \cdot 10^{-4}$	1	0.18	74	$9.33 \cdot 10^{-10}$	$9.19 \cdot 10^{-4}$	1
32	16	$2.24 \cdot 10^{-15}$	$4.86 \cdot 10^{-5}$	1	$2.19 \cdot 10^{-2}$	28	$2.16 \cdot 10^{-15}$	$1.09 \cdot 10^{-4}$	1
33	4	$1.18 \cdot 10^{-15}$	$9.35 \cdot 10^{-15}$	1	$3.77 \cdot 10^{-3}$	4	$1.9 \cdot 10^{-15}$	$1.22 \cdot 10^{-14}$	1
34	6	$2.53 \cdot 10^{-16}$	$2.64 \cdot 10^{-9}$	1	$7.55 \cdot 10^{-3}$	12	$1.48 \cdot 10^{-16}$	$2.07 \cdot 10^{-13}$	1
35	8	$8.49 \cdot 10^{-17}$	$1.07 \cdot 10^{-8}$	1	$9.09 \cdot 10^{-3}$	8	$2.66 \cdot 10^{-16}$	$1.07 \cdot 10^{-8}$	1
36	90	$4.55 \cdot 10^{-14}$	$9.7 \cdot 10^{-3}$	1	0.6	162	$4.43 \cdot 10^{-14}$	$1.27 \cdot 10^{-2}$	0
37	4	$7.12 \cdot 10^{-16}$	$1.35 \cdot 10^{-15}$	1	$8.81 \cdot 10^{-3}$	4	$2.09 \cdot 10^{-16}$	$1.22 \cdot 10^{-15}$	1
38	3	$1.58 \cdot 10^{-15}$	$2.37 \cdot 10^{-15}$	1	$4.58 \cdot 10^{-3}$	3	$1.57 \cdot 10^{-15}$	$2 \cdot 10^{-15}$	1
39	2	$2.03 \cdot 10^{-16}$	$1.45 \cdot 10^{-8}$	1	$5.62 \cdot 10^{-3}$	2	$4.2 \cdot 10^{-16}$	$1.45 \cdot 10^{-8}$	1
40	3	$2.87 \cdot 10^{-16}$	$4.4 \cdot 10^{-15}$	1	$5.41 \cdot 10^{-3}$	6	$1.47 \cdot 10^{-16}$	$4.39 \cdot 10^{-15}$	1
41	6	$2.53 \cdot 10^{-16}$	$2.64 \cdot 10^{-9}$	1	$6.56 \cdot 10^{-3}$	12	$1.48 \cdot 10^{-16}$	$1.76 \cdot 10^{-13}$	1
42	8	$8.4 \cdot 10^{-13}$	$1.39 \cdot 10^{-8}$	1	$1.04 \cdot 10^{-2}$	10	$8.39 \cdot 10^{-13}$	$1.39 \cdot 10^{-8}$	1
43	18	$2.42 \cdot 10^{-15}$	$2.88 \cdot 10^{-13}$	1	$2.28 \cdot 10^{-2}$	36	$5.22 \cdot 10^{-14}$	$2.88 \cdot 10^{-13}$	1
44	24	$1.08 \cdot 10^{-12}$	$7.37 \cdot 10^{-9}$	1	$4.32 \cdot 10^{-2}$	96	$2.58 \cdot 10^{-9}$	$5.19 \cdot 10^{-9}$	1
45	49	$4.32 \cdot 10^{-7}$	$2.31 \cdot 10^{-9}$	1	0.18	50	$2.2 \cdot 10^{-11}$	$9.95 \cdot 10^{-9}$	1
46	48	$1.35 \cdot 10^{-11}$	$1.72 \cdot 10^{-6}$	1	0.17	48	$1.39 \cdot 10^{-12}$	$1.72 \cdot 10^{-6}$	1
47	4	$9.95 \cdot 10^{-16}$	$3.85 \cdot 10^{-15}$	1	$3.47 \cdot 10^{-2}$	4	$1.35 \cdot 10^{-15}$	$3.86 \cdot 10^{-15}$	1
48	10	$1.32 \cdot 10^{-15}$	$2.11 \cdot 10^{-13}$	1	$9.33 \cdot 10^{-3}$	20	$1.67 \cdot 10^{-15}$	$2.11 \cdot 10^{-13}$	1
49	8	$3.29 \cdot 10^{-16}$	$4.47 \cdot 10^{-9}$	1	$1.15 \cdot 10^{-2}$	24	$3.72 \cdot 10^{-16}$	$5.56 \cdot 10^{-14}$	1
50	22	$7.76 \cdot 10^{-15}$	$7.87 \cdot 10^{-15}$	1	$5.41 \cdot 10^{-2}$	22	$3.05 \cdot 10^{-15}$	$4.72 \cdot 10^{-15}$	1
51	55	$6.6 \cdot 10^{-13}$	$1.83 \cdot 10^{-13}$	1	0.16	55	$2.95 \cdot 10^{-12}$	$4.82 \cdot 10^{-13}$	1
52	49	$4.32 \cdot 10^{-7}$	$2.31 \cdot 10^{-9}$	1	0.19	50	$2.2 \cdot 10^{-11}$	$9.95 \cdot 10^{-9}$	1
53	52	$2.91 \cdot 10^{-11}$	$4.04 \cdot 10^{-5}$	1	0.21	156	$1.53 \cdot 10^{-10}$	$4.15 \cdot 10^{-5}$	1
54	30	$2.58 \cdot 10^{-12}$	$1.5 \cdot 10^{-8}$	1	$5.24 \cdot 10^{-2}$	60	$4.03 \cdot 10^{-13}$	$1.5 \cdot 10^{-8}$	1
55	32	$4.69 \cdot 10^{-14}$	$1.46 \cdot 10^{-3}$	1	0.18	184	$4.73 \cdot 10^{-14}$	$4.9 \cdot 10^{-6}$	1
56	56	$1.03 \cdot 10^{-7}$	$1.82 \cdot 10^{-9}$	1	0.19	60	$1.03 \cdot 10^{-7}$	$1.82 \cdot 10^{-9}$	1
57	8	$7.96 \cdot 10^{-16}$	$6.62 \cdot 10^{-15}$	1	$1.3 \cdot 10^{-2}$	8	$7.49 \cdot 10^{-16}$	$6.62 \cdot 10^{-15}$	1
58	52	$5.09 \cdot 10^{-11}$	$9.61 \cdot 10^{-11}$	1	0.15	52	$1.37 \cdot 10^{-13}$	$5.23 \cdot 10^{-14}$	1
59	12	$8.09 \cdot 10^{-9}$	$1.2 \cdot 10^{-8}$	1	$1.08 \cdot 10^{-2}$	36	$8.09 \cdot 10^{-9}$	$1.23 \cdot 10^{-12}$	1
60	6	$5.07 \cdot 10^{-16}$	$1.66 \cdot 10^{-8}$	1	$6.75 \cdot 10^{-3}$	7	$3.05 \cdot 10^{-12}$	$1.66 \cdot 10^{-8}$	1

Table F.3: Numerical results for the methods described in 5.1.1 (EV), 5.1.2 (Syl) and 5.1.3 ($L(x, y)$), part 1.

		Syl		$L(x,y)$				
p	# refsol	time (s)	# sol	$r_{\max}$	$\epsilon_{\max}$	s	time (s)	# sol
1	32	0.11	122	$7.27 \cdot 10^{-7}$	$6.86 \cdot 10^{-5}$	1	0.21	140
2	32	0.11	86	$2.72 \cdot 10^{-8}$	$1.24 \cdot 10^{-3}$	1	0.25	100
3	12	$1.44 \cdot 10^{-2}$	48	$2.52 \cdot 10^{-16}$	$7.89 \cdot 10^{-14}$	1	$1.83 \cdot 10^{-2}$	48
4	2	$7.46 \cdot 10^{-3}$	2	$1.52 \cdot 10^{-14}$	$3.52 \cdot 10^{-14}$	1	$9.41 \cdot 10^{-3}$	2
5	4	$4.6 \cdot 10^{-3}$	8	$5.88 \cdot 10^{-17}$	$5.65 \cdot 10^{-14}$	1	$7.59 \cdot 10^{-3}$	8
6	12	$2.18 \cdot 10^{-2}$	30	$4.79 \cdot 10^{-8}$	$1.74 \cdot 10^{-12}$	1	$1.42 \cdot 10^{-2}$	42
7	22	$3.44 \cdot 10^{-2}$	116	$6.58 \cdot 10^{-16}$	$2.16 \cdot 10^{-12}$	1	$3.21 \cdot 10^{-2}$	116
8	72	$3.34 \cdot 10^{-2}$	648	$4.43 \cdot 10^{-15}$	$3.27 \cdot 10^{-12}$	1	$3.32 \cdot 10^{-2}$	648
9	2	$6.33 \cdot 10^{-3}$	2	$2.81 \cdot 10^{-16}$	$1.91 \cdot 10^{-15}$	1	$5.58 \cdot 10^{-3}$	2
10	4	$1.26 \cdot 10^{-2}$	8	$1.21 \cdot 10^{-16}$	$7.02 \cdot 10^{-9}$	1	$6.47 \cdot 10^{-3}$	8
11	6	$1.21 \cdot 10^{-2}$	6	$3.95 \cdot 10^{-15}$	$1.95 \cdot 10^{-6}$	1	$9.34 \cdot 10^{-3}$	6
12	4	$9.74 \cdot 10^{-3}$	4	$3.76 \cdot 10^{-15}$	$5.41 \cdot 10^{-16}$	1	$7.56 \cdot 10^{-3}$	4
13	4	$4.2 \cdot 10^{-3}$	8	$5.88 \cdot 10^{-17}$	$3.08 \cdot 10^{-15}$	1	$4.11 \cdot 10^{-3}$	8
14	8	$2.43 \cdot 10^{-2}$	24	$9.39 \cdot 10^{-15}$	$2.42 \cdot 10^{-8}$	1	$1.44 \cdot 10^{-2}$	24
15	48	$9.48 \cdot 10^{-2}$	48	$3.4 \cdot 10^{-13}$	$1.88 \cdot 10^{-6}$	1	0.22	48
16	48	$8.95 \cdot 10^{-2}$	48	$3.03 \cdot 10^{-13}$	$1.88 \cdot 10^{-6}$	1	0.22	48
17	10	$1.57 \cdot 10^{-2}$	20	$3.43 \cdot 10^{-15}$	$1.29 \cdot 10^{-12}$	1	$9.76 \cdot 10^{-3}$	20
18	8	$1.6 \cdot 10^{-2}$	24	$6.19 \cdot 10^{-16}$	$1.35 \cdot 10^{-13}$	1	$7.88 \cdot 10^{-3}$	24
19	22	$3.97 \cdot 10^{-2}$	26	$4.14 \cdot 10^{-9}$	$1.22 \cdot 10^{-9}$	1	$4.54 \cdot 10^{-2}$	28
20	55	$9.48 \cdot 10^{-2}$	55	$1.97 \cdot 10^{-10}$	$3.22 \cdot 10^{-11}$	1	0.92	55
21	48	$8.76 \cdot 10^{-2}$	48	$3.4 \cdot 10^{-13}$	$1.88 \cdot 10^{-6}$	1	0.26	48
22	52	$9.81 \cdot 10^{-2}$	144	$2.72 \cdot 10^{-9}$	$5.34 \cdot 10^{-9}$	1	0.24	188
23	30	$5.04 \cdot 10^{-2}$	60	$1.4 \cdot 10^{-7}$	$3.53 \cdot 10^{-8}$	1	$5.75 \cdot 10^{-2}$	68
24	8	$1.56 \cdot 10^{-2}$	8	$9.14 \cdot 10^{-16}$	$2.81 \cdot 10^{-14}$	1	$4.05 \cdot 10^{-2}$	22
25	52	$7.48 \cdot 10^{-2}$	52	$2.81 \cdot 10^{-8}$	$5.32 \cdot 10^{-8}$	1	0.31	52
26	12	$1.56 \cdot 10^{-2}$	30	$3.2 \cdot 10^{-8}$	$1.57 \cdot 10^{-8}$	1	$8.8 \cdot 10^{-3}$	36
27	6	$1.05 \cdot 10^{-2}$	7	$1.72 \cdot 10^{-7}$	$1.37 \cdot 10^{-8}$	1	$7 \cdot 10^{-3}$	10
28	38	$7.3 \cdot 10^{-2}$	57	$4.41 \cdot 10^{-13}$	$7.15 \cdot 10^{-8}$	1	0.12	57
29	11	$2.03 \cdot 10^{-2}$	27	$1.07 \cdot 10^{-7}$	$6.39 \cdot 10^{-8}$	1	$8.23 \cdot 10^{-3}$	31
30	32	$7.43 \cdot 10^{-2}$	78	$1.25 \cdot 10^{-7}$	$2.75 \cdot 10^{-6}$	1	0.17	98
31	32	$9.11 \cdot 10^{-2}$	74	$1.35 \cdot 10^{-7}$	$1.35 \cdot 10^{-3}$	1	0.17	95
32	16	$4.35 \cdot 10^{-2}$	28	$2.98 \cdot 10^{-15}$	$3.99 \cdot 10^{-5}$	1	$2.46 \cdot 10^{-2}$	28
33	4	$9.78 \cdot 10^{-3}$	4	$2.48 \cdot 10^{-14}$	$1.32 \cdot 10^{-13}$	1	$3.83 \cdot 10^{-3}$	4
34	6	$1.13 \cdot 10^{-2}$	12	$2.22 \cdot 10^{-16}$	$4.86 \cdot 10^{-9}$	1	$4.24 \cdot 10^{-3}$	12
35	8	$1.71 \cdot 10^{-2}$	8	$5.01 \cdot 10^{-16}$	$8.34 \cdot 10^{-9}$	1	$7.1 \cdot 10^{-3}$	8
36	90	0.21	162	$3.07 \cdot 10^{-14}$	$5.98 \cdot 10^{-3}$	1	0.98	162
37	4	$2.07 \cdot 10^{-2}$	4	$1.48 \cdot 10^{-16}$	$1.24 \cdot 10^{-15}$	1	$9.35 \cdot 10^{-3}$	4
38	3	$6.1 \cdot 10^{-3}$	3	$2.38 \cdot 10^{-15}$	$2.48 \cdot 10^{-15}$	1	$3.79 \cdot 10^{-3}$	3
39	2	$5.35 \cdot 10^{-3}$	2	$2.05 \cdot 10^{-16}$	$7.82 \cdot 10^{-9}$	1	$5.56 \cdot 10^{-3}$	2
40	3	$4.12 \cdot 10^{-3}$	6	$1.64 \cdot 10^{-16}$	$4.48 \cdot 10^{-15}$	1	$6.26 \cdot 10^{-3}$	6
41	6	$6.68 \cdot 10^{-3}$	12	$2.22 \cdot 10^{-16}$	$4.86 \cdot 10^{-9}$	1	$4.88 \cdot 10^{-3}$	12
42	8	$1.73 \cdot 10^{-2}$	10	$2.54 \cdot 10^{-13}$	$7.85 \cdot 10^{-8}$	1	$1 \cdot 10^{-2}$	10
43	18	$2.95 \cdot 10^{-2}$	36	$2.88 \cdot 10^{-14}$	$2.88 \cdot 10^{-13}$	1	$2.87 \cdot 10^{-2}$	36
44	24	$4.32 \cdot 10^{-2}$	88	$1.31 \cdot 10^{-12}$	$6.26 \cdot 10^{-9}$	1	$4.89 \cdot 10^{-2}$	96
45	49	$9.46 \cdot 10^{-2}$	49	$5.77 \cdot 10^{-7}$	$1.51 \cdot 10^{-9}$	1	0.2	52
46	48	$8.71 \cdot 10^{-2}$	48	$2.01 \cdot 10^{-13}$	$1.41 \cdot 10^{-6}$	1	0.22	48
47	4	$2.48 \cdot 10^{-2}$	4	$9.36 \cdot 10^{-14}$	$1.89 \cdot 10^{-14}$	1	$2.94 \cdot 10^{-2}$	4
48	10	$1.5 \cdot 10^{-2}$	20	$2.66 \cdot 10^{-14}$	$2.11 \cdot 10^{-13}$	1	$9.49 \cdot 10^{-3}$	20
49	8	$1.57 \cdot 10^{-2}$	24	$4.31 \cdot 10^{-16}$	$5.56 \cdot 10^{-14}$	1	$1 \cdot 10^{-2}$	24
50	22	$5.02 \cdot 10^{-2}$	22	$4.3 \cdot 10^{-14}$	$2.5 \cdot 10^{-14}$	1	$5.79 \cdot 10^{-2}$	22
51	55	$8.79 \cdot 10^{-2}$	55	$1.3 \cdot 10^{-10}$	$2.13 \cdot 10^{-11}$	1	0.86	55
52	49	$9.46 \cdot 10^{-2}$	49	$5.77 \cdot 10^{-7}$	$1.51 \cdot 10^{-9}$	1	0.2	52
53	52	$9.76 \cdot 10^{-2}$	148	$2.75 \cdot 10^{-9}$	$1.22 \cdot 10^{-8}$	1	0.24	188
54	30	$5.24 \cdot 10^{-2}$	60	$1.35 \cdot 10^{-7}$	$3.07 \cdot 10^{-8}$	1	$5.91 \cdot 10^{-2}$	68
55	32	$8.64 \cdot 10^{-2}$	184	$1.19 \cdot 10^{-14}$	$4.2 \cdot 10^{-6}$	1	0.15	184
56	56	$9.77 \cdot 10^{-2}$	60	$1.03 \cdot 10^{-7}$	$1.82 \cdot 10^{-9}$	1	0.36	60
57	8	$1.48 \cdot 10^{-2}$	8	$7.89 \cdot 10^{-16}$	$6.62 \cdot 10^{-15}$	1	$3.8 \cdot 10^{-2}$	22
58	52	$7.37 \cdot 10^{-2}$	52	$1.84 \cdot 10^{-9}$	$3.47 \cdot 10^{-9}$	1	0.3	52
59	12	$1.35 \cdot 10^{-2}$	36	$2.69 \cdot 10^{-9}$	$1.21 \cdot 10^{-8}$	1	$8.64 \cdot 10^{-3}$	36
60	6	$9.98 \cdot 10^{-3}$	7	$1.34 \cdot 10^{-7}$	$3.95 \cdot 10^{-9}$	1	$6.45 \cdot 10^{-3}$	10

Table F.4: Numerical results for the methods described in 5.1.1 (EV), 5.1.2 (Syl) and 5.1.3 ($L(x,y)$), part 2.

