The Bézout Construction of the Resultant

Dissertation

P. Bikker

Research Institute for Symbolic Computation
Johannes Kepler University Linz
Austria

Acknowledgments

I would like to thank all who encouraged me to finish this work. I am especially grateful to Prof. Dr. Franz Winkler, head of the Research Institute for Symbolic Computation in Hagenberg, whose support was essential. With appreciation I mention the name of Prof. Dr. Aleksej Uteshev from the University of St. Petersburg with whom I worked on a joint publication on the subject of this thesis.

Pieter Bikker

Curriculum vitae

Pieter Bikker

Personal

Born October 14, 1969, Noordeloos, The Netherlands
Residence Oosterstraat 23, NL-9203 BA, Drachten, The Netherlands
Citizen of The Netherlands
Married since 1997

Education

M.Sc. studies, University of Groningen, The Netherlands, 1988-1994
M.Sc., Mathematics, University of Groningen, 1992
M.Sc., Computer Science, University of Groningen, 1993
M.Sc., Applied Mathematics, University of Groningen, 1994
Ph.D. studies, Research Institute for Symbolic Computation, Johannes Kepler University Linz, Austria, 1994-1996

Positions

Teaching Assistant, University of Groningen, The Netherlands, 1991-1992
Software Engineer, Asset Control, Beetsterzwaag, The Netherlands, 1996-2001

Publications

A modular Gröbner basis algorithm, In *Algoritmen in de Algebra, A Seminar on Algebraic Algorithms*, University of Nijmegen, 1993

On Bézout's Method for Computing the Resultant, RISC-Linz Technical Report 95-36, 1995

Non-Admissible Orderings and Involutive Bases, RISC-Linz Technical Report 97-16, 1997

On the Bézout construction of the Resultant (with A.Yu. Uteshev), RISC-Linz Technical Report 99-1, 1999

On the Bézout construction of the Resultant (with A.Yu. Uteshev), *Journal of Symbolic Computation*, 28(1), 1999

Eidesstattliche Erklärung

Ich versichere, daß ich die Dissertation selbständig verfaßt habe, andere als die angegebenen Quellen und Hilfsmittel nicht verwendet und mich auch sonst keiner unerlaubten Hilfe bedient habe.

Pieter Bikker

Drachten, Niederlande

Januar 2001

Abstract

A fast and robust algorithm is presented for the construction of the resultant of a system of multivariate polynomial equations. The resultant of the typically overdetermined system of $m+1$ polynomials in m indeterminates,

$$f_1(x_1,\ldots,x_m)=0,\ldots,f_m(x_1,\ldots,x_m)=0,f_{m+1}(x_1,\ldots,x_m)=0,$$

gives the necessary and sufficient condition, in terms of the coefficients, for the existence of a common solution.

Based on the ideas of Bézout and Poisson for solving polynomial systems by resultants, a new resultant matrix construction is treated by two different approaches.

One approach is based on a sequence of determinantal conditions that are guaranteed to be satisfied in case of a generic polynomial system. These conditions can be written as minors of the Macaulay matrix of the leading forms of $f_1,\ldots,f_m$. As a result a set of reducers is obtained. These are used to eliminate monomials from non-constant multiples of f_{m+1}, ending in a linear system that is consistent if and only if the polynomial system possesses solutions.

The other approach uses Gröbner basis techniques for the generation of the resultant matrix. First, the Gröbner basis of the ideal generated by $f_1,\ldots,f_m$ is constructed. With the help of the Gröbner basis a set of monomials that are linearly independent of the solutions is determined. The resultant matrix is the coefficient matrix of the linear system obtained by the normal forms of the multiplication of this basis by f_{m+1}. The size of the matrix is the smallest among the ones generated by the various resultant methods.

The obtained resultant matrix is called the Bézout matrix. It can be viewed as a generalization of the Bézoutiant different from the Dixon resultant matrix. It is shown how to use the Bézout matrix to obtain all the components of the solutions.

Whereas other resultant formulations suffer from the sparsity and the algebraic dependencies between the coefficients of the polynomial system giving so-called extraneous factors and either non-square or singular matrices, this approach controls all the exceptional cases. The suspicious factors that correspond to the determinantal conditions or leading coefficients of Gröbner basis elements are used to exclude factors of the resultant from being extraneous. By this, all factors of the true eliminant can be determined. The resultant matrix is always square and its singularity is shown to correspond to a vanishing resultant: no additional investigation among the minors is needed.

The algorithm with the second approach was implemented for Maple and Asir. Computational experiments show that the method is competitive with other fast methods for solving polynomial systems, like FGLM and Dixon's resultant. Especially for dense systems the method is superior. It turns out that the time needed for the computation of the Gröbner basis and the matrix construction is not the new bottleneck. Still most of the time is used for the final determinant evaluation.

It is shown how Involutive bases can be defined with respect to non-admissible orderings. These come very close to the reduction algorithm of the determinantal approach and can be used as such.

Zusammenfassung

Wir stellen einen schnellen und robusten Algorithmus vor für die Konstruktion der Resultanten eines Systems multivariater polynomialer Gleichungen. Die Resultante des typischerweise überbestimmten Systems von $m+1$ Polynomen in m Variablen,

$$f_1(x_1,\ldots,x_m)=0,\ldots,f_m(x_1,\ldots,x_m)=0,f_{m+1}(x_1,\ldots,x_m)=0,$$

drückt die notwendige und hinreichende Bedingung in den Koeffizienten aus für die Existenz gemeinsamer Lösungen.

Aufbauend auf Ideen von Bézout und Poisson zur Lösung polynomialer Systeme mittels Resultanten, beschreiben wir zwei neue Zugänge zur Konstruktion der Resultantenmatrix.

Der erste Zugang basiert auf einer Abfolge von Determinantenbedingungen, welche für ein generisches System von Polynomen erfüllt ist. Diese Bedingungen können geschrieben werden als Minoren der Macaulay Matrix der führenden Formen von $f_1,\ldots,f_m$. Daraus erhält man eine Menge von Reduktoren, mithilfe derer man Monome aus Vielfachen von f_{m+1} eliminieren kann. Das daraus resultierende lineare System is konsistent genau dann wenn das polynomiale System lösbar ist.

Der zweite Zugang benutzt Techniken aus der Theorie der Gröbnerbasen zur Erzeugung der Resultantenmatrix. Zuerst wird die Gröbnerbasis des von $f_1,\ldots,f_m$ erzeugten Ideals konstruiert. Mithilfe dieser Gröbnerbasis bestimmt man dann eine Menge von Monomen, welche linear unabhängig sind von den Lösungen. Die Resultantenmatrix ist die Koeffizientenmatrix des linearen Systems, welche sich ergibt aus den Normalformen der Multiplikation dieses Systems mit f_{m+1}. Von allen Resultantenmethoden erhält man auf diese Weise die kleinste Matrix.

Diese Resultantenmatrix heißt die Bézout Matrix. Sie kann betrachtet werden als Verallgemeinerung ursprünglichen Bézout'schen Bedingungen, verschieden von der Resultantenmatrix Dixons. Wir zeigen, wie man aus der Bézout Matrix alle Komponenten der Lösungen erhält.

Während bei anderen Formulierungen der Resultante aufgrund dünner Besetztheit und algebraischer Abhängigkeit der Koeffizienten Probleme auftreten können, wie zum Beispiel mit Spurfaktoren und nichtquadratischen oder singulären Matrizen, kontrolliert der hier betrachtete Zugang alle diese Spezielfälle. Die aus den Determinantenbedingungen oder führenden Koeffizienten der Gröbnerbasis hervorgehenden Faktoren dienen zur Eliminierung von Spurfaktoren der Resultante. Auf diese Weise können alle Faktoren der wahren Eliminante bestimmt werden. Die Resultantenmatrix ist immer quadratisch und ihre Singularität korrespondiert exakt mit dem Verschwinden der Resultante: keine zusätzlichen Untersuchungen der Minoren sind notwendig.

Der aus dem zweiten Zugang hervorgehende Algorithmus wurde in Maple und Asir implementiert. Berechnungsexperimente zeigen, daß die Methode durchaus konkurrenzfähig ist zu anderen schnellen Methoden zur Lösung polynomialer Systeme, wie zum Beispiel FGLM oder der Dixon Resultante. Insbesondere für

dicht besetzte Systeme ist diese Methode überlegen. Es stellt sich heraus, daß die Berechnung der Gröbnerbasis und die Konstruktion der Matrix keinen neuen Engpaß darstellen. Der Großteil der Rechenzeit wird weiterhin verbraucht für die Auswertung der Determinante.

Wir zeigen auch, wie involutive Basen mittels nicht-zulässiger Ordnungen definiert werden können. Diese Basen haben starken Bezug zum Reduktionsalgorithmus des Zugangs mittels Determinanten, und können auch so verwendet werden.

Contents

Overview **1**

1 Solving polynomial systems and resultants **5**
1.1 The univariate case 5
1.2 The multivariate case 10
1.3 Macaulay's resultant 14
1.4 The Cayley and Dixon resultant 18
1.5 Solving polynomial systems by the u-resultant 21
1.6 Solving polynomial systems by Gröbner bases 22

2 The Bézout resultant of two polynomials **27**
2.1 The construction 27
2.2 The Bézoutiant 31
2.3 Root separation 32

3 Constructing the Bézout resultant: determinantal approach **37**
3.1 The Bézout resultant in the multivariate case 37
3.2 Determination of all the solutions by means of the Bézout matrix . 40
3.3 Reducibility in the bivariate case 41
3.3.1 Example workout 41
3.3.2 The reduction algorithm 44
3.3.3 Theoretical results 48
3.4 Reducibility in the trivariate case 53
3.5 Reducibility in the m-variate case 59
3.6 Comparison to the Macaulay matrix 62

4 Constructing the Bézout resultant: Gröbner basis approach **65**
4.1 The construction 65
4.2 Extraneous factors 71
4.2.1 Definitions 71
4.2.2 Evaluating the eliminant 74
4.2.3 The multiparameter case 78

4.3 The infinite Poisson product 82
4.4 The u-resultant and Bézout matrix eigenvalues 85
4.5 Experiments 87

5 Non-admissible orderings and Involutive bases **99**
5.1 Terminology 99
5.2 The Pommaret reduction relation made Noetherian 101
5.3 Involutive bases with respect to $\prec_{part}$ 103
5.4 Applications 104

A Maple code Bézout resultant **107**

B Maple code Dixon resultant **109**

Bibliography **113**

Overview

Resultants

Systems of multivariate polynomial equations arise from the mathematical modelling of a range of problems appearing in science and industry. These are often geometric problems, e.g. the implicitization of a curve and the description of the configuration space of a robot.

Among the methods for solving a multivariate polynomial system of equations symbolically, the use of resultants caught new attention in the last decade. After the Gröbner basis algorithms were highly optimized they turned out, in practice, to be able to solve small problems only. Therefore, the resultant techniques, developed long before Gröbner bases were invented, have been investigated again for the purpose of solving the larger non-linear systems.

The main idea of resultants is the transformation of the original non-linear system of equations into a linear system of equations, usually of much bigger size. The determinant of the corresponding matrix, the resultant matrix, is called the resultant (or eliminant) – the result of the elimination of some of the variables. The resultant is used as the solvability condition or to obtain, either numerically or symbolically, one of the components of the solutions of the system, namely for one variable that was not eliminated.

There are various methods for building the resultant matrix, e.g. Macaulay's and Dixon's. The concept of resultants dates from Newton and Euler. They discuss how to eliminate the variable from a system of two univariate polynomials. Bézout considers the problem of solving systems of multivariate polynomial equations. His theorem on the maximum number of solutions – the product of the degrees – is famous. In the nineteenth century Sylvester revived the method of the elimination of one variable. Around the same time a different formulation for the resultant of two univariate polynomials was given by Cayley. Macaulay was the first, at the beginning of the twentieth century, with a complete construction for the resultant of a system of multivariate equations. Dixon generalized the Cayley resultant to the bi- and multivariate case.

In recent years each of these different resultants have been given attention to: subresultants for computational optimization, sparse resultants and perturbation techniques for avoiding singular resultant matrices and optimizations.

In this thesis is dealt with the ideas of Bézout and Poisson for solving polynomial systems by resultants. A new resultant matrix construction is presented. This matrix has excellent properties.

Thesis overview

This thesis consists of five chapters.

Chapter 1 is an introduction to resultants. The definition of the resultant is given by the Poisson product for the univariate case, followed by its generalization to the multivariate case. The formulations of Sylvester, Macaulay, Cayley and Dixon for the evaluation of the resultant are presented in brief. Two sections follow on how to solve polynomial systems, one section on the use of the u-resultant, the other section on the use of Gröbner bases. The presented theory is used in Chapter 4.

The main results from Chapter 2, 3 and 4 have been published before in [14], a joint article with Uteshev.

In Chapter 2 we present, for the univariate case, a new construction for the resultant matrix, called the Bézout matrix. We show the relation between the (minors of the) Bézout matrix and the (subresultants of the) Sylvester matrix. The direct relation to the Bézoutiant is demonstrated and illustrated by its applications for separating the roots of a polynomial.

The idea for the construction of the resultant matrix from Chapter 2 is generalized in Chapter 3 and 4, taking two approaches: using determinants and using Gröbner bases. In Chapter 3 the resultant matrix is constructed via a sequence of determinantal conditions, that are shown to be fulfilled for generic systems. It corrects and supplies a deficiency of the attempt by Laurent.

We discuss the relation to the Macaulay matrix. The results regarding the relation between the Macaulay matrix of the leading forms and the reducibility conditions in the trivariate case were not published earlier.

Chapter 4 deals with the Gröbner basis approach of the construction of the Bézout matrix. This approach is new and uses, for the first time, Gröbner bases as part of the construction of the resultant. It is shown that the construction is applicable to non-generic polynomial systems as well. The resultant matrix appears to be the smallest among the ones generated by the various resultant methods. We show the matrix does not suffer from singularity for the possible sparse systems (other resultant matrices do).

The problem of extraneous factors is discussed extensively. Sufficient conditions are given on the factors of the resultant in order to exclude them from being extraneous. We show how the Bézout matrix can be used to obtain all the components of the solutions. The investigation on the existence of the resultant in case of a non-zero dimensional ideal is new and has not been published before.

The chapter closes with an investigation of several practical examples. Computational experiments show that the method is competitive with other fast methods for solving polynomial systems symbolically like FGLM and Dixon's resultant. The comparisons to the FGLM method and Dixon's resultant have not been published before. Especially for dense systems the method turns out to be superior.

In the last chapter we expose a self-contained subject regarding Involutive bases. We demonstrate the existence of Involutive basis with respect to non-admissible orderings. The application to the construction of the Bézout matrix appears to be very close to the determinantal approach from Chapter 3. This chapter is based on the results we reported in [13]. The application to the construction of the Bézout resultant is new.

Chapter 1

Solving polynomial systems and resultants

In this chapter we give an overview on resultants and on the use of resultants for solving systems of polynomial equations. For subsequent chapters, this chapter serves as introduction.

By solving a polynomial system is meant to determine all the components of all the solutions. With resultants information on the existence of solutions of the system is obtained. Most resultant formulations also contain information to obtain the components of the solutions.

In the first section the definition of the resultant for two polynomials in one indeterminate and the associated Sylvester matrix are reviewed.

In the second section this definition is being generalized to systems consisting of more than two polynomials. This leads naturally to a definition of the resultant by the Poisson product. Following classic works, e.g. by Netto [65], the resultant is defined for a system of generic polynomial equations. This means that the polynomials are dense and their coefficients are algebraically independent. In two subsequent sections the Macaulay and Dixon resultant formulations – methods for evaluating the resultant for generic polynomials – are discussed.

A section follows on how to find all the components of the solutions via resultants, e.g. by using the u-resultant.

In the last section we focus on solving a polynomial system with a technique other than resultants, namely Gröbner basis. The Gröbner basis technique is used in Chapter 4 for the construction of the Bézout resultant.

1.1 The univariate case

The first case to consider is a system of two polynomials in one variable. Let K be a field. Let $f(x)$ and $g(x)$ be non-zero polynomials from $K[x]$ with degrees n

and m, respectively:

$$\begin{array}{rcl} f(x) & = & a_0x^n + a_1x^{n-1} + \cdots + a_n, \\ g(x) & = & b_0x^m + b_1x^{m-1} + \cdots + b_m. \end{array} \tag{1.1}$$

We call a_0 and b_0 the leading coefficients of $f(x)$ and $g(x)$. Assuming $a_0 \neq 0$, we know that $f(x)$ has n (distinct) roots in $\overline{K}$, the algebraic closure of K. Denote these roots by $\lambda_1, \ldots, \lambda_n$, listing multiple roots multiple times such that we can write $f(x) = a_0(x - \lambda_1)\cdots(x - \lambda_n)$. The resultant is defined as a symmetric product over these roots and will be denoted by $\Re(f, g)$:

Definition 1.1 *The resultant of $f(x)$ and $g(x)$ is defined as the product*

$$\Re(f, g) = {a_0}^m g(\lambda_1)\cdots g(\lambda_n). \tag{1.2}$$

We will use the notation $\Re_x(f, g)$ to make explicit relative to which indeterminate the polynomials are viewed. This is necessary in case the polynomials can be viewed to be univariate in more than one indeterminate. It indicates which indeterminate is so-called eliminated from f and g as $\Re_x(f, g)$ is independent of x.

Proposition 1.1 *$\Re(f, g)$ has the following properties.*

- *$\Re(f, g)$ is a polynomial expression in the coefficients of $f(x)$ and $g(x)$. Hence, $\Re(f, g) \in K$.*
- *$f(x)$ and $g(x)$ have a common root if and only if $\Re(f, g) = 0$.*
- *$\Re(g, f) = (-1)^{nm}\Re(f, g)$.*

Proof. The first statement can be demonstrated by using the symmetric functions in the roots of $f(x)$ [90]. The second is obvious from the definition, as we have assumed $a_0 \neq 0$. To show that the resultant is, up to a factor -1, independent of the order of f and g, write $g(x) = b_0(x - \mu_1)\cdots(x - \mu_m)$, where $\mu_1, \ldots, \mu_m$ are the roots of $g(x)$. Substituting this into $\Re(f, g)$ gives

$$\Re(f, g) = {a_0}^m \prod_{j=1}^{n} \left(b_0 \prod_{k=1}^{m} (\lambda_j - \mu_k) \right) =$$

$$(-1)^{nm} {b_0}^n \prod_{k=1}^{m} \left(a_0 \prod_{j=1}^{n} (\mu_k - \lambda_j) \right) = (-1)^{nm} \Re(g, f).$$

□

It is, at first sight, not a trivial task to evaluate expression (1.2). The so-called Sylvester matrix [84, 90] is one of the matrix representations of the resultant of two polynomials in the sense that its determinant coincides with (1.2). This is an

$(n+m) \times (n+m)$ matrix, partially filled with coefficients of $f(x)$ and $g(x)$. It consists of m rows with coefficients of $f(x)$, where row j starts to list the coefficients of $f(x)$ at column j. The remaining n rows are filled with the coefficients of $g(x)$: row $m+j$ starts to list the coefficients of $g(x)$ at column j.

The following Sylvester matrix, denoted by $\mathcal{S}(f,g)$, suggests a situation where $n > m$ (hence there are more rows filled with coefficients of $g(x)$):

$$\mathcal{S}(f,g) = \left(\begin{array}{cccccccc} a_0 & a_1 & a_2 & \cdots & a_n & & & \\ & a_0 & a_1 & \cdots & & a_n & & \\ & & \ddots & & & & \ddots & \\ & & & a_0 & a_1 & \cdots & & a_n \\ b_0 & b_1 & \cdots & b_m & & & & \\ & b_0 & \cdots & & b_m & & & \\ & & \ddots & & & \ddots & & \\ & & & \ddots & & & \ddots & \\ & & & & b_0 & \cdots & & b_m \end{array}\right) \begin{array}{l} \left.\vphantom{\begin{array}{c}1\\1\\1\\1\end{array}}\right\} m \text{ rows} \\ \left.\vphantom{\begin{array}{c}1\\1\\1\\1\\1\end{array}}\right\} n \text{ rows} \end{array} . \qquad (1.3)$$

The Sylvester matrix can be viewed as the coefficient matrix of the equational system given by:

$$\left\{ \begin{array}{ll} x^j f(x) = 0, & (j = 0, \ldots, m-1) \\ x^j g(x) = 0, & (j = 0, \ldots, n-1) \end{array} \right.$$

associating column j with the x-power x^{m+n-j} $(j = 1, \ldots, m+n)$. Sylvester called this the dialytic method [84]. We will see this technique of linearization of a polynomial system throughout all the theory of resultants.

The following proposition is the most important one for the applicability of the Sylvester matrix. It relates to the second property of the resultant, listed in Proposition 1.1.

Proposition 1.2 *For every specialization of the coefficients such that $a_0 \neq 0$ or $b_0 \neq 0$, the Sylvester matrix is singular if and only if $f(x)$ and $g(x)$ have a common root.*

Proof. If $f(x)$ and $g(x)$ have a common root λ, then we can write $f(x) = (x - \lambda)f'(x)$ and $g(x) = (x-\lambda)g'(x)$. This gives a linear dependency between the rows of the Sylvester matrix as follows. Take the coefficients of $g'(x)$ and multiply these with rows 1 to m of (1.3). We do the same with the coefficients of $-f'(x)$: multiply these with rows $m+1$ up to $m+n$. Interpreting the matrix as a coefficient matrix, the sum of the resulting rows gives $g'(x)f(x) - f'(x)g(x) = 0$.

On the other hand, if the Sylvester matrix is singular we can find, with the same mechanism, non-zero polynomials $f'(x)$ and $g'(x)$ of degree at most $m-1$ and $n-1$, respectively, such that $g'(x)f(x) - f'(x)g(x) = 0$. Take the $f'(x)$ and $g'(x)$ with minimal degree. We may assume that $f'(x)$ and $g'(x)$ do not have a common root. If they would have, we can choose another pair $f'(x)$ and $g'(x)$, with lower degree. In the equation $g'(x)f(x) = f'(x)g(x)$ both sides consist of the same factors. If $f(x)$ and $g(x)$ would not have a common factor, then it follows that $f'(x) = g(x)$ and $g'(x) = f(x)$. This contradicts the degree bound on $f'(x)$ and $g'(x)$. However, it may not contradict in the case where both a_0 and b_0 are set to zero, but this is precisely the case that we have excluded. $\square$

The first part of the proof can be shortened (we have given the long version to make the second part more readable). If we multiply the Sylvester matrix by a vector

$$\mathcal{V} = [\lambda^{n+m-j}]_{j=1}^{m+n}$$

for one of the common roots λ, then $\mathcal{S}(f,g)\cdot\mathcal{V}$ equals the 0 vector. As $\mathcal{V} \neq 0$ (the last entry equals $\lambda^0 = 1$), it follows that $\det \mathcal{S}(f,g) = 0$. This necessary condition for the existence of a common root was already known by Euler [90].

The Sylvester matrix and the resultant are related even stronger:

Theorem 1.1 *The determinant of the Sylvester matrix is equal to the resultant.*

Proof. As both vanish if and only if $f(x)$ and $g(x)$ have a common root, they consist of the same factors. For algebraically independent coefficients, both the determinant of the Sylvester matrix and $\Re(f,g)$ are irreducible [61, 65, 90]. Hence, they are identical and they remain so after specializing coefficients provided $a_0 \neq 0$. $\square$

Therefore, the determinant of the Sylvester matrix is called the Sylvester resultant. The resultant often is defined by means of the Sylvester matrix [28].

For efficient evaluation of the determinant of the Sylvester matrix the so-called subresultants can be used [25, 26].

Proposition 1.2 shows that the resultant should be used with care, because of the exceptional case when both leading coefficients vanish.

Example 1.1 *Apply Proposition 1.2 for the following polynomials:*

$$\begin{aligned} f(x) &= ax^3 + x^2 + x + a, \\ g(x) &= bx^2 + x + c. \end{aligned}$$

Obviously the coefficients are algebraically dependent: the leading coefficient a appears in the constant term of $f(x)$ as well. The resultant of $f(x)$ and $g(x)$ equals

$$\Re(f,g) = (b+c-1)(a^2(b^2+c^2-bc+b+c+1) + a(-b-c-2bc) + bc).$$

We can not simply conclude that for each pair (a,b) for which the resultant $\Re(f,g)$ vanishes, $f(x)$ and $g(x)$ have a common root. The case $(a,b) = (0,0)$ may have to be excluded and is to be considered separately. In that case the resultant becomes

$$\Re(x^2 + x, x + c) = c(c-1). \tag{1.4}$$

So in case $c = 0$ or $c = 1$, we have a common root if $(a, b) = (0, 0)$.

Let us look at these exceptional cases a more in detail. Let a, b be the leading coefficients of $f(x)$ and $g(x)$, respectively. We will use the following notation: S_a is the substitution operator, sending a to 0. S_b sends b to 0 and $S_{a,b}$ sends both a and b to 0. From (1.2) we can see that in case b is sent to 0 the degree of a drops. In fact

$$S_b(\Re(f, g)) \;=\; a\,\Re(S_b(f), S_b(g)). \tag{1.5}$$

If the degree of $S_b(g)$ compared to the degree of g drops by more than 1, the factor a in the right-hand side of (1.5) gets an exponent equal to the drop in the degrees. Similarly, for S_a, we have

$$S_a(\Re(f, g)) \;=\; b\,\Re(S_a(f), S_a(g)).$$

Substituting both $a = 0$ and $b = 0$ into the resultant gives $S_{a,b}(\Re(f, g)) = 0$, although $\Re(S_{a,b}(f), S_{a,b}(g))$ may not vanish. Example 1.1 is an illustration of this: taking the resultant of the polynomials after substitution, gives the non-zero result (1.4).

The resultant gives an excellent way to solve systems of two equations in two indeterminates. Let $f, g \in K[x, y]$. Translating these into the form as in (1.1), we have polynomials from $K[y]$ for the symbols a_i, b_j. The Sylvester matrix $\mathcal{S}$ becomes a matrix with entries in $K[y]$. Hence, the resultant $\Re_x(f, g)$, the determinant of $\mathcal{S}$, is a polynomial in y.

For each root $\bar{y}$ of $\Re_x(f, g)$ we can apply Proposition 1.2. As long as for such a root a_0 and b_0 do not both vanish, we find solutions $(\bar{x}, \bar{y})$ for the equations $f(x, y) = g(x, y) = 0$. The $\bar{x}$ can be determined via Cramer's rule.

Example 1.2 *Find all common solutions for*

$$\begin{aligned} f &= x^2 + xy + y^2 + x + y + 1, \\ g &= 2\,x^2 + 3\,xy + 5\,y^2 + 7\,x + 11\,y + 13. \end{aligned}$$

The Sylvester matrix of f and g, considered as univariate polynomials in x, is

$$\mathcal{S} = \begin{pmatrix} 1 & y+1 & y^2+y+1 & 0 \\ 0 & 1 & y+1 & y^2+y+1 \\ 2 & 3\,y+7 & 5\,y^2+11\,y+13 & 0 \\ 0 & 2 & 3\,y+7 & 5\,y^2+11\,y+13 \end{pmatrix}.$$

Its determinant evaluates to

$$\Re_x(f, g) = 7\,y^4 + 38\,y^3 + 103\,y^2 + 122\,y + 91. \tag{1.6}$$

This is an irreducible polynomial. Hence it has 4 distinct roots. The corresponding x value can be found by subtracting $f(x)$ by $1/2\,g(x)$. This gives a degree 1 polynomial in x, from which we immediately derive that

$$x = -\frac{3\,y^2 + 9\,y + 11}{y + 5}. \tag{1.7}$$

The factor $y+5$ in the denominator is no problem: $y=-5$ is not a root of (1.6). This expression for x can also be determined applying Cramer's rule. This is the normal procedure as we usually do not have such a fast trick.

Cramer's rule says that for the system of n linear equations $A \cdot X = c$ where A is an $n \times n$ matrix, c is a vector, and X is a vector of variables, if $\det A \neq 0$, we can express the j-th entry of X, X_j, by

$$X_j = \frac{c_1 A_{1j} + \cdots + c_n A_{nj}}{\det A},$$

where A_{ij} is the cofactor to the ij-th element of A: $(-1)^{i+j}$ times the determinant of the minor of A deleting row i and column j. The numerator is equal to the determinant of A after replacing the j-th column by the vector c.

Now let y be one of the roots of the resultant (1.6). As the leading coefficients of $f(x)$ and $g(x)$ are constants, Proposition 1.2 tells $f(x)$ and $g(x)$ have a common solution. Assuming that the rank of $\mathcal{S}$, after deleting the last row and column, is 3, we can solve

$$\begin{pmatrix} 1 & y+1 & y^2+y+1 \\ 0 & 1 & y+1 \\ 2 & 3\,y+7 & 5\,y^2+11\,y+13 \end{pmatrix} \begin{pmatrix} x^3 \\ x^2 \\ x \end{pmatrix} = - \begin{pmatrix} 0 \\ y^2+y+1 \\ 0 \end{pmatrix}.$$

The last column of $\mathcal{S}$, associated with the constant power product 1, was moved to the right-hand side. With Cramer's rule we find

$$x = \frac{y^3+6\,y^2+6\,y+5}{2\,y^2+3\,y+6}.$$

This expression is equivalent to (1.7) precisely under the condition $\Re_x(f,g)=0$.

1.2 The multivariate case

We now generalize the definition of the resultant to the case of polynomial systems with more than one indeterminate and more than two polynomials. We start out by giving an example with three polynomials.

Suppose we slightly change the system given in Example 1.2. In the first polynomial we replace the constant 1 in f by the new indeterminate z. Moreover, we add a third polynomial to the system:

Example 1.3 *Find all the common solutions for*

$$\begin{aligned} f &= x^2+xy+y^2+x+y+z, \\ g &= 2\,x^2+3\,xy+5\,y^2+7\,x+11\,y+13, \\ h &= x+y+z. \end{aligned}$$

First, we determine all solutions $(\bar{x}, \bar{y})$ with $\bar{x}, \bar{y} \in \overline{\mathbb{Q}(z)}$, the algebraic closure of $\mathbb{Q}[z]$, for the equations $f(x, y) = g(x, y) = 0$. The resultant of f and g, considered as polynomials in x, is equal to

$$\Re_x(f, g) = 7\,y^4 + 38\,y^3 + (-9\,z + 112)\,y^2 + 4\,z^2 + (-14\,z + 136)\,y - 17\,z + 104. \quad (1.8)$$

If we compare this to (1.6), as polynomial in y, they differ in the coefficients of the terms of degree less than or equal to two. Substituting $z = 1$ in (1.8), gives (1.6).

It follows that there are four different solutions $(\bar{x}, \bar{y})$ for the system $f(x, y) = g(x, y) = 0$. Denote the set containing these four solutions by Λ.

The next step is to substitute all these solutions in $h(x, y)$, and build their product. This is similar to what is done in (1.2). We denote their product by $\Re^h(f, g)$:

$$\Re^h(f, g) = \prod_{(\bar{x}, \bar{y}) \in \Lambda} h(\bar{x}, \bar{y}). \quad (1.9)$$

We can evaluate this expression by using the Sylvester matrix for (1.8) and h, both considered as polynomials in y. First, we have to substitute x in h. From $f - 1/2\,g$ we can derive

$$x = -\frac{3\,y^2 + 9\,y - 2\,z + 13}{y + 5}.$$

Let h' be the substitution of this x in h. We need to multiply h' by the factor $y + 5$ in order to obtain a polynomial in y. We can now determine its resultant with (1.8):

$$\Re_y(\Re_x(f, g), (y + 5)h') = (2\,z - 43)^2(7\,z^4 - 9\,z^3 + 80\,z^2 - 234\,z + 169). \quad (1.10)$$

Dividing this by

$$\Re_y(\Re_x(f, g), y + 5) = (2\,z - 43)^2 \quad (1.11)$$

gives, using definition (1.2), expression (1.9) with an additional factor:

$$c \prod_{(\bar{x}, \bar{y}) \in \Lambda} h(\bar{x}, \bar{y}) = 7\,z^4 - 9\,z^3 + 80\,z^2 - 234\,z + 169. \quad (1.12)$$

The factor c is equal to 7, the leading coefficient of $\Re_x(f, g)$. This factor appears with exponent two in (1.10) because $(y + 5)h'$ is of degree two, and with exponent one in (1.11) as $y + 5$ is of degree one. Dividing these gives $c = 7^2/7 = 7$.

Now we have completed this example, we continue with the definition of the resultant of more than two multivariate polynomials.

The multivariate system we are going to look at consists of $m + 1$ polynomial equations in m indeterminates:

$$\begin{aligned} &f_1 = 0, \ldots, f_m = 0, f_{m+1} = 0, \\ &f_1, \ldots, f_m, f_{m+1} \in K[x_1, \ldots, x_m], \end{aligned} \quad (1.13)$$

K is the coefficient field. These polynomials do not necessarily have the same (total) degree. We denote by n_j the (total) degree of f_j in the indeterminates $x_1, \ldots, x_m$. The indeterminates are also denoted as X.

An important assumption is that the polynomials are generic, i.e. all the coefficients are algebraically independent. Moreover, generic means dense, i.e. f_j contains all the power products of degree up to n_j.

As the system is overdetermined, having $m + 1$ equations in m indeterminates, it does not, in general, possess a solution. The idea behind the definition of the resultant is to obtain a condition, in terms of all the coefficients of the polynomials, for which the system will possess a solution. For instance, specializing the coefficients of all degree-zero monomials to zero gives the solution $(x_1, \ldots, x_m) = (0, \ldots, 0)$. The resultant aims to give the minimal condition, i.e. the condition that is necessary and sufficient.

By the celebrated Bézout bound [12, 74, 80, 89] we know that, for a system with generic polynomials, there are $N = n_1 \cdots n_m$ common solutions for $f_1, \ldots, f_m$. Denote these by $\Lambda_1, \ldots, \Lambda_N$. We can generalize (1.9) into the product

$$\Re^{f_{m+1}} = \prod_{j=1}^{N} f_{m+1}(\Lambda_j). \tag{1.14}$$

An additional factor like we have seen in (1.2) needs to be determined. Equation (1.12) shows that such an additional factor is necessary, in order not to have denominators in the evaluation of the product (1.14) in terms of the coefficients of $f_1, \ldots, f_m$. Schläfli [79] proved that the symmetric expression (1.14) is a rational function in the coefficients of $f_1, \ldots, f_m$.

Going back to Example 1.3, we see that this factor equals to the leading coefficient of $\Re(f, g)$ as polynomial in y (with exponent one, the degree of h). This leading coefficient can be derived from the leading forms of f and g as follows [74]. From the properties of the determinant (see also the proof of Proposition 1.2) we know that there exist A and B, both polynomials in x of degree less than two, such that

$$Af + Bg = \Re_x(f, g).$$

In fact A and B can be chosen as polynomials in x and y. We can homogenize both sides of this equation, considering them as polynomials in x and y. Substituting for the homogenizing indeterminate the value 0 and for y the value 1, gives, at the right-hand side, the leading coefficient of $\Re_x(f, g)$. The left-hand side can be translated back to a resultant, namely the resultant of the leading forms of f and g, considered as polynomials in x and y and substituting for y the value 1:

$$\Re_x(x^2 + x + 1, 2\,x^2 + 3\,x + 5) = 7.$$

Generalizing this result, we can determine the additional factor, which we will denote by $\mathcal{A}$, as follows. Denote by $\bar{f}_j$ the leading form of f_j, i.e. all the terms of f_j with degree equal to n_j. Substituting in $\bar{f}_j$ the value 1 for any of the indeterminates, say x_1, gives a system with m equations in $m-1$ indeterminates. We assume that we have defined a resultant for this system and let

$$\mathcal{A} = \Re(\bar{f}_1|_{x_1=1}, \ldots, \bar{f}_m|_{x_1=1}). \tag{1.15}$$

We can now generalize the resultant definition.

Definition 1.2 *The resultant of $f_1, \ldots, f_m, f_{m+1}$ is defined as the product*

$$\Re(f_1, \ldots, f_m, f_{m+1}) = \mathcal{A}^{n_{m+1}} \prod_{j=1}^{N} f_{m+1}(\Lambda_j). \tag{1.16}$$

This definition of the resultant is referred to as the Poisson product [65]. We will also use the 'pure' Poisson product denoted by $\Re^{f_{m+1}}(f_1, \ldots, f_m)$:

$$\Re^{f_{m+1}}(f_1, \ldots, f_m) = f_{m+1}(\Lambda_1) \cdots f_{m+1}(\Lambda_N). \tag{1.17}$$

If necessary, we will use the notations $\Re_X(f_1, \ldots, f_m, f_{m+1})$ and $\Re_X^{f_{m+1}}(f_1, \ldots, f_m)$, X standing for the indeterminates $(x_1, \ldots, x_m)$, to indicate which indeterminates are eliminated.

In the resultant definition of $m+1$ polynomials in m indeterminates we use $\mathcal{A}$ as the resultant of m polynomials in $m-1$ indeterminates. This gives a valid inductive definition (in this section $m > 1$ is an assumption) with the definition of the resultant of two polynomials at the basis.

We will list some of the properties of the resultant as a generalization of Proposition 1.1.

Proposition 1.3 *The resultant $\Re$, defined as in* (1.16)*, has the following properties.*

- *$\Re$ is a polynomial in the coefficients of $f_1, \ldots, f_m, f_{m+1}$.*
- *$\Re$ is irreducible.*
- *$\Re$ vanishes if and only if $f_1, \ldots, f_m, f_{m+1}$ have a common solution.*
- *$\Re$ is, up to a factor -1, independent of the order of $f_1, \ldots, f_m, f_{m+1}$.*

Proof. See, for instance, the treatments of Netto [65] and Perron [74]. □

These properties only hold for generic polynomials. For instance, the fact that $\Re$ is irreducible does not hold in case f_{m+1} is reducible. We have the product formula:

$$\Re(f_1, \ldots, f_m, gh) = \Re(f_1, \ldots, f_m, g)\Re(f_1, \ldots, f_m, h). \tag{1.18}$$

To show how the third property from the proposition behaves in case the polynomials are not generic, i.e. have algebraic dependent coefficients, we can give a generalization of Proposition 1.2.