		C					VP		
p	# refsol	$r_{\max}$	$\epsilon_{\max}$	s	time (s)	# sol	$r_{\max}$	$\epsilon_{\max}$	s
1	32	$4.22 \cdot 10^{-14}$	$5.72 \cdot 10^{-13}$	1	0.24	32	$1.06 \cdot 10^{-14}$	$2.34 \cdot 10^{-4}$	1
2	32	$8 \cdot 10^{-5}$	$1.35 \cdot 10^{-3}$	1	$8.24 \cdot 10^{-2}$	32	$1.14 \cdot 10^{-14}$	$6.93 \cdot 10^{-4}$	1
3	12	$4.32 \cdot 10^{-16}$	$7.89 \cdot 10^{-14}$	1	0.1	12	$1.15 \cdot 10^{-16}$	$7.27 \cdot 10^{-7}$	1
4	2	$1.11 \cdot 10^{-15}$	$6.29 \cdot 10^{-15}$	1	$4.34 \cdot 10^{-2}$	2	$3.49 \cdot 10^{-17}$	$1.29 \cdot 10^{-15}$	1
5	4	$8.82 \cdot 10^{-17}$	$5.65 \cdot 10^{-14}$	1	$4 \cdot 10^{-2}$	4	$2.31 \cdot 10^{-16}$	$5.65 \cdot 10^{-14}$	1
6	12	$4.75 \cdot 10^{-16}$	$1.74 \cdot 10^{-12}$	1	$3.62 \cdot 10^{-2}$	12	$3.25 \cdot 10^{-17}$	$1.44 \cdot 10^{-9}$	1
7	22	$4.03 \cdot 10^{-16}$	$2.35 \cdot 10^{-12}$	1	$4.21 \cdot 10^{-2}$	22	$1.18 \cdot 10^{-16}$	$1.3 \cdot 10^{-3}$	1
8	72	$1.35 \cdot 10^{-15}$	$3.27 \cdot 10^{-12}$	1	0.11	72	$6.03 \cdot 10^{-16}$	$2.99 \cdot 10^{-3}$	1
9	2	$3.79 \cdot 10^{-16}$	$1.89 \cdot 10^{-15}$	1	$9.46 \cdot 10^{-2}$	2	$1.02 \cdot 10^{-16}$	$1.96 \cdot 10^{-15}$	1
10	4	$6.77 \cdot 10^{-17}$	$1.66 \cdot 10^{-13}$	1	$3.19 \cdot 10^{-2}$	4	$5.72 \cdot 10^{-17}$	$1.07 \cdot 10^{-9}$	1
11	6	$1.4 \cdot 10^{-15}$	$7.08 \cdot 10^{-14}$	1	$2.19 \cdot 10^{-2}$	6	$3.55 \cdot 10^{-17}$	$2.28 \cdot 10^{-6}$	1
12	4	$2.26 \cdot 10^{-14}$	$3.98 \cdot 10^{-15}$	1	$3.84 \cdot 10^{-2}$	4	$1.2 \cdot 10^{-16}$	$2.58 \cdot 10^{-16}$	1
13	4	$8.82 \cdot 10^{-17}$	$3.08 \cdot 10^{-15}$	1	$1.16 \cdot 10^{-2}$	4	$2.51 \cdot 10^{-17}$	$3.08 \cdot 10^{-15}$	1
14	8	$2.31 \cdot 10^{-15}$	$4.56 \cdot 10^{-13}$	1	$2.28 \cdot 10^{-2}$	8	$7.16 \cdot 10^{-15}$	$4.56 \cdot 10^{-13}$	1
15	48	$4.51 \cdot 10^{-13}$	$1.91 \cdot 10^{-7}$	1	0.14	48	$1.2 \cdot 10^{-16}$	$1.9 \cdot 10^{-7}$	1
16	48	$8.26 \cdot 10^{-13}$	$3.94 \cdot 10^{-10}$	1	0.13	48	$2.25 \cdot 10^{-16}$	$4.45 \cdot 10^{-8}$	1
17	10	$4.31 \cdot 10^{-15}$	$1.29 \cdot 10^{-12}$	1	$1.43 \cdot 10^{-2}$	10	$6.45 \cdot 10^{-16}$	$2.84 \cdot 10^{-5}$	1
18	8	$5.62 \cdot 10^{-16}$	$1.35 \cdot 10^{-13}$	1	$1.5 \cdot 10^{-2}$	8	$5.01 \cdot 10^{-16}$	$1.59 \cdot 10^{-9}$	1
19	22	$2.08 \cdot 10^{-14}$	$8.42 \cdot 10^{-15}$	1	$3.2 \cdot 10^{-2}$	22	$3.87 \cdot 10^{-16}$	$1.68 \cdot 10^{-9}$	1
20	55	$2.58 \cdot 10^{-11}$	$6.15 \cdot 10^{-12}$	1	0.32	55	$4.95 \cdot 10^{-4}$	0.63	0
21	48	$4.51 \cdot 10^{-13}$	$4.69 \cdot 10^{-11}$	1	0.13	48	$1.63 \cdot 10^{-16}$	$3.4 \cdot 10^{-8}$	1
22	52	$1.81 \cdot 10^{-14}$	$3.4 \cdot 10^{-12}$	1	$8.34 \cdot 10^{-2}$	52	$3.13 \cdot 10^{-14}$	$3.27 \cdot 10^{-5}$	1
23	30	$7.82 \cdot 10^{-7}$	$5.04 \cdot 10^{-5}$	1	$8.87 \cdot 10^{-2}$	30	$3.65 \cdot 10^{-15}$	$3.9 \cdot 10^{-8}$	1
24	8	$1.06 \cdot 10^{-15}$	$2.81 \cdot 10^{-14}$	1	$2.06 \cdot 10^{-2}$	8	$7.7 \cdot 10^{-17}$	$4.15 \cdot 10^{-9}$	1
25	52	$2.31 \cdot 10^{-6}$	$2.83 \cdot 10^{-6}$	1	0.15	52	$1.76 \cdot 10^{-13}$	$2.02 \cdot 10^{-13}$	1
26	12	$1.23 \cdot 10^{-15}$	$3.07 \cdot 10^{-13}$	1	$1.41 \cdot 10^{-2}$	12	$8.63 \cdot 10^{-17}$	$3.59 \cdot 10^{-9}$	1
27	6	$6.77 \cdot 10^{-16}$	$3.45 \cdot 10^{-13}$	1	$1.03 \cdot 10^{-2}$	6	$6.4 \cdot 10^{-17}$	$7.25 \cdot 10^{-7}$	1
28	38	$1.76 \cdot 10^{-12}$	$3.82 \cdot 10^{-8}$	1	$4.42 \cdot 10^{-2}$	38	$2.03 \cdot 10^{-7}$	$1.68 \cdot 10^{-5}$	1
29	11	$1.04 \cdot 10^{-15}$	$9.76 \cdot 10^{-14}$	1	$2.95 \cdot 10^{-2}$	11	$1.16 \cdot 10^{-16}$	$3.64 \cdot 10^{-9}$	1
30	32	$1.2 \cdot 10^{-12}$	$3.07 \cdot 10^{-13}$	1	$6.57 \cdot 10^{-2}$	32	$3.46 \cdot 10^{-14}$	$4.92 \cdot 10^{-5}$	1
31	32	$5.9 \cdot 10^{-6}$	$1.35 \cdot 10^{-3}$	1	$6.04 \cdot 10^{-2}$	32	$1.29 \cdot 10^{-8}$	$6.99 \cdot 10^{-4}$	1
32	16	$6.41 \cdot 10^{-15}$	$2.61 \cdot 10^{-12}$	1	$1.88 \cdot 10^{-2}$	16	$3.72 \cdot 10^{-16}$	$2.61 \cdot 10^{-12}$	1
33	4	$3.59 \cdot 10^{-15}$	$1.34 \cdot 10^{-14}$	1	$1.07 \cdot 10^{-2}$	4	$4.88 \cdot 10^{-17}$	$3.66 \cdot 10^{-15}$	1
34	6	$1.01 \cdot 10^{-15}$	$2.07 \cdot 10^{-13}$	1	$1.17 \cdot 10^{-2}$	6	$1.5 \cdot 10^{-16}$	$6.98 \cdot 10^{-9}$	1
35	8	$1.11 \cdot 10^{-16}$	$1.16 \cdot 10^{-12}$	1	$1.28 \cdot 10^{-2}$	8	$6.92 \cdot 10^{-17}$	$4.64 \cdot 10^{-5}$	1
36	90	$3.1 \cdot 10^{-14}$	$5.93 \cdot 10^{-12}$	1	0.14	90	$6.83 \cdot 10^{-16}$	$7.14 \cdot 10^{-3}$	1
37	4	$2.08 \cdot 10^{-15}$	$9.51 \cdot 10^{-16}$	1	$2.07 \cdot 10^{-2}$	4	$1.53 \cdot 10^{-16}$	$4.01 \cdot 10^{-16}$	1
38	3	$1.56 \cdot 10^{-15}$	$9.71 \cdot 10^{-16}$	1	$7.26 \cdot 10^{-3}$	3	$5.51 \cdot 10^{-17}$	$2.34 \cdot 10^{-16}$	1
39	2	$1.94 \cdot 10^{-16}$	$2.51 \cdot 10^{-12}$	1	$2.55 \cdot 10^{-2}$	2	$5.68 \cdot 10^{-18}$	$6.23 \cdot 10^{-9}$	1
40	3	$3.34 \cdot 10^{-16}$	$4.12 \cdot 10^{-15}$	1	$2.84 \cdot 10^{-2}$	3	$4.07 \cdot 10^{-17}$	$1.9 \cdot 10^{-9}$	1
41	6	$1.01 \cdot 10^{-15}$	$1.76 \cdot 10^{-13}$	1	$1.19 \cdot 10^{-2}$	6	$1.36 \cdot 10^{-16}$	$5.27 \cdot 10^{-9}$	1
42	8	$6.38 \cdot 10^{-14}$	$1.18 \cdot 10^{-13}$	1	$3.66 \cdot 10^{-2}$	8	$6.41 \cdot 10^{-17}$	$1.78 \cdot 10^{-9}$	1
43	18	$1.61 \cdot 10^{-15}$	$2.88 \cdot 10^{-13}$	1	$2.46 \cdot 10^{-2}$	18	$1.9 \cdot 10^{-16}$	$4.18 \cdot 10^{-5}$	1
44	24	$3.15 \cdot 10^{-17}$	$2.83 \cdot 10^{-13}$	1	$4.07 \cdot 10^{-2}$	24	$6.35 \cdot 10^{-7}$	$1.03 \cdot 10^{-6}$	1
45	49	$4.33 \cdot 10^{-12}$	$7.75 \cdot 10^{-10}$	1	0.66	49	$7.75 \cdot 10^{-17}$	$1.91 \cdot 10^{-12}$	1
46	48	$3.05 \cdot 10^{-12}$	$1.55 \cdot 10^{-9}$	1	0.13	48	$1.33 \cdot 10^{-16}$	$3.6 \cdot 10^{-8}$	1
47	4	$1.01 \cdot 10^{-15}$	$3.85 \cdot 10^{-15}$	1	$4.96 \cdot 10^{-2}$	4	$3.78 \cdot 10^{-15}$	$5.25 \cdot 10^{-14}$	1
48	10	$3.32 \cdot 10^{-15}$	$2.11 \cdot 10^{-13}$	1	$1.16 \cdot 10^{-2}$	10	$2.54 \cdot 10^{-16}$	$1.38 \cdot 10^{-5}$	1
49	8	$4.31 \cdot 10^{-16}$	$5.56 \cdot 10^{-14}$	1	$2.85 \cdot 10^{-2}$	8	$2.55 \cdot 10^{-16}$	$1.9 \cdot 10^{-9}$	1
50	22	$3.38 \cdot 10^{-15}$	$4.72 \cdot 10^{-15}$	1	$2.79 \cdot 10^{-2}$	22	$1.45 \cdot 10^{-16}$	$1.54 \cdot 10^{-9}$	1
51	55	$2.24 \cdot 10^{-10}$	$5.16 \cdot 10^{-11}$	1	0.31	55	$3.29 \cdot 10^{-4}$	0.4	0
52	49	$4.33 \cdot 10^{-12}$	$7.75 \cdot 10^{-10}$	1	0.64	49	$1.3 \cdot 10^{-16}$	$1.33 \cdot 10^{-12}$	1
53	52	$9.8 \cdot 10^{-15}$	$2.83 \cdot 10^{-12}$	1	$6.91 \cdot 10^{-2}$	52	$2.1 \cdot 10^{-14}$	$2.52 \cdot 10^{-5}$	1
54	30	$4.07 \cdot 10^{-7}$	$5.04 \cdot 10^{-5}$	1	$7.22 \cdot 10^{-2}$	30	$4.05 \cdot 10^{-15}$	$3.46 \cdot 10^{-8}$	1
55	32	$3.95 \cdot 10^{-15}$	$3.03 \cdot 10^{-12}$	1	$5.06 \cdot 10^{-2}$	32	$2.56 \cdot 10^{-13}$	$1.03 \cdot 10^{-3}$	1
56	56	$5.17 \cdot 10^{-8}$	$3.69 \cdot 10^{-12}$	1	$6.74 \cdot 10^{-2}$	56	$2.02 \cdot 10^{-12}$	$1.14 \cdot 10^{-3}$	1
57	8	$1.03 \cdot 10^{-15}$	$6.62 \cdot 10^{-15}$	1	$1.32 \cdot 10^{-2}$	8	$6.42 \cdot 10^{-17}$	$2.48 \cdot 10^{-9}$	1
58	52	$1.19 \cdot 10^{-6}$	$1.53 \cdot 10^{-6}$	1	0.15	52	$1.34 \cdot 10^{-12}$	$1.54 \cdot 10^{-12}$	1
59	12	$3.97 \cdot 10^{-16}$	$1.23 \cdot 10^{-12}$	1	$1.17 \cdot 10^{-2}$	12	$3.32 \cdot 10^{-17}$	$2.65 \cdot 10^{-9}$	1
60	6	$1.11 \cdot 10^{-15}$	$1.33 \cdot 10^{-13}$	1	$8.87 \cdot 10^{-3}$	6	$2.8 \cdot 10^{-17}$	$5.35 \cdot 10^{-7}$	1

Table F.5: Numerical results for the methods described in 5.2 (C), 5.3 (VP) and 5.1.2 with transformation of variables (Syl), part 1.

		VP		Syl				
p	# refsol	time (s)	# sol	$r_{\max}$	$\epsilon_{\max}$	s	time (s)	# sol
1	32	5.33	32	$1.53 \cdot 10^{-6}$	$1.36 \cdot 10^{-7}$	1	0.41	32
2	32	5.04	32	$1.11 \cdot 10^{-6}$	$9.87 \cdot 10^{-4}$	1	0.17	32
3	12	0.18	12	$8.9 \cdot 10^{-14}$	$1.41 \cdot 10^{-5}$	1	$4.04 \cdot 10^{-2}$	12
4	2	$2.94 \cdot 10^{-2}$	2	$3.08 \cdot 10^{-16}$	$1.27 \cdot 10^{-15}$	1	$3.25 \cdot 10^{-2}$	2
5	4	$5.54 \cdot 10^{-2}$	4	$2.2 \cdot 10^{-16}$	$5.65 \cdot 10^{-14}$	1	$1.87 \cdot 10^{-2}$	4
6	12	0.15	12	$1.24 \cdot 10^{-15}$	$2.29 \cdot 10^{-12}$	1	$5.77 \cdot 10^{-2}$	12
7	22	1.11	22	$1.32 \cdot 10^{-12}$	$2.36 \cdot 10^{-12}$	1	$8.7 \cdot 10^{-2}$	22
8	72	12.05	72	$1.74 \cdot 10^{-15}$	$5.15 \cdot 10^{-3}$	1	0.2	72
9	2	$5.18 \cdot 10^{-2}$	2	$2.6 \cdot 10^{-14}$	$2.4 \cdot 10^{-14}$	1	$2 \cdot 10^{-2}$	2
10	4	$5.8 \cdot 10^{-2}$	4	$3.2 \cdot 10^{-15}$	$3.88 \cdot 10^{-8}$	1	$1.33 \cdot 10^{-2}$	4
11	6	$5.54 \cdot 10^{-2}$	6	$1.98 \cdot 10^{-14}$	$1.14 \cdot 10^{-5}$	1	$1.58 \cdot 10^{-2}$	6
12	4	0.15	4	$5.73 \cdot 10^{-12}$	$5.17 \cdot 10^{-12}$	1	$2.74 \cdot 10^{-2}$	4
13	4	$5.28 \cdot 10^{-2}$	4	$1.06 \cdot 10^{-15}$	$3.08 \cdot 10^{-15}$	1	$1.2 \cdot 10^{-2}$	4
14	8	0.16	8	$2.93 \cdot 10^{-13}$	$9.24 \cdot 10^{-8}$	1	$2.59 \cdot 10^{-2}$	8
15	48	4.1	48	$6.06 \cdot 10^{-5}$	$4.58 \cdot 10^{-2}$	0	0.16	48
16	48	4.1	48	$1.5 \cdot 10^{-10}$	$3.31 \cdot 10^{-6}$	1	0.12	48
17	10	0.14	10	$4.17 \cdot 10^{-13}$	$1.29 \cdot 10^{-12}$	1	$2.65 \cdot 10^{-2}$	10
18	8	0.14	8	$9.99 \cdot 10^{-15}$	$1.35 \cdot 10^{-13}$	1	$2.42 \cdot 10^{-2}$	8
19	22	1.09	22	$5.15 \cdot 10^{-15}$	$6.3 \cdot 10^{-15}$	1	$6.78 \cdot 10^{-2}$	22
20	55	36.05	48	1.99	0.59	0	0.33	82
21	48	4.08	48	$6.32 \cdot 10^{-5}$	$4.47 \cdot 10^{-2}$	0	0.12	48
22	52	5.38	52	$1.14 \cdot 10^{-6}$	$1.32 \cdot 10^{-5}$	1	0.13	52
23	30	1.07	30	$6.23 \cdot 10^{-4}$	1	0	$6.72 \cdot 10^{-2}$	10
24	8	0.35	8	$6.42 \cdot 10^{-16}$	$2.81 \cdot 10^{-14}$	1	$6.32 \cdot 10^{-2}$	8
25	52	11.31	52	$1.69 \cdot 10^{-11}$	0.98	0	0.19	24
26	12	0.14	12	$8.37 \cdot 10^{-15}$	$1.44 \cdot 10^{-8}$	1	$2.5 \cdot 10^{-2}$	12
27	6	$5.27 \cdot 10^{-2}$	6	$7.1 \cdot 10^{-13}$	$1.44 \cdot 10^{-5}$	1	$1.59 \cdot 10^{-2}$	6
28	38	2.49	38	$7.58 \cdot 10^{-11}$	$3.9 \cdot 10^{-5}$	0	$9.07 \cdot 10^{-2}$	38
29	11	0.14	11	$7.75 \cdot 10^{-5}$	$2 \cdot 10^{-2}$	0	$2.81 \cdot 10^{-2}$	11
30	32	5.11	32	$3.74 \cdot 10^{-8}$	0.33	0	0.11	28
31	32	4.9	32	$2.79 \cdot 10^{-7}$	$1.35 \cdot 10^{-3}$	1	0.12	32
32	16	0.44	16	$5 \cdot 10^{-15}$	$2.4 \cdot 10^{-5}$	1	$3.57 \cdot 10^{-2}$	16
33	4	$3.23 \cdot 10^{-2}$	4	$9.33 \cdot 10^{-14}$	$2.46 \cdot 10^{-13}$	1	$1.31 \cdot 10^{-2}$	4
34	6	$8.41 \cdot 10^{-2}$	6	$2.12 \cdot 10^{-15}$	$1.06 \cdot 10^{-8}$	1	$1.71 \cdot 10^{-2}$	6
35	8	$9.68 \cdot 10^{-2}$	8	$1.71 \cdot 10^{-14}$	$1.6 \cdot 10^{-8}$	1	$2.06 \cdot 10^{-2}$	8
36	90	22.05	90	$1.93 \cdot 10^{-13}$	$9.67 \cdot 10^{-3}$	1	0.27	90
37	4	0.15	4	$3.85 \cdot 10^{-12}$	$6.63 \cdot 10^{-12}$	1	$2.24 \cdot 10^{-2}$	4
38	3	$3.06 \cdot 10^{-2}$	3	$8.41 \cdot 10^{-14}$	$2.48 \cdot 10^{-14}$	1	$1.11 \cdot 10^{-2}$	3
39	2	$2.96 \cdot 10^{-2}$	2	$2.55 \cdot 10^{-16}$	$1.01 \cdot 10^{-8}$	1	$9.98 \cdot 10^{-3}$	2
40	3	$8.12 \cdot 10^{-2}$	3	$2.55 \cdot 10^{-13}$	$6.51 \cdot 10^{-9}$	1	$2.01 \cdot 10^{-2}$	3
41	6	$8.32 \cdot 10^{-2}$	6	$5.62 \cdot 10^{-15}$	$1.24 \cdot 10^{-8}$	1	$1.48 \cdot 10^{-2}$	6
42	8	$8.27 \cdot 10^{-2}$	8	$1.1 \cdot 10^{-12}$	$1.06 \cdot 10^{-8}$	1	$2.17 \cdot 10^{-2}$	8
43	18	0.44	18	$1.02 \cdot 10^{-11}$	$4.53 \cdot 10^{-13}$	1	$4.1 \cdot 10^{-2}$	18
44	24	0.8	24	$1.66 \cdot 10^{-9}$	$2.22 \cdot 10^{-8}$	1	$5.5 \cdot 10^{-2}$	24
45	49	5.49	49	$8.46 \cdot 10^{-11}$	$6.08 \cdot 10^{-9}$	1	0.13	49
46	48	4.1	48	$6.44 \cdot 10^{-5}$	0.21	0	0.12	46
47	4	0.45	4	$3.89 \cdot 10^{-12}$	$8.37 \cdot 10^{-12}$	1	$4.16 \cdot 10^{-2}$	4
48	10	0.14	10	$2.04 \cdot 10^{-12}$	$1.03 \cdot 10^{-12}$	1	$2.24 \cdot 10^{-2}$	10
49	8	0.14	8	$3.48 \cdot 10^{-15}$	$5.56 \cdot 10^{-14}$	1	$2.21 \cdot 10^{-2}$	8
50	22	1.11	22	$4.01 \cdot 10^{-13}$	$9.45 \cdot 10^{-13}$	1	$5.73 \cdot 10^{-2}$	22
51	55	36.31	41	2	0.8	0	0.38	33
52	49	5.49	49	$8.92 \cdot 10^{-12}$	$6.44 \cdot 10^{-9}$	1	0.13	49
53	52	5.39	52	$1.45 \cdot 10^{-6}$	$7.26 \cdot 10^{-6}$	1	0.13	52
54	30	1.05	30	$2.44 \cdot 10^{-11}$	1	0	$6.76 \cdot 10^{-2}$	14
55	32	5.29	32	$8.53 \cdot 10^{-11}$	$3.23 \cdot 10^{-5}$	1	0.13	32
56	56	10.9	56	1.73	0.82	0	0.17	56
57	8	0.35	8	$9.21 \cdot 10^{-16}$	$6.62 \cdot 10^{-15}$	1	$3.98 \cdot 10^{-2}$	8
58	52	11.29	52	$3.84 \cdot 10^{-9}$	0.98	0	0.18	39
59	12	0.14	12	$5.44 \cdot 10^{-14}$	$1.58 \cdot 10^{-8}$	1	$2.68 \cdot 10^{-2}$	12
60	6	$5.09 \cdot 10^{-2}$	6	$3.86 \cdot 10^{-14}$	$8.87 \cdot 10^{-6}$	1	$1.49 \cdot 10^{-2}$	6

Table F.6: Numerical results for the methods described in 5.2 (C), 5.3 (VP) and 5.1.2 with transformation of variables (Syl), part 2.

		PHClab					Bertini DP		
p	# refsol	$r_{\max}$	$\epsilon_{\max}$	s	time (s)	# sol	$r_{\max}$	$\epsilon_{\max}$	s
1	32	$5.08 \cdot 10^{-12}$	$3.43 \cdot 10^{-4}$	1	0.52	32	$6.85 \cdot 10^{-13}$	0.54	0
2	32	$3.79 \cdot 10^{-11}$	$1.08 \cdot 10^{-3}$	1	0.5	32	$4.72 \cdot 10^{-13}$	0.54	0
3	12	$8.88 \cdot 10^{-16}$	$7.91 \cdot 10^{-14}$	0	0.16	4	$9.65 \cdot 10^{-13}$	$3.58 \cdot 10^{-13}$	1
4	2	$1.15 \cdot 10^{-15}$	$1.66 \cdot 10^{-14}$	1	0.15	2	$5.93 \cdot 10^{-16}$	$1.3 \cdot 10^{-15}$	1
5	4	$3.82 \cdot 10^{-16}$	$5.65 \cdot 10^{-14}$	0	0.15	3	$7.65 \cdot 10^{-14}$	$8.2 \cdot 10^{-14}$	1
6	12	$1.67 \cdot 10^{-16}$	$1.74 \cdot 10^{-12}$	0	0.16	7	$3.13 \cdot 10^{-15}$	$1.75 \cdot 10^{-12}$	0
7	22	$6.52 \cdot 10^{-16}$	$2.35 \cdot 10^{-12}$	0	0.17	5	$7.25 \cdot 10^{-16}$	$1.92 \cdot 10^{-12}$	1
8	72	$1.43 \cdot 10^{-15}$	$1.41 \cdot 10^{-11}$	0	0.22	9	$1.27 \cdot 10^{-11}$	$2.99 \cdot 10^{-12}$	1
9	2	$1.82 \cdot 10^{-16}$	$1.96 \cdot 10^{-15}$	1	0.15	2	$6.37 \cdot 10^{-17}$	$1.96 \cdot 10^{-15}$	1
10	4	$8.14 \cdot 10^{-16}$	$1.66 \cdot 10^{-13}$	1	0.16	4	$2.98 \cdot 10^{-14}$	$1.61 \cdot 10^{-13}$	1
11	6	$6.45 \cdot 10^{-15}$	$7.08 \cdot 10^{-14}$	1	0.17	6	$3.11 \cdot 10^{-16}$	$9.62 \cdot 10^{-14}$	0
12	4	$5.38 \cdot 10^{-16}$	$3.75 \cdot 10^{-16}$	1	0.15	4	$3.36 \cdot 10^{-16}$	$3.49 \cdot 10^{-16}$	1
13	4	$3.53 \cdot 10^{-16}$	$3.08 \cdot 10^{-15}$	0	0.15	3	$3.3 \cdot 10^{-15}$	$5.47 \cdot 10^{-15}$	1
14	8	$2.31 \cdot 10^{-17}$	$4.56 \cdot 10^{-13}$	1	0.17	8	$1.04 \cdot 10^{-13}$	$5.28 \cdot 10^{-13}$	1
15	48	$1.75 \cdot 10^{-15}$	$1.91 \cdot 10^{-7}$	1	0.4	48	$2 \cdot 10^{-14}$	0.3	0
16	48	$7.45 \cdot 10^{-16}$	$9.1 \cdot 10^{-8}$	1	0.39	48	$1.01 \cdot 10^{-14}$	$1.05 \cdot 10^{-8}$	0
17	10	$1.3 \cdot 10^{-15}$	$1.29 \cdot 10^{-12}$	0	0.18	7	$9.08 \cdot 10^{-14}$	$1.12 \cdot 10^{-12}$	1
18	8	$3.75 \cdot 10^{-16}$	$1.35 \cdot 10^{-13}$	0	0.17	7	$2.01 \cdot 10^{-14}$	$1.44 \cdot 10^{-13}$	1
19	22	$2.87 \cdot 10^{-15}$	$6.36 \cdot 10^{-15}$	0	0.2	21	$9 \cdot 10^{-15}$	$1.05 \cdot 10^{-14}$	1
20	55	$6.75 \cdot 10^{-15}$	$7.5 \cdot 10^{-15}$	1	0.6	55	$5.04 \cdot 10^{-15}$	$1.44 \cdot 10^{-14}$	1
21	48	$2.46 \cdot 10^{-15}$	$9.66 \cdot 10^{-8}$	1	0.4	48	$2.55 \cdot 10^{-11}$	0.48	0
22	52	$2.53 \cdot 10^{-8}$	$3.43 \cdot 10^{-4}$	1	0.4	52	$3.07 \cdot 10^{-12}$	0.56	0
23	30	$1.86 \cdot 10^{-15}$	$5.78 \cdot 10^{-8}$	1	0.29	30	$1.91 \cdot 10^{-13}$	$2.09 \cdot 10^{-6}$	1
24	8	$6.52 \cdot 10^{-16}$	$2.81 \cdot 10^{-14}$	0	0.15	7	$9.62 \cdot 10^{-15}$	$8.16 \cdot 10^{-14}$	1
25	52	$6.01 \cdot 10^{-15}$	$3.55 \cdot 10^{-15}$	1	0.32	52	$7.94 \cdot 10^{-16}$	$3.32 \cdot 10^{-15}$	1
26	12	$3.35 \cdot 10^{-15}$	$3.07 \cdot 10^{-13}$	0	0.17	7	$2.25 \cdot 10^{-14}$	$4.12 \cdot 10^{-13}$	1
27	6	$2 \cdot 10^{-20}$	$3.45 \cdot 10^{-13}$	1	0.17	6	$6.35 \cdot 10^{-15}$	0.71	0
28	38	$2.3 \cdot 10^{-11}$	$6.86 \cdot 10^{-5}$	0	0.42	29	$8.81 \cdot 10^{-15}$	0.19	0
29	11	$2.3 \cdot 10^{-15}$	$9.75 \cdot 10^{-14}$	0	0.16	6	$6.1 \cdot 10^{-16}$	$1.04 \cdot 10^{-13}$	1
30	32	$9.36 \cdot 10^{-14}$	$3.08 \cdot 10^{-13}$	1	0.48	32	$2.14 \cdot 10^{-14}$	0.54	0
31	32	$2.48 \cdot 10^{-10}$	$1.11 \cdot 10^{-3}$	1	0.51	32	$1.17 \cdot 10^{-14}$	0.54	0
32	16	$5.9 \cdot 10^{-15}$	$2.61 \cdot 10^{-12}$	1	0.21	16	$1.39 \cdot 10^{-15}$	0.74	0
33	4	$1.36 \cdot 10^{-15}$	$2.52 \cdot 10^{-14}$	1	0.16	4	$2.17 \cdot 10^{-16}$	$3.9 \cdot 10^{-15}$	1
34	6	$1.85 \cdot 10^{-16}$	$2.07 \cdot 10^{-13}$	0	0.16	5	$8.63 \cdot 10^{-16}$	$1.99 \cdot 10^{-13}$	1
35	8	$5.41 \cdot 10^{-16}$	$1.16 \cdot 10^{-12}$	0	0.17	5	$2.05 \cdot 10^{-13}$	$1.93 \cdot 10^{-12}$	1
36	90	$3.14 \cdot 10^{-14}$	$1.23 \cdot 10^{-3}$	1	0.44	90	$6.45 \cdot 10^{-15}$	0.55	0
37	4	$6.89 \cdot 10^{-16}$	$4.99 \cdot 10^{-15}$	1	0.18	4	$1.74 \cdot 10^{-15}$	$7.56 \cdot 10^{-15}$	1
38	3	$2.83 \cdot 10^{-16}$	$4.91 \cdot 10^{-16}$	1	0.15	3	$1.86 \cdot 10^{-16}$	$2.01 \cdot 10^{-16}$	1
39	2	$2.15 \cdot 10^{-17}$	$2.51 \cdot 10^{-12}$	1	0.15	2	$3.33 \cdot 10^{-16}$	$2.47 \cdot 10^{-12}$	1
40	3	$2 \cdot 10^{-20}$	$4.48 \cdot 10^{-15}$	1	0.15	3	$6.03 \cdot 10^{-15}$	$1.41 \cdot 10^{-14}$	1
41	6	$1.85 \cdot 10^{-16}$	$1.76 \cdot 10^{-13}$	0	0.16	5	$1.51 \cdot 10^{-14}$	$1.7 \cdot 10^{-13}$	1
42	8	$2.39 \cdot 10^{-16}$	$1.18 \cdot 10^{-13}$	1	0.17	8	$2.47 \cdot 10^{-14}$	$7.14 \cdot 10^{-14}$	1
43	18	$3.64 \cdot 10^{-15}$	$2.88 \cdot 10^{-13}$	0	0.17	11	$3.07 \cdot 10^{-15}$	$8.32 \cdot 10^{-13}$	1
44	24	$1.19 \cdot 10^{-13}$	$2.84 \cdot 10^{-13}$	1	0.25	24	0	0	0
45	49	$1.62 \cdot 10^{-15}$	$2.23 \cdot 10^{-12}$	1	0.4	49	$2.93 \cdot 10^{-16}$	$1.72 \cdot 10^{-12}$	1
46	48	$2.41 \cdot 10^{-15}$	$1.08 \cdot 10^{-7}$	1	0.42	48	$3.19 \cdot 10^{-15}$	0.36	0
47	4	$2.71 \cdot 10^{-15}$	$1.96 \cdot 10^{-14}$	1	0.15	4	$1.34 \cdot 10^{-16}$	$2.54 \cdot 10^{-15}$	1
48	10	$3.3 \cdot 10^{-15}$	$2.11 \cdot 10^{-13}$	0	0.18	7	$6.01 \cdot 10^{-14}$	$1.74 \cdot 10^{-14}$	1
49	8	$2.54 \cdot 10^{-16}$	$5.56 \cdot 10^{-14}$	0	0.15	7	$5.06 \cdot 10^{-15}$	$5.09 \cdot 10^{-14}$	1
50	22	$4.42 \cdot 10^{-15}$	$4.72 \cdot 10^{-15}$	0	0.21	21	$1.77 \cdot 10^{-15}$	$2.02 \cdot 10^{-15}$	1
51	55	$6.6 \cdot 10^{-15}$	0.5	0	0.58	54	$9.25 \cdot 10^{-15}$	$5.5 \cdot 10^{-15}$	1
52	49	$1.09 \cdot 10^{-15}$	$1.54 \cdot 10^{-12}$	1	0.39	49	$3.02 \cdot 10^{-16}$	$1.91 \cdot 10^{-12}$	1
53	52	$1.25 \cdot 10^{-9}$	$3.29 \cdot 10^{-4}$	1	0.39	52	$7.12 \cdot 10^{-14}$	0.56	0
54	30	$1.29 \cdot 10^{-15}$	$5.7 \cdot 10^{-8}$	1	0.3	30	$6.08 \cdot 10^{-11}$	$9.17 \cdot 10^{-6}$	1
55	32	$9.38 \cdot 10^{-12}$	$3.03 \cdot 10^{-12}$	0	0.28	15	$7.13 \cdot 10^{-14}$	0.59	0
56	56	$2.75 \cdot 10^{-15}$	$1.82 \cdot 10^{-9}$	0	0.27	52	$3 \cdot 10^{-14}$	$1.64 \cdot 10^{-12}$	1
57	8	$1.03 \cdot 10^{-15}$	$6.62 \cdot 10^{-15}$	0	0.16	7	$8.33 \cdot 10^{-16}$	$1.15 \cdot 10^{-14}$	1
58	52	$5.25 \cdot 10^{-15}$	$3.27 \cdot 10^{-15}$	1	0.28	52	$2.42 \cdot 10^{-15}$	$2.62 \cdot 10^{-15}$	1
59	12	$4.13 \cdot 10^{-15}$	$1.23 \cdot 10^{-12}$	0	0.17	7	$4.29 \cdot 10^{-14}$	$1.15 \cdot 10^{-12}$	1
60	6	$2 \cdot 10^{-20}$	$1.33 \cdot 10^{-13}$	1	0.17	6	$7.57 \cdot 10^{-15}$	0.71	0

Table F.7: Numerical results for PHClab, Bertini and PNLA, part 1.