Proposition 1.4 *For any specialization of the coefficients such that $\mathcal{A}$ does not vanish, the resultant $\Re$ vanishes if and only if $f_1, \ldots, f_m, f_{m+1}$ have a common solution.*

Note that this definition of the resultant differs from the so-called resultant systems, originating from Kronecker [89]. The resultant systems can be used to give a resultant definition and to prove some of its properties [52]. It is based on the resultant of two univariate polynomials. Evaluation of the resultant systems leads to a set of multiples of the resultant of which a gcd has to be determined. Due to the growth of intermediate expressions this method is of less practical interest [52]. However, if the number of equations in the polynomial system is larger than $m+1$, the ideas of resultant systems are used to combine two or more equations to one.

In subsequent sections we will describe the Macaulay and Dixon formulation for evaluation of the resultant.

1.3 Macaulay's resultant

The resultant defined as in (1.16) can be evaluated with the so-called Macaulay matrix. Macaulay [60] has defined this as a generalized Sylvester matrix [83]. In our exhibition we deviate from the normal treatment as we are considering a non-homogenized system of $m+1$ polynomials in m indeterminates. If desired, we can homogenize it and proceed with a homogeneous system of m' equations in m' indeterminates ($m' = m+1$).

The Macaulay matrix is, like the Sylvester matrix, a coefficient matrix of the given system. A coefficient matrix represents a set of polynomial equations viewed as a system of linear equations where the power products serve as the indeterminates.

For sets of power products $\mathcal{T}_1, \ldots, \mathcal{T}_{m+1}$ we can form the system

$$\tau f_j = 0 \text{ for all } \tau \in \mathcal{T}_j \quad (j = 1, \ldots, m+1) \tag{1.19}$$

and denote its coefficient matrix by

$$\begin{pmatrix} \mathcal{T}_1 f_1 \\ \mathcal{T}_2 f_2 \\ \vdots \\ \mathcal{T}_{m+1} f_{m+1} \end{pmatrix}. \tag{1.20}$$

The number of columns of this matrix is equal to the number of different power products in the equations (1.19). The number of rows is equal to the number of equations. As we will see now, the coefficient matrix is not necessarily a square matrix.

One of the obvious choices is

$$\mathcal{T}_j = \{\tau(X) = x_1^{p_1} \cdots x_m^{p_m} \mid \deg \tau(X) \leq n - n_j\} \quad (j = 1, \ldots, m+1) \tag{1.21}$$

for a certain value of n [22]. All the equations in (1.19) are then of degree n. From this a coefficient matrix can be constructed with the number of columns equal to $\binom{n+m}{m}$, the number of power products of degree less than or equal n. The following table shows the number of rows and columns of the coefficient matrix for different choices of n, in case $m = 2$, $n_1 = 3$, $n_2 = 4$ and $n_3 = 5$:

n	#rows	#columns
6	19	28
7	31	36
8	46	45
9	64	55
10	85	66

For very large n the number of equations grows to $m+1$ times the number of power products. For small n the number of equations is less than the number of power products. We would like to have a square coefficient matrix.

It is Macaulay who determined the appropriate value for n, together with the sets $\mathcal{T}_j$, such that the corresponding coefficient matrix is square [60]. The appropriate value for n is the so-called critical degree:

$$n = 1 + \sum_{j=1}^{m+1} (n_j - 1) = \sum_{j=1}^{m+1} n_j \; - m. \tag{1.22}$$

We define the sets of power products

$$\mathcal{T}_j = \{\tau(X) = x_1^{p_1} \cdots x_m^{p_m} \mid \deg \tau(X) \leq n - n_j, \text{ if } j > 1, p_1 < n_1, \ldots, p_{j-1} < n_{j-1}\} \quad (j = 1, \ldots, m+1). \tag{1.23}$$

The sum of the number of power products in the $\mathcal{T}_j$'s is equal to $\binom{n+m}{m}$, the total number of power products in m indeterminates of degree less than or equal n. This can be understood by considering the sets of power products

$$x_j^{n_j} \mathcal{T}_j = \{x_j^{n_j} \tau \mid \tau \in \mathcal{T}_j\} \quad (j = 1, \ldots, m).$$

Together with $\mathcal{T}_{m+1}$ they form disjoint sets of power products. This means that the coefficient matrix (1.20), for this choice of the $\mathcal{T}_j$'s, is square.

Definition 1.3 *The Macaulay matrix is defined to be the coefficient matrix* (1.20) *with the* $\mathcal{T}_j$ *defined as in* (1.23).

The determinant of the Macaulay matrix is denoted by $\mathcal{D}$. From its construction it follows immediately that

Proposition 1.5 *$\mathcal{D}$ vanishes for every common solution of $f_1, \ldots, f_m, f_{m+1}$. $\mathcal{D}$ is a polynomial in the coefficients of $f_1, \ldots, f_m, f_{m+1}$. For generic coefficients $\mathcal{D}$ is a non-zero polynomial of degree $|\mathcal{T}_j|$ in the coefficients of f_j, for every j.*

To show that $\mathcal{D}$ is not the zero polynomial we can use the so-called Kronecker system [65]. This is the system given by

$$\left\{ \begin{array}{ll} f_j = x_j^{n_j}, & (j = 1, \ldots, m) \\ f_{m+1} = 1. & \end{array} \right. \tag{1.24}$$

The determinant of the corresponding Macaulay matrix is equal to 1 or -1.

It turns out that $\mathcal{D}$, for generic coefficients, is not irreducible. It is divisible by the determinant of a minor of the Macaulay matrix. To construct this minor we select these columns which corresponding power product τ has the property

$$\left\{ \begin{array}{l} \deg \tau > n - n_{m+1} \text{ and } \tau \text{ divisible by } x_1^{n_1} \cdots x_m^{n_m}, \text{ or,} \\ \deg \tau \leq n - n_{m+1} \text{ and } \tau \text{ divisible by } x_1^{n_1} \cdots x_m^{n_m} / x_j^{n_j} \text{ for some } 1 \leq j \leq m. \end{array} \right. \tag{1.25}$$

The rows of the minor have to be chosen such that, for rows containing coefficients of f_j, the coefficient of $x_j^{n_j}$ (for $j = m+1$, read 1, in other words the constant coefficient) of f_j appears in the selected columns of the minor.

In the univariate case ($m = 1$) this minor is empty. This is what was expected because, in this instance, the Macaulay matrix coincides with the Sylvester matrix and the determinant of the Sylvester matrix is irreducible for generic coefficients.

It is not necessary to use the minor. As the construction of the Macaulay matrix depends on the order of indeterminates, we could construct $m!$ matrices, one for each permutation of the indeterminates. It can be proved that the gcd of all these matrices is irreducible [60]. But obviously constructing one matrix with the division by the minor is preferable.

Theorem 1.2 *The resultant equals $\mathcal{D}$ divided by the minor indicated by* (1.25).

Proof. We refer to the threatments of Macaulay [60, 61]. $\square$

Example 1.4 *Evaluate the resultant of the system given by*

$$\begin{array}{rcl} f_1 & = & 2\,x^2 + axy + y^2 + 4\,x + c, \\ f_2 & = & -x^2 + bxy + y^2 - 2\,x - 3\,y + 5, \\ f_3 & = & x^2 - y^2. \end{array}$$

The critical degree of this system is 4. We multiply f_1 by the power products 1, x, y, x^2, xy and y^2. The same power products, except x^2, are the multipliers of

f_2. The multipliers of f_3 are 1, x, y and xy. The Macaulay matrix equals to:

$$\begin{pmatrix}
[x^4 & x^3y & x^2y^2 & xy^3 & y^4 & x^3 & x^2y & xy^2 & y^3 & x^2 & xy & y^2 & x & y & 1] \\
\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & a & 1 & 4 & 0 & c \\
0 & 0 & 0 & 0 & 0 & 2 & a & 1 & 0 & 4 & 0 & 0 & c & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 2 & a & 1 & 0 & 4 & 0 & 0 & c & 0 \\
2 & a & 1 & 0 & 0 & 4 & 0 & 0 & 0 & c & 0 & 0 & 0 & 0 & 0 \\
0 & 2 & a & 1 & 0 & 0 & 4 & 0 & 0 & 0 & c & 0 & 0 & 0 & 0 \\
0 & 0 & 2 & a & 1 & 0 & 0 & 4 & 0 & 0 & 0 & c & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & b & 1 & -2 & -3 & 5 \\
0 & 0 & 0 & 0 & 0 & -1 & b & 1 & 0 & -2 & -3 & 0 & 5 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -1 & b & 1 & 0 & -2 & -3 & 0 & 5 & 0 \\
0 & -1 & b & 1 & 0 & 0 & -2 & -3 & 0 & 0 & 5 & 0 & 0 & 0 & 0 \\
0 & 0 & -1 & b & 1 & 0 & 0 & -2 & -3 & 0 & 0 & 5 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}.$$

The power products corresponding to the columns are listed within the square brackets and are called the column labels. The determinant has two non-constant factors:

$$\begin{aligned}\mathcal{D} = -6\ &\left(b^2c^2 - 10\,abc + 25\,a^2 + 25\,ac - 10\,bc + 250\,a + 80\,b + 75\,c + 525\right) \\ &\left(b^2c^2 - 10\,abc + 25\,a^2 - ac + 34\,bc - 130\,a - 80\,b + 3\,c + 165\right).\end{aligned}$$

This factorization is according to the product formula (1.18). The first corresponds to the factor $x - y$ of f_3, the second to the factor $x + y$ of f_3.

The minor to divide by is given by the columns labelled with x^2y^2, x^2 and y^2, intersected by the rows 1, 6, and 7:

$$\begin{pmatrix} 0 & 2 & 1 \\ 2 & 0 & c \\ 0 & -1 & 1 \end{pmatrix} \text{ with determinant } -6.$$

Unfortunately, for non-generic polynomial systems the Macaulay matrix (or the minor by which to divide) often becomes singular. Moreover, the Macaulay matrix is very large, even for small problems. There have been numerous publications on how to improve the applicability of the Macaulay resultant. The structure of the

Macaulay matrix can be exploited to evaluate its determinant quicker [47]. If the polynomials are sparse, submatrices of the Macaulay matrix can serve as a resultant matrix [21, 32]. These matrices are also called sparse resultant matrices and Newton matrices [33]. Perturbation of the original system may be used in case of a singular matrix [20].

1.4 The Cayley and Dixon resultant

In his original work [22] Cayley describes how to obtain the resultant for two univariate polynomials of the same degree. However, the method also works for pairs of polynomials that are not of the same degree. The formulation of the resultant by Dixon [30] is a generalization of the work of Cayley to the case of a system of three polynomials with two indeterminates. We will describe the univariate case first.

Let $f(x)$ and $g(x)$ be non-zero polynomials in $K[x]$ of degree n and m, respectively. Let α be a new indeterminate. Define the polynomial

$$\mathcal{C}(x, \alpha) = \frac{f(x)g(\alpha) - f(\alpha)g(x)}{x - \alpha}. \tag{1.26}$$

Because $f(x)g(\alpha) - f(\alpha)g(x)$ is divisible by $x - \alpha$, $\mathcal{C}(x, \alpha)$ is a polynomial in x and α. It is a symmetric polynomial in x and α. It is of degree $n - 1$ (or $m - 1$ if $m > n$) in x or α.

$\mathcal{C}$ evaluates to zero for every common root λ of $f(x)$ and $g(x)$. This means that every coefficient of $\mathcal{C}$, considered as polynomial in α, evaluates to zero for the common roots of $f(x)$ and $g(x)$. This gives n equations. These n equations can be written as a coefficient matrix. A necessary condition for the existence of a common root for $f(x)$ and $g(x)$ is singularity of this matrix.

To prove that the matrix is non-singular for generic polynomials $f(x)$ and $g(x)$, we have to prove this for at least one instance of $f(x)$ and $g(x)$. A system like (1.24) can be tested. For example, if $f(x) = x^n$ and $g(x) = 1$ then

$$\mathcal{C}(x, \alpha) = x^{n-1} + x^{n-2}\alpha + \ldots + x\alpha^{n-2} + \alpha^{n-1}.$$

This gives the unit diagonal matrix of order n.

The constructed matrix is called the Cayley matrix. Its determinant is called the Cayley resultant.

It can be shown that, if $n > m$, the Cayley resultant is not irreducible for algebraically independent coefficients. It has an additional factor equal to a^{n-m}, where a is the leading coefficient of $f(x)$. For example, choosing $f(x) = ax^n$ and $g(x) = 1$ gives a diagonal matrix with a on the diagonal. Its determinant is equal to a^n.

Dixon's generalization [30] to the bivariate case works as follows. Let f_1, f_2 and f_3 be polynomials in x_1 and x_2. Using two new indeterminates α_1 and α_2, construct the matrix

$$M = \begin{pmatrix} f_1(x_1,x_2) & f_2(x_1,x_2) & f_3(x_1,x_2) \\ f_1(\alpha_1,x_2) & f_2(\alpha_1,x_2) & f_3(\alpha_1,x_2) \\ f_1(\alpha_1,\alpha_2) & f_2(\alpha_1,\alpha_2) & f_3(\alpha_1,\alpha_2) \end{pmatrix}. \tag{1.27}$$

This matrix is singular for $x_1 = \alpha_1$ and for $x_2 = \alpha_2$. Hence, its determinant is divisible by $(x_1 - \alpha_1)(x_2 - \alpha_2)$. We define the polynomial

$$\mathcal{C}(x_1, x_2, \alpha_1, \alpha_2) = \frac{\det M}{(x_1 - \alpha_1)(x_2 - \alpha_2)}.$$

From the matrix M we see that $\mathcal{C}$ vanishes for every common solution (x_1, x_2) of f_1, f_2 and f_3. This means that vanishing of all the coefficients of $\mathcal{C}$, considered as polynomial in α_1 and α_2, is a necessary condition for the existence of a common solution for f_1, f_2 and f_3. Again, this can be transformed into a linear system, in which the power products in x_1 and x_2 serve as indeterminates. We will now look at the number of equations and indeterminates this constructed linear system consists of.

Dixon considered the following case. Let f_1, f_2 and f_3 be of degree n_1 in x_1 and of degree n_2 in x_2. Then the degree of $\mathcal{C}$ is $2\,n_1 - 1$ in α_1, $n_2 - 1$ in α_2, $n_1 - 1$ in x_1 and $2\,n_2 - 1$ in x_2. This means that $\mathcal{C}$, considered as polynomial in α_1 and α_2, has $2\,n_1 n_2$ terms. This is the number of linear equations. The number of indeterminates, the power products in x_1 and x_2, is the same. Hence, the corresponding coefficient matrix is square.

It is possible to generalize this construction to systems with $m+1$ equations in m indeterminates [77]. After introducing m new indeterminates $\alpha_1, \alpha_2, \ldots, \alpha_m$, a matrix similar to (1.27), called the Cancellation matrix, can always be constructed:

$$M = \begin{pmatrix} f_1(x_1, x_2, \ldots, x_m) & \cdots & f_{m+1}(x_1, x_2, \ldots, x_m) \\ f_1(\alpha_1, x_2, \ldots, x_m) & \cdots & f_{m+1}(\alpha_1, x_2, \ldots, x_m) \\ \vdots & \ddots & \vdots \\ f_1(\alpha_1, \alpha_2, \ldots, \alpha_m) & \cdots & f_{m+1}(\alpha_1, \alpha_2, \ldots, \alpha_m) \end{pmatrix}.$$

Its determinant can be divided by m linear factors. Let this result be denoted by $\mathcal{C}$:

$$\mathcal{C}(x_1, x_2, \ldots, x_m, \alpha_1, \alpha_2, \ldots, \alpha_m) = \frac{\det M}{(x_1 - \alpha_1)(x_2 - \alpha_2)\cdots(x_m - \alpha_m)}. \tag{1.28}$$

Each coefficient of a power product in the new m indeterminates gives one equation in the linear system that is extracted. The power products in the original m indeterminates serve as indeterminates. This leads to the Dixon matrix.

Definition 1.4 *The Dixon matrix is defined to be the coefficient matrix of the coefficients of $\mathcal{C}$ viewed as polynomial in $\alpha_1, \alpha_2, \ldots, \alpha_m$.*

With a system like in (1.24) it can be shown that the coefficient matrix is not singular for systems with generic polynomials.

Theorem 1.3 *[77] For systems with generic polynomials the determinant of the Dixon matrix evaluates to the resultant.*

The drawback of the Dixon resultant formulation is that the constructed coefficient matrix may not be square for non-generic systems. If the matrix is square, it may happen that the Dixon resultant matrix becomes singular, although the resultant is not zero. Some methods were developed to surmount these difficulties, applying pseudo inverse techniques and rank submatrix computations [75, 77]. These problems especially occur for sparse polynomial systems. The sparsity of the polynomial system can be incorporated with some modifications to the Dixon resultant [77].

Example 1.5 *Build the Dixon resultant for the system as in Example 1.4.*

Using the intermediate matrix M from (1.27) we find the polynomial

$$\begin{aligned}
\mathcal{C}(x, y, \alpha_1, \alpha_2) &= 3\,c\alpha_2 + 2\,c\alpha_1 + 15\,\alpha_2\,\alpha_1 - {\alpha_2}^2 cb + 5\,a{\alpha_2}^2 + 20\,\alpha_1 \\
&\quad + \left(-\alpha_2\,cb - 6\,\alpha_2\,\alpha_1 - 2\,a{\alpha_2}^2 - 4\,b{\alpha_2}^2 + 3\,c + 12\,\alpha_2 + 15\,\alpha_1 + 5\,a\alpha_2\right) x \\
&\quad + \left(2\,c - 12\,\alpha_1 - 3\,a{\alpha_2}^2 - cb\alpha_1 + 5\,a\alpha_1 - 9\,\alpha_2\,\alpha_1 + 15\,\alpha_2 + 20\right) y \\
&\quad + \left(5\,a - 3\,a\alpha_1 - 2\,a\alpha_2 - 3\,b{\alpha_2}^2 - 4\,b\alpha_2 - cb\right) y^2 \\
&\quad + \left(3\,b\alpha_2\,\alpha_1 - 6\,\alpha_2 - 9\,\alpha_1 + 15\right) xy,
\end{aligned}$$

from which the Dixon matrix can be constructed (here we use the column labels ${\alpha_2}^2$, $\alpha_1\alpha_2$, α_2, α_1 and 1):

$$\begin{pmatrix}
5\,a - bc & 15 & 3\,c & 2\,c + 20 & 0 \\
-2\,a - 4\,b & -6 & 12 + 5\,a - bc & 15 & 3\,c \\
-3\,a & -9 & 15 & -12 + 5\,a - bc & 2\,c + 20 \\
-3\,b & 0 & -2\,a - 4\,b & -3\,a & 5\,a - bc \\
0 & 3\,b & -6 & -9 & 15
\end{pmatrix}.$$

This example also illustrates the fact that the order of the Dixon matrix is much smaller than the order of the Macaulay matrix: 5 instead of 15.

Formulations that mix the Dixon with the Macaulay formulation have been developed [29, 33].

1.5 Solving polynomial systems by the u-resultant

By Proposition 1.3 we know that vanishing of the resultant is a necessary and sufficient condition for the existence of a common solution of the system. We have seen in Section 1.1 how this can be used to solve a polynomial system with two equations in two indeterminates with the Sylvester resultant. This method can be generalized to be applicable for the Macaulay resultant of systems with more than two equations.

The so-called u-resultant, originally due to Liouville [59, 65, 80], is proposed as the way to find all the components of the solutions. First, we transform the system (1.13) into a system in which the number of polynomials is equal to the number of indeterminates (usually the u-resultant is introduced within such a context). We assume $K = K'(t)$, K' a field and t a parameter of the system. Furthermore, we assume that all the polynomials are written as polynomials in $K'[x_1, \ldots, x_m, t]$. This gives a polynomial system with $m' = m + 1$ equations in m' indeterminates. For the u-resultant this system is extended with one more polynomial

$$f_{m'+1} = -u + u_1 x_1 + \cdots + u_m x_m + u_{m'} t \tag{1.29}$$

and we have to adjoin the new indeterminates $u, u_1, \ldots, u_{m'}$ to the coefficient field. This gives a system in the form (1.13) with $m = m'$ and $K = K'(u, u_1, \ldots, u_{m'})$.

The resultant of this extended system is called the u-resultant and is a polynomial in u, $u_1, \ldots, u_{m'}$ over K'. Substituting $u_1 = 1$ and $u_2 = \cdots = u_{m'} = 0$ in the resultant gives a polynomial in u and we find the x_1-component of the solutions as the roots of this polynomial. By similar substitutions we can find all the components of the solutions. The individual components have to be linked together to constitute one solution using substitutions as well [28, 62].

Example 1.6 *Find all the real solutions for*

$$f = x^2 + xy + y^2 + x + y - 5, \quad g = 2\,x^2 + 3\,xy + 5\,y^2 + 7\,x + 11\,y - 28.$$

The u-resultant equals

$$\begin{aligned}
&7\,u^4 + (-16\,u_1 + 38\,u_2)\,u^3 + \left(18\,{u_1}^2 + 54\,u_1\,u_2 - 48\,{u_2}^2\right)u^2 \\
&\quad + \left(144\,{u_1}^3 - 486\,{u_1}^2 u_2 + 444\,u_1\,{u_2}^2 - 286\,{u_2}^3\right)u \\
&\quad -153\,{u_1}^4 + 234\,{u_1}^3 u_2 - 144\,{u_1}^2 {u_2}^2 - 242\,u_1\,{u_2}^3 + 289\,{u_2}^4.
\end{aligned} \tag{1.30}$$

Substituting $u_1 = 1$ and $u_2 = 0$ gives $7\,u^4 - 16\,u^3 + 18\,u^2 + 144\,u - 153$ with real roots 1 and approximately -2.297, the x-components of two real solutions. Substituting $u_1 = 0$ and $u_2 = 1$ gives $7\,u^4 + 38\,u^3 - 48\,u^2 - 286\,u + 289$ with roots 1 and approximately 2.211. For $u_1 = u_2 = 1$ we obtain the roots 2 and approximately -0.086, making sure the correspondence between the components. In fact, expression (1.30) has a factor $u - u_1 - u_2$.

The u-resultant is to be evaluated by any resultant method, e.g. Macaulay resultant, and is not a resultant formulation itself. The drawback of this method [28] for finding all the components of the solutions is that for the evaluation of the resultant the number of equations and indeterminates is one more. This can lead to resultant matrices that are much bigger in size. Moreover, the arithmetic operations in the coefficient field are more expensive as the number of parameters of the system increases by m'.

One section in Chapter 4 is devoted to the properties of the Bézout resultant matrix when applied to the u-resultant, especially solving the problem of how to link the individual components to one solution.

1.6 Solving polynomial systems by Gröbner bases

Among the other methods to solve polynomial systems algebraically – for a comprehensive overview we refer to [27, 28, 54] – we mention the so-called characteristic set. This method was introduced by Ritt and has been further developed by Wu [96] and others [92]. It is often applied to polynomial systems from algebraic geometric problems [53]. The method focuses on describing the variety of solutions by sets of polynomials in triangular form. Because we are not going to use the method and quite some terminology has to be introduced [3], we refer to the treatment of Wang [93].

In this section we focus on the Gröbner basis technique for solving polynomial systems. We will do this in brief as already several good treatments on this subject exist [9, 18]. The results mentioned in this section can all be found there. Gröbner bases, originally due to Buchberger [17] – he also provided the algorithm for constructing them – were only recognized to be very powerful for solving polynomials systems years later [18, 50, 85].

In Chapter 4 we will use Gröbner bases for the construction of the resultant. An introduction to the similar concept of Involutive bases can be found in Chapter 5.

A set of polynomials $\mathcal{F} = \{f_1, \ldots, f_s\}$ from $K[x_1, \ldots, x_m]$ constitutes an ideal $\mathcal{I} \subset K[x_1, \ldots, x_m]$:

$$\mathcal{I} = \{a_1 f_1 + \cdots + a_s f_s \mid a_1, \ldots a_s \in K[x_1, \ldots, x_m]\}.$$

The set of solutions of the corresponding polynomial system

$$f_1 = 0, f_2 = 0, \ldots, f_s = 0,$$

is called the affine variety $V(f_1, \ldots, f_s)$, also denoted by $V(\mathcal{I})$:

$$V(f_1, \ldots, f_s) = \left\{ \Lambda \in \overline{K}^m \mid f_1(\Lambda) = 0, \ldots, f_s(\Lambda) = 0 \right\}.$$

The ideal is called a zero-dimensional ideal if $V(\mathcal{I})$ is a non-empty and finite set. This means that the polynomial system is solvable and there are finitely many solutions. The ideal is called a radical ideal if $\mathcal{I}$ coincides with the ideal

$$\mathcal{I}(V) = \{f \in K[x_1, \ldots, x_m] \mid f(\Lambda) = 0 \text{ for all } \Lambda \in V(\mathcal{I})\} .$$

This latter ideal always is a radical ideal itself.

The ideal $\mathcal{I}$ is said to be generated by $\mathcal{F}$ and $\mathcal{F}$ is called a basis of $\mathcal{I}$. A Gröbner basis of $\mathcal{I}$ is a basis of $\mathcal{I}$ satisfying certain properties. A Gröbner basis is defined with respect to an ordering $\prec$ on all the power products in $x_1, \ldots, x_m$. The ordering $\prec$ should be a so-called admissible ordering.

Definition 1.5 *An ordering $\prec$ is called an admissible ordering if*

- *$\prec$ is a total ordering, i.e. for any power products τ_1, τ_2, either $\tau_1 \prec \tau_2$, $\tau_2 \prec \tau_1$ or $\tau_1 = \tau_2$,*
- *for all power products τ_1, τ_2, τ_3: $\tau_1 \prec \tau_2$ implies $\tau_1\tau_3 \prec \tau_2\tau_3$,*
- *$\prec$ is a well-ordering, i.e. every non-empty set of power products has a smallest element with respect to $\prec$.*

This ordering extends to an ordering on all the monomials: $c_1\tau_1 \prec c_2\tau_2$ if and only if $\tau_1 \prec \tau_2$.

The leading monomial of a polynomial f, denoted by $\mathrm{lm}(f)$, is defined to be the monomial of f that is the maximal monomial with respect to $\prec$. The leading monomial consists of the leading coefficient, $\mathrm{lc}(f)$, multiplied by the leading power product, $\mathrm{lpp}(f)$.

Two well-known admissible orderings are the lexicographic ordering, denoted by $\prec_{\mathrm{lex}}$ and the total degree ordering $\prec_{\mathrm{deg}}$. Denote by $d_j(\tau)$ the degree of τ in x_j. The (total) degree of τ, denoted by $\deg\tau$, equals $d_1(\tau) + \cdots + d_m(\tau)$.

Definition 1.6 *Let τ_1 and τ_2 be power products in $x_1, \ldots, x_m$. Then*

- *$\tau_1 \prec_{\mathrm{lex}} \tau_2$ if, for a $j \geq 1$, $d_j(\tau_1) < d_j(\tau_2)$ and, if $j > 1$, $d_1(\tau_1) = d_1(\tau_2)$, ..., $d_{j-1}(\tau_1) = d_{j-1}(\tau_2)$.*
- *$\tau_1 \prec_{\mathrm{deg}} \tau_2$ if $\deg\tau_1 < \deg\tau_2$, or $\deg\tau_1 = \deg\tau_2$ and $\tau_1 \prec_{\mathrm{lex}} \tau_2$.*

For example: if $x \prec_{\mathrm{lex}} y$, then $x^d \prec_{\mathrm{lex}} y$ for any d.

Note that the (total) degree of a polynomial f is not related to admissible orderings but is defined as the maximum total degree of the monomials. It equals the total degree of $\mathrm{lpp}(f)$ with respect to $\prec_{\mathrm{deg}}$.

We now have enough material to give the definition of the Gröbner basis.

Definition 1.7 *$\mathcal{G}$ is called a Gröbner basis of an ideal $\mathcal{I}$ with respect to $\prec$ if $\mathcal{G}$ is a finite basis of $\mathcal{I}$ and for every $f \in \mathcal{I}$ there exists a $g \in \mathcal{G}$ such that $\mathrm{lm}(f)$ is divisible by $\mathrm{lm}(g)$.*

It can be proved that for every ideal there exists a Gröbner basis for every admissible ordering, and that it can be constructed in finitely many steps. For an ideal given by a set of generators, the so-called Buchberger algorithm can be used to construct a Gröbner basis with respect to any admissible ordering. In the past years many optimizations to this algorithm were invented [19, 35, 36, 43, 95].

For the purpose of solving a polynomial system the lexicographic ordering $\prec_{\text{lex}}$ is used. The Gröbner basis with respect to this ordering has a special property.

Proposition 1.6 *Let $\mathcal{G}$ be a Gröbner basis with respect to $x_1 \prec_{\text{lex}} \cdots \prec_{\text{lex}} x_m$ of an ideal $\mathcal{I}$. Then $\mathcal{G} \cap K[x_1, \ldots, x_j]$ is a Gröbner basis of the ideal $\mathcal{I} \cap K[x_1, \ldots, x_j]$.*

With $\mathcal{G} \cap K[x_1]$ the x_1-component of the solutions of the polynomial system associated with the ideal can be determined.

Constructing the Gröbner basis with respect to a lexicographic ordering is computationally hard. With the so-called FGLM algorithm [24, 36, 67] this Gröbner basis can be obtained more efficient by conversion of the Gröbner basis with respect to a total degree ordering. The latter can be constructed easier as the given generators of the ideal are usually 'close' to it.

The so-called Shape Lemma [7] gives the structure of some zero-dimensional ideals [9]:

Proposition 1.7 *Let $\mathcal{I}$ be a zero-dimensional radical ideal. If the x_1-components of the points in $V(\mathcal{I})$ are all distinct then $\mathcal{I}$ has a basis of the form*

$$\{g_1(x_1), x_2 - g_2(x_1), \ldots, x_m - g_m(x_1)\}. \tag{1.31}$$

This basis is a Gröbner basis with respect to the lexicographic ordering $x_1 \prec_{\text{lex}} \cdots \prec_{\text{lex}} x_m$. For a so-called reduced Gröbner basis the $g_2, \ldots, g_m$ can be made of degree less than the degree of g_1.

The form (1.31) also exists for non-radical ideals, for instance the ideal given by the basis $\{x_1^2, x_2 + x_1\}$. We refer to [7] for an exact characterization. Similar to the Shape Lemma form in (1.31) there exists the univariate rational form [2, 68, 76], representing $V(\mathcal{I})$, with $g_2(x_1), \ldots, g_m(x_1)$ rational functions in x_1.

Obviously, if the Shape Lemma is applicable, we easily find all the solutions of the system. Let λ be a root of $g_1(x_1)$, then the solution with all its components equals $(\lambda, g_2(\lambda), \ldots, g_m(\lambda))$.

In the other cases, Proposition 1.6 can be applied. Because the ideal is zero-dimensional, $\mathcal{I} \cap K[x_1]$ is generated by one univariate polynomial in x_1. The roots of this polynomial can be substituted into the other elements of $\mathcal{G}$. In order to find the x_2-component, we only have to substitute into the elements of $\mathcal{G} \cap K[x_1, x_2]$ and check whether their leading coefficients do not vanish all:

Theorem 1.4 *The solution Λ_j for $\mathcal{I} \cap K[x_1, \ldots, x_j]$ can be extended with an x_{j+1}-component from $\overline{K}$ such that this is a solution for $\mathcal{I} \cap K[x_1, \ldots, x_j, x_{j+1}]$ if there exists a $g \in \mathcal{G} \cap K[x_1, \ldots, x_j, x_{j+1}]$ that has, considered as polynomial in x_{j+1}, a leading coefficient that does not vanish at Λ_j.*

This theorem is usually called the Extension Theorem. It can be proved using the u-resultant and Proposition 1.2 [27]. The limitation of this theorem is that, in case all the mentioned leading coefficients vanish, it does not tell us whether Λ_j can be extended or not. This makes the theorem similar to Proposition 1.2.

Example 1.7 *Solve by means of Gröbner bases the polynomial system as given in Example* (1.4) *with* $b = 1$ *and* $c = -4$*.*

The Gröbner basis with respect to the lexicographic ordering $x \succ_{\text{lex}} y \succ_{\text{lex}} a$ equals

$$\begin{aligned} \mathcal{G} = \{ \quad & (5\,a+19)(25\,a^2-86\,a-47), -2136\,x+2136\,y+875\,a^2+1440\,a-7163, \\ & (5\,a+19)\,(24\,y+25\,a-31)\,, 6408\,y^2-32040\,y-5375\,a^2-21560\,a+27727 \quad \}. \end{aligned}$$

The Shape Lemma is not applicable: the partial solution $a = -\frac{19}{5}$, after checking with the Extension Theorem, can be extended to the two solutions with $x = y = \frac{5}{2} \pm \frac{1}{2}\sqrt{5}$. The other roots of the univariate polynomial in a are $a = \frac{43}{25} \pm \frac{12}{25}\sqrt{21}$ and extend to one solution with $x = \frac{25\,a-31}{24}$ and $y = -x$.

In Chapter 4 we will use the concept of the normal form of a polynomial with regard to a Gröbner basis:

Definition 1.8 *A polynomial f is in normal form with regard to a Gröbner basis $\mathcal{G}$ if* $\text{lm}(f)$ *is not divisible by any* $\text{lm}(g)$*,* $g \in \mathcal{G}$*.*

If a polynomial is not in normal form the leading monomial can be divided out as follows. Let $g \in \mathcal{G}$ be such that $\text{lm}(f)$ is divisible by $\text{lm}(g)$. Then

$$f' = f - \frac{\text{lm}(f)}{\text{lm}(g)} g$$

is 'closer' to a normal form in the sense that the leading monomial becomes smaller: $\text{lm}(f') \prec \text{lm}(f)$. After finitely many of such division steps a polynomial in normal form remains, the so-called normal form of f. The Gröbner basis definition can be reformulated in this way that the normal form of a polynomial f is zero if and only if $f \in \mathcal{I}$. Gröbner bases solve the ideal membership problem: we can determine whether $f \in \mathcal{I}$ by checking whether its normal form with regard to the Gröbner basis of $\mathcal{I}$ is zero.

Chapter 2

The Bézout resultant of two polynomials

In this chapter we discuss the resultant of two univariate polynomials according to the construction of Bézout. In subsequent chapters this construction is being generalized to the multivariate case. In the first section we give the construction of a resultant matrix, called the Bézout matrix. We show there exists a strong relation to the Sylvester matrix. In Section 2.2 the direct relation to the Bézoutiant is explained. This is continued in Section 2.3, showing how to use the Bézout matrix for root separation problems.

2.1 The construction

Let $f(x)$ and $g(x)$ be non-zero polynomials in $K[x]$ with degrees n and m, respectively:

$$\begin{aligned} f(x) &= a_0x^n + a_1x^{n-1} + \cdots + a_n, \\ g(x) &= b_0x^m + b_1x^{m-1} + \cdots + b_m. \end{aligned}$$

The construction of the Bézout resultant is based on the division by $f(x)$ of a properly chosen multiple of $g(x)$ [58, 80]. Because of the division we assume $a_0 \neq 0$. In the sequel, if we speak about remainder, we mean remainder on dividing by $f(x)$. The multiple of $g(x)$ is to be chosen such that its remainder is independent of x.

We denote the remainder of $g(x)$, by $g_0(x)$:

$$g_0(x) = b_{00}x^{n-1} + \cdots + b_{0,n-1}.$$

For $k = 1, \ldots, n-1$, define $g_k(x)$ as the remainder of $x^k g(x)$:

$$g_k(x) = b_{k0}x^{n-1} + \cdots + b_{k,n-1}. \tag{2.1}$$

It is easy to see that $g_k(x)$ is equal to the remainder of $xg_{k-1}(x)$. The coefficients of $g_k(x)$ can be obtained from those of $g_{k-1}(x)$ by the formulae

$$b_{kj} = b_{k-1,j+1} - b_{k-1,0}\frac{a_{j+1}}{a_0} \quad (j = 0, \ldots, n-1;\; b_{k-1,n} = 0). \tag{2.2}$$

From these remainders we construct the (square) coefficient matrix $\mathcal{B}$:

$$\mathcal{B} = [b_{kj}]_{k,j=0}^{n-1}. \tag{2.3}$$

In the sequel, this matrix will be referred to as the Bézout matrix. We call the determinant of the Bézout matrix the Bézout resultant.

Example 2.1 *Construct the Bézout matrix for the system given in Example 1.2:*

$$\begin{aligned} f(x) &= x^2 + xy + y^2 + x + y + 1, \\ g(x) &= 2\,x^2 + 3\,xy + 5\,y^2 + 7\,x + 11\,y + 13. \end{aligned}$$

We can compute the remainders of $g(x)$ and $xg(x)$ easily:

$$g_0(x) = g(x) - 2\,f(x) = (y+5)\,x + 3\,y^2 + 9\,y + 11,$$

$$g_1(x) = xg(x) - 2\,xf(x) - (y+5)f(x) = \left(2\,y^2 + 3\,y + 6\right)x - y^3 - 6\,y^2 - 6\,y - 5.$$

The Bézout matrix is a 2×2 matrix (half the size of the Sylvester matrix):

$$\mathcal{B} = \begin{pmatrix} y+5 & 3\,y^2 + 9\,y + 11 \\ 2\,y^2 + 3\,y + 6 & -y^3 - 6\,y^2 - 6\,y - 5 \end{pmatrix}.$$

The determinant coincides with the Sylvester resultant. This is due to the fact that a_0, the leading coefficient of $f(x)$, is 1, as we will see in the following theorem. In this theorem we will use the concept of leading principal minor [37].

Definition 2.1 *The k-th leading principal minor of a matrix M is the order k submatrix of M consisting of the elements from the first k rows and the first k columns.*

The following two results can be found in [6] and [44]. A new proof is supplied for the first theorem:

Theorem 2.1 *Let $\mathcal{B}_k$ be the k-th leading principal minor of $\mathcal{B}$. Let $\Re_k(f,g)$ be the k-th subresultant of the Sylvester matrix of $f(x)$ and $g(x)$. We have*

$$a_0^{m-k} \det \mathcal{B}_{n-k} = \Re_k(f,g),$$

which gives for $k = 0$

$$a_0^m \det \mathcal{B} = (-1)^{n(n-1)/2} \Re(f,g).$$

Proof. We can write the k-th subresultant of the Sylvester matrix as

$$\Re_k(f,g) = \begin{array}{l} \left|\begin{array}{llllllll} a_0 & a_1 & a_2 & \dots & & & \dots & a_{n+m-2k-1} \\ 0 & a_0 & a_1 & a_2 & \dots & & \dots & a_{n+m-2k-2} \\ & & & \dots & & & & \dots \\ 0 & \dots & 0 & a_0 & a_1 & a_2 & \dots & a_{n-k} \\ 0 & \dots & 0 & 0 & b_0 & b_1 & \dots & b_{m-k} \\ 0 & \dots & 0 & b_0 & b_1 & \dots & b_{m-k} & b_{m-k+1} \\ & & & \dots & & & & \dots \\ b_0 & b_1 & \dots & & & & \dots & b_{n+m-2k-1} \end{array}\right| \end{array} \begin{array}{l} \left.\vphantom{\begin{array}{l} a \\ a \\ a \\ a \end{array}}\right\} m-k \text{ rows} \\ \left.\vphantom{\begin{array}{l} a \\ a \\ a \\ a \end{array}}\right\} n-k \text{ rows} \end{array}$$

while we assume, for definiteness, that $m \leq n$, and set $a_j = 0$ for $j > n$, $b_k = 0$ for $k > m$. We have $\Re_0(f,g) = (-1)^{n(n-1)/2}\Re(f,g)$. The determinant of the subresultant matrix is unaltered by multiplying it from the left by

$$\begin{pmatrix} I_{m-k} & \mathbf{O} \\ \mathcal{B}' & I_{n-k} \end{pmatrix},$$

where I is the identity matrix of corresponding order, and

$$\mathcal{B}' = -\frac{1}{a_0}\begin{pmatrix} 0 & 0 & \dots & & 0 \\ \dots & & & & \dots \\ 0 & 0 & \dots & & 0 \\ 0 & 0 & \dots & 0 & b_{n-m-1,0} \\ 0 & 0 & \dots & b_{n-m-1,0} & b_{n-m,0} \\ \dots & & & & \dots \\ 0 & b_{n-m-1,0} & \dots & b_{n-k-4,0} & b_{n-k-3,0} \\ b_{n-m-1,0} & b_{n-m,0} & \dots & b_{n-k-3,0} & b_{n-k-2,0} \end{pmatrix} \begin{array}{l} \left.\vphantom{\begin{array}{l} a \\ a \\ a \end{array}}\right\} n-m \text{ rows} \\ \left.\vphantom{\begin{array}{l} a \\ a \\ a \\ a \\ a \end{array}}\right\} m-k \text{ rows} \end{array}$$

($b_{-1,0} = b_0$). Using the equalities (2.2), the result of this multiplication (which is equivalent to elementary linear operations applied to the rows of the subresultant matrix) will be

$$\begin{pmatrix} a_0 & a_1 & a_2 & \dots & & & & \\ 0 & a_0 & a_1 & \dots & & & & \\ \dots & & & & & & & \\ 0 & \dots & a_0 & a_1 & a_2 & \dots & & \\ 0 & \dots & 0 & b_{00} & b_{0,1} & \dots & b_{0,n-k-1} \\ 0 & \dots & 0 & b_{1,0} & b_{1,1} & \dots & b_{1,n-k-1} \\ \dots & & \dots & \dots & & & \dots \\ 0 & \dots & 0 & b_{n-k-1,0} & b_{n-k-1,1} & \dots & b_{n-k-1,n-k-1} \end{pmatrix},$$

whose determinant equals to $a_0^{m-k} \det \mathcal{B}_{n-k}$. □

Theorem 2.2 *The degree of* $\gcd(f,g)$ *equals* D *if and only if*

$$\mathcal{B}_n = \mathcal{B}_{n-1} = \ldots = \mathcal{B}_{n-D+1} = 0, \ \mathcal{B}_{n-D} \neq 0.$$

In this case, $\gcd(f,g)$ *equals the determinant obtained from* $\mathcal{B}_{n-D}$ *on replacing the entries in the last column by*

$$\sum_{j=n-D-1}^{n-1} b_{0j}x^{n-j-1},\ \sum_{j=n-D-1}^{n-1} b_{1j}x^{n-j-1},\ldots,\ \sum_{j=n-D-1}^{n-1} b_{n-D-1,j}x^{n-j-1}.$$

For $n > m$*, the polynomials* $p(x)$ *and* $q(x)$ *giving the linear representation of* $\gcd(f,g)$*:*

$$p(x)\,f(x) + q(x)\,g(x) = \gcd(f,g) \tag{2.4}$$

can be expressed as the determinants obtained on replacing the entries in the last column of $\mathcal{B}_{n-D}$ *by*

$$0, -p_0(x), \ldots, -p_{n-D-2}(x) \quad \text{and} \quad 1, x, \ldots, x^{n-D-1},$$

correspondingly, with $p_k(x) = (b_{00}x^k + b_{1,0}x^{k-1} + \cdots + b_{k,0})/a_0$.

Proof. The first part of the theorem follows from the properties of the subresultants [1, 57, 65]. Concerning the linear representation of $\gcd(f,g)$, we can use the expression

$$x^k g(x) = p_{k-1}(x)f(x) + g_k(x), \quad (k = 0,\ldots,n-D-1;\ p_{-1}(x) = 0)$$

which can be deduced easily. Multiply the k-th equality by the cofactor to the element of $\mathcal{B}_{n-D}$ standing in its last column and in the k-th row, and add these products. Using the properties of the determinant expansion and the derived expression for the gcd, we have the desired result. □

Example 2.2 *Determine the* gcd *of the polynomials*

$$\begin{aligned} f(x) &= x^3 + yx^2 + y^2 - yx + 3\,y - 3\,x + 2, \\ g(x) &= x^3 + 2\,yx^2 + x^2 + y^2x + 3\,y^2 + 5\,yx - 3\,y - 5\,x + 3, \end{aligned}$$

for various values of y.

The Bézout matrix equals

$$\mathcal{B} = \begin{pmatrix} y+1 & y^2+6\,y-2 & 2\,y^2-6\,y+1 \\ 5\,y-2 & 3\,y^2-2\,y+4 & -y^3-4\,y^2-5\,y-2 \\ -2\,y^2+4 & -y^3+y^2+8\,y-8 & -5\,y^3-13\,y^2-4\,y+4 \end{pmatrix}$$

and its determinant evaluates to

$$\det \mathcal{B} = y^3\,(y-1)\,(y+4)\,(y+9)\,(y+16)\,.$$

The values $y = 0$, $y = 1$, $y = -4$, $y = -9$ and $y = -16$ give a singular Bézout matrix. For all other values of y the gcd is 1. Substituting $y = 0$ in the Bézout matrix gives

$$\begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 4 & -8 & 4 \end{pmatrix} \text{ with } \mathcal{B}_3 = \mathcal{B}_2 = 0.$$

Thus, the gcd of $f(x)$ and $g(x)$ is of degree 2. The last column of the principal minor $\mathcal{B}_1$ is a single element. This has to be replaced by $x^2 - 2\,x + 1$, which equals the gcd. Substituting for instance $y = 1$ gives a Bézout matrix of rank 2:

$$\begin{pmatrix} 2 & 5 & -3 \\ 3 & 5 & -12 \\ 2 & 0 & -18 \end{pmatrix} \text{ and } \mathcal{B}_2 = \begin{pmatrix} 2 & 5\,x - 3 \\ 3 & 5\,x - 12 \end{pmatrix}$$

after replacing the last column as described. Its determinant $-5(x+3)$ is equal to the gcd. Substituting for y the values -4, -9 and -16 also gives a gcd of degree 1.

Note that the corresponding $p(x)$ and $q(x)$ can not be used directly, as the degrees of $f(x)$ and $g(x)$ are equal. But they can always be applied to $f(x)$ and $g_0(x)$, as the degree of $g_0(x)$ is less than the degree of $f(x)$ and $g_0(x)$ and $f(x)$ have the same gcd. In this example $g_0(x) = g(x) - f(x)$. For $y = 0$, we have $p(x) = 0$ and $q(x) = 1$. This gives $\gcd(f(x), g(x)) = g_0(x) = g(x) - f(x)$. For $y = 1$, we have

$$p(x) = \begin{vmatrix} 2 & 0 \\ 3 & -2 \end{vmatrix} = -4, \quad q(x) = \begin{vmatrix} 2 & 1 \\ 3 & x \end{vmatrix} = 2\,x - 3,$$

giving $\gcd(f(x), g(x)) = -4f(x) + (2\,x-3)g_0(x) = (-2\,x-1)f(x) + (2\,x-3)g(x) = -5(x+3)$.

2.2 The Bézoutiant

If, instead of the remainder of $x^k g(x)$, one takes for $g_k(x)$ the remainder of $(a_0 x^k + \cdots + a_k)g(x)$, the corresponding coefficient matrix $\mathcal{B}$ is referred to as the Bézoutiant. This name is due to Sylvester [57]. The name Bézoutian is often used as well [31, 33].

Proposition 2.1 *The determinant of the Bézoutiant is equal to*

$$(-1)^{n(n-1)/2} a_0^{n-m} \Re(f, g).$$

Proof. After simple row manipulations, we have a matrix that corresponds to a Bézout matrix with $g_k(x)$ the remainder of $a_0 g(x)$. This gives in the determinant an additional factor a_0^n. $\square$

This means that there is an extra factor in case $n > m$, like for the Cayley resultant. Moreover, the Bézoutiant determinant is not symmetric in its parameters $f(x)$ and $g(x)$. We have, up to a factor -1,

$$\mathcal{B}(f,g) = (a_0 b_0)^{n-m}\mathcal{B}(g,f).$$

To obtain the symmetry, the polynomials are often made of the same degree. If $n > m$, the Bézoutiant is considered for $f(x)$ and $x^{n-m}g(x)$.

An important observation is that the Bézoutiant is symmetric. This is only the case if the first column corresponds to the power product of degree $n-1$, and so on.

Having $n = m$, the coefficients of $g_0(x)$ are

$$b_{0j} = a_0 b_j - b_0 a_j \quad (j = 1, \ldots, n).$$

For $g_1(x)$ we have to compute the remainder of $(a_0 x + a_1)g(x)$, which is the same as the remainder of $xg_0(x) + a_1 g(x)$. The latter has leading coefficient $a_0 b_1 - b_0 a_1 + a_1 b_0 = a_0 b_1$. This means that we have to subtract $b_1 f(x)$ to obtain the remainder. This gives as leading coefficient the coefficient of power product x^{n-1}: $a_0 b_2 - b_0 a_2 + a_1 b_1 - b_1 a_1 = a_0 b_2 - b_0 a_2$. It turns out that each leading coefficient of $g_k(x)$ is of this form:

$$b_{k1} = a_0 b_k - b_0 a_k \quad (k = 0, \ldots, n-1).$$

This shows the symmetry between the first row and the first column. Similarly, one can prove symmetry between coefficients in the second row and column.

Cayley already observed [23, 57] that the coefficients coincide with the coefficients of (1.26). The Cayley resultant can be seen as a different formulation for the Bézoutiant. Expression (1.26) is called the generating function of the Bézoutiant [57].

2.3 Root separation

The Bézout matrix constructed as in (2.3) can be used for solving root separation problems.

Construct the following Laurent expansions [57, 58]:

$$\frac{1}{f(x)} = \sum_{k=n-1}^{\infty} \frac{d_k}{x^{k+1}}, \qquad \frac{g(x)}{f(x)} = Q(x) + \sum_{k=0}^{\infty} \frac{c_k}{x^{k+1}},$$

where $Q(x)$ is the quotient of the division of $g(x)$ by $f(x)$. In case the $n = m$, $Q(x)$ is equal to b_0/a_0. Recurrence formulae can be used to find the coefficients d_k and c_k out of the coefficients of $f(x)$ and $g(x)$. On the other hand, they can be expressed by the Lagrange formulae as symmetric functions in the roots of $f(x)$.

Denote the roots of $f(x)$ by $\lambda_1,\ldots,\lambda_n$. Provided that the roots are distinct, we have

$$c_k=\sum_{j=1}^{n}\frac{\lambda_j^k g(\lambda_j)}{f'(\lambda_j)},\quad (k=0,1,\ldots)$$

and

$$d_k=\sum_{j=1}^{n}\frac{\lambda_j^k}{f'(\lambda_j)}\quad (k=n-1,n,\ldots).$$

We will use these expressions to prove a relation with the Bézout matrix $\mathcal{B}$.

Construct the following Hankel matrices of order n:

$$D=[d_{j+k}]_{j,k=0}^{n-1},\qquad C=[c_{j+k}]_{j,k=0}^{n-1},\tag{2.5}$$

with $d_k=0$ for $k<n-1$.

Theorem 2.3 *The matrices $\mathcal{B}$, C and D are connected by the equality*

$$\mathcal{B}D=C.\tag{2.6}$$

Proof. Substituting the roots of $f(x)$ in the equations (2.1), we obtain

$$g_j(\lambda_j)=b_{k,n-1}+b_{k,n-2}\lambda_j+\cdots+b_{k0}\lambda_j^{n-1}=\lambda_j^k g(\lambda_j)\quad (k=0,\ldots,n-1;j=1,\ldots,n).$$

These equalities can be rewritten into the matrix form:

$$\mathcal{B}\mathcal{V}=\mathcal{V}\mathrm{diag}\,(g(\lambda_1),\ldots,g(\lambda_n)),\tag{2.7}$$

where $\mathcal{V}$ is the Vandermonde matrix

$$\mathcal{V}=\left[\lambda_j^{k-1}\right]_{j,k=1}^{n}.$$

Assuming now that the roots are distinct, we can multiply both sides of the equation (2.7) by the transposition of

$$\left[\frac{\lambda_j^{k-1}}{f'(\lambda_j)}\right]_{j,k=1}^{n},$$

and, using the Lagrange formulae, the claimed result is obtained. So we proved the validity of (2.6) for the case of distinct roots. The proof can be extended to the case of the existence of a multiple root. The existence of a multiple root is equivalent to the vanishing of the discriminant: $D(f)=0$. The discriminant is an algebraic expression in the coefficients of $f(x)$. And so is the equation (2.6). Applying Lemma 2.1 (see below), it follows that this equation, being valid for $D(f)\neq 0$, is also valid for $D(f)=0$. $\square$

In the proof we have used the following auxiliary result [34].

Lemma 2.1 *Let K be an integral domain with an infinite number of elements. Let g_1, g_2 and h be polynomials in $K[x_1, \ldots, x_s]$, $h \neq 0$. If $g_1(\bar{x}_1, \ldots, \bar{x}_s) = g_2(\bar{x}_1, \ldots, \bar{x}_s)$ at every point $(\bar{x}_1, \ldots, \bar{x}_s) \in K^s$ where $h(\bar{x}_1, \ldots, \bar{x}_s) \neq 0$, then $g_1 = g_2$.*

Proof. Consider the polynomial $(g_1 - g_2)h$. This polynomial is zero everywhere because whenever $h \neq 0$, one has $g_1 = g_2$. As K has an infinite number of elements, this means that $(g_1 - g_2)h$ is identically zero (the zero element in $K[x_1, \ldots, x_s]$) [34]. As h is not identically zero, and $K[x_1, \ldots, x_s]$ is an integral domain, this implies $g_1 = g_2$. $\square$

Lemma 2.2 *The leading principal minors of the matrices $\mathcal{B}$ and C are related with the equality*

$$\det \mathcal{B}_j = a_0^j \det C_j.$$

Proof. This follows from equation (2.6) and the fact that the 'Hankel diagonal' elements of the matrix D, d_{n-1}, are equal to $1/a_0$. $\square$

Let us now assume the coefficients of $f(x)$ and $g(x)$ are real. Let r be the rank of $\mathcal{B}$. We can introduce analogues of the inertia indices and the signature for $\mathcal{B}$ [57]:

$$\begin{cases} n_+(\mathcal{B}) = \mathcal{P}(1, \det \mathcal{B}_1, \ldots, \det \mathcal{B}_r), \\ n_-(\mathcal{B}) = \mathcal{V}(1, \det \mathcal{B}_1, \ldots, \det \mathcal{B}_r), \\ \sigma(\mathcal{B}) = n_+(\mathcal{B}) - n_-(\mathcal{B}), \end{cases}$$

where $\mathcal{P}$ and $\mathcal{V}$ are the number of permanences and the number of variations of sign, respectively. If some of the minors are singular, one may use an analogue of Frobenius' rule [37, 72] for computing the signature. In fact, we still have to correct with the factor a_0^j. Because we are only looking at the signs, replacing $\det \mathcal{B}_j$ with $(-1)^j \det \mathcal{B}_j$, in case $a_0 < 0$, would suffice.

To obtain the number of distinct real roots (nrz), we can choose $g(x) = f'(x)$ and have an analogue with Jacobi's theorem [57]: the signature of the corresponding Bézout matrix gives the number of distinct real roots.

Other analogues can be given. Constructing the Bézout matrices for polynomials $f(x)$ and $f'(x)g(x)$:

$$\mathcal{B}(f, f'g) = \mathcal{B}(f, f')\mathcal{B}(f, g) = \mathcal{B}(f, g)\mathcal{B}(f, f').$$

Then, provided that $\det \mathcal{B}(f, f'g) \neq 0$, one has

$$n_+(\mathcal{B}(f, f'g)) - n_-(\mathcal{B}(f, f')) = \mathrm{nrz}\{f(x) = 0 \mid g(x) > 0\},$$

an analogue of the Hermite–Sylvester theorem [57]. Taking here $g(x) = x - t$, we have a counterpart of the Joachimstahl theorem [57, 88], i.e. a system of polynomials which can be treated as the Sturm sequence for $f(x)$:

Theorem 2.4 *Let the Bézout matrix be constructed for $g(x) = f'(x)$, and r be its rank. Then*

$$\mathrm{nrz}\{f(x) = 0 \mid a < x < b\} = \mathcal{V}(1, q_1(a), \ldots, q_r(a)) - \mathcal{V}(1, q_1(b), \ldots, q_r(b)).$$

Here, $q_k(x)$ is defined as in Theorem 2.2:

$$q_k(x) = \begin{vmatrix} b_{00} & \ldots & b_{0,k-1} & 1 \\ b_{1,0} & \ldots & b_{1,k-1} & x \\ \ldots & & & \ldots \\ b_{k,0} & \ldots & b_{k,k-1} & x^k \end{vmatrix}$$

and $q_n(x) = f(x) \det \mathcal{B}/a_0$.

We apply the results for determining the minimum and maximum of a polynomial [87]:

Example 2.3 *Determine the minimum of*

$$f(x) = 3\,x^4 + 4\,x^3 + 6\,x^2 - 36\,x + 22.$$

We construct the Bézout matrix of $f(x) - z$ and $f'(x)$ and find as determinant

$$\det \mathcal{B} = -20736\,(z+1)\left(z^2 - 102\,z + 4649\right).$$

The minimum of its real roots is the minimum of $f(x)$, so the minimum of $f(x)$ is -1. The signature of the Bézout matrix after substituting $z = 0$ equals 2 and shows that $f(x)$ has two distinct real roots. These are approximately 0.827 and 1.161. The minimum is reached in between these two values. Applying Theorem 2.4, the Bézout matrix after substituting $z = -1$, can be used to check whether for a specific interval the minimum is reached. In fact the minimum is reached for $x = 1$, the only real root of $f'(x)$.

For higher degree polynomials or for polynomials with parametric coefficients the results become of more practical interest.

Chapter 3

Constructing the Bézout resultant: determinantal approach

In this chapter we expose the Bézout resultant in the multivariate case. We give a constructive version of the Bézout reduction process, which is the essential link in building the resultant. This fills the void in the work of Laurent [58]. The reduction process for the bivariate, trivariate and m-variate case are discussed separately. We show there is a fundamental relation between the conditions following from the reduction process and the Macaulay matrix of the leading forms. The last section compares the Bézout resultant to Macaulay's.

3.1 The Bézout resultant in the multivariate case

In the definition of the Bézout resultant for two univariate polynomials $f(x)$ and $g(x)$, an a-symmetry was introduced: the order of the Bézout matrix is equal to the degree of $f(x)$. A sequence of remainders of multiples of $g(x)$ on dividing by $f(x)$ were taken and represented by its square coefficient matrix. In the generalization the remainder of multiples of one polynomial on dividing by a set of polynomials is used.

Let a given system consists of $m+1$ polynomial equations in m indeterminates:

$$\begin{aligned} &f_1 = 0, \ldots, f_m = 0, f_{m+1} = 0, \\ &f_1, \ldots, f_m, f_{m+1} \in K[x_1, \ldots, x_m], \end{aligned} \tag{3.1}$$

K is the coefficient field. We denote by n_j the degree of f_j in the indeterminates $x_1, \ldots, x_m$. The indeterminates are denoted by X.

Our assumption is that the polynomials are generic, i.e. dense and with algebraically independent coefficients. This means that the Bézout bound on the subsystem given by $f_1, \ldots, f_m$ is exact. In other words, there exist $N = n_1 \cdots n_m$ common zeroes for $f_1, \ldots, f_m$.

Consider the following set of N power products:

$$\mathcal{T} = \{\tau_j(X)\}_{j=1}^{N} = \{x_1^{p_1} \cdots x_m^{p_m} \mid 0 \leq p_1 < n_1, \ldots, 0 \leq p_m < n_m\}. \tag{3.2}$$

We call a polynomial $h(X) \in K[X]$ reduced with respect to $\mathcal{T}$ if it can be represented as a linear combination, with coefficients in K, of power products from $\mathcal{T}$. Bézout proved [12, 65, 80] that for generic $f_1, \ldots, f_m$ it is possible to reduce any $h(X) \in K[X]$ modulo $f_1, \ldots, f_m$, i.e. to find $a_1(X), \ldots, a_m(X) \in K[X]$ such that the polynomial

$$h_{red}(X) = h(X) - a_1(X)f_1(X) - \cdots - a_m(X)f_m(X)$$

is reduced with repect to $\mathcal{T}$. We will denote this by

$$h(X) \to_{f_1,\ldots,f_m}^{\mathcal{T}} h_{red}(X) \quad \text{or simply by } h(X) \to^{\mathcal{T}} h_{red}(X).$$

We call $h_{red}(X)$ the remainder of $h(X)$ on dividing by $f_1, \ldots, f_m$. The reducibility problem will be discussed in subsequent sections, starting with the bivariate case.

We can now give the analogue of the sequence (2.1). For $k = 1, \ldots, N$, define $g_k(X)$ as the remainder of $\tau_k(X)f_{m+1}(X)$ on dividing by $f_1, \ldots, f_m$:

$$\tau_k(X)f_{m+1}(X) \to^{\mathcal{T}} g_k(X).$$

This means there exist $\{a_{kj}(X)\}_{1\leq j\leq m}$, polynomials from $K[X]$, such that

$$g_k(X) = \tau_k(X)f_{m+1}(X) - a_{k1}(X)f_1(X) - \cdots - a_{km}(X)f_m(X). \tag{3.3}$$

As $g_k(X)$ is reduced, there exist $\{b_{kj}\}_{1\leq j\leq N}$, coefficients from K, such that

$$g_k(X) = b_{k1}\tau_1(X) + \cdots + b_{k,N}\tau_N(X). \tag{3.4}$$

We have constructed now N equations (3.4). From these equations, viewing the power products $\tau_1, \ldots, \tau_N$ as indeterminates, we can extract the $N \times N$ coefficient matrix

$$\mathcal{B} = [b_{kj}]_{k,j=1}^{N}. \tag{3.5}$$

This matrix will be referred to as the Bézout matrix.

To demonstrate the link of the Bézout matrix with the resultant, we need some auxiliary results. Consider the set of common zeroes of $f_1, \ldots, f_m$. For every common zero Λ of $f_1, \ldots, f_m$ it is evident, by using (3.3), that $g_k(\Lambda) = \tau_k(\Lambda)f_{m+1}(\Lambda)$. For generic polynomials $f_1, \ldots, f_m$ we have N common zeroes (exact Bézout bound). We denote these by $\Lambda_1, \ldots, \Lambda_N$. Substituting $X = \Lambda_j$ in (3.3), we obtain

$$b_{k1}\tau_1(\Lambda_j) + \cdots + b_{k,N}\tau_N(\Lambda_j) = \tau_k(\Lambda_j)f_{m+1}(\Lambda_j) \quad (1 \leq k \leq N). \tag{3.6}$$

Rewriting these equalities for $j = 1, \ldots, N$ in matrix form gives

$$\mathcal{B}\mathcal{V} = \mathcal{V} \begin{pmatrix} f_{m+1}(\Lambda_1) & & \mathbf{O} \\ & \ddots & \\ \mathbf{O} & & f_{m+1}(\Lambda_N) \end{pmatrix}, \tag{3.7}$$

where $\mathcal{V}$ is the generalized Vandermonde matrix:

$$\mathcal{V} = \begin{pmatrix} \tau_1(\Lambda_1) & \cdots & \tau_1(\Lambda_N) \\ \vdots & & \vdots \\ \tau_N(\Lambda_1) & \cdots & \tau_N(\Lambda_N) \end{pmatrix}. \tag{3.8}$$

The structure of the generalized Vandermonde matrix has been discussed, for instance by Macaulay [60]. It can be proved that

$$(\det \mathcal{V})^2 = \Upsilon \mathcal{J}(\Lambda_1) \cdots \mathcal{J}(\Lambda_N),$$

where $\mathcal{J}$ is the Jacobian of $f_1, \ldots, f_m$. The factor Υ depends on the leading forms of $f_1, \ldots, f_m$:

$$\Upsilon = \frac{\Omega}{\mathcal{A}^{n_1+\cdots+n_m-m+1}},$$

where Ω is a product of $n_1 + \cdots + n_m - m$ polynomials in the coefficients of the leading forms of $f_1, \ldots, f_m$ and $\mathcal{A}$ is the resultant of the leading forms, defined as in (1.15). Thus, Υ is a rational function in the coefficients of the leading forms of $f_1, \ldots, f_m$. We will give more attention to the structure of Υ when dealing with the reducibility problem. Note that $(\det \mathcal{V})^2$ is a symmetric polynomial in the common zeroes of $f_1, \ldots, f_m$.

We are now ready to give this theorem, due to Laurent [58].

Theorem 3.1 *The determinant of the Bézout matrix equals the Poisson product* (1.17)*:*

$$\det \mathcal{B} = \Re^{f_{m+1}}(f_1, \ldots, f_m).$$

Moreover, if $\det \mathcal{V} \neq 0$*, the rank of* $\mathcal{B}$ *equals the number of common zeroes of* $f_1, \ldots, f_m$ *for which* $f_{m+1} \neq 0$*.*

Proof. If $\det \mathcal{V} \neq 0$, both statements follow from the relation (3.7). For the second statement one also needs to know the fact that the rank of a matrix is unchanged after multiplication by a non-singular square matrix. To prove that the first statement is also valid for $\det \mathcal{V} = 0$, we can apply Lemma 2.1, choosing $h = (\det \mathcal{V})^2$, $g_1 = \det \mathcal{B}$ and $g_2 = \Re^{f_{m+1}}(f_1, \ldots, f_m)$, as these three are all algebraic with regard to the coefficients of $f_1, \ldots, f_m, f_{m+1}$. □

From the construction of the Bézout resultant it immediately follows that:

Theorem 3.2 *There exist polynomials $P_1, \ldots, P_m, P_{m+1}$, elements in $K[X]$, giving a linear representation for $\Re^{f_{m+1}}(f_1, \ldots, f_m)$:*

$$P_1(X)f_1(X)+\cdots+P_m(X)f_m(X)+P_{m+1}(X)f_{m+1}(X) = \Re^{f_{m+1}}(f_1, \ldots, f_m). \quad (3.9)$$

Proof. Assume the Bézout matrix has been constructed with $\tau_1(X) = 1$. For $k = 1, \ldots, N$, multiply equatilities (3.3) by the cofactor $\mathcal{B}_{k1}$, i.e. $(-1)^{i+j}$ times the determinant of the minor of $\mathcal{B}$ deleting row k and column 1, and sum these to

$$P_j(X) = -\mathcal{B}_{11}a_{1j}(X) - \cdots - \mathcal{B}_{N1}a_{Nj}(X), \quad (j = 1, \ldots, m)$$

$$P_{m+1}(X) = \mathcal{B}_{11}\tau_1(X) + \cdots + \mathcal{B}_{N1}\tau_N(X).$$

By the determinant expansion property we find the desired equation. □

Note the similarity of this result with Theorem 2.2. Equation (3.9) is called the Bézout identity or Bézout equation [10, 11, 60]. Normally the resultant (1.16) is used as the right-hand side of (3.9). Missing the additional factor $\mathcal{A}$, the coefficients of the $P_j(X)$, $j = 1, \ldots, m+1$, are rational functions (instead of polynomials) in the coefficients of $f_1, \ldots, f_m, f_{m+1}$.

3.2 Determination of all the solutions by means of the Bézout matrix

As Theorem 3.1 points out, there are solutions for the system $f_1, \ldots, f_m, f_{m+1}$ if the Bézout matrix is singular. Moreover, from the matrix the solutions can be derived.

First, let rank $\mathcal{B} = N - 1$ and let Λ be a common zero. Substituting $X = \Lambda$ in the equations (3.4), and using (3.3), gives

$$0 = b_{k1}\tau_1(\Lambda) + \cdots + b_{k,N}\tau_N(\Lambda) \quad (k = 1, \ldots, N). \quad (3.10)$$

This can be considered as a linear system of equations with $\tau_1(\Lambda), \ldots, \tau_N(\Lambda)$ as its indeterminates. Its rank equals $N - 1$. Assume now, for definiteness, that the cofactor $\mathcal{B}_{N1}$ does not vanish. We obtain the expressions

$$\tau_j(\Lambda) = \frac{\mathcal{B}_{Nj}}{\mathcal{B}_{N1}}, \quad (j = 2, \ldots, N) \quad (3.11)$$

that give, for appropriate values of j, the components of the common zero Λ (in fact as rational expression in the coefficients of $f_1, \ldots, f_m, f_{m+1}$). The expressions for $\tau_j(\Lambda)$ can also be obtained by applying Cramer's rule to the linear system (3.10), like we have done in Example 1.2.

In case the rank of the Bézout matrix is less than $N-1$, the procedure to find the components of the common zero is slightly different. For example, let rank $\mathcal{B} = N-2$ and let the order $N-2$ minor

$$\mathcal{B}\left[\begin{array}{cc} N-1 & N \\ 1 & 2 \end{array}\right]$$

be non-singular. The upper indices are the rows and the lower indices are the columns to be deleted from $\mathcal{B}$. The first $N-2$ equations from (3.10), considered as linear system with respect to $\tau_3(\Lambda), \ldots, \tau_N(\Lambda)$, can be solved by Cramer's rule. For instance,

$$\tau_3(\Lambda) = -\left(\mathcal{B}\left[\begin{array}{cc} N-1 & N \\ 2 & 3 \end{array}\right]\tau_1(\Lambda) + \mathcal{B}\left[\begin{array}{cc} N-1 & N \\ 1 & 3 \end{array}\right]\tau_2(\Lambda)\right)/\mathcal{B}\left[\begin{array}{cc} N-1 & N \\ 1 & 2 \end{array}\right]. \tag{3.12}$$

So $\tau_3(\Lambda), \ldots, \tau_N(\Lambda)$ are expressed as linear combinations of $\tau_1(\Lambda)$ and $\tau_2(\Lambda)$. As we can choose $\tau_1 = 1$, only $\tau_2(\Lambda)$ remains to be determined. As $\tau_2(\Lambda)$ can not be determined from the matrix, the non-linear relations between τ_2 and $\tau_3, \ldots, \tau_N$ may be used. If for instance $\tau_2 = x$, $\tau_3 = y$ and $\tau_4 = xy$, we have $\tau_4 = \tau_2\tau_3$. After replacing τ_4 and τ_3 by the linear combinations in 1 and τ_2, a degree two polynomial in τ_2 remains.

We will give an example of these in the next section.

3.3 Reducibility in the bivariate case

In this section we will show how the reducibility can be performed in the bivariate case. We start out with an example and illustrate the results of previous sections. The reduction algorithm is formulated followed by theorems on the exact conditions under which the algorithm works.

3.3.1 Example workout

Example 3.1 *Find the Bézout resultant for the polynomial system*

$$f_1 = 5\,x^2 - 6\,xy + 5\,y^2 - 16, \quad f_2 = 2\,x^2 - xy + y^2 - x - y - 4, \quad g = x^2 - y + \epsilon \tag{3.13}$$

with ϵ considered as a parameter.

First, we can confirm that f_1, f_2 satisfy the exact Bézout bound, by evaluating the resultant of its leading forms:

$$\mathcal{A} = \Re(5\,x^2 - 6\,x + 5, 2\,x^2 - x + 1) = 1.$$

As $n_1 = 2$ and $n_2 = 2$, we have, according to (3.2), $\mathcal{T} = \{1, x, y, xy\}$. Hence, we have to reduce the polynomials $g, x\,g, y\,g$ and $xy\,g$ with regard to $\mathcal{T}$. In order to do this, it would suffice to have the reductions of the power products $x^2, y^2, x^3, x^2y, xy^2$ and x^3y, as these are the power products appearing in the polynomials $g, x\,g, y\,g$ and $xy\,g$ that are not contained in $\mathcal{T}$.