		Bertini DP		PNLA				
p	# refsol	time (s)	# sol	$r_{\max}$	$\epsilon_{\max}$	s	time (s)	# sol
1	32	0.51	16	$1.78 \cdot 10^{-2}$	$6.8 \cdot 10^{-3}$	0	2.55	32
2	32	0.53	16	$1.51 \cdot 10^{-2}$	$6.41 \cdot 10^{-2}$	0	3.27	32
3	12	0.14	12	$1.08 \cdot 10^{-14}$	$6.23 \cdot 10^{-6}$	1	0.11	12
4	2	0.12	2	$4.44 \cdot 10^{-16}$	$5.3 \cdot 10^{-15}$	1	$1.46 \cdot 10^{-2}$	2
5	4	0.12	4	$9.33 \cdot 10^{-16}$	$5.65 \cdot 10^{-14}$	1	$5.3 \cdot 10^{-2}$	4
6	12	0.2	11	$8.4 \cdot 10^{-5}$	$5 \cdot 10^{-5}$	1	0.13	12
7	22	0.25	22	1.8	0.48	0	0.57	22
8	72	0.57	72	$5.33 \cdot 10^{-15}$	$3.27 \cdot 10^{-12}$	1	1.52	72
9	2	0.14	2	$1.78 \cdot 10^{-15}$	$1.94 \cdot 10^{-15}$	1	$5.02 \cdot 10^{-2}$	2
10	4	0.13	4	$9.68 \cdot 10^{-16}$	$4.86 \cdot 10^{-9}$	1	$4.96 \cdot 10^{-2}$	4
11	6	0.18	4	$9.35 \cdot 10^{-15}$	$8.91 \cdot 10^{-6}$	1	$5.99 \cdot 10^{-2}$	6
12	4	0.14	4	$1.72 \cdot 10^{-13}$	$2.44 \cdot 10^{-14}$	1	0.24	4
13	4	0.12	4	$6.31 \cdot 10^{-16}$	$3.08 \cdot 10^{-15}$	1	$4.75 \cdot 10^{-2}$	4
14	8	0.15	8	$2.35 \cdot 10^{-11}$	$4.22 \cdot 10^{-8}$	1	0.17	8
15	48	1.77	35	$6.36 \cdot 10^{-10}$	$4.17 \cdot 10^{-6}$	1	2.13	48
16	48	1.22	46	$2.47 \cdot 10^{-10}$	$3.44 \cdot 10^{-6}$	1	2.13	48
17	10	0.2	10	$4.88 \cdot 10^{-13}$	$1.29 \cdot 10^{-12}$	1	0.15	10
18	8	0.13	8	$8.64 \cdot 10^{-14}$	$1.35 \cdot 10^{-13}$	1	0.14	8
19	22	0.24	22	$1.8 \cdot 10^{-9}$	$1.65 \cdot 10^{-9}$	1	1.4	22
20	55	0.63	55	$5.14 \cdot 10^{-5}$	$8.76 \cdot 10^{-5}$	1	17.39	55
21	48	1.56	39	$2.64 \cdot 10^{-9}$	$3.48 \cdot 10^{-6}$	1	2.14	48
22	52	0.46	27	1.85	$1.39 \cdot 10^{-3}$	0	2.05	52
23	30	1.11	30	$3.62 \cdot 10^{-3}$	$8.19 \cdot 10^{-3}$	1	0.85	30
24	8	0.39	8	$1.12 \cdot 10^{-15}$	$2.81 \cdot 10^{-14}$	1	0.12	8
25	52	0.32	52	1.35	0.97	0	70.7	40
26	12	0.2	12	$7.56 \cdot 10^{-4}$	$1.24 \cdot 10^{-3}$	1	0.13	12
27	6	0.15	3	$4.48 \cdot 10^{-14}$	$4.13 \cdot 10^{-6}$	1	$6.48 \cdot 10^{-2}$	6
28	38	0.38	25	$8.91 \cdot 10^{-3}$	$1.83 \cdot 10^{-2}$	0	2.14	38
29	11	0.18	11	$3.64 \cdot 10^{-3}$	$2.1 \cdot 10^{-3}$	1	0.16	11
30	32	0.59	16	$1.92 \cdot 10^{-3}$	$8.31 \cdot 10^{-3}$	0	3.03	32
31	32	0.43	14	$4.03 \cdot 10^{-3}$	0.16	0	3.49	32
32	16	0.17	12	$2.86 \cdot 10^{-11}$	$6.01 \cdot 10^{-5}$	1	0.92	16
33	4	0.15	4	$2.5 \cdot 10^{-14}$	$1.57 \cdot 10^{-13}$	1	$4.03 \cdot 10^{-2}$	4
34	6	0.14	6	$2.4 \cdot 10^{-14}$	$6.83 \cdot 10^{-8}$	1	$8.93 \cdot 10^{-2}$	6
35	8	0.13	8	1.9	0.41	0	$7.56 \cdot 10^{-2}$	8
36	90	0.43	72	$9.22 \cdot 10^{-12}$	$9.57 \cdot 10^{-3}$	1	2.79	90
37	4	0.14	4	$4.94 \cdot 10^{-15}$	$3.83 \cdot 10^{-15}$	1	0.45	4
38	3	0.12	3	$1.17 \cdot 10^{-14}$	$2.19 \cdot 10^{-14}$	1	$2.33 \cdot 10^{-2}$	3
39	2	0.12	2	$2.53 \cdot 10^{-16}$	$1.83 \cdot 10^{-8}$	1	$2.36 \cdot 10^{-2}$	2
40	3	0.13	3	$5.12 \cdot 10^{-16}$	$7.46 \cdot 10^{-9}$	1	$7.96 \cdot 10^{-2}$	3
41	6	0.13	6	$3.66 \cdot 10^{-15}$	$1.5 \cdot 10^{-8}$	1	$8.74 \cdot 10^{-2}$	6
42	8	0.13	8	$1.09 \cdot 10^{-13}$	$1.12 \cdot 10^{-8}$	1	$9.74 \cdot 10^{-2}$	8
43	18	0.17	18	0.52	$1.2 \cdot 10^{-4}$	0	0.24	18
44	24	0.48	0	0.26	$7.39 \cdot 10^{-5}$	0	0.42	24
45	49	0.36	49	$1.04 \cdot 10^{-10}$	$8.57 \cdot 10^{-9}$	1	3.07	49
46	48	1.89	42	$9.29 \cdot 10^{-11}$	$4.42 \cdot 10^{-6}$	1	2.14	48
47	4	0.18	4	$1.39 \cdot 10^{-11}$	$2.68 \cdot 10^{-11}$	1	0.82	4
48	10	0.15	10	$6.43 \cdot 10^{-13}$	$4.28 \cdot 10^{-13}$	1	0.15	10
49	8	0.2	8	$1.24 \cdot 10^{-13}$	$5.56 \cdot 10^{-14}$	1	0.14	8
50	22	0.2	22	$8.18 \cdot 10^{-14}$	$3.86 \cdot 10^{-14}$	1	0.99	22
51	55	0.55	55	$1.24 \cdot 10^{-4}$	$2.09 \cdot 10^{-4}$	1	17.3	55
52	49	0.34	49	$1.02 \cdot 10^{-7}$	$2.69 \cdot 10^{-5}$	1	3.11	49
53	52	0.55	25	0.24	$1.32 \cdot 10^{-3}$	0	2.07	52
54	30	1.23	30	$3.75 \cdot 10^{-4}$	$3.1 \cdot 10^{-4}$	1	0.88	30
55	32	0.42	25	1.47	0.4	0	1.82	32
56	56	0.34	56	1.99	0.82	0	2.32	56
57	8	0.13	8	$7.5 \cdot 10^{-16}$	$6.62 \cdot 10^{-15}$	1	0.13	8
58	52	0.34	52	1.31	0.98	0	70.42	40
59	12	0.14	12	1.51	0.61	0	0.13	12
60	6	0.14	3	$1.08 \cdot 10^{-14}$	$5.24 \cdot 10^{-6}$	1	$6.71 \cdot 10^{-2}$	6

Table F.8: Numerical results for PHClab, Bertini and PNLA, part 2.

		PHClab				Bertini DP			
δ	δ^2	$r_{\max}$	s	time (s)	# sol	$r_{\max}$	s	time (s)	# sol
1	1	$1.24 \cdot 10^{-15}$	1	0.15	1	$2.03 \cdot 10^{-16}$	1	0.22	1
2	4	$2.36 \cdot 10^{-15}$	1	0.16	4	$2.92 \cdot 10^{-16}$	1	0.13	4
3	9	$6.42 \cdot 10^{-15}$	1	0.18	9	$5.5 \cdot 10^{-16}$	1	0.14	9
4	16	$3.69 \cdot 10^{-15}$	1	0.21	16	$6.14 \cdot 10^{-16}$	1	0.16	16
5	25	$2.53 \cdot 10^{-15}$	1	0.25	25	$5.18 \cdot 10^{-16}$	1	0.51	25
6	36	$3.95 \cdot 10^{-15}$	1	0.33	36	$1.12 \cdot 10^{-15}$	1	0.23	36
7	49	$8.3 \cdot 10^{-15}$	1	0.39	49	$1.41 \cdot 10^{-15}$	1	0.26	49
8	64	$6.36 \cdot 10^{-15}$	1	0.44	64	$5.83 \cdot 10^{-15}$	1	0.3	64
9	81	$8.21 \cdot 10^{-15}$	1	0.57	81	$5.34 \cdot 10^{-15}$	1	0.41	81
10	100	$6.11 \cdot 10^{-15}$	1	0.68	100	$5.29 \cdot 10^{-15}$	1	0.5	100
11	121	$6.49 \cdot 10^{-15}$	1	0.86	121	$4.38 \cdot 10^{-15}$	1	0.54	121
12	144	$1.34 \cdot 10^{-14}$	1	1.12	144	$9.11 \cdot 10^{-15}$	1	0.81	144
13	169	$1.04 \cdot 10^{-14}$	1	1.38	169	$4.98 \cdot 10^{-15}$	1	1.23	169
14	196	$9.75 \cdot 10^{-15}$	1	1.52	196	$5.33 \cdot 10^{-15}$	1	1.66	196
15	225	$1.4 \cdot 10^{-14}$	1	2.12	225	$6.76 \cdot 10^{-15}$	1	2.32	225
16	256	$1.07 \cdot 10^{-14}$	1	2.41	256	$5.78 \cdot 10^{-15}$	1	3.08	256
17	289	$1.38 \cdot 10^{-14}$	1	3.07	288	$3.54 \cdot 10^{-15}$	1	4.29	289
18	324	$9.91 \cdot 10^{-15}$	0	3.47	320	$7.57 \cdot 10^{-15}$	1	6.03	323
19	361	$1.21 \cdot 10^{-14}$	1	3.96	358	$4 \cdot 10^{-15}$	1	8.34	361
20	400	$1.08 \cdot 10^{-14}$	1	5.01	399	$5.45 \cdot 10^{-15}$	1	10.19	400
21	441	$1.19 \cdot 10^{-14}$	1	5.34	440	$6.54 \cdot 10^{-15}$	1	15.45	441
22	484	$1.11 \cdot 10^{-14}$	1	6.93	484	$3.67 \cdot 10^{-15}$	1	16.65	484
23	529	$8.94 \cdot 10^{-15}$	1	9.6	527	$4.08 \cdot 10^{-15}$	1	24.31	528
24	576	$1.39 \cdot 10^{-14}$	0	10.63	569	$6.07 \cdot 10^{-15}$	1	25.62	575
25	625	$1.44 \cdot 10^{-14}$	1	12.29	621	$6.19 \cdot 10^{-15}$	1	34.59	625
26	676	$9.97 \cdot 10^{-15}$	0	12.27	669	$4.91 \cdot 10^{-15}$	1	42.68	676
27	729	$1.35 \cdot 10^{-14}$	1	13.6	726	$7.49 \cdot 10^{-15}$	1	49.14	728
28	784	$1.38 \cdot 10^{-14}$	0	15.56	769	$8.68 \cdot 10^{-15}$	1	55.14	782
29	841	$1.25 \cdot 10^{-14}$	1	17.73	835	$7.26 \cdot 10^{-15}$	1	72.2	838
30	900	$1.25 \cdot 10^{-14}$	0	18.84	885	$5.57 \cdot 10^{-15}$	1	87.74	899
31	961	$1.62 \cdot 10^{-14}$	1	21.73	959	$5.78 \cdot 10^{-15}$	1	83.8	959
32	1,024	$1.13 \cdot 10^{-14}$	0	24.73	1,007	$5.18 \cdot 10^{-15}$	1	115.09	1,022
33	1,089	$1.27 \cdot 10^{-14}$	0	27.93	1,078	$3.88 \cdot 10^{-15}$	1	133.37	1,088
34	1,156	$1.24 \cdot 10^{-14}$	0	30.56	1,144	$1.01 \cdot 10^{-14}$	1	151.06	1,150
35	1,225	$1.61 \cdot 10^{-14}$	0	35.43	1,205	$3.62 \cdot 10^{-15}$	1	188.87	1,225
36	1,296	$1.81 \cdot 10^{-14}$	0	37.21	1,279	$5.23 \cdot 10^{-15}$	0	214.63	1,241
37	1,369	$1.2 \cdot 10^{-14}$	1	44.46	1,357	$6.12 \cdot 10^{-15}$	0	226.86	1,325
38	1,444	$2.36 \cdot 10^{-14}$	0	51.83	1,422	$4.94 \cdot 10^{-15}$	0	243.93	1,393
39	1,521	$1.33 \cdot 10^{-14}$	0	51.3	1,494	$5.15 \cdot 10^{-15}$	1	284.5	1,516
40	1,600	$1.22 \cdot 10^{-14}$	0	58.67	1,575	$4.51 \cdot 10^{-15}$	0	297.39	1,471

Table F.9: Numerical results for PHClab, Bertini and PNLA for the random test problems of degree 1 up to 40, part 1.

		PNLA				Our solver			
δ	δ^2	$r_{\max}$	s	time (s)	# sol	$r_{\max}$	s	time (s)	# sol
1	1	$7.88 \cdot 10^{-17}$	1	$8.19 \cdot 10^{-2}$	1	$1.02 \cdot 10^{-16}$	1	$2.42 \cdot 10^{-2}$	1
2	4	$2.42 \cdot 10^{-11}$	1	0.2	4	$1.78 \cdot 10^{-15}$	1	$2.11 \cdot 10^{-2}$	4
3	9	$2.95 \cdot 10^{-14}$	1	0.15	9	$2.26 \cdot 10^{-15}$	1	$1.93 \cdot 10^{-2}$	9
4	16	$1.18 \cdot 10^{-10}$	1	0.38	16	$3.88 \cdot 10^{-15}$	1	$2.56 \cdot 10^{-2}$	16
5	25	$1.47 \cdot 10^{-12}$	1	0.88	25	$2.99 \cdot 10^{-14}$	1	$3.77 \cdot 10^{-2}$	25
6	36	$1.74 \cdot 10^{-11}$	1	1.73	36	$7.88 \cdot 10^{-15}$	1	$5.91 \cdot 10^{-2}$	36
7	49	$8.94 \cdot 10^{-2}$	0	3.17	49	$3.2 \cdot 10^{-14}$	1	$8.76 \cdot 10^{-2}$	49
8	64	0.44	0	5.43	64	$4.14 \cdot 10^{-14}$	1	0.16	64
9	81	0.54	0	8.85	81	$1.88 \cdot 10^{-14}$	1	0.24	81
10	100	0.65	0	32.38	98	$1.84 \cdot 10^{-12}$	1	0.4	100
11	121	$4.59 \cdot 10^{-7}$	1	21.24	121	$2.15 \cdot 10^{-13}$	1	0.66	121
12	144	$5.83 \cdot 10^{-3}$	0	31.46	144	$6.23 \cdot 10^{-14}$	1	1.03	144
13	169	0.83	0	45.84	169	$2.14 \cdot 10^{-13}$	1	1.93	169
14	196	1.58	0	425.22	184	$7.99 \cdot 10^{-14}$	1	2.09	196
15	225	$2.88 \cdot 10^{-4}$	0	92.59	225	$8.06 \cdot 10^{-13}$	1	3.31	225
16	256	0.37	0	129.37	256	$3.25 \cdot 10^{-13}$	1	4.46	256
17	289	1.25	0	8,894.44	268	$9.62 \cdot 10^{-12}$	1	6.42	289
18	324	0.99	0	10,583.7	314	$2.72 \cdot 10^{-12}$	1	8.76	324
19	361	0.33	0	329.17	361	$2.73 \cdot 10^{-11}$	1	8.75	361
20	400	1.65	0	24,630.13	387	$6.73 \cdot 10^{-13}$	1	14.39	400
21	441	1.43	0	28,244.92	420	$1.19 \cdot 10^{-12}$	1	19.16	441
22	484	1.15	0	19,180.61	455	$4.16 \cdot 10^{-13}$	1	28.59	484
23	529	1.11	0	26,086.08	463	$1.89 \cdot 10^{-12}$	1	26	529
24	576	1.31	0	29,557.26	548	$1.24 \cdot 10^{-11}$	1	32.79	576
25	625	1.3	0	89,050.39	593	$1.8 \cdot 10^{-12}$	1	48.31	625
26	676	–	–	–	–	$1.92 \cdot 10^{-12}$	1	59.92	676
27	729	–	–	–	–	$3.24 \cdot 10^{-12}$	1	69.79	729
28	784	–	–	–	–	$2.03 \cdot 10^{-11}$	1	86.97	784
29	841	–	–	–	–	$6.18 \cdot 10^{-11}$	1	119.08	841
30	900	–	–	–	–	$1.17 \cdot 10^{-10}$	1	133.13	900
31	961	–	–	–	–	$5.29 \cdot 10^{-12}$	1	157.11	961
32	1,024	–	–	–	–	$2.81 \cdot 10^{-11}$	1	186.6	1,024
33	1,089	–	–	–	–	$3.55 \cdot 10^{-11}$	1	249.41	1,089
34	1,156	–	–	–	–	$7.56 \cdot 10^{-12}$	1	351.18	1,156
35	1,225	–	–	–	–	$1.69 \cdot 10^{-10}$	1	348.09	1,225
36	1,296	–	–	–	–	$4.18 \cdot 10^{-10}$	1	417.45	1,296
37	1,369	–	–	–	–	$8.82 \cdot 10^{-12}$	1	587.35	1,369
38	1,444	–	–	–	–	$1.78 \cdot 10^{-10}$	1	584.98	1,444
39	1,521	–	–	–	–	$5.34 \cdot 10^{-10}$	1	664.51	1,521
40	1,600	–	–	–	–	$1.37 \cdot 10^{-8}$	1	735.08	1,600

Table F.10: Numerical results for PHClab, Bertini and PNLA for the random test problems of degree 1 up to 40, part 2.

Appendix G

An Example in $\mathbb{C}^3$

In this appendix we will propose a way to generalize the approach presented in this text for bivariate polynomial systems to the multivariate case with a number of variables $s > 2$. A complete analysis of the method is beyond the scope of this text. We will present an example in $\mathbb{C}^3$ to fix the ideas. The example system is given by

$$\begin{cases} p_1(x, y, z) = -1 + x^2 + y^2 + z^2 = 0 \\ p_2(x, y, z) = x^2 + y^2 - z = 0 \\ p_3(x, y, z) = x + y - z = 0 \end{cases} \tag{G.1}$$

and we want to find all points $(x, y, z) \in \mathbb{C}^3$ that satisfy all three equations. The system can be solved as a bivariate problem using the substitution $z = x + y$ and the four solutions are found using our bivariate solver. They are given in Table G.1.

G.1 A linear pencil in x, y and z

Recall that in the bivariate case the first step of our solution method was to construct a linear pencil $L(x, y)$ consisting out of a coefficient matrix and a part that defines the used basis. We had

$$L(x, y)\boldsymbol{v}(x, y) = \begin{pmatrix} \Phi_p \\ \Phi_q \\ \mathcal{B}_x - x\mathcal{C}_x \\ \mathcal{B}_y - y\mathcal{C}_y \end{pmatrix} \boldsymbol{v}(x, y) = \begin{pmatrix} p(x, y) \\ q(x, y) \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix},$$

$\Re(x)$	$\Im(x)$	$\Re(y)$	$\Im(y)$	$\Re(z)$	$\Im(z)$
$-8.090170 \cdot 10^{-1}$	$-1.209763 \cdot 10^{0}$	$-8.090170 \cdot 10^{-1}$	$1.209763 \cdot 10^{0}$	$-1.618034 \cdot 10^{0}$	$0.000000 \cdot 10^{0}$
$-8.090170 \cdot 10^{-1}$	$1.209763 \cdot 10^{0}$	$-8.090170 \cdot 10^{-1}$	$-1.209763 \cdot 10^{0}$	$-1.618034 \cdot 10^{0}$	$0.000000 \cdot 10^{0}$
$7.711052 \cdot 10^{-1}$	$0.000000 \cdot 10^{0}$	$-1.530712 \cdot 10^{-1}$	$0.000000 \cdot 10^{0}$	$6.180340 \cdot 10^{-1}$	$0.000000 \cdot 10^{0}$
$-1.530712 \cdot 10^{-1}$	$0.000000 \cdot 10^{0}$	$7.711052 \cdot 10^{-1}$	$0.000000 \cdot 10^{0}$	$6.180340 \cdot 10^{-1}$	$0.000000 \cdot 10^{0}$

Table G.1: Numerical solutions to (G.1) found using our bivariate solver after the substitution $z = x + y$. Every row of the table represents one solution.

where $\boldsymbol{v}(x,y)$ represents the bivariate monomial vector of the appropriate degree. Analogously, in the 3-dimensional case we can construct $L(x,y,z)$ such that

$$L(x,y,z)\boldsymbol{v}(x,y,z) = \begin{pmatrix} \Phi_{p_1} \\ \Phi_{p_2} \\ \Phi_{p_3} \\ \mathcal{B}_x - x\mathcal{C}_x \\ \mathcal{B}_y - y\mathcal{C}_y \\ \mathcal{B}_z - z\mathcal{C}_z \end{pmatrix} \boldsymbol{v}(x,y,z) = \begin{pmatrix} p_1(x,y,z) \\ p_2(x,y,z) \\ p_3(x,y,z) \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix}.$$

For our example problem (G.1) we have

$$\left(\begin{array}{cccccc|cccc|c} -1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 \\ \hline -x & 1 & & & & & & & & \\ & -x & 1 & & & & & & & \\ & & & -x & 1 & & & & & \\ & & & & & & -x & 1 & & \\ \hline -y & & & 1 & & & & & & \\ & & & -y & & 1 & & & & \\ & & & & & & -y & & 1 & \\ \hline -z & & & & & & 1 & & & \\ & & & & & & -z & & & 1 \end{array}\right) \begin{pmatrix} 1 \\ x \\ x^2 \\ y \\ xy \\ y^2 \\ z \\ xz \\ yz \\ z^2 \end{pmatrix} = \left(\begin{array}{c} p_1(x,y,z) \\ p_2(x,y,z) \\ p_3(x,y,z) \\ \hline 0 \\ 0 \\ 0 \\ 0 \\ \hline 0 \\ 0 \\ 0 \\ \hline 0 \\ 0 \end{array}\right). \tag{G.2}$$

Again we used the variable x as much as possible for the linear recurrences in the basis definition, followed by the variable y and we avoided the variable z. Note that this can be written as a square 3-parameter eigenvalue problem of size 10×10 using only one coefficient row for each equation.

G.2 Degree extension

Equation (G.2) does not provide us with enough equations in x to obtain a tall rectangular eigenvalue problem by maintaining only the coefficient rows and the x-rows. We can only hope to find a finite number of eigenvalues if the pencil contains at least as much rows as columns (Appendix E). We will perform a degree extension to obtain a tall REP in x. For a total shift degree $\Delta\delta$, all shifts of the form

$$m(y,z)p_i(x,y,z) = 0, \quad i = 1,2,3$$

with $m(y,z)$ any bivariate monomial in y and z of degree $\leq \Delta\delta$ are added to (G.1)[1]. Let us determine the shift degree $\Delta\delta$ that is needed to obtain a tall pencil in x. The

[1] Again, we do not use shifts in the variable x. This can be understood intuitively by noting that

$$\begin{pmatrix} \hat{\Phi} \\ \hat{\mathcal{B}}_x - x\hat{\mathcal{C}}_x \end{pmatrix} \boldsymbol{v} = \mathbf{0} \Rightarrow \Psi_x \boldsymbol{v} = \mathbf{0}$$

where Ψ_x contains the coefficients of $xp_i(x,y,z)$, $i = 1,2,3$. Therefore, the x-shifts do not contain any new information.

number of monomials in three variables of degree $\leq \hat{\delta}$ (number of columns of the extended pencil $\hat{\Pi}_x(x)$) is given by

$$M = \sum_{i=0}^{\hat{\delta}} \sum_{j=1}^{i+1} j = \frac{\hat{\delta}(\hat{\delta}+1)(2\hat{\delta}+1)}{12} + \frac{3\hat{\delta}(\hat{\delta}+1)}{4} + \hat{\delta} + 1.$$

The number of x-recurrences (number of rows of $\hat{\mathcal{B}}_x - x\hat{\mathcal{C}}_x$) for an extended degree $\hat{\delta}$ is

$$N_x = \sum_{i=1}^{\hat{\delta}} \sum_{j=1}^{i} j = \frac{\hat{\delta}(\hat{\delta}+1)(2\hat{\delta}+1)}{12} + \frac{\hat{\delta}(\hat{\delta}+1)}{4}.$$

The number of rows in $\hat{\Phi}$ is $N_\phi = 3$ and the number of shifts of degree $\leq \Delta\delta$ and > 0 is equal to

$$N_\psi = 3 \sum_{i=2}^{\Delta\delta+1} i = 3\left(\frac{(\Delta\delta+1)(\Delta\delta+2)}{2} - 1\right).$$

In order to obtain a tall pencil, we need

$$\begin{aligned} N_\phi + N_\psi + N_x &\geq M \\ \frac{\hat{\delta}(\hat{\delta}+1)}{4} + \frac{3(\Delta\delta+1)(\Delta\delta+2)}{2} &\geq \frac{3\hat{\delta}(\hat{\delta}+1)}{4} + \hat{\delta} + 1 \\ \delta^2 + 2\delta\Delta\delta + 3\delta - 2\Delta\delta^2 - 6\Delta\delta - 4 &\leq 0 \end{aligned}$$

where we used $\delta = \delta + \Delta\delta$. Let

$$f(\delta, \Delta\delta) = M - N_x - N_\psi - N_\phi = \frac{1}{2}(\hat{\delta}^2 + 2\delta\Delta\delta + 3\delta - 2\Delta\delta^2 - 6\Delta\delta - 4).$$

The function f is plotted in Figure G.1, along with the resulting minimal shift degree corresponding to $\delta = 1, \ldots, 8$. We conclude that for our example of degree $\delta = 2$ we need $\Delta\delta = 2$. The extended system is given by

$$\begin{aligned} p_1 = p_2 = p_3 = yp_1 = yp_2 = yp_3 = zp_1 = zp_2 = zp_3 = y^2p_1 = y^2p_2 = y^2p_3 \ldots \\ \ldots = yzp_1 = yzp_2 = yzp_3 = z^2p_1 = z^2p_2 = z^2p_3 = 0. \end{aligned}$$

The nonzero structure of the resulting extended x-pencil $\hat{\Pi}_{x,r}(x)$ is shown in Figure G.2. Note that the difference between the number of columns and the number of rows is equal to $f(2,2) = -3$. Again it is possible to reduce the pencil size by shortening the x-recurrence chains from the right side. Applying the reduction algorithm we get

$$\hat{\Pi}_x(x) \underset{R}{\longrightarrow} \hat{\Pi}_{x,r}(x)$$

and the nonzero structure of the reduced pencil $\hat{\Pi}_{x,r}(x)$ is shown in the right part of Figure G.2. The eigenvalues of the REP

$$\hat{\Pi}_{x,r}(x)\boldsymbol{v} = \mathbf{0}$$

are found as explained in Appendix E. The set of eigenvalues will be denoted by $\mathcal{X}$. The set $\mathcal{X}$ contains the candidate x-values for the isolated solutions of (G.1).

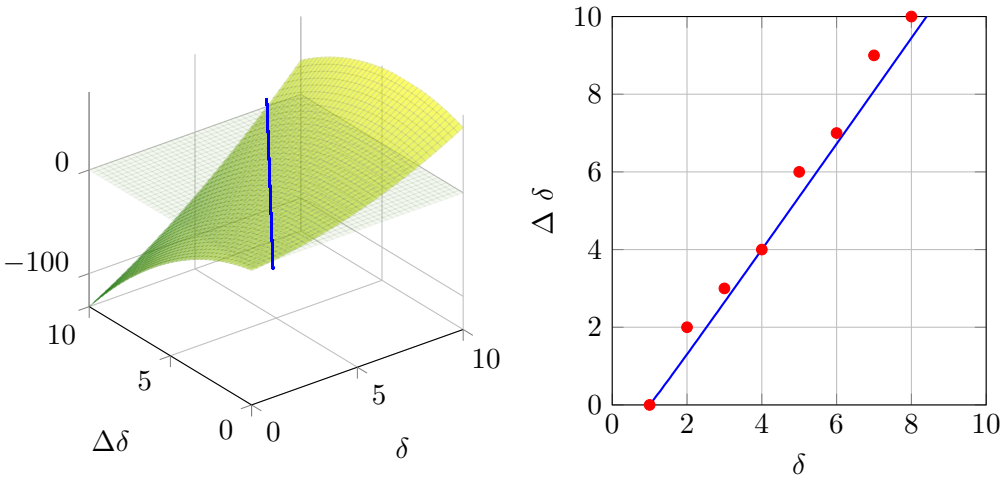


Figure G.1: Left: illustration of the zero level set of $f(\delta, \Delta\delta)$ ( ). Right: resulting shift degrees (indicated by ●) for $\delta = 1, \ldots, 8$.

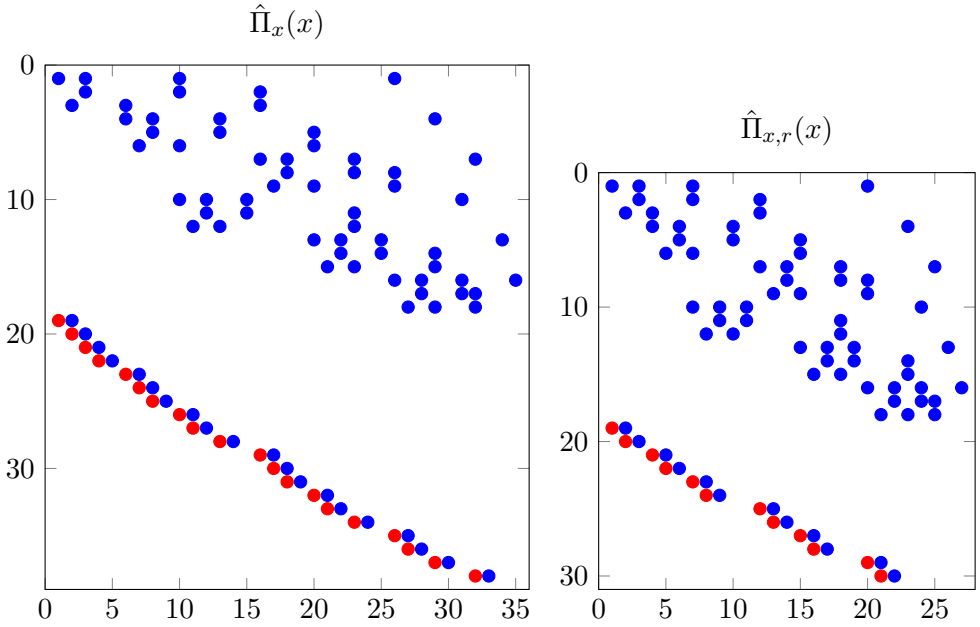


Figure G.2: Left: nonzero structure of the extended x-pencil $\hat{\Pi}_x(x)$. Red dots (●) represent the element $-x$ coming from the matrix $\hat{\mathcal{C}}_x$, blue dots (●) correspond to nonzero elements of $\hat{\Pi}_x(0)$. Right: same representation of $\hat{\Pi}_{x,r}(x)$.

G.3 Obtaining the isolated solutions

We now return to the original linear pencil $L(x, y, z)$ from (G.2). Note that L can be reduced by shortening the recurrence chains from the right:

$$L(x,y,z) \underset{R}{\longrightarrow} L_r(x,y,z) = \left(\begin{array}{ccccc|cc} -1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 & -1 & 0 \\ \hline -x & 1 & & & & & \\ & -x & 1 & & & & \\ \hline -y & & & 1 & & & \\ & & & -y & 1 & & \\ \hline -z & & & & & 1 & \\ & & & & & -z & 1 \end{array}\right)$$

We observe that for each $x^* \in \mathcal{X}$, the first three block rows of $L_r(x^*, y, z)$ form a square linear pencil in y^2. The candidate y-values corresponding to each $x^* \in \mathcal{X}$ are found as the eigenvalues of this GEP. If for some $x^* \in \mathcal{X}$ no finite y-value is found, it is discarded. The other couples (x^*, y^*) are collected and for each such couple, we find the possible z-coordinates as the eigenvalues of the tall REP

$$L_r(x^*, y^*, z)\boldsymbol{v} = \mathbf{0}.$$

The procedure for finding the isolated roots of a 0-dimensional polynomial system in three variables is summarized in Algorithm 3.

Algorithm 3. *Let $p_1, p_2, p_3 \in \mathcal{P}^3$ be given polynomials that define a 0-dimensional system of equations.*

Construct the pencil $L(x, y, z)$.
Determine the appropriate shift degree for this problem and perform the degree extension $L(x,y,z) \underset{E}{\longrightarrow} \hat{L}(x,y,z)$.
Reduce the extended x-pencil $\hat{\Pi}_x(x)$, which is defined as the rows of $\hat{L}$ that do not contain y or z: $\hat{\Pi}_x(x) \underset{R}{\longrightarrow} \hat{\Pi}_{x,r}(x)$.
Find the set $\mathcal{X}$ as the eigenvalues of $\hat{\Pi}_{x,r}(x)$.
Perform the appropriate degree extension $L(x,y,z) \underset{E'}{\longrightarrow} \hat{L}'(x,y,z)$ such that the block row of $\hat{L}'$ that does not contain z is tall.
Reduce $\hat{L}'$: $\hat{L}'(x,y,z) \underset{R}{\longrightarrow} \hat{L}'_r(x,y,z)$ and denote the rows that do not contain z by $\hat{\Pi}'_{xy}(x,y)$.
for** every $x^* \in \mathcal{X}$ **do
    Let $\mathcal{Y}$ be the set of eigenvalues of $\hat{\Pi}'_{xy}(x^*, y)$.

[2] This is not a general result. For $\delta > 2$ a degree extension is needed to obtain a tall problem in y. This shift degree will be smaller than the one used for the first extension step $\underset{E}{\longrightarrow}$. We will denote $L(x,y,z) \underset{E'}{\longrightarrow} \hat{L}'(x,y,z)$.

$\Re(x)$	$\Im(x)$	$\Re(y)$	$\Im(y)$	$\Re(z)$	$\Im(z)$
$-8.090170 \cdot 10^{-1}$	$1.209763 \cdot 10^{0}$	$-8.090170 \cdot 10^{-1}$	$-1.209763 \cdot 10^{0}$	$-1.618034 \cdot 10^{0}$	$1.782122 \cdot 10^{-15}$
$-8.090170 \cdot 10^{-1}$	$-1.209763 \cdot 10^{0}$	$-8.090170 \cdot 10^{-1}$	$1.209763 \cdot 10^{0}$	$-1.618034 \cdot 10^{0}$	$-1.782122 \cdot 10^{-15}$
$7.711052 \cdot 10^{-1}$	$0.000000 \cdot 10^{0}$	$-1.530712 \cdot 10^{-1}$	$0.000000 \cdot 10^{0}$	$6.180340 \cdot 10^{-1}$	$0.000000 \cdot 10^{0}$
$-1.530712 \cdot 10^{-1}$	$0.000000 \cdot 10^{0}$	$7.711052 \cdot 10^{-1}$	$0.000000 \cdot 10^{0}$	$6.180340 \cdot 10^{-1}$	$0.000000 \cdot 10^{0}$

Table G.2: Numerical solutions to (G.1) found by using Algorithm 3. Every row of the table represents one solution.

for *every* $y^* \in \mathcal{Y}$ ***do***
Store the couple (x^*, y^*) *in the set* $\mathcal{S}_{xy}$.
end for
end for
Perform the reduction $L(x, y, z) \xrightarrow[R]{} L_r(x, y, z)$.
for *every couple* $(x^*, y^*) \in \mathcal{S}_{xy}$ ***do***
Let $\mathcal{Z}$ *be the set of eigenvalues of* $L_r(x^*, y^*, z)$.
for *every* $z^* \in \mathcal{Z}$ ***do***
Add (x^*, y^*, z^*) *to the solution set* $\mathcal{S}$.
end for
end for
Output the set $\mathcal{S}$.

The result for the example problem is found using Matlab and it is given in Table G.2. The solutions are equal to the ones in Table G.1 up to machine precision. Figure G.3 shows a graphical representation of the result in $\mathbb{R}^3$.

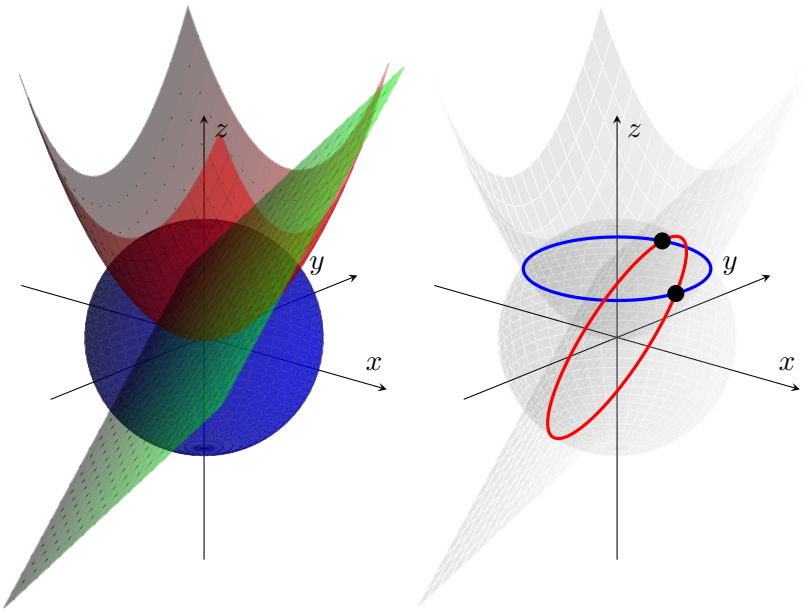


Figure G.3: Left: real zero level sets of p_1 (■), p_2 (■) and p_3 (■) in $\mathbb{R}^3$. Right: the intersection set of p_1 and p_2 (—) and the intersection set of p_1 and p_3 (—). The real numerically found solutions (using Algorithm 3) are represented by black dots (●).

Appendix H

Some Examples in Matlab

In this appendix, we illustrate how to use the Matlab implementations of the concepts that are introduced in this text. All implementations are done in Matlab R2015b. To work with bivariate polynomials, their matrix representation is used. Recall that the matrix representation P of a bivariate polynomial $p(x, y)$ is introduced in Chapter 2 as the matrix that satisfies $p(x, y) = \begin{pmatrix} 1 & y & \dots & y^{\delta_p^y} \end{pmatrix} P \begin{pmatrix} 1 & x & \dots & x^{\delta_p^x} \end{pmatrix}^\top$. In this appendix, we will not make any distinction between $p(x, y)$ and its matrix representation P. Once the folder containing the software is downloaded, it should be selected as the "current folder" in Matlab. As a first step, enter

```
>> addpath('bivar_systems')
```

in order to use the programs. We will briefly discuss how to use the programs to solve any self-defined bivariate problem, how to generate generic systems, how to solve one of the example systems, how to perform an affine transformation of variables and how to evaluate the results. Some of these options are also illustrated by the demo that is implemented, which can be activated by entering

```
>> demo_bivar
```

in the command line.

H.1 Solving a user defined system

Suppose we want to solve the system

$$\begin{cases} p(x, y) = -1 + x^2 + y^2 = 0 \\ q(x, y) = 2 + 6x + 4x^2 + y^2 = 0 \end{cases}$$

which is the system given in Example 2.2.3 with a 4-fold zero in $(-1, 0)$. First, define the polynomials P and Q as follows.

```
>> P = [-1 0 1 ; 0 0 0 ; 1 0 0] ;
>> Q = [2 6 4 ; 0 0 0 ; 1 0 0] ;
```

Several versions of our method have been implemented. An overview is given in Table H.1. We can pick any of these versions to solve the system. For example

`polyrootsEV`	Method described in Subsection 5.1.1 using the eigenvectors of the square pencil.
`polyrootsSyl`	Method described in Subsection 5.1.2 based on the Sylvester matrix of two univariate polynomials.
`polyrootsLxy`	Method described in Subsection 5.1.3 using the linear pencil $L(x, y)$.
`polyrootsC`	Method that uses the coupling based on the residual matrix and connection diagrams.
`polyrootsVP`	Variable precision method, can only be used if the Multiprecision Computing Toolbox for Matlab is installed: `http://www.advanpix.com/`.