The reductions for x^2 and y^2 can be obtained by considering $f_1 = 0$ and $f_2 = 0$ as linear equations in x^2 and y^2. Resolving this linear system gives the equalities

$$x^2 = \left[-\frac{1}{5}f_1 + f_2\right] - \frac{1}{5}xy + x + y + \frac{4}{5}, \tag{3.14}$$

$$y^2 = \left[\frac{2}{5}f_1 - f_2\right] + \frac{7}{5}xy - x - y + \frac{12}{5}. \tag{3.15}$$

Multiplying then (3.14) by y and (3.15) by x, we obtain a linear system in x^2y and xy^2. Resolving it and afterwards replacing x^2 and y^2 by (3.14) and (3.15) gives

$$x^2y = \left[\frac{1}{32}(-2\,x - 5\,y + 9)f_1 + \frac{5}{32}(x + 5\,y - 4)f_2\right] + 2\,xy - x + 2, \tag{3.16}$$

$$xy^2 = \left[\frac{1}{32}(10\,x - 7\,y + 19)f_1 + \frac{5}{32}(-5\,x + 7\,y - 12)f_2\right] + 2\,xy - y + 2. \tag{3.17}$$

Multiplying (3.14) by x and substituting successively (3.16), (3.14) and (3.15), we obtain the expression for x^3:

$$x^3 = \left[\frac{1}{160}(-30\,x + 5\,y - 41)f_1 + \frac{1}{32}(31\,x - 5\,y + 36)f_2\right] + \frac{2}{5}xy + 2\,x + y + \frac{2}{5}.$$

Finally, multiply the derived reduction for x^3 by y, replace xy^2 by formulae (3.17), and then y^2 by (3.15) (the replacement should not be applied to the multipliers of f_1 or f_2 containing y^2):

$$\begin{aligned} x^3y \;=\; & \left[\frac{1}{160}(5\,y^2 - 30\,xy + 20\,x - 55\,y + 102)f_1 \right.\\ & \left. +\frac{1}{32}(-5\,y^2 + 31\,xy - 10\,x + 50\,y - 56)f_2\right] + \frac{21}{5}xy - x - y + \frac{16}{5}. \end{aligned}$$

If one multiplies (3.16) by x, replaces x^2y by (3.16) and afterwards x^2 by (3.14) (the replacement should not be applied to the multipliers of f_1 or f_2 containing x^2), one will obtain another formula for x^3y:

$$\begin{aligned} x^3y \;=\; & \left[\frac{1}{160}(-10\,x^2 - 25\,xy + 25\,x - 50\,y + 122)f_1 \right.\\ & \left. +\frac{1}{32}(5\,x^2 + 25\,xy - 10\,x + 50\,y - 72)f_2\right] + \frac{21}{5}xy - x - y + \frac{16}{5}. \end{aligned}$$

With the obtained formulae we can easily reduce $g, x\,g, y\,g$ and $xy\,g$ modulo f_1 and f_2:

$$g = \left[-\frac{1}{5}f_1 + f_2\right] - \frac{1}{5}xy + x + \varepsilon + \frac{4}{5},$$

$$
\begin{aligned}
xg &= \frac{1}{32}\left[\frac{1}{5}(-30\,x+5\,y-41)f_1+(31\,x-5\,y+36)f_2\right] \\
&\quad -\frac{3}{5}xy+(\varepsilon+2)x+y+\frac{2}{5}, \\
yg &= \frac{1}{32}\left[\frac{1}{5}(-10\,x-25\,y-19)f_1+(5\,x+25\,y+12)f_2\right] \\
&\quad +\frac{3}{5}xy+(\varepsilon+1)y-\frac{2}{5}, \qquad (3.18)\\
xyg &= \frac{1}{32}\left[\frac{1}{5}(5\,y^2-30\,xy-30\,x-20\,y+7)f_1\right. \\
&\quad \left.+(-5\,y^2+31\,xy+15\,x+15\,y+4)f_2\right]+\left(\varepsilon+\frac{11}{5}\right)xy-x+\frac{6}{5}.
\end{aligned}
$$

Out of these we now extract the coefficient matrix $\mathcal{B}$ with respect to the power products $1, x, y, xy$, and compute its determinant:

$$
\Re^g(f_1,f_2)=\det\mathcal{B}=\begin{vmatrix} \varepsilon+4/5 & 1 & 0 & -1/5 \\ 2/5 & \varepsilon+2 & 1 & -3/5 \\ -2/5 & 0 & \varepsilon+1 & 3/5 \\ 6/5 & -1 & 0 & \varepsilon+11/5 \end{vmatrix}=\varepsilon(\varepsilon+2)^3.
$$

As the system given by f_1 and f_2 has the four solutions $(x,y)=(-1,1)$, $(2,2)$ and $(1,-1)$ (multiplicity 2), we can easily verify that the Poisson product (1.17) equals the determinant of the Bézout matrix:

$$
\Re^g(f_1,f_2)=g(-1,1)g(2,2)g(1,-1)^2=\varepsilon(\varepsilon+2)^3.
$$

To complete this example we illustrate the results from previous sections to find the Bézout identity and all the components of the solutions. To determine the Bézout identity we need the cofactors

$$
\mathcal{B}_{11}=\frac{1}{5}(\varepsilon+2)(5\,\varepsilon^2+16\,\varepsilon+8), \quad \mathcal{B}_{21}=-(\varepsilon+2)(\varepsilon+1),
$$

$$
\mathcal{B}_{31}=\epsilon+2, \quad \mathcal{B}_{41}=\frac{1}{5}(\varepsilon+2)(\varepsilon-2)
$$

and sum the products with (3.18). For $\varepsilon=-1$ this gives

$$
P_1(x,y)=\frac{1}{160}(3\,y^2-18\,xy-8\,x+13\,y+4),
$$

$$
P_2(x,y)=\frac{1}{160}(-15\,y^2+93\,xy+20\,x-80\,y+48),
$$

$$
P_3(x,y)=\frac{1}{5}(-3\,xy+5\,y-3) \text{ with } P_1f_1+P_2f_2+P_3(x^2-y^2-1)=-1.
$$

To find the solutions that exist when ε is equal to 0 or -2, we inspect the rank of $\mathcal{B}$. For $\varepsilon = 0$, rank $\mathcal{B} = 3$ and we apply equation (3.11):

$$x = \frac{\mathcal{B}_{42}}{\mathcal{B}_{41}} = \frac{(\varepsilon+2)(3\,\varepsilon+2)}{(\varepsilon+2)(\varepsilon-2)} = -1, \quad y = \frac{\mathcal{B}_{43}}{\mathcal{B}_{41}} = -\frac{(\varepsilon+2)(3\,\varepsilon+2)}{(\varepsilon+2)(\varepsilon-2)} = 1.$$

For $\varepsilon = -2$, rank $\mathcal{B} = 2$ and we apply equation (3.12):

$$\begin{vmatrix} \varepsilon+4/5 & -1/5 \\ 2/5 & -3/5 \end{vmatrix} + x \begin{vmatrix} 1 & -1/5 \\ \varepsilon+2 & -3/5 \end{vmatrix} + y \begin{vmatrix} 0 & -1/5 \\ 1 & -3/5 \end{vmatrix} = 0$$

giving the equation $4 - 3\,x + y = 0$. Similarly we can find $6 - 5\,x + xy = 0$. These combine to the equation $3\,x^2 - 9\,x + 6 = 3(x-1)(x-2) = 0$.

3.3.2 The reduction algorithm

We are now ready to formulate how the reduction, in the case of a bivariate generic system, can be performed by solving a succession of linear systems. The coefficients in these linear systems are rationally dependent on the coefficients of the leading forms of f_1 and f_2.

The reduction process can be represented by the following diagram (case $n_1 = n_2 = n$):

$$\begin{array}{l} f_1 = 0 \\ f_2 = 0 \end{array} \Longrightarrow \left\{ \begin{array}{c} x^n \\ y^n \end{array} \right\} \hookrightarrow \left\{ \begin{array}{c} x^n y \\ xy^n \end{array} \right\} \hookrightarrow \left\{ \begin{array}{c} x^n y^2 \\ x^2 y^n \end{array} \right\} \hookrightarrow \cdots \hookrightarrow \left\{ \begin{array}{c} x^n y^{n-1} \\ x^{n-1} y^n \end{array} \right\}.$$

Each pair of power products represent a linear system with respect to these power products. Solving the system gives new reductions. For instance the first pair, representing $f_1 = 0$, $f_2 = 0$ as system of linear equations with respect to x^n and y^n,

$$\left\{ \begin{array}{c} x^n \\ y^n \end{array} \right\} \text{ gives } \left\{ \begin{array}{l} x^n \to^{\mathcal{T}} \cdot x^{n-1}y + \cdots + \cdot xy^{n-1} + \text{ terms of deg} < n \\ y^n \to^{\mathcal{T}} \cdot x^{n-1}y + \cdots + \cdot xy^{n-1} + \text{ terms of deg} < n \end{array} \right. \tag{3.19}$$

as new reductions (under the condition the system is consistent). The dot $\cdot$ before a power product stands for a coefficient in K. All power products at the right-hand sides of the reductions are contained in $\mathcal{T}$.

The next pair is constructed by multiplication of the obtained reductions from the previous pair, by y and x, respectively. This multiplication operation gives the equations

$$\left\{ \begin{array}{l} x^n y = \cdot x^{n-1}y^2 + \cdots + \cdot xy^n + \text{ terms of deg} \leq n \\ xy^n = \cdot x^n y + \cdots + \cdot x^2 y^{n-1} + \text{ terms of deg} \leq n \end{array} \right., \text{ written as } \left\{ \begin{array}{c} x^n y \\ xy^n \end{array} \right\}$$

in the diagram notation where these equations are viewed as linear equations in $x^n y$ and xy^n. Solving the linear system

$$\left\{ \begin{array}{c} x^n y \\ xy^n \end{array} \right\} \text{ gives } \left\{ \begin{array}{l} x^n y \to \cdot x^{n-1}y^2 + \cdots + \cdot x^2 y^{n-1} + \text{ terms of deg} \leq n \\ xy^n \to \cdot x^{n-1}y^2 + \cdots + \cdot x^2 y^{n-1} + \text{ terms of deg} \leq n \end{array} \right. . \tag{3.20}$$

All the terms of degree equal to n are either contained in $\mathcal{T}$ or equal to x^n or y^n. In the latter case, the terms can be further reduced using (3.19) after which all the terms are contained in $\mathcal{T}$. Comparing (3.19) with (3.20), we observe that the number of terms in the leading forms decreases by one, from $n-1$ to $n-2$.

Using a multiplication of the obtained reductions (3.19) by x and y, respectively, gives two new reductions without solving a linear system. For instance, multiplying the reduction for x^n by x gives

$$x^{n+1} \to \cdot x^n y + \cdots + \cdot x^2 y^{n-1} + \text{ terms of deg} \leq n$$

which can be further reduced using the reduction for $x^n y$. After applying the reductions for x^n and y^n, all the terms are contained in $\mathcal{T}$. In the same way the reduction for y^{n+1} can be determined.

After performing n steps, we have reached:

$$\left\{ \begin{array}{l} x^n y^{n-1} \\ x^{n-1} y^n \end{array} \right\} \text{ giving the reductions } \left\{ \begin{array}{l} x^n y^{n-1} \to \text{ terms of deg} < 2\,n-1 \\ x^{n-1} y^n \to \text{ terms of deg} < 2\,n-1 \end{array} \right. . \tag{3.21}$$

We assumed the n linear systems of two equations in two variables to be consistent. This gives n conditions in terms of the coefficients of the leading forms of f_1 and f_2.

All the remaining reductions for power products of degree $2\,n-1$ can be obtained by 'back substitution' in multiples of reductions obtained in previous steps. The remaining reductions for power products of degree less than $2\,n-1$ can be determined in a similar way. Reductions for power products of degree greater than $2\,n-1$ can be directly retrieved out of the reductions for power products of degree $2\,n-1$.

We will give an illustration with $n=4$. We have the sequence

$$\begin{array}{l} f_1 = 0 \\ f_2 = 0 \end{array} \Longrightarrow \left\{ \begin{array}{c} x^4 \\ y^4 \end{array} \right\} \hookrightarrow \left\{ \begin{array}{c} x^4 y \\ x y^4 \end{array} \right\} \hookrightarrow \left\{ \begin{array}{c} x^4 y^2 \\ x^2 y^4 \end{array} \right\} \hookrightarrow \left\{ \begin{array}{c} x^4 y^3 \\ x^3 y^4 \end{array} \right\},$$

with the corresponding pairs of reductions

$$\left\{ \begin{array}{l} x^4 \to \cdot x^3 y + \cdot x^2 y^2 + \cdot x y^3 + \cdots \\ y^4 \to \cdot x^3 y + \cdot x^2 y^2 + \cdot x y^3 + \cdots \end{array} \right\} \hookrightarrow \left\{ \begin{array}{l} x^4 y \to \cdot x^3 y^2 + \cdot x^2 y^3 + \cdots \\ x y^4 \to \cdot x^3 y^2 + \cdot x^2 y^3 + \cdots \end{array} \right\}$$

$$\hookrightarrow \left\{ \begin{array}{l} x^4 y^2 \to \cdot x^3 y^3 + \cdots \\ x^2 y^4 \to \cdot x^3 y^3 + \cdots \end{array} \right\} \hookrightarrow \left\{ \begin{array}{l} x^4 y^3 \to \cdots \\ x^3 y^4 \to \cdots \end{array} \right\},$$

where the terms of lower degree are denoted by the dots. Multiplying the pair of reductions from step three by x and y, respectively, we have

$$\left\{ \begin{array}{l} x^5 y^2 \to \cdot x^4 y^3 + \cdots \\ x^2 y^5 \to \cdot x^3 y^4 + \cdots \end{array} \right. .$$

These can be further reduced using the reductions for $x^4 y^3$ and $x^3 y^4$ (the 'back substitution') derived in the last step. This gives another two reductions for power products of degree seven. Multiplying the pair of reductions from step two by x^2 and y^2, respectively, gives another two reductions, namely for x^6y and xy^6. The remaining reductions for x^7 and y^7 can be retrieved from step one, using multiplication by x^3 and y^3, respectively.

The remaining reductions of the degree six power products, x^6, x^5y, xy^5 and y^6 can be determined from step one using multiplication by x^2 and y^2, and from step two using multiplication by x and y. The remaining reductions of the degree five power products, x^5 and y^5, can be computed from the reductions from step one, using multiplication by x and y. Power products of degree eight can be determined via multiplication of any reduction of a power product of less degree and reducing this further, using the already obtained reductions.

Summarizing, if the power products of degree $n+k-1$ are reducible and the linear system (as showed in the diagram) with respect to x^ny^k and x^ky^n is consistent, then all the power products of degree $n+k$ are reducible. For the complete reducibility n steps are necessary, which gives n conditions in terms of the coefficients of the leading forms of f_1 and f_2.

For the case $n_1 < n_2$, a modification of the reduction procedure will be illustrated via an example with $n_1 = 3$ and $n_2 = 5$. This means we have

$$\mathcal{T} = \left\{1, x, y, x^2, xy, y^2, x^2y, xy^2, y^3, x^2y^2, xy^3, y^4, x^2y^3, xy^4, x^2y^4\right\}.$$

Let

$$f_1 = ax^3 + \cdot x^2y + \cdot xy^2 + \cdot y^3 + \cdots.$$

Under the condition $a \neq 0$, the reduction for x^3 is given, as all other terms are contained in $\mathcal{T}$:

$$x^3 \to \frac{\cdot x^2y + \cdot xy^2 + \cdot y^3 + \cdots}{-a}.$$

Multiplying this by y, we have the reduction for x^3y:

$$x^3y \to \cdot x^2y^2 + \cdot xy^3 + \cdot y^4 + \cdots.$$

Multiplication by x gives a reduction for x^4 (terms with x^3y and x^3 can be removed using the just derived reductions). The reduction for x^3y, multiplied by y, gives

$$x^3y^2 \to \cdot x^2y^3 + \cdot xy^4 + \cdot y^5 + \cdots,$$

while multiplied by x we have (after using the reduction for x^3y^2)

$$x^4y \to \cdot x^2y^3 + \cdot xy^4 + \cdot y^5 + \cdots.$$

The reduction of x^3, multiplied by x^2 gives (after removing the terms with x^4y and x^3y^2):

$$x^5 \to \cdot x^2y^3 + \cdot xy^4 + \cdot y^5 + \cdots.$$

We can now remove all the degree five terms from f_2, except x^2y^3, xy^4 and y^5. This gives a linear relation between these. Under the condition that the coefficient of y^5 does not vanish, we have derived the reduction for y^5. Hence, the term with y^5 in the reductions of x^3y^2 and x^4y can be removed. So we have obtained all the reductions for power products up to degree five. The next part is similar to the case $n_1 = n_2 = 5$, although we now start with the pair

$$\left\{ \begin{array}{c} x^3y^2 \\ y^5 \end{array} \right\}, \text{ instead of } \left\{ \begin{array}{c} x^5 \\ y^5 \end{array} \right\}.$$

Only three steps more (instead of five) are needed in the reduction procedure.

Thus, for the complete reducibility in the general case when $n_1 < n_2$, n_1 conditions have to be fulfilled, together with the condition that the leading coefficient of f_1 does not vanish.

In the following example the first condition is not fulfilled. We are forced to use another definition for $\mathcal{T}$, which we will denote by $\mathcal{T}_1$, in order to have complete reducibility.

Example 3.2 *Find the reductions in case* $f_1 = x^2 + xy + y^2 + \cdots$ *and* $f_2 = 5\,x^2 - 6\,xy + 5\,y^2 + \cdots$.

It is impossible to find reducers for x^2 and y^2 simultaneously. After doing some of the usual manipulations, we can find out that $\mathcal{T}_1 = \{1, x, y, y^2\}$ gives no problems for establishing complete reducibility. In general we may use

$$\mathcal{T}_1 = \{x^p y^q \mid 0 \le p < n_1, 0 \le q \le n_1 + n_2 - 2\,p - 2\}\,, \tag{3.22}$$

containing $n_1 n_2$ elements as well.

However, the set $\mathcal{T}_1$ does not help in every situation, as we can illustrate by the following example.

Example 3.3 *Find the Bézout matrix for the system given by*

$$\begin{array}{rcl} f_1 &=& 3\,x^2 - 4\,xy + y^2 - 2\,xz - x + yz + 2\,y + z^2 + z - 2, \\ f_2 &=& -6\,x^2 - xy + y^2 - xz + x - yz - y + 2\,z + 2\,z^2 + 2, \\ f_3 &=& 3\,x^2 + 2\,xy - y^2 - 2\,xz - 3\,x - 2\,yz + 3\,y + z^2 + 2\,z - 1 \end{array} \tag{3.23}$$

considering z *as a parameter.*

We have to reduce multiples of f_3 with respect to $\mathcal{T} = \{1, x, y, xy\}$. In the second step of the reduction algorithm we obtain the system

$$x^2y = \frac{1}{3}\,xy^2 + \cdots, \quad xy^2 = 3\,x^2y + \cdots.$$

Considering this system as linear in x^2y and xy^2, it can not be solved. Using the set $\mathcal{T} = \{1, x, y, y^2\}$ does not help. In the first step we can obtain reducers

$$x^2 = \frac{1}{9}y^2 + \cdots, \quad xy = \frac{1}{3}y^2 + \cdots,$$

where the dots mean terms of degree one in x and y. In the second step we obtain the system, after multiplying these reducers by y and x, respectively,

$$x^2y = \frac{1}{9}y^3 + \cdots, \quad x^2y = \frac{1}{3}xy^2 + \cdots.$$

The y^3 in the first equation has to be solved by the reducer of xy, multiplied by y, as y^3 does not appear in the reducer of x^2 multiplied by x. Substituting $y^3 = 3\,xy^2 + \cdots$ gives a new system

$$x^2y = \frac{1}{3}y^3 + \cdots, \quad x^2y = \frac{1}{3}xy^2 + \cdots$$

which can not be solved. Using the set $\mathcal{T} = \{1, x, x^2, y\}$ does not help either.

The reason that we can not find an appropriate set $\mathcal{T}$ is that the resultant of the leading forms (1.15) vanishes:

$$\mathcal{A} = \Re(3\,x^2 - 4\,xy + y^2|_{y=1}, -6\,x^2 - xy + y^2|_{y=1}) = 0. \tag{3.24}$$

For Example 3.2 the resultant of the leading forms does not vanish, whereas one of its minors does.

3.3.3 Theoretical results

The determinantal conditions for the reducibility can be characterized by the resultant of the leading forms and its minors.

Theorem 3.3 *Let $f_1(x,y)$, $f_2(x,y) \in K[x,y]$ of degree n_1 and n_2, respectively, with $n_1 \le n_2$. Writing their leading forms as*

$$\bar{f}_1 = \sum_{j=0}^{n_1} a_{n_1-j,j}x^{n_1-j}y^j \text{ and } \bar{f}_2 = \sum_{j=0}^{n_2} b_{n_2-j,j}x^{n_2-j}y^j\,,$$

the resultant of $\bar{f}_1$ and $\bar{f}_2$, denoted by $\mathcal{A}$ like in the definition (1.15), can be written as $(-1)^{n_1(n_1-1)/2}\Re(\mathcal{A})$, where $\Re(\mathcal{A})$ is defined as

$$\begin{vmatrix} a_{n_1,0} & a_{n_1-1,1} & & \dots & a_{0,n_1} & 0 & \dots & 0 \\ & & & \dots & & & \dots & \\ 0 & \dots & 0 & 0 & a_{n_1,0} & a_{n_1-1,1} & \dots & a_{0,n_1} \\ 0 & \dots & 0 & b_{n_2,0} & b_{n_2-1,1} & b_{n_2-2,2} & \dots & b_{0,n_2} \\ & \dots & & & \dots & & & \\ b_{n_2,0} & b_{n_2-1,1} & b_{n_2-2,2} & \dots & \dots & b_{0,n_2} & \dots & 0 \end{vmatrix} \begin{array}{l} \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} n_2 \text{ rows} \\ \left.\vphantom{\begin{matrix}a\\a\\a\end{matrix}}\right\} n_1 \text{ rows} \end{array}.$$

Denote the k-th subresultant of $\Re(\mathcal{A})$ by $\Re_k(\mathcal{A})$. The system of inequalities

$$\mathcal{A} \ne 0, \Re_1(\mathcal{A}) \ne 0, \dots, \Re_{n_1-1}(\mathcal{A}) \ne 0, a_{n_1,0} \ne 0$$

provides a necessary and sufficient condition for every polynomial $f_3(x,y) \in K[x,y]$ to be reducible modulo f_1 and f_2 with respect to $\mathcal{T}_1$.

Proof. The case $n_1 = n_2 = n$. The algorithm to determine all the reductions with respect to $\mathcal{T}_1$ is as follows. We can write f_1 and f_2 as:

$$f_1 = a_{n,0}x^n + a_{n-1,1}x^{n-1}y + a_{n-2,2}x^{n-2}y^2 + \cdots + a_{0,n}y^n + \cdots,$$
$$f_2 = b_{n,0}x^n + b_{n-1,1}x^{n-1}y + b_{n-2,2}x^{n-2}y^2 + \cdots + b_{0,n}y^n + \cdots.$$

Except the first two terms, all the terms are reduced with respect to $\mathcal{T}_1$. The power products x^n and $x^{n-1}y$ can be expressed from the equations $f_1 = 0$, $f_2 = 0$, if the smallest subresultant does not vanish:

$$\Re_{n-1}(\mathcal{A}) = \begin{vmatrix} a_{n,0} & a_{n-1,1} \\ b_{n,0} & b_{n-1,1} \end{vmatrix} \neq 0.$$

If this condition is fulfilled, then every power product of degree n is reducible with respect to $\mathcal{T}_1$.

Note that in case we would work with $\mathcal{T}$, we have to express the power products x^n and y^n, giving a different condition:

$$\begin{vmatrix} a_{n,0} & a_{0,n} \\ b_{n,0} & b_{0,n} \end{vmatrix} \neq 0.$$

Assume now that all power products of degree less than or equal $n+k-1$ ($1 \leq k \leq n-1$) are reducible. Consider the system of equations given by multiplying f_1 and f_2 by the power products of degree k:

$$x^k f_1 = 0, x^{k-1}yf_1 = 0, \ldots, y^k f_1 = 0, x^k f_2 = 0, x^{k-1}yf_2 = 0, \ldots, y^k f_2 = 0$$

as a linear system with respect to the power products $x^{n+k}, x^{n+k}y, \ldots, x^{n-k-1}y^{2k+1}$. All other power products in these equalities are either reducible, or contained in $\mathcal{T}_1$. The coefficient matrix of the linear system coincides with the subresultant $\Re_{n-k-1}(\mathcal{A})$. In order to have reducibility for the power products of degree $n+k$, the matrix should be non-singular. For $k = n-1$, we find the condition $\Re_0(\mathcal{A}) = \mathcal{A} \neq 0$. Power products of degree greater than or equal $2\,n$ are reducible without additional assumptions. Hence, for complete reducibility the conditions $\mathcal{A} \neq 0, \Re_1(\mathcal{A}) \neq 0, \ldots, \Re_{n-1}(\mathcal{A}) \neq 0$ have to be fullfilled.

The case $n_1 < n_2$ is treated similarly, with a preprocess that can only be performed under the condition $a_{n,0} \neq 0$. The power products

$$x^{n_1}, x^{n_1}y, x^{n_1+1}, \ldots, x^{n_1}y^{n_2-n_1-1}, x^{n_1+1}y^{n_2-n_1-2}, \ldots, x^{n_2-1}$$

can be successively reduced using the equations

$$f_1 = 0, yf_1 = 0, xf_1 = 0, \ldots, y^{n_2-n_1-1}f_1 = 0, xy^{n_2-n_1-2}f_1 = 0, \ldots, x^{n_2-n_1-1}f_1 = 0$$

under the condition $a_{n_1,0} \neq 0$. The remaining power products of degree less than n_2 are contained in $\mathcal{T}_1$.

To reduce the power products of degree n_2 not contained in $\mathcal{T}_1$ – these are $x^{n_2}, x^{n_2-1}y, \ldots, x^{n_1-1}y^{n_2-n_1+1}$ – consider the equations

$$x^{n_2-n_1}f_1 = 0, x^{n_2-n_1-1}yf_1 = 0, \ldots, y^{n_2-n_1}f_1 = 0, f_2 = 0$$

as a linear system in these power products. The corresponding determinant coincides with $\Re_{n_1-1}(\mathcal{A})$. Hence, the second condition obtained is $\Re_{n_1-1}(\mathcal{A}) \neq 0$. For the power products of degree $n_2 + k$, $k \leq n_1 - 1$, the condition $\Re_{n_1-k-1}(\mathcal{A}) \neq 0$ can be derived as it coincides with the determinant of the equations

$$x^{n_2-n_1+k}f_1 = 0, x^{n_2-n_1+k-1}yf_1 = 0, \ldots, y^{n_2-n_1+k}f_1 = 0,$$

$$y^k f_2 = 0, xy^{k-1}f_2 = 0, \ldots, x^k f_2 = 0$$

as a linear system with respect to the power products

$$x^{n_2+k}, x^{n_2+k-1}y, \ldots, x^{n_1-k-1}y^{n_2-n_1+2k+1}.$$

For the reduction of power products of degree higher than n_1+n_2-1 the procedure is the same as for the case $n_1 = n_2$. □

The same subresultants appear in the relation between the Vandermonde matrix and the Jacobian.

Theorem 3.4 *Let $\mathcal{A} \neq 0$ and let $\mathcal{V}$ be the generalized Vandermonde matrix defined as in* (3.8)*. We have the equality*

$$(\det \mathcal{V})^2 = (-1)^{n_1(n_1-1)/2} \frac{[\Re_1(\mathcal{A}) \cdots \Re_{n_1-1}(\mathcal{A})]^2}{\mathcal{A}^{n_1+n_2-1}} \mathcal{J}(\Lambda_1) \cdots \mathcal{J}(\Lambda_{n_1 n_2}),$$

where $\mathcal{J}$ is the Jacobian of f_1 and f_2 and $\Lambda_1, \ldots, \Lambda_{n_1 n_2}$ their common zeroes.

For an idea of the proof we refer to [86, 87].

In Theorem 3.3 we used definition (3.22) for $\mathcal{T}$ because of the nice relation with the subresultants. If we would use $\mathcal{T}$ defined as in (3.2), the statements are similar, with this difference that instead of the normal subresultant we have to use different minors of $\Re(\mathcal{A})$, namely

$$\Re_k(\mathcal{A}) = \Re(\mathcal{A}) \left[\begin{array}{ccccccc} 1 & \ldots & k & n_1+n_2-k+1 & n_1+n_2-k+2 & \ldots & n_1+n_2 \\ 1 & \ldots & k & n_2+1 & n_2+2 & \ldots & n_2+k \end{array} \right],$$

where the upper (lower) indices are the rows (columns) to be deleted. For example, in case $n_1 = n_2 = 3$ this gives the minors:

$$\Re_2(\mathcal{A}) = \begin{vmatrix} a_{30} & a_{03} \\ b_{30} & b_{03} \end{vmatrix} \quad \text{and} \quad \Re_1(\mathcal{A}) = \begin{vmatrix} a_{30} & a_{21} & a_{03} & 0 \\ 0 & a_{30} & a_{12} & a_{03} \\ 0 & b_{30} & b_{12} & b_{03} \\ b_{30} & b_{21} & b_{03} & 0 \end{vmatrix}. \tag{3.25}$$

The way we have derived the conditions in terms of subresultants or minors of $\Re(\mathcal{A})$ differs from the conditions following directly from the sketched algorithm to determine all the reductions of power products. For instance, in case $n_1 = n_2 = 3$ the second condition – for the first condition, with the smallest minor, there is no difference – we face with is

$$\Re_1^*(\mathcal{A}) = \begin{vmatrix} \begin{vmatrix} a_{30} & a_{03} \\ b_{30} & b_{03} \end{vmatrix} & \begin{vmatrix} a_{12} & a_{03} \\ b_{12} & b_{03} \end{vmatrix} \\ & \\ \begin{vmatrix} a_{30} & a_{21} \\ b_{30} & b_{21} \end{vmatrix} & \begin{vmatrix} a_{30} & a_{03} \\ b_{30} & b_{03} \end{vmatrix} \end{vmatrix}.$$

Its value is equal to the one from (3.25). The last condition following from the reduction algorithm, $\Re_0^*(\mathcal{A})$ (in Theorem 3.3 this is just $\mathcal{A}$), is equal to

$$\mathcal{A} \begin{vmatrix} a_{30} & a_{03} \\ b_{30} & b_{03} \end{vmatrix}.$$

This additional factor is already contained in the first condition. For higher degree cases we have the extra factors

$$\Re_k^*(\mathcal{A}) = \Re_k(\mathcal{A}) \prod_{j=k+2}^{n-1} \Re_j(\mathcal{A}) \quad (0 \leq k \leq n-3).$$

So the bivariate reducibility conditions are formulated in terms of the resultant of the leading forms and its minors. If we do not fix the set $\mathcal{T}$, then non-vanishing of the resultant of the leading forms is a sufficient condition for reducibility [65]:

Theorem 3.5 *If $\mathcal{A} \neq 0$, we can always choose a set of $n_1 n_2$ power products with respect to which any polynomial is reducible.*

Proof. As the determinant of the Sylvester resultant matrix of the leading forms is non-singular, there exists a sequence of non-zero minors with decreasing order. Each minor corresponds to the search for reducers of power products of a certain degree. □

Example 3.4 *Find a proper set $\mathcal{T}$ in case $f_1 = x^3 + 2\,x^2y + 7\,xy^2 + 3\,y^3 + \cdots$ and $f_2 = x^3 + 2\,x^2y + 11\,xy^2 + 3\,y^3 + \cdots$.*

Obviously, choosing $\mathcal{T}$ as in (3.2) or (3.22) does not work. To start with, we can find reducers for x^3 and xy^2:

$$x^3 \to -2\,x^2y - 3\,y^3 + \cdots, \quad xy^2 \to \cdots.$$

This means that $\mathcal{T}$ is going to contain x^2y and y^3 as degree three power products. Multiplying the reducer for xy^2 by x and y gives reducers for x^2y^2 and xy^3. Multiplying the reducer for x^3 by x and y gives reducers for x^4 and x^3y:

$$x^4 \to 6\,y^4 + \cdots, \quad x^3y \to -3\,y^4 + \cdots,$$

and we can complete $\mathcal{T}$ with y^4:

$$\mathcal{T} = \{1, x, y, x^2, xy, y^2, x^2y, y^3, y^4\}.$$

Multiplying the reducers of degree four by x and y gives reducers for all power products of degree five.

To end this section, we illustrate how the algorithm to obtain the reducers has to be modified to have it work even in case $\mathcal{A}$ vanishes. The idea is to use the non-trivial representation of $\mathcal{A}$, which exists according to Theorem 3.2.

Example 3.5 *Find the Bézout resultant of the system as given in* (3.23).

Applying Theorem 3.2 to equation (3.24) we can find the representation

$$\mathcal{A} = (2\,x + 1)(3\,x^2 - 4\,x + 1) + (x - 1)(-6\,x^2 - x + 1) = 0.$$

After making the factors homogeneous with y and using them with the corresponding polynomials gives

$$(2\,x+y)f_1+(x-y)f_2 = (-1-5\,t)x^2+xy+(3+2\,t)y^2+(4\,t^2+4\,t-2)x-(t^2+t+4)y.$$

Hence, we have, besides $f_1 = 0$ and $f_2 = 0$, one additional degree two equation. Solving these as a linear system in x^2, xy, y^2:

$$\begin{pmatrix} 3 & -4 & 1 \\ -6 & -1 & 1 \\ -1-5\,t & 1 & 3+2\,t \end{pmatrix} \begin{pmatrix} x^2 \\ xy \\ y^2 \end{pmatrix} = -\begin{pmatrix} (-2\,t-1)x + (2+t)y + t^2 + t - 2 \\ (1-t)x - (1+t)y + 2\,t^2 + 2\,t + 2 \\ (4\,t^2 + 4\,t - 2)x - (t^2 + t + 4)y \end{pmatrix}$$

is possible provided the determinant

$$\begin{vmatrix} 3 & -4 & 1 \\ -6 & -1 & 1 \\ -1-5\,t & 1 & 3+2\,t \end{vmatrix} = -3(13\,t + 29)$$

does not vanish. We denote this determinant by $\mathcal{A}^*$. Hence, if $t \neq -29/13$, then x^2, xy and y^2 can be expressed as linear combinations of the power products from $\mathcal{T}^* = \{1, x, y\}$, for instance

$$x^2 = -\frac{9\,y - 23\,x + 26\,ty + 7\,tx + 7\,yt^2 + 16\,xt^2 - 36\,t^2 - 42\,t - 34 - 14\,t^3}{3(13\,t + 29)}.$$

It gives no further difficulties to find reducers for the degree three power products. After reducing f_3, xf_3 and yf_3 we have the 3×3 Bézout matrix $\mathcal{B}$ with

$$\det \mathcal{B} = \frac{2277\,t^7 + 3873\,t^6 - 2752\,t^5 + 4810\,t^4 + 2521\,t^3 - 4407\,t^2 - 2406\,t - 3916}{9(13\,t + 29)}.$$

Like $\mathcal{A}$ was defined to remove the denominator from the pure Poisson product, $\mathcal{A}^*$ has the same rôle in this example. In fact, for this example we have that $\mathcal{A}^*$ is the leading coefficient of

$$\Re_x(f_1, f_2) = 9(13\,t + 29)y^3 + \left(177\,t^2 - 312 + 261\,t\right) y^2 \\ + \left(69\,t^2 - 141 - 69\,t^3 - 237\,t\right) y + 99\,t^4 + 189\,t^3 - 153\,t^2 - 171\,t + 36$$

considered as polynomial in y. The number of solutions for $f_1 = 0, f_2 = 0$ is three, provided $\mathcal{A}^* \neq 0$. For $t = -29/13$, the number of solutions drops to two.

Chapter 4 gives a complete solution for obtaining the reducers independent of the value of $\mathcal{A}$. It is shown that the Bézout resultant can be defined properly even in case $\mathcal{A} = 0$.

3.4 Reducibility in the trivariate case

Some essential differences appear when trying to generalize the algorithm to determine the reducers to the trivariate case. This section serves as an illustration for the next section, in which we will show how the m-variate case works.

Let us start with an example of the most simple non-trivial trivariate case: $n_1 = n_2 = n_3 = 2$. For readability we use the indeterminates x, y and z. We use the set $\mathcal{T}$ given by (3.2):

$$\mathcal{T} = \{1, x, y, z, xy, xz, yz, xyz\}.$$

Let the system be given by

$$\begin{array}{rcl} f_1 & = & a_{200}x^2 + a_{110}xy + a_{101}xz + a_{020}y^2 + a_{011}yz + a_{002}z^2 + \cdots, \\ f_2 & = & b_{200}x^2 + b_{110}xy + b_{101}xz + b_{020}y^2 + b_{011}yz + b_{002}z^2 + \cdots, \\ f_3 & = & c_{200}x^2 + c_{110}xy + c_{101}xz + c_{020}y^2 + c_{011}yz + c_{002}z^2 + \cdots, \end{array}$$

where the dots stand for terms of lower degree (here just linear and constant terms). Solving the equations $f_1 = f_2 = f_3 = 0$ as a linear system in the power products x^2, y^2 and z^2 is possible only if the condition

$$\begin{vmatrix} a_{200} & a_{020} & a_{002} \\ b_{200} & b_{020} & b_{002} \\ c_{200} & c_{020} & c_{002} \end{vmatrix} \neq 0$$

is fulfilled. This gives the reductions

$$\begin{array}{lll} x^2 & \to^{\mathcal{T}} & A_{110}xy + A_{101}xz + A_{011}yz + \cdots, \\ y^2 & \to^{\mathcal{T}} & B_{110}xy + B_{101}xz + B_{011}yz + \cdots, \\ z^2 & \to^{\mathcal{T}} & C_{110}xy + C_{101}xz + C_{011}yz + \cdots. \end{array}$$

Multiplying the reducer for x^2 by y and z, the reducer for y^2 by x and z, and the reducer for z^2 by x and y, gives six equations of degree three (displayed as reducers):

$$\begin{aligned}
x^2y &\rightarrow A_{110}xy^2 + A_{101}xyz + A_{011}y^2z + \cdots, \\
x^2z &\rightarrow A_{110}xyz + A_{101}xz^2 + A_{011}yz^2 + \cdots, \\
xy^2 &\rightarrow B_{110}x^2y + B_{101}x^2z + B_{011}xyz + \cdots, \\
xz^2 &\rightarrow C_{110}x^2y + C_{101}x^2z + C_{011}xyz + \cdots, \\
y^2z &\rightarrow B_{110}xyz + B_{101}xz^2 + B_{011}yz^2 + \cdots, \\
yz^2 &\rightarrow C_{110}xy^2 + C_{101}xyz + C_{011}y^2z + \cdots.
\end{aligned}$$

The power product xyz is contained in $\mathcal{T}$. Degree two terms x^2, y^2 and z^2 can be replaced using their previously determined reductions.

Consider these six reducers as linear equations with respect to the power products x^2y, x^2z, xy^2, y^2z, xz^2 and yz^2. This linear system is solvable under the condition

$$\begin{vmatrix}
1 & 0 & -A_{110} & 0 & -A_{011} & 0 \\
0 & 1 & 0 & -A_{101} & 0 & -A_{011} \\
-B_{110} & -B_{101} & 1 & 0 & 0 & 0 \\
-C_{110} & -C_{101} & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & -B_{101} & 1 & -B_{011} \\
0 & 0 & -C_{110} & 0 & -C_{011} & 1
\end{vmatrix} \neq 0, \qquad (3.26)$$

giving six reducers for the power products x^2y, x^2z, xy^2, y^2z, xz^2 and yz^2. Reductions for x^3, y^3 and z^3 can be found now. For instance, for x^3, multiply the reducer for x^2 by x and replace the power products x^2y and x^2z using the just obtained reducers. To derive the reductions for power products of degree four, one extra condition is required. It suffices to have the reductions for x^2yz, xy^2z and xyz^2: multiplying any reducer of degree three by x, y or z gives as degree four power products at the right-hand side x^2yz, xy^2z or xyz^2. The first two can be reduced to the third, using the reductions for x^2y and xy^2, multiplied by z. The third, xyz^2, can be solved by multiplying

$$xy^2 \rightarrow^{\mathcal{T}} Dxyz + \cdots, \quad xz^2 \rightarrow^{\mathcal{T}} Exyz + \cdots$$

by z and y, respectively. This gives a linear system in xy^2z and xyz^2, solvable if the condition

$$\begin{vmatrix} 1 & -D \\ -E & 1 \end{vmatrix} \neq 0$$

is satisfied. Like in the bivariate case, finding reductions for higher degree power products gives no extra conditions.

For the case $n_1 = n_2 = n_3 = n$, we can display the algorithm by the following diagram.

$$\begin{array}{l} f_1 = 0 \\ f_2 = 0 \\ f_3 = 0 \end{array} \Longrightarrow \left\langle \left\{ \begin{array}{c} x^n \\ y^n \\ z^n \end{array} \right\}_3 \hookrightarrow \left\{ \begin{array}{c} x^n y, x^n z \\ xy^n, y^n z \\ xz^n, yz^n \end{array} \right\}_6 \hookrightarrow \left\{ \begin{array}{c} x^n y^2, x^n yz, x^n z^2 \\ y^n x^2, y^n xz, y^n z^2 \\ z^n x^2, z^n xy, z^n y^2 \end{array} \right\}_9 \right.$$

$$\hookrightarrow \cdots \hookrightarrow \left\{ \begin{array}{l} x^n y^{n-1}, x^n y^{n-2} z, \ldots, x^n z^{n-1} \\ y^n x^{n-1}, y^n x^{n-2} z, \ldots, y^n z^{n-1} \\ z^n x^{n-1}, z^n x^{n-2} y, \ldots, z^n y^{n-1} \end{array} \right\}_{3n} \Bigg\rangle \tag{3.27}$$

$$\hookrightarrow \left\langle \left\{ \begin{array}{l} y^n x^{n-1} z, y^n x^{n-2} z^2, \ldots, y^n x z^{n-1} \\ z^n x^{n-1} y, z^n x^{n-2} y^2, \ldots, z^n x y^{n-1} \end{array} \right\}_{2(n-1)} \hookrightarrow \right.$$

$$\left. \left\{ \begin{array}{l} y^n x^{n-1} z^2, y^n x^{n-2} z^3, \ldots, y^n x^2 z^{n-1} \\ z^n x^{n-1} y^2, z^n x^{n-2} y^3, \ldots, z^n x^2 y^{n-1} \end{array} \right\}_{2(n-2)} \hookrightarrow \cdots \hookrightarrow \left\{ \begin{array}{l} x^{n-1} y^n z^{n-1} \\ x^{n-1} y^{n-1} z^n \end{array} \right\}_2 \right\rangle .$$

Here $\{\ \}$ represents the linear system with respect to the power products included. The subscript indicates the number of power products to solve, i.e. the order of the determinant of the corresponding linear system. These systems are collected into two cascades denoted by $\langle\ \rangle$. The sign $\hookrightarrow$ denotes multiplication of the reductions obtained on the previous step by x, y, or z if the system belongs to the first cascade and by z or y if it belongs to the second one. The second cascade can be seen as a combination of $n-1$ diagrams corresponding to a bivariate case with indeterminates y and z (and the x-power as constant).

Theorem 3.6 *If the power products of degree $n+k$ are reducible and the linear system with respect to the indicated power products of degree $n+k+1$ is consistent, then all power products of degree $n+k+1$ are reducible. For total reducibility it is sufficient that $2n-1$ determinants are non-singular.*

Proof. If $k < n-1$, let $k_2 + k_3 = k$, $0 \le k_2, k_3 < n-1$, and consider the power product reduction

$$x^n y^{k_2} z^{k_3} \to^{\mathcal{T}} \sum_{\substack{K_2+K_3=k+1 \\ K_j<n}} \cdot x^{n-1} y^{K_2} z^{K_3} + \sum_{\substack{L_1+L_3=k+1 \\ L_j<n}} \cdot x^{L_1} y^{n-1} z^{L_3}$$

$$+ \sum_{\substack{M_1+M_2=k+1 \\ M_j<n}} \cdot x^{M_1} y^{M_2} z^{n-1} + \sum_{\substack{p_1+p_2+p_3=n+k \\ p_j<n-1}} \cdot x^{p_1} y^{p_2} z^{p_3} + \cdots,$$

where the dots stand for terms of degree less then $n+k$ and the single dot stands for a coefficient. Multiplying this reduction by y and z and applying the reduction formulae deduced earlier, linear relations are obtained between the power products of degree $n+k+1$. These are either contained in $\mathcal{T}$ or in the $k+1$-th system of the diagram. More relations can be obtained by multiplying the reductions for $x^{l_1} y^n z^{l_3}$, $0 \le l_1, l_3 < n-1$, by x and z, and by multiplying the reductions for $x^{m_1} y^{m_2} z^n$, $0 \le m_1, m_2 < n-1$, by x and y. As assumed, the system of these linear relations can be solved with respect to the power products of the $k+1$-th system. With their reductions one can find the reductions for any power product of degree $n+k+1$ via multiplication of the corresponding power product of degree $n+k$ by x, y or z.

For $k = n-1$, the last system of the first cascade, the reduction

$$x^{l_1} y^n z^{l_3} \to^{\mathcal{T}} \sum_{\substack{K_2+K_3=n \\ K_j<n}} \cdot x^{n-1} y^{K_2} z^{K_3} + \sum_{\substack{L_1+L_3=n \\ L_j<n}} \cdot x^{L_1} y^{n-1} z^{L_3} \tag{3.28}$$

$$+ \sum_{\substack{M_1+M_2=n \\ M_j<n}} \cdot x^{M_1} y^{M_2} z^{n-1} + \sum_{\substack{p_1+p_2+p_3=2n-1 \\ p_j<n-1}} \cdot x^{p_1} y^{p_2} z^{p_3} + \cdots$$

where $l_1 + l_3 = n-1$, $0 \leq l_1, l_3 < n-1$, multiplied by z gives linear relations between the power products of degree $2\,n$, contained in either $\mathcal{T}$ or in the $n+1$-th system of the diagram. More relations can be obtained by multiplying by y all the reductions for $x^{m_1} y^{m_2} z^n$, $m_1 + m_2 = n-1$, $0 \leq m_1, m_2 < n-1$. If these relations are solvable, then any power product $x^{p_1} y^{p_2} z^{p_3}$, $p_1 + p_2 + p_3 = 2\,n$, $p_1 < n$, can be reduced via y and z multiples of reductions of lower degree power products. For $p_1 \geq n$ successive reductions have to be used. For instance, multiply (3.28) by x and rewrite the right-hand side:

$$x^{l_1+1} y^n z^{l_3} \to^{\mathcal{T}} \sum_{\substack{K_2+K_3=n \\ K_j<n}} \cdot x^n y^{K_2} z^{K_3} + \cdots = \sum_{\substack{K_2+K_3=n \\ K_j<n}} \cdot \left(x^n y^{K_2-1} z^{K_3}\right) y + \cdots.$$

As we have assumed that any power product of degree $2\,n-1$ is reducible and multiplication of its reduction by y only gives reducible terms, then every power product of degree $2\,n$ is reducible.

This procedure continues in a similar way and terminates when we either meet a degenerate linear system or reach the degree $3\,n-2$. In the latter case, any power product of degree greater than $3\,n-2$ will be reducible because the reductions of the power products only contain the terms of degree $\leq 3\,n-3$. □

To show that the reduction algorithm at least works in the generic case, we can use the system

$$f_1 = x^n + \cdots, f_2 = y^n + \cdots, f_3 = z^n + \cdots, \tag{3.29}$$

where the dots stand for terms of degree less than n. The matrices of the constructed linear system are all identity matrices. The system as given in (3.29) is called a Pham system [45].

Like in the bivariate case we investigate the relation to the Macaulay matrix.

Theorem 3.7 *The first n conditions of the $2\,n-1$ mentioned in the trivariate reduction algorithm can be rewritten in terms of the nested minors of the Macaulay matrix of the leading forms of f_1, f_2 and f_3 for finding the resultant $\mathcal{A}$ defined by equation (1.15). The order of the minors is $\left\{3\binom{k+2}{2}\right\}_{k=0}^{n-1}$.*

Proof. Assume that any monomial of degree less than $n+k$, $0 \le k < n-1$, is reducible with respect to $\mathcal{T}$. Multiply f_1, f_2 and f_3 by all possible power products of degree k and use the reductions obtained in previous steps to reduce all reducible terms. This gives $3\binom{k+2}{2}$ linear equations with the $\binom{n+k+2}{2}$ power products of degree $n+k$ as unknowns. From these, $3\binom{k+2}{2}$ are not contained in $\mathcal{T}$ as there are, for instance, $\binom{k+2}{2}$ power products with a degree in x greater than or equal n.

Hence, the number of equations coincides with the number of power products to be reduced. To resolve this system, the determinant of the corresponding coefficient matrix should be non-zero. This condition is equivalent to the $k+1$-th condition. The coefficient matrix coincides with a minor of the Macaulay matrix of the leading forms as the Macaulay matrix at least consists of the rows corresponding with the multiplications (of the leading forms) by power products up to degree $k < n$ (Definition 1.3), or more precisely the multiplications by power products up to degree $k \le 2n-2$ and with the degree in x, y and z that is less than n. $\square$

For the example worked out above – the case $n=2$ – the condition (3.26) can be replaced by

$$\begin{vmatrix}
a_{200} & a_{110} & a_{101} & a_{020} & a_{002} & 0 & 0 & 0 & 0 \\
0 & a_{200} & 0 & a_{110} & 0 & a_{020} & a_{011} & a_{002} & 0 \\
0 & 0 & a_{200} & 0 & a_{101} & 0 & a_{020} & a_{011} & a_{002} \\
b_{200} & b_{110} & b_{101} & b_{020} & b_{002} & 0 & 0 & 0 & 0 \\
0 & b_{200} & 0 & b_{110} & 0 & b_{020} & b_{011} & b_{002} & 0 \\
0 & 0 & b_{200} & 0 & b_{101} & 0 & b_{020} & b_{011} & b_{002} \\
c_{200} & c_{110} & c_{101} & c_{020} & c_{002} & 0 & 0 & 0 & 0 \\
0 & c_{200} & 0 & c_{110} & 0 & c_{020} & c_{011} & c_{002} & 0 \\
0 & 0 & c_{200} & 0 & c_{101} & 0 & c_{020} & c_{011} & c_{002}
\end{vmatrix} \neq 0.$$

The $n-1$ conditions in the second cascade do not correspond directly to minors of the Macaulay matrix of the leading forms. For example, consider the case $n=3$. At the end of the first cascade the reductions for all degree five power products are determined, except for x^2y^2z, x^2yz^2 and xy^2z^2, which are contained in $\mathcal{T}$. These reductions can be written as

$$\tau \to \sum_{p_1+p_2+p_3 \le 2} x^{p_1}y^{p_2}z^{p_3}(\cdot f_1 + \cdot f_2 + \cdot f_3) + \sum_{\tau' \in \mathcal{T}, \deg \tau' \le 5} \cdot \tau'$$

where the dots stand for coefficients in K. In the generic case none of these coefficients vanishes. In the first step of the second cascade two of these reductions are multiplied by y and two by z. This means that the first condition of the second cascade is based on multiples of f_1, f_2 and f_3 that consist of power products $x^{p_1}y^{p_2+1}z^{p_3}$ and $x^{p_1}y^{p_2}z^{p_3+1}$, $p_1+p_2+p_3 \le 2$. After successive elimination we can write these four reductions of degree six – these are xy^2z^3, xy^3z^2, x^2yz^3 and x^2y^3z – as

$$\tau \to \sum_{p_1+p_2+p_3\leq 2, q_2+q_3\leq 1} x^{p_1}y^{p_2+q_2}z^{p_3+q_3}(\cdot f_1 + \cdot f_2 + \cdot f_3) + \sum_{\tau'\in\mathcal{T}, \deg \tau' \leq 6} \cdot\tau'.$$

Looking at the leading forms of this expression, we have on the right-hand side

$$\sum_{p_1+p_2+p_3=2, q_2+q_3=1} x^{p_1}y^{p_2+q_2}z^{p_3+q_3}(\cdot \bar{f}_1 + \cdot \bar{f}_2 + \cdot \bar{f}_3) + \sum_{\tau'\in\mathcal{T}, \deg \tau' = 6} \cdot\tau'.$$

This means that the y^3 and z^3 multiples of f_3 are used. This is not according to rows of the Macaulay matrix. The Macaulay matrix of the leading forms of f_1, f_2 and f_3, choosing z as homogenizing indeterminate, contains rows corresponding to $x^{p_1}y^{p_2}\bar{f}_3$ with $p_1 < 2, p_2 < 2$. The other rows correspond to multiples of $\bar{f}_1$ and $\bar{f}_2$. Hence, there is no row corresponding to $y^3\bar{f}_3$.

The difficulty in generalizing Theorem 3.5 is that for the trivariate case we do not have a representation of $\mathcal{A}$ as a determinant in the coefficients of the leading forms of f_1, f_2 and f_3. On the other hand, the Macaulay resultant of the leading forms of f_1, f_2 and f_3 is a multiple of $\mathcal{A}$ [60, 65].

Theorem 3.8 *If the Macaulay matrix of the leading forms of f_1, f_2 and f_3 after substituting $z = 1$ is non-singular, there exists a choice of $n_1n_2n_3$ power products from the set $\{x^{p_1}y^{p_2}z^{p_3} \mid p_1 + p_2 + p_3 < n_1 + n_2 + n_3 - 2\}$ with respect to which any polynomial is reducible.*

Proof. If the Macaulay matrix is non-singular, we can derive reductions of power products in x, y and z of degree equal to the critical degree defined as in (1.22). This can be done by homogenizing the sets (1.23) with indeterminate z. The critical degree is equal to $n_1 + n_2 + n_3 - 2$. As the Macaulay matrix is non-singular, $\mathcal{A}$ is non-zero. Hence, there exist exactly $n_1n_2n_3$ linear independent power products modulo f_1, f_2 and f_3 (Bézout bound). □

This proof throws light upon the question how to formulate the conditions of a reduction algorithm in terms of minors of the Macaulay matrix of the leading forms. For example, consider again the case $n_1 = n_2 = n_3 = 3$. The critical degree is seven. The size of the Macaulay matrix of the leading forms is 36. Because of the non-singularity we can derive reductions for all degree seven power products. Taking only the degree three power products from the sets (1.23) we have a minor with 27 rows. After homogenizing with z we have 27 equations in 28 power products of degree six. Hence, at least one power product of degree six is not reducible. This can be $x^2y^2z^2$ in case the appropriate order 27×27 minor is non-singular. Taking only the degree two power products from the sets (1.23) we have a minor with 18 rows. After homogenizing with z we have 18 equations in 21 power products of degree five. Hence, at least three power products of degree five are not reducible. These can coincide with the ones from the presented reduction algorithm, in case an appropriate 18×18 minor is non-singular. This process can be continued with the degree one and degree zero power products from the sets (1.23). This gives minors of order 9 and 3, respectively.

These minors can be the basis for a reformulation of the reduction algorithm.

Theorem 3.9 *The reduction algorithm can be reformulated based on the minors of the Macaulay matrix of the leading forms of f_1, f_2 and f_3, after substituting $z = 1$. The order of the minors, in case $n_1 = n_2 = n_3 = n$, equals $\left\{3\binom{k+2}{2}\right\}_{k=0}^{n-1}$ for the first n minors and $\left\{3\binom{n+k+2}{2} - 3\binom{k+2}{2}\right\}_{k=0}^{n-2}$ for the last $n-1$ minors. The last minor coincides with the Macaulay matrix.*

Proof. Consider the rows of the Macaulay matrix corresponding to power product multipliers of degree k as a linear system with respect to the power products of degree $n+k$. The number of these rows is $3\binom{k+2}{2}$ for $0 \leq k \leq n-1$ and $3\binom{k+2}{2} - 3\binom{k-n+2}{2}$ for $n \leq k \leq 2n-2$. Let $\mathcal{T}$ contain the power products that can not be solved from these systems after homogenizing them with z. The sum of the number of rows of the linear systems is $3\binom{2n+1}{3} - 3\binom{n+1}{3}$. The total number of power products up to degree $3n-2$ in three indeterminates is $\binom{3n+1}{3}$. It can be verified easily that the difference between these, the number of elements of $\mathcal{T}$, is exactly n^3. $\square$

In the last section of this chapter we discuss the relation of the reduction algorithm to the Macaulay matrix more in detail.

Now we have formulated the conditions that should be fulfilled to let a reduction algorithm work, we can consider what to do in case $\mathcal{A}$ vanishes. We can use the trick illustrated in Example 3.5 for the bivariate case, exploiting the linear dependencies that follow from the singularity of the resultant matrix corresponding to $\mathcal{A}$. Actually, this is done – although with a different approach – in the solution presented in the next chapter.

3.5 Reducibility in the m-variate case

For $m > 3$, the reduction algorithm is similar to the trivariate case. It can be performed by solving a succession of linear systems with coefficients that rationally depend on the coefficients of the leading forms of $f_1, \ldots, f_m$.

The linear systems to be solved can be displayed by the following diagram for the case $n_1 = \cdots = n_m = n$:

$$
\begin{array}{c}
\begin{array}{c} f_1 = 0 \\ \cdots \\ f_m = 0 \end{array} \Longrightarrow \left\langle \begin{array}{c} x_1^n \tau(x_2, \ldots, x_m) \\ \cdots \\ x_m^n \tau(x_1, \ldots, x_{m-1}) \end{array} \;\middle|\; 0 \le \deg \tau \le n-1 \right\rangle_{(n)} \\
\downarrow \\
\left\langle \begin{array}{c} x_2^n \tau(x_1, x_3, \ldots, x_m) \\ \cdots \\ x_m^n \tau(x_1, \ldots, x_{m-1}) \end{array} \;\middle|\; \begin{array}{c} n-1 < \deg \tau \le 2n-2 \\ \\ \tau \in \mathcal{T} \end{array} \right\rangle_{(n-1)} \\
\downarrow \\
\left\langle \begin{array}{c} x_3^n \tau(x_1, x_2, x_4, \ldots, x_m) \\ \cdots \\ x_m^n \tau(x_1, \ldots, x_{m-1}) \end{array} \;\middle|\; \begin{array}{c} 2n-2 < \deg \tau \le 3n-3 \\ \\ \tau \in \mathcal{T} \end{array} \right\rangle_{(n-1)} \\
\downarrow \\
\vdots \\
\downarrow \\
\left\langle \begin{array}{c} x_k^n \tau(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_m) \\ \cdots \\ x_m^n \tau(x_1, \ldots, x_{m-1}) \end{array} \;\middle|\; \begin{array}{c} (k-1)(n-1) < \deg \tau \le k(n-1) \\ \\ \tau \in \mathcal{T} \end{array} \right\rangle_{(n-1)} \\
\downarrow \\
\vdots \\
\downarrow \\
\left\langle \begin{array}{c} x_{m-1}^n \tau(x_1, \ldots, x_{m-2}, x_m) \\ \\ x_m^n \tau(x_1, \ldots, x_{m-1}) \end{array} \;\middle|\; \begin{array}{c} (m-2)(n-1) < \deg \tau \le (m-1)(n-1) \\ \\ \tau \in \mathcal{T} \end{array} \right\rangle_{(n-1)} .
\end{array}
$$

The $\langle\ \rangle$ stands for a cascade, i.e. a finite sequence of linear systems. The subscript denotes the number of systems. For example, the first of the considered cascades has the structure:

$$
\left\{ \begin{array}{c} x_1^n \\ \cdots \\ x_m^n \end{array} \right\}_m \hookrightarrow \left\{ \begin{array}{c} x_1^n x_j,\ 1 < j \le m \\ \cdots \\ x_m^n x_j,\ 1 \le j < m \end{array} \right\}_{m(m-1)} \hookrightarrow \left\{ \begin{array}{c} x_1^n x_j x_k,\ 1 < j, k \le m \\ \cdots \\ x_m^n x_j x_k,\ 1 \le j, k < m \end{array} \right\}_{m\binom{m}{2}}
$$

$$
\hookrightarrow \cdots \hookrightarrow \left\{ \begin{array}{cc} x_1^n x_2^{j_2} \cdots x_m^{j_m}, & j_2 + \cdots + j_m = n-1 \\ \cdots & \\ x_1^{j_1} \cdots x_{n-1}^{j_{m-1}} x_m^n, & j_1 + \cdots + j_{m-1} = n-1 \end{array} \right\}_{m\binom{m+n-3}{n-1}} . \tag{3.30}
$$

Here $\{\ \}$ represents the linear system with respect to the power products included. The subscript indicates the number of power products to solve, i.e. the order of the determinant of the corresponding linear system.

If the power products of degree $n + l - 1$ are reducible with respect to $\mathcal{T}$ and the linear system in the power products of degree $n + l$ is consistent, then all monomials of degree $n+l$ are reducible with respect to $\mathcal{T}$. The transfer, denoted by $\hookrightarrow$, of the system

$$\left\{ \begin{array}{c} x_k^n \tau(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_m) \\ \ldots \\ x_m^n \tau(x_1, \ldots, x_{m-1}) \end{array} \left| \begin{array}{c} \deg \tau = l \\ \\ \tau \in \mathcal{T} \end{array} \right. \right\} \tag{3.31}$$

to the next one in the sequence is performed via multiplication of the reductions obtained in the previous step by any of $x_k, \ldots, x_m$ if the new system is contained in the same cascade and by any of $x_{k+1}, \ldots, x_m$ if the new system is the first of the next cascade.

To count the number of power products, we can use a result of Bézout [12, 65, 80].

Theorem 3.10 *The number of power products of degree d in m indeterminates equals*

$$N(m, d) = \binom{m+d-1}{d}.$$

The number of those not dividing $x_1^{d_1}$ equals

$$\Delta_{d_1} N(m, d) = N(m, d) - N(m, d - d_1).$$

The number of those not dividing $x_1^{d_1}, x_2^{d_2}$ equals

$$\Delta_{d_2} \Delta_{d_1} N(m, d) = \Delta_{d_1} N(m, d) - \Delta_{d_1} N(m, d - d_2).$$

The number of those not dividing $x_1^{d_1}, x_2^{d_2}, \ldots, x_k^{d_k}$ equals

$$\Delta_{d_k} \ldots \Delta_{d_1} N(m, d) = \Delta_{d_{k-1}} \ldots \Delta_{d_1} N(m, d) - \Delta_{d_{k-1}} \ldots \Delta_{d_1} N(m, d - d_k).$$

From these, it can be derived that the number of power products in (3.31) equals

$$(m-k+1)\left[\binom{m+l-2}{m-2} - \binom{m-1}{1}\binom{m+l-2-n}{m-2} + \binom{m-1}{2}\binom{m+l-2-2n}{m-2} - \cdots\right].$$

As the number of the cascade transfers is equal to $(m-1)(n-1)$, in total $(m-1)(n-1)+1$ conditions need to be fulfilled in order to determine all the reducers.

As in Theorem 3.7 it can be shown that the first n conditions, illustrated by (3.30), correspond to minors of the Macaulay matrix of the leading forms of $f_1, \ldots, f_m$ for finding the resultant $\mathcal{A}$ defined by equation (1.15). Similar to what is done in Theorem 3.9 the reduction algorithm can be reformulated based on minors of the Macaulay matrix of the leading forms of $f_1, \ldots, f_m$.

3.6 Comparison to the Macaulay matrix

As we have seen in previous sections there is a strong relation between the reduction algorithm and the Macaulay matrix of the leading forms. In this section we discuss this relation together with the construction of the Bézout matrix.

The main observation is that the reduction algorithm is a procedure of triangularization of a coefficient matrix.

First, consider the bivariate case with $n_1 = n_2 = n_3 = n$ and indeterminates x and y. The Macaulay matrix is constructed as

$$\begin{pmatrix} \mathcal{T}_1 f_1 \\ \mathcal{T}_2 f_2 \\ \mathcal{T}_3 f_3 \end{pmatrix} \tag{3.32}$$

with $\mathcal{T}_1$, $\mathcal{T}_2$ and $\mathcal{T}_3$ defined as

$$\begin{aligned} \mathcal{T}_1 &= \{x^p y^q \mid p + q \leq 2\,n - 2\}, \\ \mathcal{T}_2 &= \{x^p y^q \mid p + q \leq 2\,n - 2, p < n\}, \\ \mathcal{T}_3 &= \{x^p y^q \mid p + q \leq 2\,n - 2, p < n, q < n\} \end{aligned}$$

based on the critical degree $3\,n - 2$. The order of these sets is

$$\binom{2\,n}{2}, \binom{2\,n}{2} - \binom{n}{2} \text{ and } n^2,$$

respectively. The reduction algorithm can be represented by the homogenized variant of the Macaulay matrix of the leading forms of f_1 and f_2:

$$\begin{pmatrix} \bar{\mathcal{T}}_1 \bar{f}_1 \\ \bar{\mathcal{T}}_2 \bar{f}_2 \end{pmatrix}.$$

As we have seen these multipliers of the leading forms are bound by degree n. Hence, we have

$$\bar{\mathcal{T}}_1 = \bar{\mathcal{T}}_2 = \{x^p y^q \mid p + q < n\}.$$

A representation like (3.32) of the reduction algorithm together with the Bézout matrix construction is

$$\begin{pmatrix} \begin{bmatrix} \bar{\mathcal{T}}_1 f_1 \\ \bar{\mathcal{T}}_2 f_2 \end{bmatrix} \\ \mathcal{T} f_3 \end{pmatrix} \tag{3.33}$$

where [ ] represents the reduction algorithm with $\bar{f}_1$, $\bar{f}_2$ replaced by f_1 and f_2. The set $\mathcal{T}$ defined as in (3.32) coincides with $\mathcal{T}_3$ of the Macaulay matrix representation (3.32).

The power products can be ordered in such a way that we can write the result of the reduction algorithm as the result of a Gaussian elimination process:

$$\begin{array}{c} \overbrace{\hphantom{\star \quad \star \quad \ddots \quad \star \quad \star}}^{M\backslash\mathcal{T}} \; \overbrace{\hphantom{* \; \dots \; *}}^{\mathcal{T}} \\ \left[\begin{array}{ccccc|ccc} \star & & & & \mathbf{O} & * & \dots & * \\ & \star & & & & * & \dots & * \\ & & \ddots & & & \vdots & & \vdots \\ & & & \star & & * & \dots & * \\ \mathbf{O} & & & & \star & * & \dots & * \end{array}\right] \end{array} \tag{3.34}$$

with M containing all the power products up to degree $2\,n-1$. The reducibility conditions mentioned in this chapter correspond to the non-vanishing of the diagonal elements – marked with $\star$ – of the left part.

If the conditions are satisfied, the construction of the Bézout matrix is nothing else than elimination from the rows of the $\mathcal{T} f_3$-part of the matrix (3.33) the entries in the columns that do not correspond to elements in the set $\mathcal{T}$. This results in the matrix

$$\begin{pmatrix} [\star & *] \\ \mathbf{O} & \mathcal{B} \end{pmatrix},$$

where the submatrix $\mathcal{B}$ is the Bézout matrix. The upper part $[\star\ *]$ stands for the matrix (3.34) [14]. However, $\mathcal{T} f_3$ may contain power products that do not correspond to elements in M, as the degree of $\mathcal{T} f_3$ goes up to $3\,n-2$. The reducers for the power products of degree $2\,n$ up to $3\,n-2$ were not determined in the reduction algorithm, as we concentrated on finding all the conditions following from the reduction algorithm. Once the reducers for power products up to degree $2\,n-1$ are known, the ones for power products up to degree $3\,n-2$ follow without additional conditions. Hence, the matrix (3.34) can be considered as the non-redundant part.

From our considerations we conclude that the Macaulay matrix is very closely related to our approach using the reduction algorithm. The reduction algorithm can be viewed as a Gaussian elimination on those rows of the Macaulay matrix of $f_1, \dots, f_m, f_{m+1}$ that have a corresponding row in the Macaulay matrix of the leading forms of $f_1, \dots, f_m$. The order of these Macaulay matrices is, in case $n_1 = \cdots = n_{m+1} = n$,

$$\binom{mn+n}{m} \text{ and } \binom{mn}{m-1},$$

which differ by a factor between n and en.

Chapter 4

Constructing the Bézout resultant: Gröbner basis approach

In this chapter we treat another approach for constructing the Bézout resultant. This approach becomes necessary in case the polynomials are not enough generic anymore and conditions following from the reduction algorithms of Chapter 3 are not fulfilled. A Gröbner basis is used for two purposes. Firstly, to obtain a proper set $\mathcal{T}$, and secondly, to determine the normal forms: the Gröbner basis equivalent of the reduction form. This results in an effective algorithm for the evaluation of the resultant.

In the second section we retrieve sufficient conditions on the factors of the computed resultant for knowing these factors do not belong to the extraneous factors. One section follows on how to adapt the method to handle the case of an infinite Poisson product. In the fourth section we demonstrate how to use the eigenvalues of the Bézout matrix associated with the u-resultant to find the components of the solutions.

In the last section experiments show this resultant method is competitive with the fastest known methods for evaluating the resultant.

4.1 The construction

We translate the determinantal approach of the construction of the Bézout resultant into the language of ideal theory and Gröbner bases, which terminology is introduced in Section 1.6.

Let $\mathcal{I}$ be the ideal generated by $f_1, \ldots, f_m$ over $K[x_1, \ldots, x_m]$. Let $\prec$ be an admissible ordering on the power products in $x_1, \ldots, x_m$. As most of our results are not depending on the choice of $\prec$, in practice a total degree ordering has preference over, for instance, a lexicographic ordering. In practice $f_1, \ldots, f_m$ are close to the Gröbner basis with respect to a total degree ordering, which means much less computational effort.

We compute the Gröbner basis of $\mathcal{I}$ with respect to the ordering $\prec$. We denote this by $\mathcal{G}(\mathcal{I})$. The set $\mathcal{T}$ can now be chosen as a basis of the K-vector space $K[X]/\mathcal{I}$, the residue or equivalence class ring modulo $\mathcal{I}$: for $f, g \in K[X]/\mathcal{I}$, $f \cong g$ if and only if $f - g \in \mathcal{I}$. A natural choice for this basis can be read off from the Gröbner basis:

$$\mathcal{T} = \{\tau(X) = x_1^{p_1} \cdots x_m^{p_m} \mid \tau(X) \text{ is not a divisible by any } \operatorname{lpp}(g), g \in \mathcal{G}(\mathcal{I})\}. \tag{4.1}$$

We will assume this set to be finite. For now, we do not consider the degenerate case where $\mathcal{T}$ is empty. In Section 4.3 we consider the case where $\mathcal{T}$ is not a finite set.

Hence, let $\mathcal{T}$ contain N^* elements, $N^* > 0$. This number corresponds to the number of common zeroes of $f_1, \ldots, f_m$, counting multiplicities. This number is less than or equal to the Bézout bound $N = n_1 \cdots n_m$. This does not require that the given polynomials are generic, as was assumed in the previous chapter.

Theorem 4.1 *The dimension N^* of the vector space given by $K[X]/\mathcal{I}$ is equal to the number of common zeroes, counting multiplicities, of $f_1, \ldots, f_m$ in the algebraic closure of K. Moreover, if N^* is finite, then $N^* \leq N$.*

For a proof of this important theorem we refer to [28]. Essential in this proof is the proper definition of multiplicity as the dimension of the equivalence class ring of the so-called local ring R of rational functions defined at the common zero Λ, modulo the ideal $\mathcal{I}$ generated in R. The multiplicity measures how complex $\mathcal{I}$ is at Λ [28].

The concept of reducedness from Chapter 3 is translated into that of normal forms. As we have seen in Section 1.6, for any $g(X) \in K[X]$ we can define its so-called normal form $g_{red}(X) \in K[X]$ with regard to $\mathcal{G}(\mathcal{I})$. The $g_{red}(X)$ is the polynomial in the span of $\mathcal{T}$, i.e. all its power products are from $\mathcal{T}$ and $g(X) - g_{red}(X) \in \mathcal{I}$. This normal form is unique and can be computed using $\mathcal{G}(\mathcal{I})$ [28].

We now evaluate a sequence of normal forms for $\tau_j(X) f_{m+1}(X)$, $\tau_j(X) \in \mathcal{T}$:

$$(\tau_j(X) f_{m+1}(X))_{red} = b_{j1}\tau_1(X) + \cdots + b_{jN^*}\tau_{N^*}(X) \quad (j = 1, \ldots, N^*). \tag{4.2}$$

Out of these N^* normal forms we can extract the square coefficient matrix of order N^*

$$\mathcal{B} = [b_{jk}]_{j,k=1}^{N^*}, \tag{4.3}$$

which will be called the Bézout matrix. Obviously, the Bézout matrix is a matrix with all its entries in K. Similar to the equations (3.7) and (3.8) we have

$$\mathcal{B}\mathcal{V} = \mathcal{V} \begin{pmatrix} f_{m+1}(\Lambda_1) & & \mathbf{O} \\ & \ddots & \\ \mathbf{O} & & f_{m+1}(\Lambda_{N^*}) \end{pmatrix}, \tag{4.4}$$

where $\mathcal{V}$ is the generalized Vandermonde matrix given by

$$\mathcal{V} = \begin{pmatrix} \tau_1(\Lambda_1) & \cdots & \tau_1(\Lambda_{N^*}) \\ \vdots & & \vdots \\ \tau_{N^*}(\Lambda_1) & \cdots & \tau_{N^*}(\Lambda_{N^*}) \end{pmatrix}.$$

The $\Lambda_1, \ldots, \Lambda_{N^*}$ are the common zeroes of $f_1, \ldots, f_m$ over the algebraic closure of K.

Provided that the Vandermonde matrix $\mathcal{V}$ is not singular, clearly, after dividing the determinant of both sides by $\det \mathcal{V}$, the determinant of the Bézout matrix equals the Poisson product:

$$\det \mathcal{B} = f_{m+1}(\Lambda_1) \cdots f_{m+1}(\Lambda_{N^*}). \tag{4.5}$$

This Poisson product is the pure Poisson product (1.17) adapted for the non-generic case. We denote it by $\Re^{f_{m+1}}(f_1, \ldots, f_m)$.

Obviously, $\mathcal{V}$ is singular if one of the common zeroes Λ_j is a multiple zero. On the other hand, if all the common zeroes have multiplicity one, we can prove that $\mathcal{V}$ is non-singular.

Theorem 4.2 *If $\mathcal{I}$ is a radical ideal, then $\mathcal{V}$ is non-singular.*

Proof. If $\mathcal{V}$ is singular, then there exists a linear dependency between its rows. This gives a non-zero polynomial $g(X) \in K[X]$ with all its power products in $\mathcal{T}$ that vanishes on all the common zeroes $\Lambda_1, \ldots, \Lambda_{N^*}$. Therefore, $g(X)$ belongs to the radical of $\mathcal{I}$. As $\mathcal{I}$ is a radical ideal, $g(X)$ belongs to $\mathcal{I}$. This contradicts with the linear independency of $\mathcal{T}$ as basis of $K[X]/\mathcal{I}$. □

In order to prove the validity of equation (4.5) in all cases, we need the following result, often called the Stickelberger Theorem [78].

Theorem 4.3 *For any zero-dimensional ideal $\mathcal{I}$ over $K[X]$ and any $g \in K[X]$, there exists a basis for the vector space $\overline{K}[X]/\mathcal{I}$ for which the matrix of the linear mapping $K[X]/\mathcal{I} \to K[X]/\mathcal{I}$ with $h(X) \mapsto (h(X)g(X))_{red}$ has the block-diagonal form*

$$\begin{pmatrix} U_1 & & \mathbf{O} \\ & \ddots & \\ \mathbf{O} & & U_{N'} \end{pmatrix}, \text{ where } U_j = \begin{pmatrix} g(\Lambda_j) & & * \\ & \ddots & \\ \mathbf{O} & & g(\Lambda_j) \end{pmatrix} \tag{4.6}$$

is of order m_j, the multiplicity of Λ_j. Here, we order the common zeroes such that $\Lambda_1, \ldots, \Lambda_{N'}$ are all the distinct common zeroes of the generators of $\mathcal{I}$.

We can now state the main theorem, an analogue of Theorem 3.1.

Theorem 4.4 *For any basis of $K[X]/\mathcal{I}$ the determinant of the corresponding Bézout matrix $\mathcal{B}$ is equal to the pure Poisson product. Moreover, if* $\det \mathcal{V} \neq 0$*, the rank of $\mathcal{B}$ equals the number of common zeroes of $f_1, \ldots, f_m$ for which $f_{m+1} \neq 0$.*

Proof. For every basis there exists a non-singular matrix P such that $P^{-1}\mathcal{B}P$ is of the diagonal form (4.6). From this follows

$$\det \mathcal{B} = \det U_1 \cdots \det U_{N'} = \prod_{j=1}^{N'} f_{m+1}(\Lambda_j)^{m_j} = \Re^{f_{m+1}}(f_1, \ldots, f_m).$$

The second statement is identical to the one from Theorem 3.1. □

This theorem shows that the choice (4.1) for $\mathcal{T}$ is a good choice because it makes the construction of the Bézout matrix straightforward. Although different choices may give different Bézout matrices, their determinants (and eigenvalues) coincide. The choice (4.1) for $\mathcal{T}$ only depends on the ordering $\prec$, but again the determinants (and eigenvalues) are invariant with regard to the orderings.

We refer to [8, 28, 46] for results on how to use the traces of a set of Bézout matrices for computing the radical of an ideal and counting the number of real solutions.

As a corollary of the last theorem, we know that the eigenvalues of $\mathcal{B}$ that we can read off from equation (4.4) form the complete set of eigenvalues.

Theorem 4.5 *The eigenvalues of the Bézout matrix are $f_{m+1}(\Lambda_1), \ldots, f_{m+1}(\Lambda_{N^*})$. The columns of the generalized Vandermonde matrix supply at least one of the associated eigenvectors per eigenvalue.*

Note that the algebraic multiplicity of an eigenvalue $f_{m+1}(\Lambda_j)$ may be larger than the multiplicity of Λ_j, as $f_{m+1}(\Lambda_j)$ can coincide with $f_{m+1}(\Lambda_k)$ for $\Lambda_j \neq \Lambda_k$. The form (4.6) is close to the Jordan canonical form [37] of the Bézout matrix. It shows that the Jordan canonical form consists of at least N' blocks. In Section 4.4 we will consider the eigenvalues of the Bézout matrix of u-resultants more in detail.

As an illustration we repeat Example 3.13 and show a Bézout matrix together with its Jordan canonical form.