Table H.1: Implemented versions of the bivariate system solver proposed in this text.

```
>> sol = polyrootsSyl(P,Q)
```

gives as output

```
sol =

-1.0000 - 0.0000i  -0.0001 + 0.0001i
-1.0000 - 0.0000i   0.0001 - 0.0001i
-1.0000 + 0.0000i  -0.0001 - 0.0001i
-1.0000 + 0.0000i   0.0001 + 0.0001i
-1.0000 + 0.0000i   0.0000 + 0.0002i
-1.0000 + 0.0000i   0.0000 - 0.0002i
-1.0000 + 0.0000i  -0.0002 + 0.0000i
-1.0000 + 0.0000i   0.0002 + 0.0000i
```

where the fact that there are too many solutions should come as no surprise. Using

```
>> sol = polyrootsC(P,Q)
```

we find

```
sol =

-1.0000 - 0.0000i   0.0000 + 0.0000i
-1.0000 - 0.0000i   0.0000 + 0.0000i
-1.0000 - 0.0000i   0.0000 + 0.0000i
-1.0000 - 0.0000i   0.0000 + 0.0000i
```

which is the correct number of solutions. One can check that all four solutions are equal to $(-1, 0)$ up to machine precision.

H.2 Solving a generic system

A "generic" polynomial of degree d having coefficients that are normally distributed with mean 0 and standard deviation 1 corresponding to all monomials of degree $\leq d$ can be generated by the following command.

```
>> P = gen(d) ;
```

Doing the same for Q, we can generate a generic system and solve it like we did in the previous section. The BKK-bound (which is in this case equal to the Bézout number) gives the exact number of solutions that should be found (for random systems, see Appendix B). It can be calculated in the following way.

```
>> BKK = number_of_solutions(P,Q)
```

For example, the commands

```
>> P = randn(4,4); Q = randn(6,6); sol = polyrootsEV(P,Q);
>> BKK = number_of_solutions(P,Q), nsol = size(sol,1)
```

give the following output.

```
BKK =

30


nsol =

30
```

Note that in this example, P and Q are not "generic" as defined previously. For example, P is of degree 6 but its coefficient corresponding to x^6 is zero.

H.3 Solving an example problem

Some example problems are included in the software package. They are found in the directory `example_problems`. The folder contains the set of 60 problems that is used to perform the numerical experiments of Chapter 6. Perhaps the easiest way to load the example problems is the following.

```
>> lst = dir('example_problems') ;
>> files = arrayfun(@(i)lst(i).name,1:length(lst),'UniformOutput',0);
```

Then the n-th problem in the example problem set can be loaded into the workspace using

```
>> L = length(files) ;
>> load(['example_problems/' files{n+L-60}]) ; P = p ; Q = q ;
```

After that, the problem can be solved like we did in the previous sections.

H.4 Affine transformations of variables

The function `transform` can be used to perform an affine transformation of variables. For example, let

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \in \mathbb{C}^{2\times 2}$$

and let $\boldsymbol{b} = \begin{pmatrix} b_1 & b_2 \end{pmatrix}^\top \in \mathbb{C}^2$. We want to calculate the matrix representation of the polynomial $p_t(x,y) = p(A_{11}x + A_{12}y + b_1, A_{21}x + A_{22}y + b_2)$. Note that if p_T vanishes for some couple (x_T^*, y_T^*) $(p_t(x_T^*, y_T^*) = 0)$, then p vanishes for (x^*, y^*):

$$\begin{pmatrix} x^* \\ y^* \end{pmatrix} = A \begin{pmatrix} x_T^* \\ y_T^* \end{pmatrix} + \boldsymbol{b}.$$

It is shown in Chapter 5 that using such a transformation of variables the versions of our method described in Section 5.1 are also able to find all solutions with the correct multiplicity. The function `polyrootsSyl` is equipped with an option to realize this. Let us reconsider the example from Section H.1. The commands

```
>> P = [-1 0 1 ; 0 0 0 ; 1 0 0] ; Q = [2 6 4 ; 0 0 0 ; 1 0 0] ;
>> A = randn(2) ; b = zeros(2,1) ;
>> PT = transform(P,A,b) ; QT = transform(Q,A,b) ;
>> sol = polyrootsSyl(PT,QT,1) ;
>> sol = (A*sol.').'
```

give the output

```
sol =

-1.0000 + 0.0000i   0.0004 + 0.0004i
-1.0000 - 0.0000i  -0.0004 + 0.0004i
-1.0000 - 0.0000i   0.0004 - 0.0004i
-1.0000 + 0.0000i  -0.0004 - 0.0004i
```

which corresponds to all solutions with correct multiplicities. Note that the extra argument of `polyrootsSyl` is needed to ensure that only one y-value is assigned to each x-value.

H.5 Evaluating the results

After solving a system, the real solutions can be verified by plotting the real zero level sets of P and Q in a certain region of the real plane. This can be done by the command `plotprob`.

```
>> P = gen(15) ; Q = gen(15) ; sol = polyrootsC(P,Q) ;
>> realsol = [] ; % Extract the real solutions
```

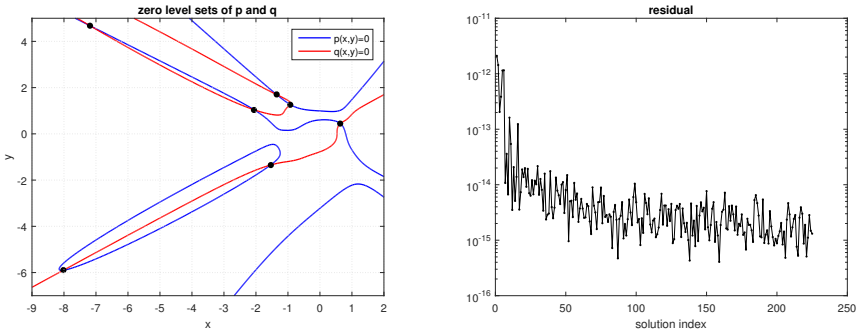


Figure H.1: Left: Matlab plot of the real zero level sets of two generic polynomials of degree 15. Real intersections are indicated with black dots. Right: residuals for all numerical solutions of the same problem.

```
>> for i = 1:size(sol,1)
if norm(imag(sol(i,:))) < 10^-6
realsol = [realsol ; sol(i,:)] ;
end
end
>> figure ; plotprob(P,Q,-9,2,-7,5) ; hold on ;
>> plot(real(realsol(:,1)), real(realsol(:,2)),'k.','markersize',20) ;
```

This gives the Matlab plot on the left side of Figure H.1 as a result. An indication of the quality of all solutions is the residual. This can be calculated by using the `inspectsol` function. For example,

```
>> res = inspectsol(sol.',P,Q) ;
>> nsol = size(sol,1) ;
>> figure ; semilogy(1:nsol, res, 'k.-') ;
>> xlabel('solution index') ; title('residual') ;
```

gives the result shown on the right side of Figure H.1. It can be seen that there are exactly $15^2 = 225$ solutions found and all residuals are small.

Appendix I

Dutch Article

Het Oplossen van Stelsels Veeltermvergelijkingen

Simon Telen
Departement Computerwetenschappen
KU Leuven

Marc Van Barel
Departement Computerwetenschappen
KU Leuven

3 juni 2016

Samenvatting

We stellen een nieuwe methode voor om 0-dimensionale bivariate veeltermstelsels op te lossen gebruik makend van numerieke lineaire algebratools. Er wordt een graadsuitbreiding toegepast op een 2-parameter eigenwaardenprobleem die resulteert in een resultant equivalent met die van Sylvester. Alle oplossingen (reële en complexe) worden berekend met de juiste multipliciteit. De methode slaagt erin om zonder Newton-Raphson verfijningen of uitbreidingen naar hogere precisie nauwkeurige resultaten te bekomen, ook voor meervoudige nulpunten.

1 Inleiding

Multivariate veeltermsystemen duiken op in vele ingenieursdisciplines. Ze vinden hun toepassingen typisch in problemen die een intrinsiek veeltermkarakter hebben of problemen die niet nauwkeurig genoeg beschreven worden door lineaire modellen. Een aantal voorbeelden van toepassingsgebieden zijn chemische ingenieurstechnieken, filterontwerp, computer aided design, robotica, ... Vrij recent zijn er verschillende algoritmische methodes ontwikkeld om veeltermsystemen op te lossen. Groebner basismethodes [12, 13, 5], homotopiemethodes [2, 16] en methodes gebaseerd op resultanten [9, 7, 3, 11, 4] zijn daarvan de belangrijkste klassen. In dit artikel ligt de focus op het tweedimensionale geval. Noteer de ring van bivariate veeltermen met $\mathcal{P}^2$. Het probleem wordt als volgt geformuleerd.

Probleem 1. *Vind alle paren $(x, y) \in \mathbb{C}^2$ die voldoen aan*

$$\begin{cases} p(x,y) = 0 \\ q(x,y) = 0 \end{cases} \tag{1}$$

met $p, q \in \mathcal{P}^2$.

We zullen er steeds vanuit gaan dat (1) een eindig aantal geïsoleerde oplossingen heeft (of ook, de beschouwde systemen zijn 0-dimensionaal). Er wordt in dit artikel een methode gebaseerd op numerieke lineaire algebra voorgesteld om (1) op te lossen. We streven ernaar om niet enkel alle (reële en complexe) oplossingen van (1) te vinden maar ook om hun multipliciteit in rekening te brengen. De methode combineert de voorstelling van (1) als een twee-parameter eigenwaardeprobleem [10] met een graadsuitbreiding (die ook eigen is aan Macaulay resultant-gebaseerde methodes [7, 3]) om een lineair vierkant pencil te verkrijgen in slechts 1 veranderlijke. Er kan aangetoond worden dat de eigenwaarden van dit pencil dezelfde zijn als die van de Sylvesterresultant [13, 12, 5, 1]. Zo kunnen de x-waarden (of y-waarden) gevonden worden als de eigenwaarden van een veralgemeend eigenwaardeprobleem (GEP). In dit artikel illustreren we de methode aan de hand van een eenvoudig voorbeeld en geven we de belangrijkste resultaten. Daarna worden er enkele numerieke experimenten toegelicht en wordt de snelheid en nauwkeurigheid ten opzichte van bestaande oplossingsmethodes geïllustreerd. Voor bewijzen, uitbreidingen naar andere basissen en meer details verwijzen we de lezer naar [14]. In [14] wordt ook een uitbreiding naar meerdere variabelen voorgesteld in bijlage.

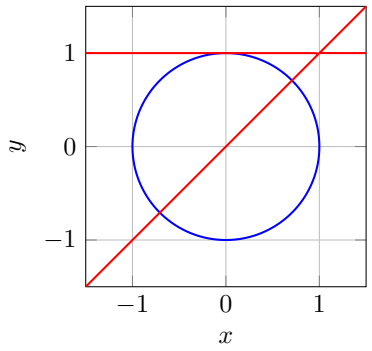
2 Een Eenvoudig Voorbeeld

Beschouw het probleem

$$\begin{cases} p(x,y) = x^2 + y^2 - 1 = 0 \\ q(x,y) = -x + y + xy - y^2 = 0 \end{cases} \tag{2}$$

met als oplossingen $(\sqrt{2}/2, \sqrt{2}/2)$, $(-\sqrt{2}/2, -\sqrt{2}/2)$, $(0,1)$ en $(0,1)$[1]. De reële nulverzamelingen van p en q zijn afgebeeld in Figuur 1. Een eerste stap van de oplossingsmethode is de overgang van (2) naar de

[1] Hier is $(0,1)$ een oplossing met multipliciteit 2. Voor meer uitleg over de multipliciteitstructuur van de oplossingen van multivariate veeltermproblemen verwijzen we naar [12].



Figuur 1: Reële nulverzamelingen van de veeltermen p (——) en q (——) uit (2).

matrixformulering

$$\underbrace{\left(\begin{array}{ccc|cc|c} -1 & 0 & 1 & 0 & 0 & 1 \\ 0 & -1 & 0 & 1 & 1 & -1 \end{array}\right)}_{\Phi} \boldsymbol{v}(x,y) = \begin{pmatrix} p(x,y) \\ q(x,y) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{3}$$

met $\boldsymbol{v}(x,y) = \left(1 \;\; x \;\; x^2 \middle| y \;\; xy \middle| y^2\right)^\top$ de vector van monomialen van graad ≤ 2 en Φ de *coëfficiëntenmatrix*. Het is duidelijk dat elke oplossing (x^*, y^*) een richting genereert in de rechtse nulruimte van Φ. Inderdaad, als $p(x^*, y^*) = q(x^*, y^*) = 0$ dan is $\Phi v(x^*, y^*) = \mathbf{0}$. Echter, niet elke vector in null(Φ) geeft informatie over een oplossing van het stelsel. Beschouw bijvoorbeeld de vector $\boldsymbol{w} = \left(1 \;\; 0 \;\; 1 \middle| 0 \;\; 0 \middle| 0\right)^\top$. Het is duidelijk dat $\Phi \boldsymbol{w} = \mathbf{0}$, maar uit $\boldsymbol{w}$ kunnen we geen informatie halen over een oplossing van (2). Er bestaat immers geen complex koppel (x^*, y^*) zodanig dat de vector $\boldsymbol{w}$ een veelvoud is van $\boldsymbol{v}(x^*, y^*)$. Met andere woorden, $\boldsymbol{w}$ heeft geen *Vandermondestructuur*.

Definitie 2.1. *Een vector $\boldsymbol{w}$ heeft een (bivariate)* Vandermondestructuur *in de monomiaalbasis gegeven door de elementen van $\boldsymbol{v}(x,y)$ als er een koppel $(x^*, y^*) \in \mathbb{C}^2$ bestaat zodanig dat $\boldsymbol{w} = C\boldsymbol{v}(x^*, y^*)$ met $C \in \mathbb{C}_0$.*

Dit leidt tot de volgende stap, waarin we aan (3) twee nieuwe blokrijen toevoegen die de Vandermonde structuur opleggen aan de nulvectoren van Φ:

$$\underbrace{\left(\begin{array}{ccc|cc|c} -1 & 0 & 1 & 0 & 0 & 1 \\ 0 & -1 & 0 & 1 & 1 & -1 \\ \hline -x & 1 & & & & \\ & -x & 1 & & & \\ \hline & & & -x & 1 & \\ \hline -y & & & 1 & & \\ & & & -y & & 1 \end{array}\right)}_{L(x,y)} \left(\begin{array}{c} 1 \\ x \\ x^2 \\ \hline y \\ xy \\ \hline y^2 \end{array}\right) = \begin{pmatrix} p(x,y) \\ q(x,y) \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{4}$$

Merk op dat in (4) de kolommen van $L(x,y)$ verdeeld zijn in blokken die overeenkomen met monomialen van toenemende graad in y. De rijen zijn verdeeld in blokrijen als volgt. De bovenste twee rijen worden gegeven door de coëfficiëntenmatrix Φ. Het volgende blok bestaat uit drie rijen die de variabele x bevatten. We gebruiken de notatie $\mathcal{B}_x - x\mathcal{C}_x$ voor deze deelmatrix. Voor de laatste twee rijen gebruiken we de analoge notatie $\mathcal{B}_y - y\mathcal{C}_y$. Elke vector in de nulruimte van $L(x^*, y^*)$ (verschillend van $\mathbf{0}$), $(x^*, y^*) \in \mathbb{C}^2$ heeft een Vandermonde structuur omwille van de twee onderste blokrijen. Elk zo'n vector komt overeen met een oplossing van (1). Immers, $L(x^*, y^*)C\boldsymbol{v}(x^*, y^*), C \in \mathbb{C}_0$ impliceert dat $p(x^*, y^*) = q(x^*, y^*) = 0$.

Theorema 2.1. *De oplossingen van (2) in $\mathbb{C}^2$ zijn de koppels (x^*, y^*) waarvoor $L(x^*, y^*)$ niet van volle kolomrang is.*

De onderste twee blokrijen moeten opgevat worden als een 'basisdefinitie' van de klassieke monomiaalbasis. Er zijn verschillende mogelijkheden om zo een basisdefinitie op te stellen. Er is gekozen om x en y enkel lineair te laten voorkomen in het resulterende pencil en om zoveel mogelijk het gebruik van y te vermijden. We noteren

$$\begin{cases} p(x,y) = 0 \\ q(x,y) = 0 \end{cases} \xrightarrow[C]{} \quad L(x,y)$$

waar de C staat voor de constructie van het lineaire pencil $L(x,y)$. Om de veranderlijken te scheiden staan er in $L(x,y)$ nog niet genoeg vergelijkingen in x. De bovenste twee blokrijen zijn immers niet van volle kolomrang, en dit voor elke mogelijke waarde van x. Een volgende stap is de graadsuitbreiding. Hier maken we gebruik van de volgende equivalentie

$$\begin{cases} p(x,y) = 0 \\ q(x,y) = 0 \end{cases} \Leftrightarrow \begin{cases} p(x,y) = 0 \\ q(x,y) = 0 \\ yp(x,y) = 0 \\ yq(x,y) = 0 \end{cases}, \forall (x,y) \in \mathbb{C}^2. \tag{5}$$

De veeltermen $yp(x,y)$ en $yq(x,y)$ geven aanleiding tot de coëfficiëntenmatrix

$$\Psi = \left(\begin{array}{cccc|cc|cc|c} 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 1 & -1 \end{array}\right)$$

in de monomiaalbasis

$$\left(1 \;\; x \;\; x^2 \;\; x^3 \middle| y \;\; xy \;\; x^2y \middle| y^2 \;\; xy^2 \middle| y^3\right)^\top.$$

In het *uitgebreide pencil* $\hat{L}(x,y)$ wordt de eerste blokrij gevormd door de coëfficiëntenmatrix $\hat{\Phi}$ van p en q

in de uitgebreide monomiaalbasis. De tweede blokrij is Ψ en de volgende twee blokrijen vormen de basisdefinitie, opnieuw opgedeeld in een x- en een y-deel. Het resultaat is

$$\hat{L}(x,y) = \left(\begin{array}{cccc|cccc|cc|c} -1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 1 & 0 & -1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 1 & -1 \\ \hline -x & 1 & & & & & & & & \\ & -x & 1 & & & & & & & \\ & & -x & 1 & & & & & & \\ & & & & -x & 1 & & & & \\ & & & & & -x & 1 & & & \\ & & & & & & & -x & 1 & \\ \hline -y & & & & 1 & & & & & \\ & & & & -y & & & 1 & & \\ & & & & & & & -y & & 1 \end{array}\right).$$

We gebruiken de notatie $L(x,y) \underset{E}{\longrightarrow} \hat{L}(x,y)$ waar E staat voor de uitbreiding (of 'extensie') van het pencil. Merk op dat de monomiaal x^3 niet voorkomt in de vergelijkingen van het uitgebreide stelsel (5). Inderdaad, we hebben de graad uitgebreid door te vermenigvuldigen met y. De hoogste graad in x is nog steeds 2. Daarom kunnen we de vierde kolom en de zevende rij uit $\hat{L}(x,y)$ schrappen. Het resultaat noemen we $\hat{L}_r(x,y)$, we vinden

$$\hat{L}_r(x,y) = \left(\begin{array}{ccc|ccc|cc|c} -1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 1 & 0 & -1 & 0 & 0 \\ \hline 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 & 1 & -1 \\ \hline -x & 1 & & & & & & & \\ & -x & 1 & & & & & & \\ & & & -x & 1 & & & & \\ & & & & -x & 1 & & & \\ & & & & & & -x & 1 & \\ \hline -y & & & 1 & & & & & \\ & & & -y & & & 1 & & \\ & & & & & & -y & & 1 \end{array}\right)$$

en we noteren $\hat{L}(x,y) \underset{R}{\longrightarrow} \hat{L}_r(x,y)$, waar R staat voor de reductie door het verwijderen van de gepaste rijen en kolommen.