Example 4.1 *Evaluate the pure Poisson product $\Re^{f_3}(f_1, f_2)$ for the system*

$$\begin{aligned} f_1 &= 5\,x^2 - 6\,xy + 5\,y^2 - 16, \\ f_2 &= 2\,x^2 - xy + y^2 - x - y - 4, \\ f_3 &= x^2 - y + \epsilon. \end{aligned}$$

Let $\mathcal{I}$ be the ideal generated by f_1 and f_2 over $\mathbb{Q}[x, y]$. We construct the Gröbner basis of $\mathcal{I}$ with respect to the total degree ordering $x \succ y$. This gives the Gröbner basis

$$\{7\,xy - 5\,y^2 - 5\,x - 5\,y + 12, \quad 7\,x^2 + y^2 - 6\,x - 6\,y - 8, \quad 7\,y^3 - 2\,y^2 - 9\,x - 16\,y + 2\}.$$

The leading power products of the elements in these generators of $\mathcal{I}$ are xy, x^2 and y^3. Hence, we can choose $\mathcal{T} = \{1, x, y, y^2\}$. The polynomials f_1 and f_2 have four common zeroes. The normal forms of f_3, $x\,f_3$, $y\,f_3$ and $y^2\,f_3$ form the rows of the Bézout matrix. The columns are corresponding to the power products $1, x, y$ and y^2, respectively:

$$\mathcal{B} = \begin{pmatrix} \epsilon + 8/7 & 6/7 & -1/7 & -1/7 \\ 10/7 & \epsilon + 11/7 & 4/7 & -3/7 \\ -10/7 & 3/7 & \epsilon + 10/7 & 3/7 \\ -6/7 & 6/7 & -1/7 & \epsilon + 13/7 \end{pmatrix}.$$

This matrix differs from the one obtained in Example 3.13, but the determinant is the same: $\epsilon(\epsilon + 2)^3$.

Using the transformation matrix from the Jordan canonical form of $\mathcal{B}$ we can obtain another choice for $\mathcal{T}$, namely

$$\mathcal{T} = \left\{1 - \frac{9\,x}{16} + \frac{5\,y}{16} - \frac{y^2}{8}, \frac{63\,x}{32} + \frac{7\,y}{32} + \frac{7\,y^2}{4}, 1 + \frac{3\,x}{4} + \frac{3\,y}{4} - y^2, x + y\right\},$$

such that the Bézout matrix is in Jordan canonical form. This Bézout matrix consists of three Jordan blocks:

$$\mathcal{B} = \begin{pmatrix} \epsilon & 0 & 0 & 0 \\ 0 & \epsilon + 2 & 1 & 0 \\ 0 & 0 & \epsilon + 2 & 0 \\ 0 & 0 & 0 & \epsilon + 2 \end{pmatrix}.$$

Another example shows how the Gröbner basis approach handles the cases where $\mathcal{A}$, defined as in (1.15), vanishes. In the determinantal approach we already described a way to overcome this, but it is difficult to give an algorithm that covers all cases. The cardinality N^* of the choice (4.1) for $\mathcal{T}$, in case $\mathcal{A} = 0$, is strictly less than the Bézout bound: $N^* < N$. The equation $\mathcal{A} = 0$ indicates there exist common zeroes at infinity [65, 69]. Therefore, the number of affine zeroes is less than the Bézout bound.

Example 4.2 *Solve Example 3.5 using the Gröbner basis approach.*

Solution. Compute the Gröbner basis of the ideal $\mathcal{I}$ generated by f_1 and f_2 in $K[y, z]$, with $K = \mathbb{Q}(x)$. Choosing the total degree monomial ordering $y \succ z$, a

Gröbner basis with respect to this ordering is given by

$$\mathcal{G} = \left\{ \begin{array}{ll} (39\,x+87)\,y^2 + \left(-2\,x^2+11\,x-93\right)y & +\left(118\,x^2+70\,x-62\right)z \\ & -35\,x^3-30\,x^2-45\,x-16, \\ (39\,x+87)\,yz + \left(-5\,x^2-19\,x-60\right)y & +\left(61\,x^2+76\,x-11\right)z \\ & -29\,x^3-66\,x^2-45\,x+14, \\ (39\,x+87)\,z^2 + \left(7\,x^2+26\,x+9\right)y & +\left(16\,x^2+7\,x-23\right)z \\ & -14\,x^3-36\,x^2-42\,x-34 \end{array} \right\}.$$

From this we find the set $\mathcal{T} = \{1, y, z\}$, giving $N^* = 3$, while the Bézout bound is $N = 4$. We observe that the leading coefficients depend on the parameter x. The normal forms of f_3, yf_3 and zf_3 may have the factor $13\,x+29$ in their denominator. In fact, the normal forms of $y\,f_3$ and $z\,f_3$ have indeed $13\,x+29$ in their denominator. After removing the denominators from the normal forms we have the fraction-free Bézout matrix $\mathcal{B}'$, the coefficient matrix of the normal forms of f_3, $27\,(13\,x+29)yf_3$ and $9\,(13\,x+29)zf_3$ (columns labelled with 1, y and z):

$$\mathcal{B}' = \begin{pmatrix} 8\,x^2+11\,x-1 & -7\,x+9 & -10\,x-8 \\ -535\,x^4-787\,x^3-1023\,x^2+73\,x+256 & b_{2,2} & b_{2,3} \\ -343\,x^4-673\,x^3-429\,x^2-173\,x-398 & b_{2,3} & b_{3,3} \end{pmatrix}$$

with $b_{2,2} = 248\,x^3+990\,x^2-584\,x+270$, $b_{2,3} = 35\,x^3+228\,x^2+49\,x+612$, $b_{2,3} = 1436\,x^3+676\,x^2-566\,x+470$ and $b_{3,3} = 899\,x^3+1306\,x^2-17\,x-172$. Its determinant equals $243\,(13\,x+29)^2 \det\mathcal{B}$, giving for $\det\mathcal{B}$ the result

$$\frac{(x-1)\left(2277\,x^6+6150\,x^5+3398\,x^4+8208\,x^3+10729\,x^2+6322\,x+3916\right)}{9\;(13\,x+29)}. \tag{4.7}$$

This determinant can be evaluated for any specialization of x, except for $x = -29/13$. In fact, all leading coefficients of the elements in the Gröbner basis vanish for $x = -29/13$. The common zeroes of f_1 and f_2 are discontinuous in x (we refer to [70] for some considerations on the continuity of solutions of algebraic equations). This also appears from the resultant of f_1 and f_2 eliminating y:

$$\begin{array}{l} \Re_y(f_1,f_2) = (39\,x+87)\,z^3 + \left(11\,x^2-12\,x-83\right)z^2 \\ \quad +\left(-27\,x^3-70\,x^2-54\,x-17\right)z + 7\,x^4+27\,x^3+27\,x+43\,x^2+22. \end{array} \tag{4.8}$$

One of the three common zeroes becomes infinite for $x = -29/13$. Recomputing the Gröbner basis under the condition $x = -29/13$ gives

$$\mathcal{G} = \left\{ \begin{array}{l} 247\,y-715\,z-627, \\ 1014\,z^2-37297\,z-33573 \end{array} \right\}.$$

The corresponding Bézout matrix, the coefficient matrix of the normal forms of f_3 and zf_3, is (columns labelled with 1 and z)

$$\mathcal{B} = \begin{pmatrix} 4324/169 & 21134/741 \\ 6223963/6591 & 1634533/1521 \end{pmatrix} \quad \text{with } \det\mathcal{B} = 144708374/257049.$$

This shows that there are no common zeroes of f_1, f_2 and f_3 in case $x = -29/13$. Therefore, the numerator of (4.7) can be taken as an eliminant in x because its vanishing is a necessary and sufficient condition for f_1, f_2 and f_3 to have a common zero. □

Although for this example we have $\mathcal{A} = 0$, we can take $\mathcal{A}$ to be the leading coefficient of (4.8) and multiply this with (4.7), according to equation (1.16), and remove the denominator. In fact, according to (1.16), we have to multiply with $\mathcal{A}^2$. This gives the factor $\mathcal{A}$ in the numerator. In the next section we will investigate these so-called extraneous factors.

4.2 Extraneous factors

Each method for constructing the resultant has to deal with the phenomenon of extraneous factors – they are inherent in the method – so the Gröbner basis approach has to. However, the extraneous factors of this approach also appear in a natural way giving particular information about the continuity of solutions of the system. In this section we derive a necessary condition for extraneous factor candidates. It is shown how to use the extraneous factors for constructing the eliminant in one parameter of the system and in the process of solving a polynomial system.

4.2.1 Definitions

In Section 1.4 we mentioned that for the Cayley resultant, in the univariate case, extraneous factors can appear if the degrees of the input polynomials are not equal. Because the extraneous factor is known in advance, we can adapt the Cayley resultant easily and remove the extraneous factor.

Consider the simple example:

$$f(x) = a\,x^3 + x^2 + x + 1, \quad g(x) = x^2 + 2\,x + 3.$$

The Cayley resultant evaluates to $a(27\,a^2 - 14\,a + 3)$ which is a factor a times the Sylvester resultant $\Re_x(f, g) = 27\,a^2 - 14\,a + 3$, defined as in (1.2). However, the Sylvester resultant for the equivalent system given by the equations $f(x) = 0, g(x) + f(x) = 0$ evaluates to $a(27\,a^2 - 14\,a + 3)$ as well. Essentially, the factor a is an extraneous factor: for $a = 0$ there is no common root for $f(x)$ and $g(x)$. In the univariate case, if the leading coefficients are equal, they appear as an extraneous factor in the resultant, in the sense that when they vanish this may not lead to a common root for $f(x)$ and $g(x)$. This directly relates to Proposition 1.2.

We can generalize this concept of extraneous factors to our Gröbner basis approach for the resultant. Going back to Example 4.2, we see that the factor $\mathcal{A}$ first appears in the denominator of the determinant of the Bézout matrix, whereas for equation (1.16) it appears in the numerator. As there were no solutions in case $\mathcal{A} = 0$, this factor is called extraneous.

Before continuing our considerations, we propose the following definition for the concept of extraneous factors.

Definition 4.1 *Let the coefficients of $f_1, \ldots, f_m, f_{m+1}$ be polynomials in the parameter t and let $K = K'(t)$, the quotient field of $K'[t]$ with K' a field. An element $C(t)$ in $K'[t]$ is called an extraneous factor in t if $C(t)$ does not have a root $\bar{t} \in \overline{K'}$, the algebraic closure of K', such that after specializing t to $\bar{t}$, $f_1, \ldots, f_m, f_{m+1}$ have a common solution over $\overline{K'}$.*

If the coefficients of $f_1, \ldots, f_m, f_{m+1}$ are polynomials in more than one parameter, there may exist extraneous factors in each of the parameters, choosing different coefficient fields K'. For instance, $K = \mathbb{Q}(t_1, t_2)$ can be viewed as the quotient field of the integral domains $\mathbb{Q}(t_1)[t_2]$ and $\mathbb{Q}(t_2)[t_1]$. In the literature extraneous factors are not defined explicitly, but are rather described as the 'strange' factor of the result of an elimination method (or so-called projection operator) [53, 55, 77].

Specializing one or more of the parameters of the system may have influence on the computation of the Gröbner basis [42, 51, 94], the choice of the set $\mathcal{T}$ and hence the Poisson product evaluation. If the elements in the Gröbner basis are made monic (leading coefficients 1), a specialization of the parameter may lead to the vanishing of some of its denominators. First doing the specialization and after that the computation, gives a different Gröbner basis. In order to handle these cases, we introduce the notion of suspicious factors.

Definition 4.2 *The set of suspicious factors $SF(\mathcal{G}(\mathcal{I}))$ associated with the Gröbner basis $\mathcal{G}$ of the ideal $\mathcal{I}$, consists of all polynomials in the parameters of the system by which a division was performed during the computation of $\mathcal{G}$. These divisions are either to make a polynomial monic, or to divide out a common multiple in the coefficients. Here, we assume that the computation of the Gröbner basis $\mathcal{G}$ is only completed after making all polynomials monic.*

Note that this definition depends on the algorithm, namely on the paths the algorithm follows to construct the Gröbner basis. We assume the ideal to be given by the fixed set of generators $f_1, \ldots, f_m$. In order to gather all suspicious factors, the Gröbner basis algorithm can be modified easily (the current implementations of the algorithm in Maple and Mathematica do not have this kind of reporting).

It is possible to give a definition of the set of suspicious factors independent of the algorithm, but this is of little value in practice. As every polynomial in the constructed Gröbner basis – we assume the polynomial is monic – can be written as

$$h_1(X)f_1(X) + \cdots + h_m(X)f_m(X),$$

where the coefficients of $h_1(X), \ldots, h_m(X)$ are rational functions in the parameters, their denominators belong to the set of suspicious factors. Theoretically we can choose the representations of Gröbner basis elements such that the set of suspicious factors contains the fewest possible number of elements. However, this does not define the set of suspicious factors in a unique way.

Note that the set of suspicious factors, like the Gröbner basis itself, depends on the monomial ordering. The so-called comprehensive Gröbner basis [94] can be used to determine the set of suspicious factors and to obtain a Gröbner basis that stays a Gröbner basis after specializing parameters (which will be done in the next section).

Applying Theorem 4.4 we have a result similar to Theorem 1.4, the so-called Extension Theorem used for solving polynomial systems by Gröbner bases of elimination ideals.

Theorem 4.6 *For any specialization of the parameters for which none of the suspicious factors vanishes, there exists a common zero for* $f_1, \ldots, f_m, f_{m+1}$ *if and only if the specialization of the Bézout matrix is singular.*

Proof. If none of the suspicious factors vanishes, the specialization can be applied to the Gröbner basis construction as well as to the normal form computations. This gives the specialization of the Bézout matrix. The result then follows from equation (1.17). □

To cover all other possible specializations of the parameters, i.e. specializations for which one or more of the suspicious factors vanish, we introduce the notion of approximate eliminant. The approximate eliminant is defined as a multiple of the determinant of the Bézout matrix:

Definition 4.3 *The approximate eliminant, denoted by* $\tilde{\mathcal{E}}$*, is defined to be*

$$\tilde{\mathcal{E}} = \operatorname{num}(\det \mathcal{B}) \prod_{s \in SF(\mathcal{G}(\mathcal{I}))} s \tag{4.9}$$

with $\operatorname{num}(\det \mathcal{B})$ *denoting the numerator of the determinant of* $\mathcal{B}$*.*

The approximate eliminant is a polynomial in the parameters of the system – we assume the original polynomials of the system do not have denominators. The approximate eliminant removes the denominator of the Poisson product adjusted for the non-generic case, and does that in such a way that no factors of interest are removed. In this sense, it corresponds to the factor $\mathcal{A}$ as part of Definition 1.2 of the resultant in the generic case. Some exact generalizations of the factor $\mathcal{A}$ for the non-generic case, actually sparse systems with algebraically independent coefficients, can be found in [73].

In case we focus on one of the parameters of the system, like is done in the definition of extraneous factors, we can define the suspicious factors as suspicious factors in the parameter, for instance t, and consequently we can define the approximate eliminant in t, which we denote by $\tilde{\mathcal{E}}(t)$.

Corollary 4.1 *Let t be a parameter of the system. If the approximate eliminant in t has an extraneous factor in t, this factor is contained in the set of suspicious factors.*

Because the approximate eliminant depends on the suspicious factors, it depends on the Gröbner basis algorithm as well. We can relate the approximate eliminant with the so-called true eliminant. The true eliminant only depends on the given polynomial system. It is defined with respect to one parameter, in the context of Definition 4.1.

Definition 4.4 *Let t be a parameter of the system, like in Definition 4.1. The true eliminant, denoted by $\mathcal{E}(t)$, is the Poisson product*

$$\mathcal{E}(t) = \prod_j t - \bar{t}_j, \tag{4.10}$$

where $\bar{t}_j$ runs over all the solutions of $f_1, \ldots, f_m, f_{m+1}$ over the algebraic closure of K'. Multiple solutions occur in the product as often as their multiplicity.

The true eliminant can be viewed as a Poisson product of the system

$$f_1 = 0, \ldots, f_m = 0, f_{m+1} = 0, T - t = 0 \tag{4.11}$$

with coefficient field $K = K'(T)$. Hence the true eliminant can be evaluated by a Bézout matrix. For a generic system it can be evaluated by the Poisson product (1.17) as well:

Corollary 4.2 *For a generic system – dense and with algebraically independent coefficients in K' – the true eliminant in t coincides, up to a factor in K', with the Poisson product* (1.17).

Proof. As the system is generic the true eliminant is an irreducible polynomial in t (Proposition 1.3). The factor $\mathcal{A}$ is independent of t because t does not appear in the leading forms of $f_1, \ldots, f_m$. Therefore the Poisson product (1.17) is a polynomial in t. Again by Propositon 1.3 we know that this polynomial is irreducible as well. As both represent the t-component of the solutions, they coincide up to a constant factor. □

Usually any polynomial whose distinct factors coincide with the distinct factors of $\mathcal{E}(t)$ is called an eliminant [80]. Such an eliminant correctly represents the necessary and sufficient condition to be satisfied in order to make the system solvable. The eliminant $\mathcal{E}(t)$ is also called the t-eliminant [71].

4.2.2 Evaluating the eliminant

To illustrate the given definitions we consider an example with one parameter.

Example 4.3 *Find all the values for the parameter* t *for which the polynomials*

$$\begin{aligned} f_1 &= y^2 z + 2\,xyt - 2\,x - z, \\ f_2 &= 2\,yzt + xt^2 - x - 2\,z, \\ f_3 &= -x^3 z + 4\,xy^2 z + 4\,x^2 yt + 2\,y^3 t - 10\,y^2 + 4\,xz - 10\,yt + 4\,x^2 + 2, \\ f_4 &= -xz^3 + 4\,yz^2 t + 4\,xzt^2 + 2\,yt^3 + 4\,xz + 4\,z^2 - 10\,yt - 10\,t^2 + 2 \end{aligned} \tag{4.12}$$

have a common zero.

Solution. We compute the Gröbner basis of the ideal generated by f_1, f_2 and f_3 over $\mathbb{Q}(t)[x, y, z]$ with respect to the total degree ordering $x \succ y \succ z$. From the denominators of the Gröbner basis elements (after making them monic) we find the set of suspicious factors

$$SF(\mathcal{G}(\mathcal{I})) = \left\{t, t-1, t+1, t^2 + 1/3\right\}.$$

We can choose $\mathcal{T} = \left\{1, x, y, z, xz, y^2, z^2, xz^2, y^3, z^3, z^4\right\}$. The determinant of the 11×11 Bézout matrix equals

$$12^9 (t-1)(t+1)(t^2+1)(t^4 - 10\,t^2 + 1)(t^2 + 1/3)^4 (t^2+3)^8. \tag{4.13}$$

Its associated approximate eliminant adds the factor t and increases the multiplicities of the factors $t-1$, $t+1$ and $t^2 + 1/3$ by one. The roots from the factors t^2+1, t^2+3 and $t^4 - 10\,t^2 + 1$,

$$\pm i,\ \pm i\sqrt{3},\ \pm\sqrt{5 + 2\sqrt{6}},\ \pm\sqrt{5 - 2\sqrt{6}}$$

give common zeroes for f_1, f_2, f_3 and f_4. The remaining roots have to be investigated separately, as the corresponding factors are contained in $SF(\mathcal{G}(\mathcal{I}))$ and their vanishing may influence the Gröbner basis computation.

We recompute the Poisson product with K being the quotient field of $\mathbb{Q}[t]/s(t)$ for every suspicious factor $s(t) \in SF(\mathcal{G}(\mathcal{I}))$. For $s(t) = t$ (in fact specializing t to 0) the 10×10 Bézout matrix is non-singular. This means that the factor t is an extraneous factor of the approximate eliminant. For $s(t) = t^2 + 1/3$ (specializing t to $\pm i/\sqrt{3}$) the 7×7 Bézout matrix is singular and has rank 3. For $s(t) = t \pm 1$ (specializing t to ± 1) the Bézout matrix can not be constructed as the number of common zeroes of f_1, f_2 and f_3 becomes infinite. In fact, there exist common zeroes for f_1, f_2, f_3 and f_4 in case $t = \pm 1$.

This means that (4.13) can be considered as an eliminant in t, free of extraneous factors. Computing the true eliminant in t gives

$$\mathcal{E}(t) = (t-1)^9 (t+1)^9 (t^2+1)(t^4 - 10\,t^2 + 1)(t^2 + 1/3)^8 (t^2+3)^8,$$

which differs from (4.13) regarding the multiplicities of the factors and regarding the constant factor. □

In order to know what happens in case one of the suspicious factors vanishes, we have to explore these cases separately, as is done in the example above. One of the methods is to recompute the Poisson product over another coefficient field K, namely the quotient field of $K'[t]/s(t)$ (K was equal to $K'(t)$), where $s(t)$ is a suspicious factor in t. This requires the suspicious factor $s(t)$ to be irreducible, otherwise $K'[t]/s(t)$ is not an integral domain and its quotient field is not defined. For considerations on arithmetic in algebraic extension fields we refer to [91].

If the recomputed Poisson product is non-zero, we know $s(t)$ is an extraneous factor in t. Otherwise $s(t)$ is a factor of the true eliminant in t. If the Poisson product can not be recomputed due to the fact that the ideal generated by $f_1, \ldots, f_m$ over the new coefficient field is not zero-dimensional (i.e. the set $\mathcal{T}$ is not finite), we can use a fall-back rule like the computation of the Gröbner basis of the ideal generated by $f_1, \ldots, f_m, f_{m+1}$ over the new coefficient field and check whether it contains 1 [18]. If it contains 1 we know $s(t)$ is an extraneous factor in t.

The algorithm to find all factors of the true eliminant can be summarized as follows. First of all, all non-suspicious factors from the approximate eliminant are factors of the true eliminant. For each suspicious factor $s(t)$, we have to check whether it is extraneous or not by recomputing the Poisson product after specializing t to the algebraic number given by $s(t)$. If the Poisson product vanishes, it is not an extraneous factor, hence it is a factor of the true eliminant. Similar to Corollary 4.1 we conclude:

Theorem 4.7 *Let t be a parameter of the system. The factors of the true eliminant in t consist of all non-suspicious factors from the approximate eliminant in t. The remaining factors are among the finitely many suspicious factors. To determine whether a suspicious factor is extraneous or not, a Poisson product computation over the algebraic extension field suffices.*

Regarding the complexity of the algorithm, as a Gröbner basis is needed at start, we are faced with a complexity polynomial in N [36]. With all polynomials of the same degree n we have $N = n^m$. Assume the polynomials of the system given as in the context of Definition 4.1 have degree n in $x_1, \ldots, x_m, t$. The degree of the eliminant in t does not exceed the Bézout bound n^{m+1}, the intrinsic complexity of the problem of evaluating the eliminant. The complexity of the FGLM algorithm is polynomial in this bound.

For the Bézout matrix $N^* \leq N$ normal forms have to be evaluated. Except for the normal form of f_{m+1}, the normal form computations can be ordered in such a way that for the normal form of τf_{m+1}, with $\tau = x_j \tau'$, we can use the x_j-multiple of the normal form of $\tau' f_{m+1}$. This limits the number of terms to reduce to N^* per normal form. Each reduction costs N^* operations, so up to ${N^*}^2$ operations are needed per normal form. This means that apart from the complexity of evaluating the normal form of f_{m+1} the computational complexity for constructing the Bézout matrix is $\mathcal{O}({N^*}^3)$. The number of monomials of f_{m+1} is limited by $\binom{n+m}{m}$, roughly giving complexity $\mathcal{O}(NN^*)$ for its normal form computation. Obviously, the determinant evaluation contributes $\mathcal{O}({N^*}^3)$. Note that this estimates the number of operations in the coefficient field, actually polynomials in the parameter. They do not exceed the bound for the Gröbner basis algorithm. We conclude:

Theorem 4.8 *The complexity of evaluating the eliminant by the Bézout matrix is polynomial in* n^m.

The experiments from Section 4.5 show that in practice the determinant evaluation takes most of the computation time. Further investigations are needed for more precise complexity analyses taking into account the growth of the coefficients (polynomials in the parameters).

Remark Regarding the multiplicities of the factors of the eliminant there is a conjecture [14] saying that the true eliminant (4.10) can be written as

$$\mathcal{E}(t) = \prod_{s(t)} s(t)^{\max\{\mathrm{mult}(s),\mathrm{dfc}(s)\}}, \tag{4.14}$$

where $s(t)$ runs over all the irreducible factors in t of the approximate eliminant. The $\mathrm{dfc}(s)$ denotes the deficiency index of the Bézout matrix after specializing t to a root of $s(t)$. For $s(t)$ being a suspicious factor, $\mathrm{mult}(s)$ is set to 0, otherwise $\mathrm{mult}(s)$ equals the multiplicity of $s(t)$ in the denominator of $\det \mathcal{B}$.

The following example shows that this formula is not correct.

Example 4.4 *Compare the multiplicities of the factors of the approximate eliminant to the true eliminant for the following system, considering* t *as a parameter:*

$$\begin{aligned} f_1 &= 2\,x^3 t - 2\,x\,y^3 + 3\,xy^3\,t^2 - y^3 t^4 + 2\,y^3 t^2, \\ f_2 &= y^2 + 4\,x - 4\,x\,t^2 + t^4 - 2\,t^2 + 1, \\ f_3 &= 2\,x^3 y^3 - 2\,x^3 - 4\,x^2 y^3 t + 5\,xy^3\,t^3 - y^3 t^5. \end{aligned}$$

The only suspicious factor in t is $s(t) = t$. The determinant of the 6×6 Bézout matrix equals

$$t^3 (t-1)^{24} (t+1)^{24} (t^2 - 2\,t - 2) P(t),$$

where $P(t)$ is an irreducible polynomial in t of degree 37. The true eliminant equals

$$t^{12} (t-1)^{18} (t+1)^{18} (t^2 - 2\,t - 2) P(t),$$

with multiplicities for the factors $t-1$ and $t+1$ that do not correspond to (4.14). The normal form of f_3 turns out to have a factor $(t-1)(t+1)$ in its content, the gcd of the coefficients. As the other normal forms used for the construction of the Bézout matrix are multiples of the normal form of f_3, they have this factor as well. This gives a factor $(t-1)^6 (t+1)^6$ in the final determinant.

4.2.3 The multiparameter case

The algorithm from above suffices in case there is one parameter. In case the system has more than one parameter, things can become more complex. Obviously, if there is no suspicious factor, Theorem 4.6 can be applied. In case there are suspicious factors, we can consider this suspicious factor as a univariate polynomial in one of the parameters. Then, however, the recomputed Poisson product, if it is non-zero, may depend on another parameter and possibly vanishes after further specializations. This gives some additional difficulties.

We first want to remark that, in the multiparameter case, there does not exist a definition for the true eliminant, like Definition 4.4: the eliminant as one polynomial in all the parameters whose vanishing is a necessary and sufficient condition for the system to have a solution. Usually the elimination ideal is being exploited. For example, the ideal generated by

$$f_1 = (a-1)x^2, f_2 = (b-1)y^2 \text{ and } f_3 = xy+1 \tag{4.15}$$

has the Gröbner basis $\{xy+1, a-1, b-1\}$ with respect to the lexicographic ordering $x, y \succ a, b$. Hence, a basis for the elimination ideal in a and b is $\{a-1, b-1\}$. An eliminant in a and b in terms of a single polynomial in a and b does not exist. The elimination ideals will be considered again at the end of this section.

We now expose a possible procedure for finding a necessary condition in terms of all the parameters for the existence of a common zero for $f_1, \ldots, f_m, f_{m+1}$. The complication in the multiparameter case is that this condition may be not sufficient.

So let the system have coefficients in $v > 1$ parameters: $K = \mathbb{Q}(t_1, \ldots, t_v)$. After the computation of the Poisson product $\Re_X^{f_{m+1}}(f_1, \ldots, f_m)$ we have to check all the appeared suspicious factors. Let $S_1(t_1, \ldots, t_v)$ be a suspicious factor. We can join this polynomial to the generators of the ideal $f_1, \ldots, f_m$ and compute the Poisson product

$$\Re_{X,t_1}^{f_{m+1}}(f_1, \ldots, f_m, S_1(t_1, \ldots, t_v))$$

eliminating $x_1, \ldots, x_m$ and t_1. This gives a new necessary condition for the existence of a common zero of $f_1, \ldots, f_m, f_{m+1}$ considered as polynomials in $t_2, \ldots, t_v$. However, for this Poisson product computation suspicious factors in $t_2, \ldots, t_v$ may appear. These can be handled in the same way. Let $S_2(t_1, \ldots, t_v)$ be such a suspicious factor. Join it to the generators of the ideal and compute

$$\Re_{X,t_1,t_2}^{f_{m+1}}(f_1, \ldots, f_m, S_1(t_1, \ldots, t_v), S_2(t_2, \ldots, t_v)).$$

This procedure can be continued and terminates in a finite number of steps, as in each step we join a suspicious factor with at least one parameter less. Hence, this can be done up to v times. Finally, the disjunction of all obtained conditions builds a necessary condition. This condition is not always a sufficient one.

Example 4.5 *Find all the values for the parameter a, b and c for which the polynomials*

$$\begin{aligned} f_1 &= 2\,x^2 + axy + y^2 + 4\,x + c, \\ f_2 &= -x^2 + bxy + y^2 - 2\,x - 3\,y + 5, \\ f_3 &= x - y \end{aligned}$$

have a common zero.

Solution. We compute a Gröbner basis for the ideal generated by f_1 and f_2 over $K[x,y]$, $K = \mathbb{Q}(a,b,c)$, with respect to the total degree ordering $x \succ y$. This gives the set of suspicious factors $SF = \{a + 2\,b, \mathcal{A}\}$ where $\mathcal{A} = 2\,b^2 - a^2 - ab + 9$, the resultant of the leading forms of f_1 and f_2. We can choose $\mathcal{T} = \{1, x, y, y^2\}$ and determine the Poisson product:

$$\Re^{f_3}_{x,y}(f_1, f_2) = \frac{\varphi_1(a,b,c)}{\mathcal{A}}, \text{with}$$

$$\varphi_1(a,b,c) = b^2c^2 + 10\,b(-ac - c + 8) + 25(a^2 + ac + 10\,a + 3\,c + 21).$$

For the suspicious factor $a + 2\,b$ we compute the Gröbner basis for f_1, f_2 and $a + 2\,b$ over $K[x,y,b]$, $K = \mathbb{Q}(a,c)$, with respect to the total degree ordering $x \succ y \succ b$. This gives no suspicious factors. We can determine the order four Bézout matrix and find its determinant coincide with the already found Poisson product substituting $b = -a/2$.

For the other suspicious factor, the case $\mathcal{A} = 0$, we compute the Gröbner basis for f_1, f_2 and $\mathcal{A}$ over $K[x,y,b]$, $K = \mathbb{Q}(a,c)$, with respect to the total degree ordering $x \succ y \succ b$. This gives the suspicious factor $a^2 - 12$. With $\mathcal{T} = \{1, x, y, b, xb, yb\}$, we determine the Bézout matrix and find the determinant:

$$\Re^{f_3}_{x,y,b}(f_1, f_2, \mathcal{A}) = \frac{\varphi_2(a,c)}{324(a^2 - 12)}, \text{with}$$

$$\begin{aligned} \varphi_2(a,c) = (a+3)(a-3)^2c^4 + 5(2\,a^3 + 21\,a^2 - 6\,a - 180)c^3 \\ +15(-5\,a^3 + 33\,a^2 + 156\,a + 120)c^2 + 100(-5\,a^3 + 42\,a^2 + 441\,a + 954)c \\ +100(25\,a^3 + 465\,a^2 + 2427\,a + 4059). \end{aligned}$$

Next, we consider the case where $\mathcal{A} = 0$ and $a^2 - 12 = 0$. Computing the Gröbner bases of the ideal generated by f_1, f_2, $\mathcal{A}$ and $a^2 - 12$ over $K[x,y,a,b]$, $K = \mathbb{Q}(c)$, with respect to the total degree ordering $x \succ y \succ b \succ a$ gives the set of suspicious factors $SF = \{c + 1, c + 10, 13\,c^2 + 116\,c + 184\}$ and $\mathcal{T} = \{1, x, y, b, a, x\,a, x\,b, y\,a, y\,b, a\,b\}$. The Poisson product equals

$$\Re^{f_3}_{x,y,b,a}(f_1, f_2, \mathcal{A}, a^2 - 12) = \frac{\varphi_3(c)}{1728(13\,c^2 + 116\,c + 184)}, \text{with}$$

$$\varphi_3(c) = (c^4 + 20\,c^3 + 460\,c^2 + 3600\,c + 8100)(c^4 + 460\,c^3 - 1860\,c^2 - 70400\,c - 167900).$$

Finally, we have to check the suspicious factors in c. It turns out that they all are extraneous factors.

Thus a necessary condition, in terms of a, b and c, for the existence of a common zero for f_1, f_2 and f_3 is given by

$$\begin{array}{rllll} \varphi_1(a,b,c)=0 & \text{and} & \mathcal{A}\neq 0, & & \\ \text{or, } \varphi_2(a,c)=0 & \text{and} & \mathcal{A}=0 & \text{and} & a^2-12\neq 0, \\ \text{or, } \varphi_3(c)=0 & \text{and} & \mathcal{A}=0 & \text{and} & a^2-12=0. \end{array} \tag{4.16}$$

This condition is not a sufficient one, as can be seen by looking closer at one of the factors of $\varphi_3(c)$:

$$c^4+20\,c^3+460\,c^2+3600\,c+8100=\left(c^2+10\,c+90(2+\sqrt{3})\right)\left(c^2+10\,c+90(2-\sqrt{3})\right)$$

gives for the last condition row

$$\varphi_3(c)=0 \quad \text{and} \quad \mathcal{A}=0 \quad \text{and} \quad a^2-12=0 \tag{4.17}$$

valid triples (a,b,c)

$$\left\{\left(2\sqrt{3},\frac{3+\sqrt{3}}{2},-5\pm\sqrt{-90\sqrt{3}-155}\right),\ \left(2\sqrt{3},\frac{3+\sqrt{3}}{2},-5\pm\sqrt{90\sqrt{3}-155}\right)\right\},$$

but a common zero only exists for the first triples.

The condition (4.17) is not a sufficient one, because we only know that $\varphi_3(c)=0$ is a sufficient (and necessary) condition to let the system

$$f_1=0, f_2=0, f_3=0, \mathcal{A}=0, a^2-12=0 \tag{4.18}$$

have a solution for x, y, a, b. Apparently the system imposes conditions on c that were ignored. A solution for (4.18) restricts the possible choices for c: in this case to two. There are 8 solutions in total, but the condition (4.17) gives $8\cdot 2\cdot 2=32$ possible choices for a, b, c.

Note that the φ_1, φ_2 and φ_3 are related by subsequent resultants:

$$\varphi_2(a,c)=\frac{\Re_b(\varphi_1,\mathcal{A})}{a+3} \text{ and } \varphi_3(c)=27\,\Re_a(\varphi_2,a^2-12).$$

The condition with $\varphi_2(a,c)$ and $\mathcal{A}$ excludes the case $a=-3$: $\varphi(-3,0,c)=0$ for all c but a common zero for f_1, f_2 and f_3 only exists for $c=-4$.

Note that the special case $a=-3$ will not be detected easily when solving the system by computing a Gröbner basis. If we want to apply the Extension Theorem, a Gröbner basis of f_1, f_2 and f_3 with respect to the lexicographic ordering $x\succ y\succ a\succ b\succ c$ is needed. This Gröbner basis equals

$$\left\{\begin{array}{c} x-y, \quad by^2-5\,y+5, \quad 5\,ay+4\,by+15\,y-5\,a+bc-15, \\ bcy+4\,by-5\,a+bc-5\,c-35, \quad \varphi_1(a,b,c) \end{array}\right\}.$$

The derived condition is not reflected in the elimination ideal. The elimination ideal can be read off from the Gröbner basis: $\mathcal{I} \cap \mathbb{Q}[a, b, c]$ is the principal ideal generated by $\varphi_1(a, b, c)$. Hence, we know that $\varphi_1(a, b, c) = 0$ is a necessary condition.

Using the Extension Theorem we can find the sufficient condition. Specializing c gives no restrictions. Specializing b and a, we find that, in order to extend to a specialization of y, there appears a restriction on b, as the element $by^2 - 5\,y + 5$ has leading coefficient b. If this vanishes, together with other leading coefficients of polynomials in y, the Extension Theorem can not be applied. In fact, in case $b = 0$ there is the solution $(x, y) = (1, 1)$ if and only if $a = -c - 7$.

This example can be solved more easily after interchanging f_1 or f_2 with f_3. In both cases we find a 2×2 Bézout matrix. Interchanging f_1 with f_3 gives the suspicious factor $a + 3$ and the Poisson product equals $\varphi_1/(a+3)^2$. Interchanging f_2 with f_3 gives the suspicious factor b. The Poisson product is equal to φ_1/b^2. The suspicious factors $a + 3$ and b exactly represent the case when $\varphi_1(a, b, c) = 0$ is not a sufficient condition.

Continuing with the case $a = -3$, we can recompute the Poisson product after substituting $a = -3$. This gives another 2×2 Bézout matrix with determinant $(bc^2 + 20\,c + 80)/b$. The new suspicious factor is b. Substituting $b = 0$, we obtain the solution $c = -4$. As before in (4.16), we can give a necessary condition in terms of a, b and c, for the existence of a common zero for f_1, f_2 and f_3. This condition is a sufficient one:

$$\begin{array}{rclcl} \varphi_1(a,b,c) = 0 & \text{and} & a \neq -3, & & \\ \text{or, } bc^2 + 20\,c + 80 = 0 & \text{and} & a = -3 & \text{and} & b \neq 0, \\ \text{or, } c = -4 & \text{and} & a = -3 & \text{and} & b = 0. \end{array}$$

This is equivalent to

$$\varphi_1(a,b,c) = 0 \text{ and } a \neq -3, \text{ or, } bc^2 + 20\,c + 80 = 0 \text{ and } a = -3.$$

□

Finally, we look at the elimination ideal $\mathcal{I} \cap \mathbb{Q}[t_1, \ldots, t_v]$, with $\mathcal{I}$ the ideal generated by $f_1, \ldots, f_m, f_{m+1}$ over the polynomial ring $\mathbb{Q}[x_1, \ldots, x_m, t_1, \ldots, t_v]$. The system (4.15) shows that the elimination ideal does not need to be a principal ideal. Example 4.5 shows that the elimination ideal does not supply the necessary and sufficient condition on the parameters to let the system have a solution. In fact, the affine variety associated with the elimination ideal, $V(\mathcal{I} \cap \mathbb{Q}[t_1, \ldots, t_v])$, holds the so-called partial solutions. Those not extending to affine solutions of $\mathcal{I}$, i.e. points in $V(\mathcal{I})$, do extend to points in the projective closure. We refer to [27] for a comprehensive treatment of the Extension Theorem.

4.3 The infinite Poisson product

In this subsection we discuss the case of $\mathcal{I}$ being a non-zero dimensional ideal, while there still exists an eliminant. Obviously, the Poisson product becomes an infinite product, but still may evaluate to a finite value. We will concentrate on the single parameter case. The results can be extended to the multiparameter case.

Let $\mathcal{I}$ be the ideal generated by $f_1, \ldots, f_m$ over $K[x_1, \ldots, x_m]$, $K = \mathbb{Q}(t)$. $\mathcal{I}$ is of positive dimension: the set of solutions $V(\mathcal{I})$ is an infinite set. The assumption of the existence of an eliminant in t can be formulated as the assumption that the ideal $\mathcal{I}_t$, generated by $f_1, \ldots, f_m, f_{m+1}$ over $\mathbb{Q}[x_1, \ldots, x_m, t]$, contains a univariate polynomial in t.

First, one may try to choose any $f_j \in \{f_1, \ldots, f_m, f_{m+1}\}$ such that the ideal generated by $f_1, \ldots, f_{j-1}, f_{j+1}, \ldots, f_{m+1}$ is zero-dimensional. This choice is not always possible as the following example shows.

Example 4.6 *Determine the eliminant in t for the system given by*

$$\begin{aligned} f_1 &= (x-1)(x-2)(x-3), \\ f_2 &= (x-2)(xy-y-3)(x^2-3), \\ f_3 &= (x-1)(xt-2t+1)(x^2-3). \end{aligned}$$

We see that none of the ideals generated by $\{f_1, f_2\}$, $\{f_1, f_3\}$, or $\{f_2, f_3\}$ is zero-dimensional: the generators have common factors $x-2$, $x-1$ and x^2-3, respectively. However, the true eliminant in t is $\mathcal{E}(t) = t+1$, as can be derived for instance from a Gröbner basis: the system only has one solution $x = 3$, $y = \frac{3}{2}$, $t = -1$.

Let $\mathcal{I}$ be the ideal generated by f_1 and f_2 over $\mathbb{Q}(t)[x, y]$. Looking closer at the Poisson product $\Re^{f_3}(f_1, f_2)$, we can write

$$\Re^{f_3}(f_1, f_2) = \prod_{\Lambda \in V(\mathcal{I})} f_3(\Lambda) = f_3(3, \frac{3}{2}) \prod_{(2,\bar{y}) \in V(\mathcal{I})} f_3(2, \bar{y}) = f_3(3, \frac{3}{2}) = 12(t+1). \tag{4.19}$$

Except the point $(3, \frac{3}{2})$, all points in $V(\mathcal{I})$ (infinitely many) have x-coordinate 2. For $x = 2$, f_3 evaluates to 1, so the Poisson product is well defined. However, the Poisson product $\Re^{2f_3}(f_1, f_2)$ evaluates to infinity: $2f_3$ evaluates to 2 for $x = 2$. We would then be interested in the non-constant factors. Obviously, that does not help for the case $\Re^{tf_3}(f_1, f_2)$. In fact, in the latter case, there exists an eliminant in t, namely $t(t+1)$, but the ideal generated by f_1, f_2, tf_3 is not zero-dimensional. There are infinitely many solutions, but the t-component is either 0 or -1.

Utilizing the Bézout matrix still turns out to be possible: a finite subset of $\mathcal{T}$ may be used to find the factors of the eliminant. Suppose we have a finite subset of $\mathcal{T}$, $\mathcal{T}'$, containing N power products $\tau_1, \ldots, \tau_N$. Similar to the equations (4.2) we have

$$(\tau_j(X) f_{m+1}(X))_{red} = b_{j1}\tau_1(X) + \cdots + b_{jN^*}\tau_{N^*}(X), \quad (j = 1, \ldots, N) \tag{4.20}$$

with $N^* \geq N$. If we can find $\mathcal{T}'$ such that $N^* = N$, we have a square Bézout matrix (4.3) and the determinant can be evaluated. According to Theorem 3.2 there exist polynomials $P_1, \ldots, P_m, P_{m+1} \in K[x_1, \ldots, x_m]$ with

$$P_1 f_1 + \cdots + P_m f_m + P_{m+1} f_{m+1} = \det \mathcal{B}.$$

After removing all denominators in t – these can only consist of suspicious factors – we have $Q_1, \ldots, Q_m, Q_{m+1} \in \mathbb{Q}[x_1, \ldots, x_m, t]$ with

$$Q_1 f_1 + \cdots + Q_m f_m + Q_{m+1} f_{m+1} = Q \det \mathcal{B},$$

where Q is a product of suspicious factors. The right-hand side of this equation is similar to the approximate eliminant (4.9). Assuming there are no suspicious factors, we know $P_1, \ldots, P_m, P_{m+1}$ as well as $\det \mathcal{B}$ belong to $\mathbb{Q}[x_1, \ldots, x_m, t]$. In that case $\det \mathcal{B}$ is a multiple of the generator of the elimination ideal, as $\det \mathcal{B}$ belongs to the elimination ideal.

A straightforward iterative algorithm can be applied in order to find a proper $\mathcal{T}'$ with $N^* = N$. We denote the set $\mathcal{T}'$, obtained at iteration k, by $\mathcal{T}'_k$. Its cardinality will be $N_k = k$. The number of power products in the normal forms (4.20) is denoted by N^*_k. For the first iteration we use $\mathcal{T}'_1 = \{1\}$. The normal form of f_{m+1} may not be a constant, hence $N^*_1 > 1$. The (non-constant) power products in this normal form are labelled with $\tau_2, \ldots, \tau_{N^*_1}$. In the second iteration let $\mathcal{T}'_2 = \{1, \tau_2\}$. The normal form of $\tau_2 f_{m+1}$ may consist of power products that are not among $\tau_2, \ldots, \tau_{N^*_1}$, giving $N^*_2 > N^*_1$. At each iteration new power products may appear, causing N^*_k to increase. The algorithm stops in case $N^*_k = k$.

Applying this to our example, the normal form of f_3 turns out to be $(12t + 11)x - 24t - 21$. Hence, for the second iteration we use $\mathcal{T}'_2 = \{1, x\}$. The normal form of xf_3 is $(36t + 34)x - 72t - 66$. The corresponding Bézout matrix is

$$\mathcal{B} = \begin{pmatrix} 12t + 11 & -24t - 21 \\ 36t + 34 & -72t - 66 \end{pmatrix}$$

with determinant $12t + 12$. This result coincides with the Poisson product (4.19).

The question to be answered is what conditions on the polynomial system there are in order to have the algorithm terminating. At the k-th iteration, we can write

$$\mathcal{T}'_k f_{m+1} + M_k \mathcal{G} = \mathcal{B}_k \mathcal{T}'_{N^*_k}, \tag{4.21}$$

with the $\mathcal{T}'$'s and $\mathcal{G}$, the Gröbner basis of $f_1, \ldots, f_m$, in vector form. The matrix M_k is a multiplication matrix containing the normal form computation divisors. The $\mathcal{B}_k$ is the coefficient matrix of the normal forms and has order k times N^*_k. The number of rows of M_k is k and the number of columns equals the number of elements in the Gröbner basis. Assuming the ideal $\mathcal{I}_t$ contains a univariate polynomial in t, $\mathcal{E}(t)$, there exist polynomials $Q_1, \ldots, Q_m, Q_{m+1} \in \mathbb{Q}[x_1, \ldots, x_m, t]$ with

$$Q_1 f_1 + \cdots + Q_m f_m + Q_{m+1} f_{m+1} = \mathcal{E}(t). \tag{4.22}$$

Viewing this equation over the polynomial ring $\mathbb{Q}(t)[x_1, \ldots, x_m]$, let $\mathcal{T}_{\mathcal{E}(t)}$ be the set of power products in $x_1, \ldots, x_m$ of Q_{m+1}. If, for a certain k, $\mathcal{T}'_k$ coincides with $\mathcal{T}_{\mathcal{E}(t)}$, this does not mean that $\mathcal{T}'_k = \mathcal{T}'_{N^*_k}$.

Example 4.7 *Determine the eliminant in t for the system given by*

$$f_1 = tx^2 + txy - t^2 + 2tx + ty + 1, \quad f_2 = (x+1)f_1, \quad f_3 = t(1 + x + y).$$

Polynomial f_3 is in normal form with regard to the Gröbner basis of f_1 and f_2. Hence, $\mathcal{T}'_{N^*_1} = \{1, x, y\}$. The normal form of xf_3 is $-tx - ty + t^2 - 1$. This means that the normal form of $f_3 + xf_3$ is $\mathcal{E}(t) = t^2 + t - 1$ and $\mathcal{T}_{\mathcal{E}(t)} = \{1, x\}$. The proposed algorithm will not terminate: the normal form of yf_3 is $txy + ty^2 + ty$, giving $\mathcal{T}'_{N^*_3} = \{1, x, y, y^2\}$. The normal form of $y^k f_3$ is $y^k t(x + y + 1)$, giving $\mathcal{T}'_k = \{1, x, y, \ldots, y^{k-2}\}$ and $\mathcal{T}'_{N^*_k} = \{1, x, y, \ldots, y^k\}$. Note that, although an eliminant in t exists, the ideal $\mathcal{I}_t$ is not zero-dimensional. We have not found any example of a zero-dimensional $\mathcal{I}_t$ for which the algorithm did not terminate.

To guarantee the termination, the algorithm has to be modified. One of the possibilities is to add all obtained normal forms as new reducers to the Gröbner basis. In the normal form computation at iteration k, multiples of these normal forms are used in order to decrease the number of power products of $\mathcal{T}'_{N^*_k}$. By Dickson's lemma [27] we know this algorithm terminates: the ideal generated by the leading power products can only increase finitely many times. The eliminant can be found as the last non-zero normal form, assuming fraction-free normal form computations are used. The leading coefficients of the normal forms that are added to the Gröbner basis have to be considered as suspicious factors as well.

Applying this to Example 4.7, in the first iteration the normal form f_3 is added as a reducer to the Gröbner basis. The normal form of xf_3 with regard to this new basis is $t^2 + t - 1$, the normal form of yf_3 is 0. The last non-zero normal form is $t^2 + t - 1$, the eliminant we were looking for.

This algorithm is close to a Gröbner basis algorithm for $f_1, \ldots, f_m, f_{m+1}$ using fraction-free normal form computations.

Another modification to ensure termination is based on ideas from the FGLM algorithm for conversion of Gröbner bases [36]. We can have the power products of the $\mathcal{T}'$'s in increasing order with respect to a total degree ordering $\prec$ and complete, i.e. if $\tau \in \mathcal{T}'_{N^*_k}$, then $\mathcal{T}'_{N^*_k}$ contains every power product $\tau_v \prec \tau$ that is in normal form. Applying fraction-free Gaussian elimination [5] we find the eliminant as first entry in the row that is zero everywhere except in the first column. The Bézout equation (4.22) ensures termination, namely at iteration k with $\mathcal{T}_{\mathcal{E}(t)} \subset \mathcal{T}'_k$.

4.4 The u-resultant and Bézout matrix eigenvalues

As we have seen in Theorem 4.3 and 4.5 the eigenvalues of the Bézout matrix correspond one-to-one to the factors in the Poisson product. As the u-resultant is a Poisson product, it can be evaluated via a Bézout matrix. This is done in the definition of the true eliminant. To determine the components of the solutions the factorization of the u-resultant is used. This is equivalent to an eigenvalue problem of a specific Bézout matrix. This has been considered by several authors [2, 4, 56, 63, 64, 81, 97] in recent years. In this section we will present some of the results of this closely related application of the Bézout matrix.

Let $\mathcal{I} \subset K[x_1, \ldots, x_m]$ be a zero-dimensional ideal given by a set of generators $\mathcal{F} = \{f_1, \ldots, f_s\}$ from $K[x_1, \ldots, x_m]$, K the field of coefficients. Usually for zero-dimensional ideals s equals m. For any indeterminate x_j we can evaluate the Poisson product of the system

$$f_1 = 0, \ldots, f_s = 0, x_j - \lambda = 0.$$

This is possible because the ideal $\mathcal{I}$ is zero-dimensional. The result is a polynomial in λ that is equivalent to the true eliminant in x_j. The roots of this eliminant are the eigenvalues of the Bézout matrix associated with the Poisson product of the system

$$f_1 = 0, \ldots, f_s = 0, x_j = 0.$$

This Bézout matrix is called the x_j-multiplication table and is usually denoted by A_{x_j}. From these tables the u-resultant can be constructed as (see equation (1.29))

$$\det(u_1 A_{x_1} + \cdots + u_m A_{x_m}).$$

An important observation is that the matrices A_{x_j} are commutative, i.e. for every pair (j, k) we have $A_{x_j} A_{x_k} = A_{x_k} A_{x_j}$. The Poisson product (4.5) for a certain $f_{m+1}(x_1, \ldots, x_m) \in K[x_1, \ldots, x_m]$ may be determined by evaluating the polynomial over the matrices: $\det f_{m+1}(A_{x_1}, \ldots, A_{x_m})$.

The following property of commutative matrices [64] is used:

Lemma 4.1 *For commutative matrices A and B, let λ be an eigenvalue of A with eigenvector v. If the geometric multiplicity of λ is 1, then v also is an eigenvector of B.*

Proof. Bv belongs to the eigenspace of λ: $A(Bv) = B(Av) = \lambda(Bv)$. Because the geometric multiplicity of λ is 1, we have that Bv is a multiple of v. □

Assume we have, for a matrix A_{x_j}, an eigenvalue λ of geometric multiplicity 1, with eigenvector v. The eigenvalue λ is the x_j-component of a solution of the system. The other components are determined using the fact that v is a eigenvector of all matrices $A_{x_1}, \ldots, A_{x_m}$. For the eigenvalue λ_k of A_{x_k} with eigenvector v, we have to solve $A_{x_k} v = \lambda_k v$. If v_1, the first element of v, is non-zero, only the first row of A_{x_k}, $A_{x_k,1}$, is needed:

$$\lambda_k = \frac{1}{v_1} A_{x_k,1} v.$$

Any row l can be used for this purpose as long as $v_l \neq 0$.

If every eigenvalue of a matrix has geometric multiplicity 1, the matrix is called nonderogatory [64]. If the matrix A_{x_j} is nonderogatory, the algebraic multiplicities of the eigenvalues are the multiplicities of the corresponding solutions. All components of the solutions can be determined by the eigenvectors and one row of each multiplication table $A_{x_1}, \ldots, A_{x_{j-1}}, A_{x_{j+1}}, \ldots, A_{x_m}$.

In case all the matrices A_{x_j}, $j = 1, \ldots, m$, are derogatory and diagonizable, a nonderogatory matrix A, commuting with $A_{x_1}, \ldots, A_{x_m}$, can be found as $A = c_1 A_{x_1} + \cdots + c_m A_{x_m}$ for almost every choice of $c_1, \ldots, c_m$ [97]. The radical of the ideal is used [63, 64] for the case of derogatory and non-diagonizable matrices. The multiplication tables have to be recomputed relative to the radical of the ideal. The multiplicities have to be retrieved from the algebraic multiplicity of the eigenvalue of the original multiplication table.

For Example 4.1 one multiplication table is

$$A_x = \begin{pmatrix} 0 & -1 & 1 & 0 \\ 1 & 3 & -1 & 2 \\ 5 & 5 & 0 & 4 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

with the eigenvalues -1 and 2 of multiplicity 1 and eigenvalue 1 of algebraic multiplicity 2. A_x is nonderogatory: all eigenvalues are of geometric multiplicity 1. The corresponding eigenvectors are $(-1, 0, 1, 1)^T$, $(2, -2, 2, 1)^T$ and $(1, -2, -1, 1)^T$. Knowing the columns are labelled with x, ϵ, y and 1, we can read off the solutions from the eigenvectors directly. Starting with the multiplication table A_ϵ gives eigenvalue 0 of multiplicity 1 and eigenvalue -2 of algebraic multiplicity 3. The geometric multiplicity of the latter is 2. Hence, A_ϵ is derogatory. Other nonderogatory multiplication tables are, for instance, $A_{\epsilon+3y}$ and $A_{x+y+\epsilon}$.

Note that the difference with the Bézout resultant method is that for the resultant ϵ is considered as a parameter of the system. With the resultant we search for values of ϵ for which the Bézout matrix – the multiplication table $A_{x^2-y+\epsilon}$ relative to the ideal generated by f_1 and f_2 over $\mathbb{Q}(\epsilon)[x, y]$ – becomes singular. If for this eigenvalue 0 the eigenspace is of dimension 1, we can evaluate the components of the solution. For $\epsilon = 0$, the multiplication table is

$$A_{x^2-y+\epsilon}|_{\epsilon=0} = \frac{1}{7}\begin{pmatrix} 11 & 4 & -3 & 10 \\ 3 & 10 & 3 & -10 \\ 6 & -1 & 13 & -6 \\ 6 & -1 & -1 & 8 \end{pmatrix}.$$

The eigenvector corresponding to eigenvalue 0 is $(-1, 1, 1, 1)^T$. As the columns are labelled with x, y, y^2 and 1, we obtain the solution $(x, y, \epsilon) = (-1, 1, 0)$. For $\epsilon = -2$, the matrix has rank 1. The eigenspace corresponding to eigenvalue 0 is of dimension 2. Hence, the matrix is non-diagonizable and derogatory. The same is valid for the multiplication table relative to the radical ideal. This is due to the fact that $x^2 - y + \epsilon$ vanishes at two different solutions of f_1 and f_2. We can obtain the components as a linear combination of the eigenvectors $(0, -4, -2, 1)^T$ and $(1, 3, 3, 0)^T$. From the condition $a_1(0, -4, -2, 1)^T + a_2(1, 3, 3, 0)^T = (x, y, y^2, 1)^T$ it follows that $a_1 = 1$ and $a_2 \in \{1, 2\}$, giving $(x, y) \in \{(1, -1), (2, 2)\}$. The matrix $A_{x^2-y+\epsilon}$ gives no information about the fact that $(x, y, \epsilon) = (1, -1, -2)$ is a multiple solution. The multiplication table A_x has to be used additionally.

4.5 Experiments

In this section we report on experiments performed in order to compare the Bézout method using the Gröbner basis approach to other methods for solving a polynomial system. Most of the polynomial systems we discuss have been used in the literature in order to make comparisons between various solving methods. Hence, these systems can be considered as a kind of benchmark.

For each of the polynomial systems we make the following experiment. Firstly, the eliminant using the Bézout matrix is computed. The intermediate steps we distinguish are the computation of the Gröbner basis, the normal form computations, the construction of the matrix and the evaluation of the determinant. The Gröbner bases are computed with respect to a total degree ordering.

Secondly, we determine the Gröbner bases with respect to a lexicographic and a block-lexicographic ordering, both containing the eliminant. The block-lexicographic ordering consists of two blocks: one block with a total degree ordering on the indeterminates to be eliminated, another block ordering the one or more indeterminates left.

In order to obtain the Gröbner basis with respect to the lexicographic ordering, we apply the FGLM method in which we can distinguish between the computation of the Gröbner basis of the ideal with respect to a total degree ordering and the process of constructing an eliminant by basis conversion (in fact we apply the Asir algorithm that uses modular arithmetic with lifting techniques [67]).

Thirdly, the Dixon resultant is determined. This consists of the construction of the Dixon matrix and the evaluation of the determinant.

We have used Maple V release 5.1 for most of the computations. The Maple code for the Bézout and the Dixon resultant can be found in the appendices. We used Mathematica and Asir [66] to compare the different Computer Algebra systems and for those computations for which Maple fails due to an 'out of memory' or an 'object too large' exception.

The Gröbner basis with respect to a block-lexicographic ordering was computed by Mathematica, because the block-lexicographic ordering in Maple, package 'Groebner', does only work for very small systems. This seems not to be due to the 'Groebner' package: simulating the ordering by substitution of the relevant indeterminates (the ones to be eliminated) by their n-th power (n bigger than the degree of the eliminant), gives similar results with the 'grobner' package. For the Gröbner basis computations we use the notation $\mathcal{G}(f_1, \ldots, f_m)$ followed by the ordering.

All computation times are in seconds. The times for Mathematica are indicated by $^{\mathrm{M}}$, for Asir by $^{\mathrm{A}}$. The sign ∞ means that the computation did not succeed. Some computations could not finish due to memory shortage. Others were interrupted after running for more than one day.

We have used a single processor – 450 Mhz Pentium III – machine with 256 MB of internal memory and running the operating system Linux Debian 2.2. Asir was run on a 400 Mhz Pentium II machine with 128 MB of internal memory and running Windows NT.

Cyclic 5-roots system (C5)

The cyclic 5-roots system has often been used for comparing various algorithms for solving a polynomial system. We refer to [15] for more information on the cyclic n-roots systems. The cyclic 5-roots system is the first in the series of cyclic n-roots systems that is considered to be hard to solve. We want to find an eliminant in one of the indeterminates of the system. Because of the symmetry of the system, the eliminants in each indeterminate are equivalent. The system is given by

$$\begin{aligned}
f_1 &= x_1 + x_2 + x_3 + x_4 + x_5, \\
f_2 &= x_1x_2 + x_2x_3 + x_3x_4 + x_4x_5 + x_5x_1, \\
f_3 &= x_1x_2x_3 + x_2x_3x_4 + x_3x_4x_5 + x_4x_5x_1 + x_5x_1x_2, \\
f_4 &= x_1x_2x_3x_4 + x_2x_3x_4x_5 + x_3x_4x_5x_1 + x_4x_5x_1x_2 + x_5x_1x_2x_3, \\
f_5 &= x_1x_2x_3x_4x_5 - 1.
\end{aligned}$$

We will compute the eliminant in x_5. The computation of the Gröbner basis of $f_1, \ldots, f_4$ with respect to the total degree ordering $x_1 \succ x_2 \succ x_3 \succ x_4$ does not give suspicious factors. The set $\mathcal{T}$ has 18 elements, the order of the Bézout matrix. It turns out that each diagonal element is of the form $ax_5^5 - 1$ with a an integer number. The other entries are either 0 or of the form bx_5^c with b and c integer numbers and $c \geq 2$. The determinant evaluates to the true eliminant in

x_5:

$$\det \mathcal{B} = \mathcal{E}(x_5) = (x_5-1)^{10}(x_5^2+3\,x_5+1)^2(x_5^4-4\,x_5^3+6\,x_5^2+x_5+1)^2 \\ (x_5^4+x_5^3+6\,x_5^2-4\,x_5+1)^2(x_5^4+x_5^3+x_5^2+x_5+1)^{10}. \tag{4.23}$$

In fact, the system has 70 solutions, the degree of the true eliminant. The computation time is:

$$\left.\begin{array}{ll} \mathcal{G}(f_1,\ldots,f_4),\ \succ_{\deg}(x_1,\ldots,x_4) & 1.04 \\ \text{18 normal forms} & 1.29 \\ \text{determinant evaluation} & 3.27 \end{array}\right\} \text{ total } 5.63$$

The computation times for the various Gröbner bases are:

$\mathcal{G}(f_1,\ldots,f_5),\ \succ_{\deg}(x_1,\ldots,x_4,x_5)$	3.64 [0.92^{M}, 0.09^{A}]
$\mathcal{G}(f_1,\ldots,f_5),\ \succ_{\mathrm{lex}}(x_1,\ldots,x_4,x_5)$	∞[0.60^{M}, FGLM:0.28^{A}]
$\mathcal{G}(f_1,\ldots,f_5),\ x_1,\ldots,x_4 \succ_{\mathrm{lex}} x_5$	6.19 [0.35^{M}, 0.08^{A}]

The order of the Dixon resultant matrix is 22. The computation time is:

$$\left.\begin{array}{ll} \text{matrix construction} & 0.28 \\ \text{determinant evaluation} & 1.95 \end{array}\right\} \text{ total } 2.69$$

Modified cyclic 5-roots system (MC5)

We modify the cyclic 5-roots system by replacing the term $x_1x_2x_3x_4$ in f_4 by $x_2x_3x_4$. This gives the polynomials

$$\begin{aligned} f_1 &= x_1+x_2+x_3+x_4+x_5, \\ f_2 &= x_1x_2+x_2x_3+x_3x_4+x_4x_5+x_5x_1, \\ f_3 &= x_1x_2x_3+x_2x_3x_4+x_3x_4x_5+x_4x_5x_1+x_5x_1x_2, \\ f_4 &= x_2x_3x_4+x_2x_3x_4x_5+x_3x_4x_5x_1+x_4x_5x_1x_2+x_5x_1x_2x_3, \\ f_5 &= x_1x_2x_3x_4x_5-1. \end{aligned}$$

This system turns out to be much harder to solve because a part of the symmetry is lost. Again, we will determine the eliminant in x_5. The suspicious factors appearing in the Gröbner basis computation of $f_1,\ldots,f_4$ with respect to the total degree ordering $x_1 \succ x_2 \succ x_3 \succ x_4$ are x_5 and $2\,x_5-1$.

The set $\mathcal{T}$ has 15 elements. Hence the Bézout matrix is smaller than for the cyclic 5-roots system. However, the matrix entries are more complex. Some entries have denominators that are expressions in the suspicious factors. The numerators of the matrix entries are polynomials in x_5 with up to 20 terms of degree up to 22.

The determinant evaluates to the true eliminant divided by x_5^5 and a constant factor:

$$\det \mathcal{B} = 1/16384\, x_5^{-5}(x_5-1)^4(x_5^2+3\,x_5+1)(x_5^4+x_5^3+x_5^2+x_5+1)$$
$$({x_5}^{54}+6\,{x_5}^{53}+20\,{x_5}^{52}+45\,{x_5}^{51}+85\,{x_5}^{50}+433\,{x_5}^{49}+1848\,{x_5}^{48}+5696\,{x_5}^{47}+$$
$$12645\,{x_5}^{46}+24375\,{x_5}^{45}+59980\,{x_5}^{44}+167775\,{x_5}^{43}+400993\,{x_5}^{42}+799771\,{x_5}^{41}$$
$$+1373650\,{x_5}^{40}+2011168\,{x_5}^{39}+2403563\,{x_5}^{38}+2182810\,{x_5}^{37}+1048773\,{x_5}^{36}$$
$$-1163944\,{x_5}^{35}-4294962\,{x_5}^{34}-7630447\,{x_5}^{33}-10273047\,{x_5}^{32}-11427860\,{x_5}^{31}$$
$$-10489727\,{x_5}^{30}-7328421\,{x_5}^{29}-2689631\,{x_5}^{28}+2396008\,{x_5}^{27}+6825673\,{x_5}^{26}$$
$$+9512440\,{x_5}^{25}+9888418\,{x_5}^{24}+8254788\,{x_5}^{23}+5149399\,{x_5}^{22}+1527518\,{x_5}^{21}$$
$$-1483402\,{x_5}^{20}-3291731\,{x_5}^{19}-3853101\,{x_5}^{18}-3229132\,{x_5}^{17}-2000926\,{x_5}^{16}$$
$$-849057\,{x_5}^{15}-32140\,{x_5}^{14}+386635\,{x_5}^{13}+376189\,{x_5}^{12}+188483\,{x_5}^{11}$$
$$+34614\,{x_5}^{10}-59750\,{x_5}^{9}-90605\,{x_5}^{8}-57385\,{x_5}^{7}-22781\,{x_5}^{6}$$
$$-7657\,{x_5}^{5}-1615\,{x_5}^{4}+525\,{x_5}^{3}-150\,{x_5}^{2}+20\,x_5-1).$$

Note the difference with (4.23): the highest degree of irreducible factors increased from 4 to 54.

The computation time is roughly two and a half times longer than for the cyclic 5-roots system:

$$\left.\begin{array}{ll} \mathcal{G}(f_1,\ldots,f_4),\ \succ_{\deg}(x_1,\ldots,x_4) & 2.85 \\ \text{15 normal forms} & 4.83 \\ \text{determinant evaluation} & 6.68 \end{array}\right\} \text{ total } 14.6$$

The determinant evaluation was performed on a fraction-free matrix because this this gives much better performance [5]. Note that, apart from rational numbers, fractions in the parameters of the system can only appear in the final result due to suspicious factors.

Computing the eliminant in either x_1, x_2, x_3 or x_4 gives no suspicious factors. This means that the matrix entries have no denominators. The total computation time is also approximately 30 percent less. The size of the Bézout matrices for computing the eliminant in x_2 and x_3 is 17, and for x_1 and x_4 the size is 18. In each case the result is a degree 64 polynomial with an irreducible factor of degree 54.

The computation times for the Gröbner bases are more than five times longer than for the cyclic 5-roots system:

$\mathcal{G}(f_1,\ldots,f_5),\ \succ_{\deg}(x_1,\ldots,x_4,x_5)$	33.1 [4.68^{M}, 0.92^{A}]
$\mathcal{G}(f_1,\ldots,f_5),\ \succ_{\mathrm{lex}}(x_1,\ldots,x_4,x_5)$	∞[∞^{M}, FGLM:20.1^{A}]
$\mathcal{G}(f_1,\ldots,f_5),\ x_1,\ldots,x_4 \succ_{\mathrm{lex}} x_5$	∞[164^{M}, 18.5^{A}]

The order of the Dixon resultant matrix is 22, like for the cyclic 5-roots system. Compared to the cyclic 5-roots system, the computation time is only a little more:

$$\left.\begin{array}{ll} \text{matrix construction} & 0.26 \\ \text{determinant evaluation} & 2.38 \end{array}\right\} \text{ total } 3.28$$

Cyclic 6-roots system (C6)

This system is given by

$$\begin{array}{rcl} f_1 & = & x_1+x_2+x_3+x_4+x_5+x_6, \\ f_2 & = & x_1x_2+x_2x_3+x_3x_4+x_4x_5+x_5x_6+x_6x_1, \\ f_3 & = & x_1x_2x_3+x_2x_3x_4+x_3x_4x_5+x_4x_5x_6+x_5x_6x_1+x_6x_1x_2, \\ f_4 & = & x_1x_2x_3x_4+x_2x_3x_4x_5+x_3x_4x_5x_6+x_4x_5x_6x_1+x_5x_6x_1x_2+x_6x_1x_2x_3, \\ f_5 & = & x_1x_2x_3x_4x_5+x_2x_3x_4x_5x_6+x_3x_4x_5x_6x_1+x_4x_5x_6x_1x_2 \\ & & +x_5x_6x_1x_2x_3+x_6x_1x_2x_3x_4, \\ f_6 & = & x_1x_2x_3x_4x_5x_6-1. \end{array}$$

Applying the Bézout resultant, the main difference with the cyclic 5-roots system is that the ideal generated by $f_1,\ldots,f_5$, considering x_6 as a parameter, is of positive dimension. However, it turns out that the Poisson product is defined, as for all but finitely many solutions f_6 evaluates to 1.

We can apply the algorithm from Section 4.3 and we find an order 53 Bézout matrix (the same procedure for the cyclic 7-roots system results in an order 272 matrix). There are no suspicious factors. The determinant of the Bézout matrix gives the true eliminant, a degree 156 polynomial with irreducible factors of degree 1, 2, 4 and 8. The computation time is:

$\mathcal{G}(f_1,\ldots,f_5)$, $\succ_{\deg}(x_1,\ldots,x_5)$	26.8	total 2476
53 normal forms	51.1	
determinant evaluation	2398	

Interchanging the polynomials, e.g. for computing $\Re^{f_5}(f_1,\ldots,f_4,f_6)$, gives a zero-dimensional ideal. However, suspicious factors appear and the Bézout matrix entries become more complex. The computation time increases by more than a factor 10. This demonstrates how the arithmetic operations due to suspicious factors influence the performance of the method.

The computation times for the other Gröbner bases are:

$\mathcal{G}(f_1,\ldots,f_6)$, $\succ_{\deg}(x_1,\ldots,x_5,x_6)$	$\infty[73.4^{\mathrm{M}}, 2.05^{\mathrm{A}}]$
$\mathcal{G}(f_1,\ldots,f_6)$, $\succ_{\mathrm{lex}}(x_1,\ldots,x_5,x_6)$	$\infty[\infty^{\mathrm{M}}$, FGLM:$3.51^{\mathrm{A}}]$
$\mathcal{G}(f_1,\ldots,f_6)$, $x_1,\ldots,x_5 \succ_{\mathrm{lex}} x_6$	$\infty[24.3^{\mathrm{M}}, 2.90^{\mathrm{A}}]$

Gröbner basis computations with respect to the total degree ordering have been reported even for the cyclic 9-roots system by using arithmetic modulo a prime number [35].

The Dixon resultant matrix could not be constructed by Maple because of intermediate expressions that are too large. We constructed the Dixon matrix by Asir. The matrix is of order 86 and singular. The computation time is:

matrix construction	42.8^{A}	total 44.5^{A}
determinant evaluation (with result 0)	1.74^{A}	

Katsura-5 system (K5)

The Katsura-m systems occur in the determination of truncated Laurent series [16] and are given by the polynomials

$$\begin{aligned} f_1 &= x_0 - 1 + 2\sum_{j=1}^{m} x_j, \\ f_{k+2} &= x_k - \sum_{j=-m}^{m} x_{|k-j|}x_{|j|} \quad (k = 0, \ldots, m-1). \end{aligned}$$

We focus at computing the eliminant in x_0. No suspicious factors appear. The Bézout matrix is of order 2^{m-1} and its determinant equals the true eliminant in x_0. The eliminant is of degree 2^m and has factors $x_0 - 1$, $3\,x_0 - 1$ and an irreducible polynomial of degree $2^m - 2$. For example, for the Katsura-3 system given by

$$\begin{aligned} f_1 &= x_0 + 2\,x_1 + 2\,x_2 + 2\,x_3 - 1, \\ f_2 &= x_0 - x_0^2 - 2\,x_3^2 - 2\,x_2^2 - 2\,x_1^2, \\ f_3 &= x_1 - 2\,x_1x_0 - 2\,x_2x_3 - 2\,x_1x_2, \\ f_4 &= x_2 - 2\,x_2x_0 - 2\,x_1x_3 - x_1^2, \end{aligned}$$

the eliminant in x_0 equals

$$(3\,x_0-1)(x_0-1)(18711\,x_0^6-55728\,x_0^5+67743\,x_0^4-42872\,x_0^3+14805\,x_0^2-2616\,x_0+181).$$

The computation time for Katsura-5 is:

$\mathcal{G}(f_1, \ldots, f_5)$, $\succ_{\deg}(x_1, \ldots, x_4)$	4.80	total 15.5
16 normal forms	2.19	
determinant evaluation	8.42	

The computation times for the Gröbner bases are:

$\mathcal{G}(f_1, \ldots, f_5)$, $\succ_{\deg}(x_1, x_2, x_3, x_4, x_0)$	8.99 [4.31^{M}, 0.52^{A}]
$\mathcal{G}(f_1, \ldots, f_5)$, $\succ_{\mathrm{lex}}(x_1, x_2, x_3, x_4, x_0)$	∞[∞^{M}, FGLM:6.33^{A}]
$\mathcal{G}(f_1, \ldots, f_5)$, $x_1, x_2, x_3, x_4 \succ_{\mathrm{lex}} x_0$	∞[19.0^{M}, 7.82^{A}]

Constructing the Dixon resultant matrix with Maple fails: the number of terms in the intermediate expressions – polynomials in 10 indeterminates – is too large. Asir could handle these large expressions without problems. The Dixon matrix turns out to be non-square: 83×84. The time for the construction is 39.0^{A} seconds.

The computation times for the Bézout method applied to the Katsura-6 and Katsura-7 systems are 451 and 46121 seconds, respectively.