Theorema 2.2. *De oplossingen van* (2) *in* $\mathbb{C}^2$ *zijn de koppels* (x^*, y^*) *waarvoor* $\hat{L}_r(x^*, y^*)$ *niet van volle kolomrang is.*

Noteer de eerste drie blokrijen van $\hat{L}_r(x,y)$ met $\hat{\Pi}_{x,r}(x)$ (in dit deel van het pencil komt de variabele y niet voor). Een nodige voorwaarde opdat $\hat{L}_r(x,y)$ niet van volle kolomrang is, is dat $\hat{\Pi}_{x,r}(x)$ niet van volle kolomrang is. Het probleem 'zoek de waarden van x waarvoor $\hat{\Pi}_{x,r}(x)$ niet van volle kolomrang is' is een vierkant veralgemeend eigenwaardeprobleem. Immers, $\hat{\Pi}_{x,r}(x^*) \in \mathbb{C}^{9\times 9}, \forall x^* \in \mathbb{C}$.

Theorema 2.3. *Beschouw het probleem* (1). *Noteer de graad van* p *in* y *als* δ_p^y *en die van* q *als* δ_q^y. *Construeer het pencil* $\hat{L}_r(x,y)$ *waarbij voor de graadsuitbreiding van* $L(x,y)$ *alle vergelijkingen* $\{y^i p(x,y) = 0\}_{1\le i\le \delta_q^y - 1}$ *en* $\{y^i q(x,y) = 0\}_{1\le i\le \delta_p^y - 1}$ *worden toegevoegd aan het stelsel* (1). *Dan is het resulterende gereduceerde* x*-pencil* $\hat{\Pi}_{x,r}(x)$ *vierkant en er geldt*

$$\det \hat{\Pi}_{x,r}(x) = \gamma \mathrm{res}^{p,q}(x)$$

met $\gamma \in \{-1, 1\}$ *en* $\mathrm{res}^{p,q}(x)$ *de resultant van Sylvester geassocieerd met* (1).

Theorema 2.3 impliceert dat de x-coördinaten van de oplossingen van het probleem (1) eigenwaarden zijn van $\hat{\Pi}_{x,r}(x)$, geconstrueerd zoals voor dit eenvoudige voorbeeld. Meer nog, de multipliciteit van een eindige eigenwaarde x^* van $\hat{\Pi}_{x,r}(x)$ komt overeen met de som van de multipliciteiten van alle oplossingen van (1) van de vorm (x^*, y) (waarbij ook rekening gehouden dient te worden met oplossingen van de vorm (x^*, ∞)). De numerieke waarden voor x zijn voor ons voorbeeldprobleem (gebruik makend van Matlab)

```
-8.896452425993e-09 + 7.205292475198e-10i
8.896453574757e-09 - 7.205292872444e-10i
-7.071067811865e-01 + 0.000000000000e+00i
7.071067811865e-01 + 1.357635865057e-17i
```

waarnaar we zullen verwijzen als $\tilde{x}_1$, $\tilde{x}_2$, $\tilde{x}_3$ en $\tilde{x}_4$ respectievelijk.

Om de bijhorende y-waarden te vinden zijn er verschillende benaderingen mogelijk. In dit artikel illustreren we een methode die alle mogelijke y-waarden berekent en dan een koppeling vindt tussen de set x- en y-waarden. De methode slaagt er goed in om de juiste informatie wat betreft de multipliciteiten te achterhalen. De koppeling is gebaseerd op een groepering van alle gevonden x- en y-waarden.

We willen de numerieke benaderingen van eenzelfde x- of y-waarde in dezelfde groep onderbrengen. Dit gebeurt op basis van de veronderstelling dat x-waarden in dezelfde groep een klein residu[2] hebben

[2]Het *residu* is gedefiniëerd met een gemengd absoluut en relatief criterium:

$$r(x^*, y^*) = \frac{|p(x^*, y^*)|}{|p|(|x^*|, |y^*|) + 1} + \frac{|q(x^*, y^*)|}{|q|(|x^*|, |y^*|) + 1}$$

met $|p|(x,y) \triangleq \sum_{i=0}^{\delta} \sum_{j=0}^{\delta - i} |p_{ij}| x^j y^i$ en $|q|(x,y) \triangleq \sum_{i=0}^{\delta} \sum_{j=0}^{\delta - i} |q_{ij}| x^j y^i$.

gekoppeld met dezelfde set y-waarden. Voor y vinden we de numerieke waarden

```
1.000000000000e+00 + 0.000000000000e+00i
9.999999999999e-01 + 0.000000000000e+00i
-7.071067811865e-01 + 0.000000000000e+00i
7.071067811865e-01 + 0.000000000000e+00i
```

waarnaar we zullen verwijzen als $\tilde{y}_1$, $\tilde{y}_2$, $\tilde{y}_3$ en $\tilde{y}_4$ respectievelijk. De residumatrix voor dit probleem is

$$R(\tilde{\mathcal{X}}, \tilde{\mathcal{Y}}) = \begin{array}{c} \\ \tilde{y}_1 \\ \tilde{y}_2 \\ \tilde{y}_3 \\ \tilde{y}_4 \end{array} \begin{array}{c} \begin{array}{cccc} \tilde{x}_1 & \tilde{x}_2 & \tilde{x}_3 & \tilde{x}_4 \end{array} \\ \begin{pmatrix} \tilde{0} & \tilde{0} & \times & \times \\ \tilde{0} & \tilde{0} & \times & \times \\ \times & \times & \tilde{0} & \times \\ \times & \times & \times & \tilde{0} \end{pmatrix} \end{array}.$$

Daarbij stelt $R(\tilde{\mathcal{X}}, \tilde{\mathcal{Y}})_{ij}$ het residu voor van het koppel (x_j, y_i). Het symbool $\tilde{0}$ stelt een positief getal voor dat kleiner is dan een bepaalde (kleine) drempelwaarde ϵ (in dit geval geeft $\epsilon = 10^{-15}$ al de gewenste opdeling). Een '$\times$' staat voor een getal $> \epsilon$. Uit de residumatrix besluiten we dat $\tilde{x}_1$ en $\tilde{x}_2$ tot eenzelfde groep behoren (ze zorgen voor kleine residu's met dezelfde set y-waarden). Hetzelfde geldt voor $\tilde{y}_1$ en $\tilde{y}_2$. Alle andere x- en y-waarden worden apart gegroepeerd. Uit de residumatrix kunnen we afleiden dat er een koppeling moet zijn tussen de groep $\{\tilde{x}_1, \tilde{x}_2\}$ en de groep $\{\tilde{y}_1, \tilde{y}_2\}$ met multipliciteit 2. De andere koppelingen zijn $(\tilde{x}_3, \tilde{y}_3)$ en $(\tilde{x}_4, \tilde{y}_4)$. Merk op dat de enkelvoudige oplossingen gevonden zijn tot op machinenauwkeurigheid. De tweevoudige oplossing is 'opgesplitst' in twee oplossingen die op een afstand van orde 10^{-8} van de originele oplossing verwijderd liggen. Nemen we echter het gemiddelde van beide groepen $\{\tilde{x}_1, \tilde{x}_2\}$ en $\{\tilde{y}_1, \tilde{y}_2\}$ dan vinden we het punt $(0, 1)$ terug tot op machineprecisie.

3 Numerieke Experimenten

In deze paragraaf bespreken we eerst kort de resultaten van de vergelijking van onze methode met een aantal bestaande oplossingsmethodes. Daarna worden er enkele interessante numerieke voorbeelden gegeven.

3.1 Een vergelijking

De voorgestelde methode is getest op een set problemen en vergeleken met andere solvers. Een tool die uiterst geschikt is om snel een idee te geven van de prestatie van een bepaalde solver in vergelijking met andere solvers is een *performantieprofiel* [6]. Beschouw een set S van kandidaten (solvers) en een set P van problemen. De performantiecurve voor een solver $s \in S$ is gegeven door

$$\rho_s(\tau) = \frac{|\{p \in P \mid t_{p,s} \leq 2^\tau (\min_{s \in S} t_{p,s})\}|}{|P|} \qquad (6)$$

met $|\cdot|$ de cardinaliteit van een verzameling en $t_{p,s}$ de tijd die een solver s nodig had om het probleem p 'succesvol' op te lossen. Met 'succesvol' wordt hier bedoeld dat er evenveel oplossingen zijn gevonden als het aantal oplossingen in de referentieset[3] en dat aan het volgende voldaan is. Er moet een bijectieve afbeelding $b : \mathcal{S}_{\text{ref}} \to \tilde{\mathcal{S}}$ bestaan zodat voor elke oplossing $\boldsymbol{s} \in \mathcal{S}_{\text{ref}}$ geldt dat

$$\|\boldsymbol{s} - \tilde{\boldsymbol{s}}\|_2 \leq 10^{-2}(1 + \|\boldsymbol{s}\|_2),$$

met $\tilde{\boldsymbol{s}} \triangleq b(\boldsymbol{s})$. Deze test is geïmplementeerd met behulp van `bipartite_matching` van de gaimc Matlab toolbox voor graafalgoritmen [8]. Het performantieprofiel voor de vergelijking van onze methode met PHClab [16, 15], Bertini [2] en PNLA [7, 3] is te zien in Figuur 2 (links). Er is voor elke methode gebruik gemaakt van de standaardinstellingen, zonder variabele precisie of verfijningsopties voor de oplossingen[4]. Voor een eerste vergelijking is een set van 60 lage graadsproblemen met uitdagende meervoudige oplossingen gebruikt. Bemerk dat onze methode alle problemen succesvol oplost, rekening houdend met de multipliciteiten en dat voor meer dan 90% van de problemen op de snelste manier doet. De homotopiemethodes en PNLA lossen minder dan 80% van de problemen succesvol op. Een tweede vergelijking is gebeurd op basis van een set random problemen van graad 1 tot en met 40. Met een 'random' probleem van graad δ bedoelen we een probleem waarbij p en q normaalverdeelde coëfficiënten met gemiddelde 0 en standaardafwijking 1 hebben bij alle monomialen van graad $\leq \delta$. Alle oplossingen van zo een generiek stelsel zijn enkelvoudig en er zijn er δ^2 volgens de stelling van Bézout. Het criterium voor succes is in dit geval dat er meer dan 99% van alle δ^2 oplossingen moet gevonden zijn met een residu $< 10^{-6}$. Het performantieprofiel is rechts op Figuur 2 weergegeven. De testen voor PNLA zijn niet volledig voltooid omdat ze te veel tijd vergen. Voor het probleem van graad 25 waren er ongeveer 25 uren nodig. Wat opvalt is

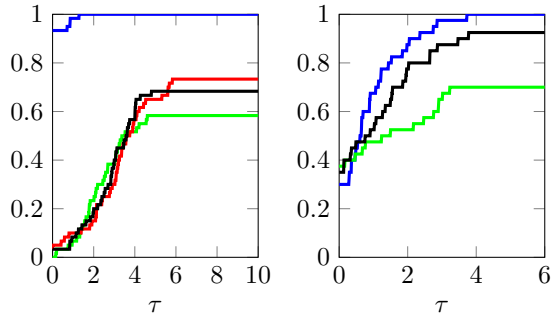
[3] De referensieset bestaat uit de oplossingen gevonden door Bertini in variabele precisie, een zeer betrouwbare solver. Omwille van de variabele precisie is deze oplossingsmethode wel beduidend trager dan de andere (met standaardinstellingen).

[4] Voor Bertini is `MPTYPE: 0` gebruikt, omdat er default `MPTYPE: 1` (adaptive precision) wordt gebruikt. Voor PNLA gebruiken we de `sparf` functie.

$\Re(x)$	$\Im(x)$	$\Re(y)$	$\Im(y)$
2	$4.592 \cdot 10^{-15}$	2	$6.480 \cdot 10^{-15}$
2	$4.592 \cdot 10^{-15}$	2	$6.480 \cdot 10^{-15}$
2	$4.592 \cdot 10^{-15}$	2	$6.480 \cdot 10^{-15}$
$1.571 \cdot 10^{0}$	$1.207 \cdot 10^{-15}$	$-1.429 \cdot 10^{-1}$	$-1.523 \cdot 10^{-16}$
1	$-5.990 \cdot 10^{-16}$	1	$9.605 \cdot 10^{-16}$
1	$-5.990 \cdot 10^{-16}$	1	$9.605 \cdot 10^{-16}$
1	$-5.990 \cdot 10^{-16}$	1	$9.605 \cdot 10^{-16}$
1	$-5.990 \cdot 10^{-16}$	1	$9.605 \cdot 10^{-16}$
1	$-5.990 \cdot 10^{-16}$	1	$9.605 \cdot 10^{-16}$

Tabel 1: Numerieke oplossingen van (7).

dat voor hogere graadsproblemen onze methode het niet haalt van Bertini of PHClab wat de snelheid betreft. Er is echter wel telkens 100% van alle oplossingen gevonden met een residu $< 10^{-6}$ terwijl alle berekeningen gebeuren in dubbele precisie en er geen Newton-Raphson verfijning gebruikt is.



Figuur 2: Links: Performance profiel voor de vergelijking van onze methode (——), PHClab (——), Bertini (——) en PNLA (——) voor de set van 60 testproblemen. Rechts: analoog voor de random problemen van graad 1 tot en met 40.

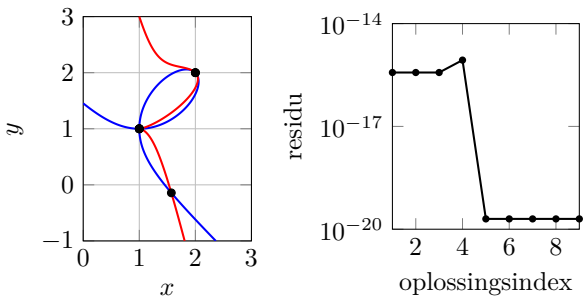
3.2 Enkele voorbeelden

- Beschouw het systeem gegeven door

$$\begin{cases} p(x,y) = & -4 + 5x - 3x^2 + x^3 + 5y - 2xy - 3y^2 \\ & +y^3 = 0 \\ q(x,y) = & -4 + x - 2x^2 + 2x^3 + 9y + 2xy - 4x^2y \\ & -8y^2 + 3xy^2 + y^3 = 0 \end{cases} \tag{7}$$

dat enkel reële oplossingen heeft, waaronder een drievoudige in het punt $(2,2)$ en een vijfvoudige in $(1,1)$. De reële nulverzamelingen van p en q zijn weergegeven in Figuur 3. De numerieke oplossingen gevonden door onze solver zijn gegeven in Tabel 1. Bemerk dat zelfs de vijfvoudige oplossing tot op machineprecisie is teruggevonden.

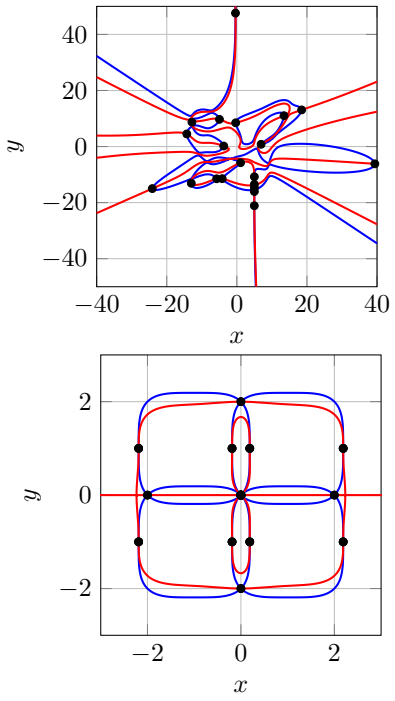
- Ter illustratie zijn in Figuur 4 twee voorbeelden gegeven van reële nulverzamelingen en de gevonden numerieke reële oplossingen. Voor het eerste probleem (bovenaan op Figuur 4) is p van graad 10 en q van graad 9. Er zijn 90 oplossingen die allemaal worden gevonden met een klein residu. Voor het tweede probleem heeft p graad 16 en q graad 15. Alle oplossingen zijn meervoudig en er zijn er in totaal 192 (meervoudigheden meegeteld). Het punt $(0,0)$ is bijvoorbeeld een 16-voudig nulpunt. Resultaten zijn weergegeven in Figuur 5. De residu's liggen hoger voor het tweede probleem ten gevolge van de hogere graad en de meervoudigheden. Voor beide problemen voldoet de numerieke oplossingsverzameling aan het succescriterium dat beschreven werd voor het opstellen van de performantieprofielen.



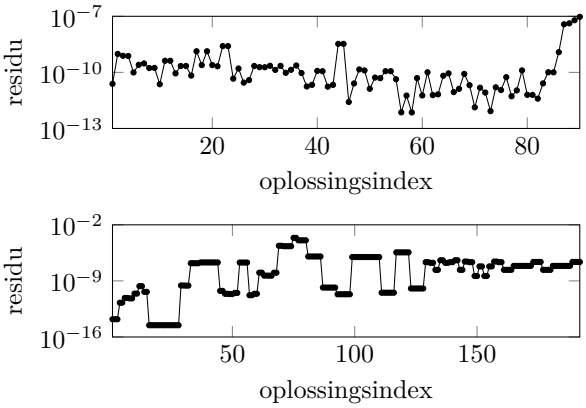
Figuur 3: Links: Reële nulverzamelingen van p (——) en q (——) uit (7) en het reële deel van de gevonden numerieke oplossingen (•). Rechts: residu voor alle 9 numerieke oplossingen.

4 Besluit

De solver die in dit artikel wordt voorgesteld gebruikt numerieke lineaire algebratools om de geïsoleerde oplossingen van bivariate stelsels veeltermvergelijkingen te vinden. Er wordt op een intuïtieve manier een algoritme opgesteld om een lineair pencil te construeren waarvan de eigenwaarden de x- of y-coördinaten zijn van de oplossingen. Inderdaad, we hebben aangetoond dat de determinant van dit intuïtief bekomen pencil sterk verbonden is met de Sylvesterresultant. Die link maakt het mogelijk voor onze solver om ook informatie over de multipliciteit van de oplossingen te geven. De coëfficienten van de gegeven veeltermen komen ongemanipuleerd voor in het veralgemeend eigenwaardeprobleem, wat de nauwkeurigheid ten goede komt. Uit de numerieke experimenten blijkt dat de solver er voor lagere graden (≤ 20) in slaagt om snel nauwkeurige oplossingen te bekomen met de juiste multipliciteiten. Voor generieke problemen met een graad tot zeker 40 worden alle oplossingen gevonden met een kleine residu en binnen een aanvaardbare tijd. Het residu groeit echter met



Figuur 4: Reële nulverzamelingen van twee voorbeeldproblemen. De zwarte punten geven de reële numerieke oplossingen aan.



Figuur 5: Bovenaan: residu voor alle 90 oplossingen van het probleem bovenaan op Figuur 4. Onderaan: residu voor alle 192 oplossingen van het probleem onderaan op Figuur 4.

de graad en homotopiegebaseerde methodes blijken sneller voor graden > 15. Deze methodes slagen er echter niet in om alle oplossingen terug te vinden voor generieke problemen van graad > 35 (bij gebruik van de standaardinstellingen en dubbele precisie).

Referenties

[1] S. Basu, R. Pollack, and M.-F. Roy. *Algorithms in Real Algebraic Geometry*. Springer, 2006.

[2] D. J. Bates, J. D. Hauenstein, A. J. Sommese, and C. W. Wampler. *Numerically solving polynomial systems with Bertini*, volume 25 of *Software, Environments and Tools*. SIAM, 2013.

[3] K. Batselier. *A Numerical Linear Algebra Framework for Solving Problems with Multivariate Polynomials*. KU Leuven - Faculty of Engineering Science, 2013. PhD thesis, promotor: Bart De Moor.

[4] L. Busé, H. Khalil, and B. Mourrain. *Resultant-Based Methods for Plane Curves Intersection Problems*. Springer-Verlag Berlin Heidelberg, 2016.

[5] D. Cox, J. Little, and D. O'Shea. *Ideals, Varieties and Algorithms*. Springer, 2007.

[6] E. D. Dolan and J. J. Moré. Benchmarking optimization software with performance profiles. *Mathematical Programming*, 91:201–213, 2002.

[7] P. Dreesen. *Back to the Roots*. KU Leuven - Faculty of Engineering Science, 2013. PhD thesis, promotor: Bart De Moor.

[8] D. Gleich. gaimc: Graph algorithms in matlab code. Matlab File Exchange, `http://www.mathworks.com/matlabcentral/fileexchange/24134-gaimc---graph-algorithms-in-matlab-code`.

[9] Y. Nakatsukasa, V. Noferini, and A. Townsend. *Computing the common zeros of two bivariate functions via Bézout resultants*. Springer-Verlag Berlin Heidelberg, 2014.

[10] B. Plestenjak and M. E. Hochstenbach. Roots of bivariate polynomial systems via determinantal representations. *SIAM J. Sci. Comput.*, 38(2):A765–A788, 2015.

[11] L. Sorber, M. Van Barel, and L. De Lathauwer. Numerical solution of bivariate and polyanalytic polynomial systems. *SIAM J. Num. Anal. 52*, pages 1551–1572, 2014.

[12] H. J. Stetter. *Numerical Polynomial Algebra*. Society for Industrial and Applied Mathematics, 2004.

[13] B. Sturmfels. *Solving Systems of Polynomial Equations*. Number 97 in CBMS Regional Conferences. Amer. Math. Soc., 2002.

[14] S. Telen. *Solving Systems of Polynomial Equations*. KU Leuven, Department of Computer Science, 2016. Master thesis, promotor: Marc Van Barel.

[15] J. Verschelde. *Homotopy Continuation Methods for Solving Polynomial Systems*. KU Leuven - Faculty of Engineering Science, 1996. PhD thesis, promotor: Ann Haegemans.

[16] J. Verschelde. Algorithm 795: Phcpack: A general-purpose solver for polynomial systems by homotopy continuation. *ACM Transactions on Mathematical Software Vol. 25, No. 2*, pages 251–276, 1999.

Appendix J

Poster

Solving Systems of Polynomial Equations

Simon Telen
Promotor: Prof. dr. ir. Marc Van Barel
KU Leuven - Faculty of Engineering Science
Master of Mathematical Engineering

2015-2016

Problem Statement

Find all vectors $(x_1, x_2, \ldots, x_s)^\top \in \mathbb{C}^s$ that satisfy

$$\begin{cases} p_1(x_1, x_2, \ldots, x_s) = 0 \\ p_2(x_1, x_2, \ldots, x_s) = 0 \\ \vdots \\ p_s(x_1, x_2, \ldots, x_s) = 0 \end{cases}$$

where $p_i(x_1, x_2, \ldots, x_s)$, $1 \leq i \leq s$ are polynomials.
Emphasis on the two-dimensional case: find all vectors $(x, y)^\top \in \mathbb{C}^2$ that satisfy

$$\begin{cases} p(x, y) = 0 \\ q(x, y) = 0 \end{cases} \tag{1}$$

where $p(x, y)$ and $q(x, y)$ are bivariate polynomials.

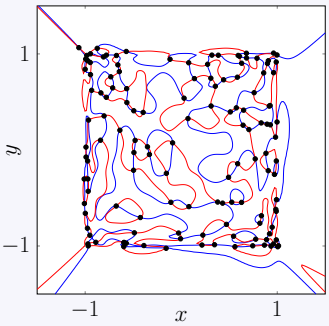


Figure: Plot of the zero level lines in $\mathbb{R}^2$ of two bivariate polynomials $p(x, y)$ (–) and $q(x, y)$ (–) of degree 20. The real solutions are indicated by black dots (•).

Objective

- Solve (1) using Numerical Linear Algebra tools.
- Find all solutions in $\mathbb{C}^2$.
- Take multiplicities into account.

Applications

Applications in chemical engineering, civil engineering, signal processing and filter design, system identification, robotics, mechanical systems design,...

- *Polynomial optimization.* Find the width x^* and the length y^* of a rectangular piece of cardboard with area 1 that minimize the diagonal length. The problem can be formulated as

$$(x^*, y^*) = \underset{x,y}{\operatorname{argmin}}\ x^2 + y^2$$
$$\text{subject to} \quad xy - 1 = 0.$$

$$\mathcal{L}(x, y, z) = x^2 + y^2 - z(xy - 1) \rightarrow \begin{cases} \frac{\partial \mathcal{L}}{\partial x} = 2x - zy = 0 \\ \frac{\partial \mathcal{L}}{\partial y} = 2y - zx = 0 \\ \frac{\partial \mathcal{L}}{\partial z} = xy - 1 = 0 \end{cases}$$

- *Equilibrium concentrations in chemical reactions.*

$$\mathrm{H_2O} \rightleftharpoons 2\mathrm{H} + \mathrm{O}$$
$$\begin{cases} x_H + 2x_{H_2O} = T_H \\ x_O + x_{H_2O} = T_O \\ K x_{H_2O} = x_H^2 x_O \end{cases}$$

Existing Methods

- Groebner bases: symbolic computation [1].
- Resultant based methods: Sylvester, Bézout, Macaulay resultants [2, 3, 5, 7].
- Two-parameter eigenvalue approach [6].
- Homotopy Continuation: a hybrid approach, homotopy and numerical continuation [4].
- Contouring algorithms [5].

[1] Hans J. STETTER, *Numerical Polynomial Algebra*. Society for Industrial and Applied Mathematics, Philadelphia, 2004.
[2] Philippe DREESEN, *Back to the Roots*. PhD thesis under supervision of prof. Bart DE MOOR. KU Leuven - Faculty of Engineering Science, 2013.
[3] Kim BATSELIER *A Numerical Linear Algebra Framework for Solving Problems with Multivariate Polynomials*. PhD thesis under supervision of prof. Bart DE MOOR. KU Leuven - Faculty of Engineering Science, 2013.
[4] Jan VERSCHELDE, *Homotopy Continuation Methods for Solving Polynomial Systems*, PhD thesis under supervision of prof. A. HAEGEMANS, KU Leuven - faculty of engineering science, 1996.
[5] Yuji NAKATSUKASA, Vanni NOFERINI, Alex TOWNSEND, *Computing the common zeros of two bivariate functions via Bézout resultants*. Springer-Verlag Berlin Heidelberg, 2014.
[6] Bor PLESTENJAK, Michiel E. HOCHSTENBACH, *Roots of bivariate polynomial systems via determinantal representations*. June 7, 2015.
[7] Laurent SORBER, Marc VAN BAREL, Lieven DE LATHAUWER, *Numerical solution of bivariate and polyanalytic polynomial systems*. SIAM J. Num. Anal. 52 (2014) 1551-1572.

A two-parameter eigenvalue approach

$$\begin{cases} p(x, y) = x^2 + y^2 - 4 = 0 \\ q(x, y) = -3 - 2x + x^2 + xy + y^2 = 0 \end{cases} \leftrightarrow \underbrace{\left(\begin{array}{ccc|ccc} -4 & 0 & 1 & 0 & 0 & 1 \\ -3 & -2 & 1 & 0 & 1 & 1 \\ \hline -x & 1 & & & & \\ & -x & 1 & & & \\ \hline & & & -x & 1 & \\ \hline -y & & & 1 & & \\ & & & -y & & 1 \end{array}\right)}_{L(x,y)} \begin{pmatrix} 1 \\ x \\ x^2 \\ y \\ xy \\ y^2 \end{pmatrix} = \mathbf{0}$$

Solutions are the couples (x, y) for which $L(x, y)$ is column rank deficient.

- *Finding $\mathcal{X}$.* Degree extension: $L(x, y) \rightarrow \hat{L}(x, y)$:

$$\hat{L}(x, y) = \left(\begin{array}{ccc|ccc|c} -4 & 0 & 1 & 0 & 0 & 1 & \\ -3 & -2 & 1 & 0 & 1 & 1 & \\ \hline & & & -4 & 0 & 1 & 0\ 0\ 1 \\ & & & -3 & -2 & 1 & 0\ 1\ 1 \\ \hline -x & 1 & & & & & \\ & -x & 1 & & & & \\ & & -x & 1 & & & \\ \hline & & & -x & 1 & & \\ & & & & -x & 1 & \\ \hline & & & & & & -x\ 1 \\ \hline -y & & & 1 & & & \\ & & & -y & & 1 & \\ & & & & -y & & 1 \end{array}\right) = \left(\begin{array}{c} \hat{\Pi}_x(x) \\ \hline \hat{\Pi}_y(y) \end{array}\right).$$

  $\mathcal{X}$ is found as the eigenvalues of $\hat{\Pi}_x(x)$ (a square GEP).
- *Finding the corresponding y-values ($\mathcal{Y}$).*
  - Construct the generalized eigenvalue problem for y in the same way and connect the values of $\mathcal{X}$ and $\mathcal{Y}$.
    - by minimizing the maximal residual.
    - by clustering $\mathcal{X}$ and $\mathcal{Y}$ values, determining the cardinality of all clusters and make the multiplicity of the connections between x- and y-clusters feasible.
  - Without solving the GEP in y
    - by using the eigenvectors of $\hat{L}(x, y)$.
    - by looking for the common zeros of $p(x^*, y)$ and $q(x^*, y)$, $\forall x \in \mathcal{X}$, by using the Sylvester matrix or the associated companion matrices.

Results

- Every problem (1) leads to a GEP in x or y by degree extension.
- The resulting square pencil is a new *resultant*, equivalent to that of Sylvester.
- Simple well-conditioned roots are calculated efficiently and with high accuracy. The method is competitive with other root finding methods.

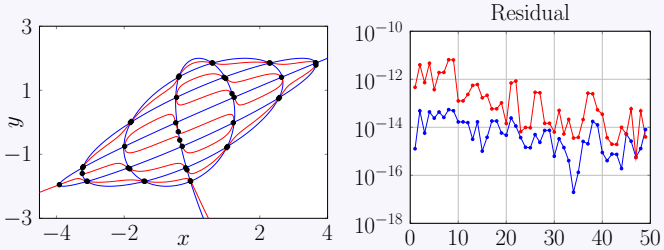


Figure: Left: plot of the zero level lines in $\mathbb{R}^2$ of two bivariate polynomials $p(x, y)$ (–) and $q(x, y)$ (–) of degree 7, the calculated real solutions are indicated by black dots (•). Right: residuals of all 49 (complex) solutions with respect to $p(x, y)$ (•) and $q(x, y)$ (•).

- Multiplicity of the solutions is taken into account and multiple solutions are calculated with reasonably small residuals.

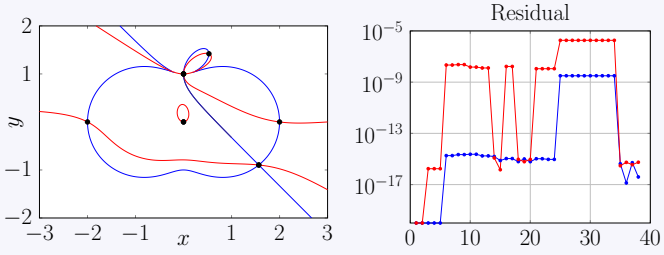


Figure: A problem with intersections of multiplicity > 1. For example, $(0, 1)$ is a 13-fold zero.

Bibliography

[1] S. Barnett. *A Companion Matrix Analogue for Orthogonal Polynomials.* Queen's University, Kingston - Mathematics Department, 1975.

[2] S. Basu, R. Pollack, and M.-F. Roy. *Algorithms in Real Algebraic Geometry.* Springer, 2006.

[3] D. J. Bates, J. D. Hauenstein, A. J. Sommese, and C. W. Wampler. *Numerically solving polynomial systems with Bertini*, volume 25 of *Software, Environments and Tools.* SIAM, 2013.

[4] K. Batselier. *A Numerical Linear Algebra Framework for Solving Problems with Multivariate Polynomials.* KU Leuven - Faculty of Engineering Science, 2013. PhD thesis, promotor: Bart De Moor.

[5] B. Buchberger. Groebner bases: A short introduction for systems theorists. *Computer Aided Systems Theory*, 2178:1–19, 2002.

[6] A. Bultheel and D. Huybrechs. *Wavelets.* CuDi VTK vzw, 2014. Course notes on Wavelets.

[7] L. Busé, H. Khalil, and B. Mourrain. *Resultant-Based Methods for Plane Curves Intersection Problems.* Springer-Verlag Berlin Heidelberg, 2016.

[8] D. Cox, J. Little, and D. O'Shea. *Ideals, Varieties and Algorithms.* Springer, 2007.

[9] I. Daubechies. *Ten Lectures on Wavelets.* SIAM, 1992.

[10] T. A. Davis, S. N. Yeralan, and S. Ranka. *User's Guide for SuiteSparseQR, a multifrontal multithreaded sparse QR factorization package (with optional GPU acceleration).* 2016.

[11] B. De Moor. *Back to the Roots: Solving Polynomial Systems with Numerical Linear Algebra Tools (slides).* KU Leuven - Department of Electrical Engineering ESAT/SCD, 2013.

[12] R. Descartes. *La Géométrie.* Jean Maire, 1637. Slide show used in Valencia.

[13] E. D. Dolan and J. J. Moré. Benchmarking optimization software with performance profiles. *Mathematical Programming*, 91:201–213, 2002.

[14] P. Dreesen. *Back to the Roots.* KU Leuven - Faculty of Engineering Science, 2013. PhD thesis, promotor: Bart De Moor.

[15] D. Gleich. gaimc: Graph algorithms in matlab code. Matlab File Exchange, `http://www.mathworks.com/matlabcentral/fileexchange/24134-gaimc---graph-algorithms-in-matlab-code`.

[16] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials.* SIAM, 2009.

[17] P. W. Lawrence, M. Van Barel, and P. Van Dooren. Backward error analysis of polynomial eigenvalue problems solved by linearization. *SIAM journal on Matrix Analysis and Applications*, 2005.

[18] D. Lazard. Groebner bases, gaussian elimination and resolution of systems of algebraic equations. *European Computer Algebra Conference London*, 162:146–156, 1983.

[19] L. Ljung. *System Identification: Theory for the User.* Prentice Hall, 2 edition, 1999.

[20] Y. Nakatsukasa, V. Noferini, and A. Townsend. *Computing the common zeros of two bivariate functions via Bézout resultants.* Springer-Verlag Berlin Heidelberg, 2014.

[21] V. Y. Pan. Solving a polynomial equation: Some history and recent progress. *Siam Review*, 39:187–220, 1997.

[22] B. Plestenjak. Multipareig. Matlab File Exchange, `http://www.mathworks.com/matlabcentral/fileexchange/47844-multipareig`.

[23] B. Plestenjak and M. E. Hochstenbach. Roots of bivariate polynomial systems via determinantal representations. *SIAM J. Sci. Comput.*, 38(2):A765–A788, 2015.

[24] A. J. Sommese and C. W. Wampler. *The Numerical Solution of Systems of Polynomials Arising in Engineering Science.* World Scientific Publishing Co. Pte. Ltd., 2005.

[25] L. Sorber, M. Van Barel, and L. De Lathauwer. Numerical solution of bivariate and polyanalytic polynomial systems. *SIAM J. Num. Anal. 52*, pages 1551–1572, 2014.

[26] H. J. Stetter. *Numerical Polynomial Algebra.* Society for Industrial and Applied Mathematics, 2004.

[27] G. Strang and T. Nguyen. *Wavelets and Filter Banks.* Wellesley-Cambridge Press, 1997.

[28] B. Sturmfels. Polynomial equations and convex polytopes. *The American Mathematical Monthly*, 105:907–922, 1998.

[29] B. Sturmfels. *Solving Systems of Polynomial Equations*. Number 97 in CBMS Regional Conferences. Amer. Math. Soc., 2002.

[30] J. Verschelde. *Homotopy Continuation Methods for Solving Polynomial Systems*. KU Leuven - Faculty of Engineering Science, 1996. PhD thesis, promotor: Ann Haegemans.

[31] J. Verschelde. Algorithm 795: Phcpack: A general-purpose solver for polynomial systems by homotopy continuation. *ACM Transactions on Mathematical Software Vol. 25, No. 2*, pages 251–276, 1999.

[32] Wikipedia. La géométrie. `https://en.wikipedia.org/wiki/La_G%C3%A9om%C3%A9trie`.

[33] Wikipedia. Polynomial. `https://en.wikipedia.org/wiki/Polynomial`.

Fiche masterproef

Student: Simon Telen

Titel: Solving Systems of Polynomial Equations

Nederlandse titel: Het Oplossen van Stelsels Veeltermvergelijkingen

UDC: 621.3

Korte inhoud:
Multivariate systems of polynomial equations find their applications in various fields of science and engineering. Some examples are filter design, parametric system identification, robotics, computer aided design, chemical engineering, ... Solving such a system is a long studied mathematical problem. The emphasis has long been on theoretical aspects. More recently several algorithmic solving methods have been developed. Groebner basis methods, resultant methods and homotopy continuation methods are the most important classes. In this text, the aim is to propose a new numerical linear algebra based method for solving bivariate 0-dimensional systems. By "solving" we mean finding all solutions, both real and complex, and taking multiplicities into account. We start from a two-parameter eigenvalue approach, similar to the one introduced by Plestenjak and Hochstenbach in 2015. We use the concept of degree extension, which is also used to construct Macaulay resultants, to construct a one-parameter square generalized eigenvalue problem directly from the coefficients of the given polynomials. The degree extension allows us to eliminate one of the variables, which is a typical aspect of resultant methods. The coefficients appear directly, without being manipulated in the pencil, which is constructed in a very intuitive manner. We show that a square generalized eigenvalue problem can be constructed for any 0-dimensional system and that the resulting eigenvalues are equal to those of the Sylvester resultant. After obtaining one of the coordinates in this way, we propose some possible approaches for finding the other coordinate of the solutions. The strong link with the Sylvester resultant allows us to give information about the multiplicity of the solutions. Results are promising. Solutions are obtained with small residuals and the computation time is competitive with other solvers. We show that the method can be generalized to other bases than the classical monomial basis and we propose a generalization for more than two dimensions.

Thesis voorgedragen tot het behalen van de graad van Master of Science in de ingenieurswetenschappen: wiskundige ingenieurstechnieken

Promotor: Prof. dr. ir. Marc Van Barel

Assessor: Prof. dr. ir. Daan Huybrechs
Prof. dr. ir. Lieven De Lathauwer

Begeleider: Prof. dr. ir. Marc Van Barel