Speer system (SP)

This is a system from chemistry due to Speer [39]:

$$\begin{aligned} f_1 &= (4 + 2\,(x_1 + x_2 + x_3 + x_4) - 8\,x_1)(x_1^2 + x_2^2 + x_3^2 + x_4^2) - x_2x_3x_4 + x_2 + x_4, \\ f_2 &= (4 + 2\,(x_1 + x_2 + x_3 + x_4) - 8\,x_2)(x_1^2 + x_2^2 + x_3^2 + x_4^2) - x_1x_3x_4 + x_1 + x_3, \\ f_3 &= (4 + 2\,(x_1 + x_2 + x_3 + x_4) - 8\,x_3)(x_1^2 + x_2^2 + x_3^2 + x_4^2) - x_1x_2x_4 + x_2 + x_4, \\ f_4 &= (4 + 2\,(x_1 + x_2 + x_3 + x_4) - 8\,x_4)(x_1^2 + x_2^2 + x_3^2 + x_4^2) - x_1x_2x_3 + x_1 + x_3. \end{aligned}$$

We determine the eliminant in x_4. The Gröbner basis of f_1, f_2, f_3 with respect to the total degree ordering $x_1 \succ x_2 \succ x_3$ gives the suspicious factor x_4. In fact all the elements in the Gröbner basis have the factor x_4 in their leading coefficient. The determinant of the Bézout matrix equals

$$\begin{gathered}
1/227864434916783554 56\, x_4^{-11}(x_4^2 - 2\,x_4 + 2)(x_4^2 + 2\,x_4 + 2) \\
(x_4^2 - 16\,x_4 - 2)(11\,x_4^2 - 11\,x_4 + 6)(215\,x_4^4 - 32\,x_4^3 + 96\,x_4^2 + 160\,x_4 + 128)^2 \\
(144\,x_4^4 - 96\,x_4^3 + 41\,x_4^2 + 96\,x_4 + 144)^2(264\,x_4^4 - 264\,x_4^3 + 95\,x_4^2 + 528\,x_4 + 1056)^2 \\
(184900\,x_4^8 + 104920\,x_4^7 - 121192\,x_4^6 + 777968\,x_4^5 - 4636676\,x_4^4 \\
+3575172\,x_4^3 + 5573487\,x_4^2 - 32261832\,x_4 + 36639441).
\end{gathered}$$

The system has solutions for $x_4 = 0$. Hence, the numerator of the determinant does not contain every factor of the eliminant. The computation time is:

$\mathcal{G}(f_1, \ldots, f_3)$, $\succ_{\deg}(x_1, \ldots, x_3)$	2.01	total 13.4
15 normal forms	3.28	
determinant evaluation	8.13	

The computation times for the other Gröbner bases are:

$\mathcal{G}(f_1, \ldots, f_4)$, $\succ_{\deg}(x_1, x_2, x_3, x_4)$	18.7 [2.36^{M},0.83^{A}]
$\mathcal{G}(f_1, \ldots, f_4)$, $\succ_{\mathrm{lex}}(x_1, x_2, x_3, x_4)$	∞[∞^{M}, FGLM:9.20^{A}]
$\mathcal{G}(f_1, \ldots, f_4)$, $x_1, x_2, x_3 \succ_{\mathrm{lex}} x_4$	∞[3968^{M}, 18.8^{A}]

The Dixon resultant matrix could not be constructed by Maple because the intermediate expressions become too large. With Asir the 43×43 Dixon matrix could be constructed. This matrix turns out to be singular. The computation time is:

matrix construction	3.69^{A}	total 65.0^{A}
determinant evaluation (with result 0)	61.3^{A}	

Implicitization problem (IM)

This system is a parameterization of a bicubic surface and is due to Hoffmann [48]. One recognizes the definition of x, y and z as rational functions in s and t:

$$\begin{aligned}
f_1 &= t^3(-85\,s^3 - 55\,s^2 - 37\,s - 35) + t^2(97\,s^3 + 50\,s^2 + 79\,s + 56) \\
&\quad +t(49\,s^3 + 63\,s^2 + 57\,s - 59) + 45\,s^3 - 8\,s^2 - 93\,s + 92 \\
&\quad -x(t^3(-53\,s^3 + 85\,s^2 + 49\,s + 78) + t^2(17\,s^3 + 72\,s^2 - 99\,s - 85) \\
&\quad +t(-86\,s^3 + 30\,s^2 + 80\,s + 72) + 66\,s^3 - 29\,s^2 - 91\,s - 53), \\
f_2 &= t^3(43\,s^3 - 62\,s^2 + 77\,s + 66) + t^2(54\,s^3 - 5\,s^2 + 99\,s - 61) \\
&\quad +t(-50\,s^3 - 12\,s^2 - 18\,s + 31) - 26\,s^3 - 62\,s^2 + s - 47 \\
&\quad -y(t^3(-53\,s^3 + 85\,s^2 + 49\,s + 78) + t^2(17\,s^3 + 72\,s^2 - 99\,s - 85) \\
&\quad +t(-86\,s^3 + 30\,s^2 + 80\,s + 72) + 66\,s^3 - 29\,s^2 - 91\,s - 53), \\
f_3 &= t^3(-91\,s^3 - 47\,s^2 - 61\,s + 41) + t^2(-58\,s^3 - 90\,s^2 + 53\,s - 1) \\
&\quad +t(94\,s^3 + 83\,s^2 - 86\,s + 23) - 84\,s^3 + 19\,s^2 - 50\,s + 88 \\
&\quad -z(t^3(-53\,s^3 + 85\,s^2 + 49\,s + 78) + t^2(17\,s^3 + 72\,s^2 - 99\,s - 85) \\
&\quad +t(-86\,s^3 + 30\,s^2 + 80\,s + 72) + 66\,s^3 - 29\,s^2 - 91\,s - 53).
\end{aligned}$$

The implicitization eliminates s and t from the system in order to obtain one equation in x, y and z. This system turns out to be very hard to solve: the coefficient manipulations in the resultant computations are operations on trivariate polynomials. Because of too large expressions Maple did not succeed in determining all the normal forms. Maple also failed to construct the Dixon matrix.

With Asir we were able to construct the matrices. Both the Bézout and the Dixon matrix are of order 18 and contain polynomials in x, y and z with coefficients up to order 150. Suspicious factors in x and y with degree up to 9 appear. The computation time for the Bézout matrix is 352^{A} seconds of which 346^{A} is needed for the normal form computations. The Dixon matrix was constructed much faster, in 0.33^{A} seconds. For both matrices we failed to evaluate the determinant. This suggests resultants are not suited to implicitization process and other methods that apply directly to these kind of problems are desirable [49].

The implicitization of the more simple surface [48] whose parametrization is given by

$$\begin{aligned} x &= 3\,t\,(t-1)^2 + (s-1)^3 + 3\,s, \\ y &= 3\,s(s-1)^2 + t^3 + 3\,t, \\ z &= -3\,s(s^2 - 5\,s + 5) - 3(s^3 + 6\,s^2 - 9\,s + 1)t^2 \\ &\quad + t(6\,s^3 + 9\,s^2 - 18\,s + 3) - 3\,s(s-1) \end{aligned}$$

– eliminating s and t – was completed in 9.27^{A} seconds with a Bézout matrix of order 9. The computation time for the Dixon resultant (Dixon matrix of order 15) was 31.5^{A} seconds.

Geometric problem (G)

The following system is the algebraic form of Heymann's geometric problem related to the question whether a triangle can be constructed from its angle bisectors [75, 77]:

$$\begin{aligned} f_1 &= a_i^2(b+c)^2 - cb(c+b-a)(c+b+a), \\ f_2 &= a_e^2(c-b)^2 - cb(a+b-c)(c-b+a), \\ f_3 &= b_e^2(c-a)^2 - ac(a+b-c)(c+b-a). \end{aligned}$$

The objective is to eliminate b and c. The Gröbner basis of f_1 and f_2 with respect to the total degree ordering $b \succ c$ gives the suspicious factor $a_i^2 + a_e^2$ (corresponding to a degenerate case of the original geometric problem). The Bézout matrix is of order 12. The computation time is:

$\mathcal{G}(f_1, f_2, f_3), \succ_{\deg}(b, c)$	0.19	
12 normal forms	0.54	total 14.0
determinant evaluation	13.3	

The numerator of the determinant is $a^8 b_e^8$ times an irreducible polynomial with 330 terms and of total degree 40. The order of the coefficients is up to 13. The suspicious factor appears in the denominator with power 4.

The Gröbner bases could only be constructed with respect to the total degree ordering. The computation times are:

$\mathcal{G}(f_1, f_2, f_3)$, $\succ_{\deg}(b, c, a, a_i, a_e, b_e)$	6.21 [∞^{M},0.15$^{\mathrm{A}}$]
$\mathcal{G}(f_1, f_2, f_3)$, $\succ_{\mathrm{lex}}(b, c, a, a_i, a_e, b_e)$	∞[∞^{M}, FGLM:∞^{A}]
$\mathcal{G}(f_1, f_2, f_3)$, $b, c \succ_{\mathrm{lex}} a, a_i, a_e, b_e$	∞[∞^{M}, ∞^{A}]

The Dixon resultant matrix is of order 13×14. Minors can be used to obtain the resultant [75].

Dense system $m = 3$ (D3)

This is a polynomial system consisting of 4 dense polynomials in 4 indeterminates. It was first mentioned in [14]:

$$\begin{aligned}
f_1 &= x_1^2 + x_1x_2 + x_2^2 - 2\,x_1x_3 - 4\,x_2x_3 + 3\,x_3^2 - 3\,x_1x_4 + 2\,x_2x_4 + x_4^2 \\
&\quad -3\,x_1 - 2\,x_2 + 3\,x_3 - 2\,x_4 - 2, \\
f_2 &= 2\,x_1^2 - x_1x_2 + x_2^2 - x_1x_3 - x_2x_3 - 6\,x_3^2 - x_1x_4 + x_2x_4 - 5x_3x_4 - 3\,x_4^2 \\
&\quad -5\,x_1 + x_2 + 5\,x_3 + 2\,x_4 + 5, \\
f_3 &= x_1^3 - x_1^2x_3 + x_1x_2x_3 + x_1x_2x_4 - 5\,x_1x_3^2 + x_1x_4^2 + x_2^3 - 5\,x_2^2x_3 + 7\,x_2x_3^2 \\
&\quad -3\,x_3^3 - 5\,x_3^2x_4 + 2\,x_4^3 + x_1^2 - x_2^2 + 2\,x_3^2 - 3\,x_4^2 - 3\,x_1x_2 + 2\,x_1x_3 - 3\,x_1x_4 \\
&\quad -2\,x_3x_4 - x_1 + x_2 + 11\,x_3 - 2\,x_4 - 3, \\
f_4 &= -x_1^3 + 6\,x_1^2x_2 - x_1^2x_3 - 12\,x_1x_2^2 - 4\,x_1x_2x_3 + 6\,x_1x_2x_4 + 5\,x_1x_3^2 - x_1x_3x_4 \\
&\quad +11\,x_1x_4^2 + 6\,x_2^3 + 6\,x_2^2x_3 + 4\,x_2x_3^2 - 7\,x_2^2x_4 + 15\,x_3^3 + 4\,x_4^3 + 3\,x_1^2 + 2\,x_1x_2 \\
&\quad +2\,x_2^2 + 2\,x_2x_4 - x_3^2 - x_3x_4 + 5\,x_4^2 - 35\,x_1 - 14\,x_2 + 4\,x_3 - 10\,x_4 - 15.
\end{aligned}$$

We compute the eliminant in x_4. The suspicious factor $85\,x_4 + 137$ appears in the Gröbner basis computation of f_1, f_2, f_3 with respect to the total degree ordering $x_1 \succ x_2 \succ x_3$. The Bézout matrix is of order 11. The degree of the polynomial entries (when the matrix is made fraction-free) in x_4 is up to 9. The determinant equals

$$\frac{x_4(x_4-1)(x_4-2)\mathcal{X}(x_4)}{272097792(85\,x_4+137)^3}$$

with $\mathcal{X}(x_4)$ an irreducible polynomial of degree 33 with all coefficients of order ≥ 37. The numerator is the true eliminant. The system has 10 real zeroes including $(0, 1, 1, 1)$, $(-1, 0, -1, 0)$ and $(1, 2, 0, 2)$. For $x_4 = -\frac{137}{85}$ the system does not have a solution. This can be concluded immediately: the non-suspicious factors have degree 36, the maximum number of solutions for the system. The computation time is:

$\mathcal{G}(f_1, \ldots, f_3)$, $\succ_{\deg}(x_1, \ldots, x_3)$	0.84	total 5.71
11 normal forms	1.68	
determinant evaluation	3.08	

The computation times for the various Gröbner bases are:

$\mathcal{G}(f_1,\ldots,f_4)$, $\succ_{\rm deg}(x_1,x_2,x_3,x_4)$	9.92 [$2.86^{\rm M}$, $0.67^{\rm A}$]
$\mathcal{G}(f_1,\ldots,f_4)$, $\succ_{\rm lex}(x_1,x_2,x_3,x_4)$	∞[$\infty^{\rm M}$, FGLM:$77.1^{\rm A}$]
$\mathcal{G}(f_1,\ldots,f_4)$, $x_1,x_2,x_3 \succ_{\rm lex} x_4$	∞[$489^{\rm M}$, $95.5^{\rm A}$]

The Dixon resultant matrix is of order 43 and turns out to be singular. The computation time is more than four times the time needed for the Bézout resultant:

matrix construction	2.05	total 24.3
determinant evaluation (with result 0)	21.5	

Dense system $m = 4$ (D4)

This system consists of 5 quadratic polynomials with coefficients either 1 or 2. All coefficients of the first polynomial are 1:

$$
\begin{aligned}
f_1 &= {x_1}^2 + x_1 x_2 + x_1 x_3 + x_1 x_4 + x_1 x_5 + {x_2}^2 + x_2 x_3 + x_2 x_4 + x_2 x_5 + {x_3}^2 \\
&\quad + x_3 x_4 + x_3 x_5 + {x_4}^2 + x_4 x_5 + {x_5}^2 + x_1 + x_2 + x_3 + x_4 + x_5 + 1, \\
f_2 &= 2\,{x_1}^2 + 2\,x_1 x_2 + x_1 x_4 + 2\,{x_2}^2 + x_2 x_3 + 2\,x_2 x_5 + 2\,{x_3}^2 + 2\,x_3 x_5 \\
&\quad + 2\,{x_4}^2 + x_4 x_5 + 2\,{x_5}^2 + x_1 + x_2 + 2\,x_3 + 2\,x_4 + 1, \\
f_3 &= {x_1}^2 + 2\,x_1 x_3 + 2\,x_1 x_5 + {x_2}^2 + 2\,x_2 x_4 + {x_3}^2 + 2\,x_3 x_4 + {x_4}^2 \\
&\quad + 2\,{x_5}^2 + x_2 + x_3 + 2\,x_4 + 2\,x_5, \\
f_4 &= 2\,{x_1}^2 + x_1 x_2 + 2\,x_1 x_4 + x_1 x_5 + x_2 x_3 + 2\,x_2 x_4 + 2\,{x_3}^2 \\
&\quad + 2\,x_3 x_4 + 2\,{x_5}^2 + 2\,x_1 + 2\,x_4 + 1, \\
f_5 &= 2\,{x_1}^2 + x_1 x_3 + {x_2}^2 + x_2 x_3 + x_2 x_4 + 2\,x_3 x_4 + 2\,{x_4}^2 \\
&\quad + 2\,x_4 x_5 + {x_5}^2 + 2\,x_1 + 2\,x_2 + x_3 + 1.
\end{aligned}
$$

The Gröbner basis of $f_1,\ldots,f_4$ with respect to the total degree ordering $x_1 \succ \cdots \succ x_4$ gives no suspicious factors. The Bézout matrix is of size 16. The degree of the polynomial entries in x_5 is up to 5. The eliminant in x_5 is an irreducible polynomial of degree 32 with coefficients of order 20. The computation time is:

$\mathcal{G}(f_1,\ldots,f_4)$, $\succ_{\rm deg}(x_1,\ldots,x_4)$	3.64	
16 normal forms	1.80	total 16.2
determinant evaluation	10.6	

The computation times for the various Gröbner bases are:

$\mathcal{G}(f_1,\ldots,f_5)$, $\succ_{\rm deg}(x_1,\ldots,x_4,x_5)$	3.24 [$4.83^{\rm M}$, $0.78^{\rm A}$]
$\mathcal{G}(f_1,\ldots,f_5)$, $\succ_{\rm lex}(x_1,\ldots,x_4,x_5)$	∞[$\infty^{\rm M}$, FGLM:$30.2^{\rm A}$]
$\mathcal{G}(f_1,\ldots,f_5)$, $x_1,\ldots,x_4 \succ_{\rm lex} x_5$	∞[$54.9^{\rm M}$, $26.4^{\rm A}$]

The Dixon resultant matrix was constructed by Asir. The matrix is of order 41 and appears to be singular. The computation time is:

$$\left.\begin{array}{ll} \text{matrix construction} & 7.28^{\mathrm{A}} \\ \text{determinant evaluation (with result 0)} & 18.7^{\mathrm{A}} \end{array}\right\} \text{ total } 26.0^{\mathrm{A}}$$

Randomly choosing the coefficients in the range 1 to 9 gives order 47 coefficients in the resultant. Although the computation time for the construction of the Bézout matrix does not increase much, the evaluation of the determinant takes twice as long.

Summary and discussion

The Gröbner basis computation as part of the Bézout resultant method contributes, in most cases, from 10 up to 30 percent of the total computation time. The same applies for the normal form computations. The determinant evaluation takes 45 up to 65 percent of the total time (for Asir this is 90 up to 95 percent). The cyclic 6-roots system, the system with the biggest matrix, is an exception: the determinant evaluation contributes 97 percent. This means that, like for other resultant methods, the determinant evaluation contributes most to the total computation time.

In the following table we have summarized the total times for the Bézout and Dixon resultant and for the Gröbner basis with respect to the block-lexicographic ordering and lexicographic ordering by the FGLM basis conversion. We have excluded the IM system from the table as this system could not be solved by any of the methods. Columns with the times for Asir have been added.

	Bézout		Dixon		Gröbner		FGLM
C5	5.63	0.46^{A}	2.69	1.48^{A}	0.35^{M}	0.08^{A}	0.28^{A}
MC5	14.6	7.95^{A}	3.28	3.14^{A}	164^{M}	18.5^{A}	20.1^{A}
C6	2476	77.0^{A}	–	$44.5^{*\mathrm{A}}$	24.3^{M}	2.90^{A}	3.51^{A}
K5	15.5	6.69^{A}	–	$39.0^{**\mathrm{A}}$	19.0^{M}	7.82^{A}	6.33^{A}
SP	13.4	6.26^{A}	–	$65.0^{*\mathrm{A}}$	3968^{M}	18.8^{A}	9.20^{A}
G	14.0	7.59^{A}	–	$0.64^{**\mathrm{A}}$	∞^{M}	∞^{A}	∞^{A}
D3	5.71	3.18^{A}	24.3^{*}	$41.6^{*\mathrm{A}}$	489^{M}	95.5^{A}	77.1^{A}
D4	16.2	9.88^{A}	–	$26.0^{*\mathrm{A}}$	54.9^{M}	26.4^{A}	30.2^{A}

The * indicates the Dixon matrix is singular, ** indicates the Dixon matrix is non-square and the time is only for the matrix construction.

For the FGLM method, according to experiments [36], the Gröbner basis computation is 15 up to 35 percent of the total time needed for this method. Except for C6, our experiments give a percentage of between 1 and 20. For the sparse systems C5 and C6 the FGLM method is fast. For these systems Asir is even faster for the Göbner basis with regard to the block-lexicographic ordering. Using

Mathematica and the block-lexicographic ordering for the systems MC5 and SP turns out to be unfortunate.

Applying the Dixon resultant was not very successful. For systems with more than three indeterminates Maple can not handle the large intermediate expressions necessary for the construction of the resultant matrix. The reason is that these are polynomials in two times the number of original indeterminates. We see that the Dixon resultant method is favourite for the sparse cyclic 5-roots systems. For the other systems, the Dixon matrix is not square or becomes singular, even for dense systems like D3 and D4. The Dixon matrix is never smaller in size than the Bézout matrix. For K5, SP, D3 and D4, the Dixon matrix is two up to five times bigger.

Remarks Comparing Maple 5 release 3 and 5, the package 'Groebner' in release 5 did not bring any performance gain. However, the normal form computations in release 5 were up to 25 percent faster.

With Asir, for the construction of the intermediate expressions (1.28), the determinant evaluation by minor expansion is much faster than by Gaussian elimination. Maple uses a combination of minor expansion and Gaussian elimination. This seems to be unfortunate for the determinant of the Bézout matrix for the cyclic 6-roots system. The Bézout matrix determinant evaluation of Asir contributes 90 up to 95 percent of the total time needed. This also may be caused by Asir's strong optimizations regarding Gröbner basis computations. Asir is two up to twenty times faster than Mathematica in computing the Gröbner basis with regard to the block-lexicographic ordering.

Chapter 5

Non-admissible orderings and Involutive bases

In this chapter we show that the monomial ordering used in Involutive bases does not necessarily need to be an admissible ordering. We supply a set of non-admissible orderings that can be used in Involutive bases computations. In the first section some terminology related to Involutive bases is introduced. The second section contains the proofs for the correctness and termination of the algorithm for Involutive bases with respect to non-admissible orderings. As an application, we show that replacing the Gröbner basis part of the Bézout resultant construction by Involutive bases brings the methods from Chapter 3 and 4 closer together.

5.1 Terminology

We define $\prec_{part}$ to be the following total ordering on power products in the indeterminates $x_1, \ldots, x_m$. Let $X_i = x_1^{i_1} \ldots x_m^{i_m}$ and $X_j = x_1^{j_1} \ldots x_m^{j_m}$ be two power products. The ordering is a total degree ordering, i.e.

$$X_i \prec_{part} X_j \text{ if } \deg X_i < \deg X_j.$$

Furthermore, if the degrees are equal, view the exponents as partitions of the number $\deg X_i = \deg X_j$. Sort the entries of the partitions in decreasing order. Then compare these lexicographically:

$$X_i \prec_{part} X_j \text{ if } \operatorname{sort}(i_1, \ldots, i_m) \prec_{lex} \operatorname{sort}(j_1, \ldots, j_m).$$

If the partitions are equal, the power products do not need to be equal and ties are broken with the normal lexicographic ordering. How ties are broken exactly is not relevant for the current purpose. The definition of this ordering is taken from [82].

A simple example shows that $\prec_{part}$ is not an admissible ordering. We know that for admissible orderings $\prec$, $X_i \prec X_j$ implies $X_k X_i \prec X_k X_j$ for power products X_i, X_j and X_k. Hence, if $\prec_{part}$ were an admissible ordering $x_1 \prec_{part} x_2$ implies $x_1^2 \prec_{part} x_1 x_2$ which is not true because for the partitions of the exponents $(1,1) \prec_{lex} (2,0)$.

We recall the main terminology of the theory of Involutive bases. For an introduction to Involutive bases we refer to [41, 98]. Following [38], we call the used division and reduction relation after Pommaret.

Let $x_1, \ldots, x_m$ be indeterminates and let the domain of polynomials we are working with be $K[x_1, \ldots, x_m]$, K is a field. Let $\prec$ be a total ordering on the power products, e.g. $\prec_{part}$ or an admissible ordering, determining for every polynomial a leading monomial. The leading monomial of a polynomial f is denoted by $\mathrm{lm}(f)$.

With every indeterminate an integer number, called its index, is associated. This number should be unique in the sense that two different indeterminates have a different index. We can choose the index of the indeterminate x_j, $1 \le j \le m$, to be j.

Definition 5.1 *The class of a non-constant monomial is the minimum among the indices of the indeterminates of the monomial. The class of a non-constant polynomial f is the class of its leading monomial and is denoted by* $\mathrm{class}(f)$.

Definition 5.2 *The monomial u is called a Pommaret divisor for the monomial w, if there exists a monomial v such that $w = u \cdot v$ and either v is a constant or both u and v are non-constant monomials and all indeterminates of v have the index less than or equal to the class of u. We write $w = u \times v$.*

Definition 5.3 *For a set of polynomials $\mathcal{F}$, called the set of reducers, the Pommaret reduction relation $\rightarrow_P^{\mathcal{F}}$ is defined as: $g \rightarrow_P^{\mathcal{F}} g'$, if there exists an $f \in \mathcal{F}$ such that* $\mathrm{lm}(f)$ *is a Pommaret divisor of some monomial τ in g and $g' = g - v \cdot f$, where $\tau = \mathrm{lm}(f) \times v$. A polynomial g is in normal form with respect to $\rightarrow_P^{\mathcal{F}}$ if there does not exist a g' such that $g \rightarrow_P^{\mathcal{F}} g'$.*

From this definition follows that the Pommaret reduction relation is completely determined by the ordering on monomials, the indices of the indeterminates and the set of reducers.

Definition 5.4 *A set of polynomials $\mathcal{F}$ is said to be autoreduced if every polynomial $f \in \mathcal{F}$ is in normal form with respect to $\rightarrow_P^{\mathcal{F}\setminus\{f\}}$ and all monomials of f, except the leading monomial, are in normal form with respect to $\rightarrow_P^{\{f\}}$.*

This definition of the term autoreduced is different from the one mentioned in [98]. There, the non-leading monomials of f may be reducible.

Our definition of the term autoreduced allows non-admissible orderings with, for instance, $x_1 x_2 \prec x_1$. Then the set $\mathcal{F} = \{x_1 + x_1 x_2, x_3 + 1\}$ is autoreduced. In the next section we show that not for every non-admissible ordering a set can be brought into an autoreduced equivalent set.

5.2 The Pommaret reduction relation made Noetherian

In this section we prove that the Pommaret reduction relation is Noetherian for every monomial ordering, if the set of reducers is autoreduced. Moreover, it is shown how to bring a set of polynomials into autoreduced form for non-admissible total degree orderings.

Let $\prec$ be a monomial ordering, not necessarily an admissible ordering. To let the Pommaret reduction relation $\rightarrow_P^{\mathcal{F}}$ be of practical value this reduction relation needs to be Noetherian, i.e. there does not exist any infinite sequence of reductions. A simple example shows that this is not the case, which was expected because the monomial ordering may not be an admissible one. Consider $\mathcal{F} = \{f_1, f_2\}$ with

$$\begin{aligned} f_1 &= x_2^2 - x_1^2, \\ f_2 &= x_1^2 - x_1 x_2. \end{aligned}$$

Let the leading monomials of f_1 and f_2 be x_2^2 and x_1^2, respectively. This can be achieved by choosing the ordering $\prec_{part}$. The following infinite sequence of reductions is possible:

$$x_1 x_2^2 \rightarrow_P^{\mathcal{F}} x_1^3 \rightarrow_P^{\mathcal{F}} x_1^2 x_2 \rightarrow_P^{\mathcal{F}} x_1 x_2^2 \rightarrow_P^{\mathcal{F}} x_1^3 \rightarrow_P^{\mathcal{F}} \cdots.$$

We observe that $\mathcal{F}$ is not autoreduced because f_1 can be reduced by f_2:

$$f_1 \rightarrow_P^{\{f_2\}} f_1 + f_2 = x_2^2 - x_1 x_2.$$

Proposition 5.1 *Let $\mathcal{F} = \{f_1, \ldots, f_d\}$. If $\mathcal{F}$ is autoreduced, then the reduction relation $\rightarrow_P^{\mathcal{F}}$ is Noetherian.*

Proof. Let τ_0 be a monomial in normal form with respect to $\rightarrow_P^{\mathcal{F}}$, for instance 1. Let v_0 be any monomial. It suffices to prove that reducing $\tau_0 v_0$ will give a sum of monomials in normal form in finitely many steps.

If $v_0\tau_0$ is in normal form, we are done. So suppose that $v_0\tau_0$ can be reduced in one step to the polynomial f' by using the reducer $f \in \mathcal{F}$. Hence, v_0 is not a constant and there exists a monomial v_1 such that $v_0\tau_0 = \mathrm{lm}(f) \times v_1$. If v_1 is a constant, we are done, because f is in normal form and therefore is f'. Another trivial case is when f consists of one monomial (then $v_0\tau_0 \rightarrow_P^{\mathcal{F}} 0$).

Let now v_1 be a non-constant monomial and let f consists of more than one monomial. So f' is a sum of monomials of the form $v_1\tau_1$ where τ_1 is a monomial in normal form. The indeterminates in v_1 have indices less than or equal class(f). Let the index of the indeterminates of v_0 be between i_{min} and i_{max}. Note that it can not happen that class$(f) > i_{max}$, because then τ_0 is reducible by f, as v_1 is divisible by v_0 and $\tau_0 = \mathrm{lm}(f) \times \frac{v_1}{v_0}$ holds. There are two cases left.

Firstly, if class$(f) \prec i_{max}$, then the maximal index of indeterminates in v_1 is smaller than the maximal index of indeterminates in v_0.

Secondly, if $\text{class}(f) = i_{max}$, then the power of $x_{i_{max}}$ in v_1 is smaller than the power of $x_{i_{max}}$ in v_0, because otherwise τ_0 is reducible by f (again $\tau_0 = \text{lm}(f) \times \frac{v_1}{v_0}$).

Therefore, $v_0\tau_0$ reduces to monomials in normal form and monomials of the form $v_1\tau_1$, where τ_1 is a monomial in normal form. We showed that for v_1 either the maximal index of its indeterminates or the power of the indeterminate with largest index drops. The same happens in every further reduction step when $v_k\tau_k$ reduces to monomials $v_{k+1}\tau_{k+1}$ with τ_{k+1} in normal form. Ultimately the v_k will drop to a constant and we are left with monomials in normal form. □

Note that during the process of reducing the leading monomial may increase with respect to the ordering. In fact, in the proof no assumption about the ordering is used. However, not every set of polynomials is autoreduced. We will now find constraints on the ordering, when deriving a so-called autoreduced form of a set of polynomials.

Definition 5.5 *$\mathcal{F}'$ is said to be an autoreduced form of $\mathcal{F}$ if $\mathcal{F}'$ is autoreduced and $\mathcal{F}'$ generates the same ideal as $\mathcal{F}$.*

For instance, for the non-admissible ordering $\prec$ with $x_1^2 \prec x_1$, the set $\mathcal{F} = \{x_1 + x_1^2\}$ does not have an autoreduced form because x_1^2 is divisible by x_1. We will now demonstrate that for every total degree ordering, not necessarily an admissible ordering, e.g. $\prec_{part}$, an autoreduced form can be derived.

Proposition 5.2 *Let $\prec$ be any total degree ordering, not necessarily an admissible ordering. Let $\mathcal{F} = \{f_1, \ldots, f_d\}$. We can derive an autoreduced form of $\mathcal{F}$ with cardinality less than or equal d.*

Proof. We prove the proposition by induction on the number of polynomials in $\mathcal{F}$. A singleton $\mathcal{F}$ is autoreduced because a total degree ordering implies that no monomial in f_1, except the leading one, is divisible by $\text{lm}(f_1)$.

Let now be $d > 1$. Without loss of generality we let f_d be a polynomial with highest degree. Due to the induction hypothesis we can derive a set $\mathcal{F}' = \{f_1', \ldots, f_{d'}'\}$ with $d' \leq d-1$, an autoreduced form of $\{f_1, \ldots, f_{d-1}\}$. We can now derive f_d', the normal form of f_d with respect to $\rightarrow_P^{\mathcal{F}'}$. If f_d' is zero, we are done with $\mathcal{F}'$ as the autoreduced form of $\mathcal{F}$. If f_d' is not zero, we distinguish three cases.

Firstly, if the degree f_d' is greater than the maximum of the degrees of $f_1', \ldots, f_{d'}'$, then no monomial in $f_1', \ldots, f_{d'}'$ is divisible by $\text{lm}(f_d')$ and we obtained as autoreduced form of $\mathcal{F}$ the set $\{f_1', \ldots, f_{d'}', f_d'\}$ with cardinality less than or equal d.

Secondly, if the degree f_d' equals the maximum of the degrees of $f_1', \ldots, f_{d'}'$, then some monomial τ of an $f_j' \in \mathcal{F}'$ may be divisible by $\text{lm}(f_d')$. But then, the power products of τ and $\text{lm}(f_d')$ are equal, because they are of the same degree. In this case f_j' reduces to $f_j' + c_j f_d'$ where c_j is a constant. Obviously, $f_j' + c_j f_d'$ is in normal form. Hence, an autoreduced form of $\mathcal{F}$ will be $\{f_1' + c_1 f_d', \ldots, f_{d'}' + c_{d'} f_d', f_d'\}$ where the c_j's are (possibly non-zero) constants.

Thirdly, if the degree of f_d' is smaller than the degree of any $f_j' \in \mathcal{F}'$, repeat the whole procedure with f_j' as the polynomial with highest degree. The sum of the degrees of the polynomials drops in this case and that can only happen finitely many times. Hence, this procedure terminates with an autoreduced form left. $\Box$

Note that our results only apply for Involutive bases with respect to the Pommaret reduction relation. For the Thomas reduction relation [40], denoted by $\to_T^{\mathcal{F}}$, we can not use the non-admissible total degree ordering with, for instance, $x_1^3 x_2 \succ x_2^4$ and $x_1 x_2^3 \succ x_1^4$. Let $\mathcal{F}$ be the set of reducers given by $\{x_1^3 x_2 - x_2^4, x_1 x_2^3 - x_1^4\}$. Although $\mathcal{F}$ is obviously autoreduced, we still have an infinite sequence of reductions for $x_1^4 x_2$. As the multiplier of the first reducer is x_1 and of the second reducer is x_2, we have

$$x_1^4 x_2 = x_1(x_1^3 x_2) \to_T^{\mathcal{F}} x_1 x_2^4 = x_2(x_1 x_2^3) \to_T^{\mathcal{F}} x_1^4 x_2 \to_T^{\mathcal{F}} \cdots.$$

It is still unknown whether there exist non-admissible orderings for Involutive basis with respect to the Thomas or Janet reduction relation.

5.3 Involutive bases with respect to $\prec_{part}$

The results from the previous section provide an algorithm for Involutive bases with respect to $\prec_{part}$. This algorithm terminates for zero-dimensional ideals, like the algorithm for Involutive bases with respect to admissible orderings. The only difference with the usual Involutive bases algorithm is that the definition of the term autoreduced is more strict.

To see what happens when we apply Involutive bases with respect to the non-admissible orderings, consider the polynomial system given by

$$\begin{array}{rcl} f_1 & = & x^2 + xy + 2\,x + 3\,y + 5, \\ f_2 & = & y^2 + 7\,xy + 11\,x + 13\,y + 17. \end{array} \tag{5.1}$$

There is no admissible ordering for which the set of irreducible power products, as the basis of the residue class ring $\mathbb{Q}[x,y]/\mathcal{I}$, with $\mathcal{I}$ the ideal generated by f_1 and f_2, using either Gröbner bases or Involutive bases, equals

$$\{1, x, y, xy\}. \tag{5.2}$$

Instead, this basis is one among

$$\{1, x, x^2, x^3\}, \{1, x, x^2, y\}, \{1, x, y, y^2\} \text{ or } \{1, y, y^2, y^3\}.$$

If the basis would be $\{1, x, y, xy\}$, then x^2 and y^2 are reducible. This is only possible by f_1 and f_2, with x^2 and y^2 the leading monomials of f_1 and f_2, respectively. As we have seen in Section 5.1, this can not happen for admissible orderings. The good news is that for the ordering $\prec_{part}$ we have x^2 and y^2 as the leading monomials of f_1 and f_2. Moreover, f_1 and f_2 are in autoreduced form with respect to this ordering. Choosing the index of x to be 2 and the index of y to be 1, the Involutive bases algorithm only adds the polynomial

$$xy^2 - \frac{45}{2}\,xy - \frac{113}{3}\,x - \frac{205}{6}\,y - \frac{151}{3}$$

to the basis.

We know that for admissible orderings every Involutive basis is also a Gröbner basis. This example showed that an Involutive basis with respect to $\prec_{part}$ may not be a Gröbner basis.

To give one more illustration, consider the ideal generated by

$$\begin{aligned}
f_1 &= x^3 + \cdot x^2 y + \cdot x^2 z + \cdot xy^2 + \cdot xyz + \cdot x^2 + \cdot y^2 + \cdot xy + \cdot xz + \cdot x + \cdot y + \cdot z + \cdot,\\
f_2 &= y^3 + \cdot x^2 y + \cdot x^2 z + \cdot xy^2 + \cdot xyz + \cdot x^2 + \cdot y^2 + \cdot xy + \cdot xz + \cdot x + \cdot y + \cdot z + \cdot,\\
f_3 &= z^2 + \cdot x^2 + \cdot y^2 + \cdot xy + \cdot xz + \cdot x + \cdot y + \cdot z + \cdot.
\end{aligned}$$

The dot $\cdot$ stands for parametric coefficients that are algebraically independent. Only for Involutive bases with respect to the ordering $x \prec_{part} y \prec_{part} z$ together with the choice of indices 1, 2 and 3 for x, y, and z, respectively, the set of irreducible power products equals

$$\mathcal{T} = \{x^p y^q z^r \mid 0 \le p, q < 3, \quad 0 \le r < 2\}.$$

5.4 Applications

Because Involutive bases can be used to compute normal forms, we can modify the algorithm for the construction of the resultant by Bézout's method, replacing the Gröbner basis and normal form computation by the Involutive basis and the corresponding normal form with regard to this basis.

To see how this works out we consider the following system similar to (5.1):

$$\begin{aligned}
f_1 &= x^2 + axy + 2\,x + 3\,y + 5,\\
f_2 &= y^2 + bxy + 11\,x + 13\,y + 17.
\end{aligned}$$

Computing a Gröbner basis for the system with respect to the total degree ordering with $x \succ y$ gives the set

$$\left\{\begin{array}{l}
bx^2 - ay^2 + (-11\,a + 2\,b)\,x + (-13\,a + 3\,b)\,y - 17\,a + 5\,b,\\
bxy + y^2 + 11\,x + 13\,y + 17\\
\left(ab^2 - b\right)y^3 + \left(13\,ab^2 - 3\,b^3 + 2\,b^2 - 26\,b + 11\right)y^2 + \left(17\,b^2 - 143\,b + 121\right)x\\
\qquad + \left(17\,ab^2 - 5\,b^3 - 7\,b^2 - 186\,b + 143\right)y - 21\,b^2 - 221\,b + 187
\end{array}\right\}.$$

Using this basis in the Bézout method we have to deal with the suspicious factors b and $ab-1$. Choosing the total degree ordering with $x \prec y$, the basis has elements of similar complexity and the suspicious factors are a and $ab - 1$.

On the other hand, computing an Involutive basis with respect to $\prec_{part}$ gives, besides f_1 and f_2, the element

$$-\,(ab-1)\,y^2x+\left(3\,b^2-11\,a-2\,b+13\right)yx+(34\,b-33)\,y+(33\,b-5)\,x+51\,b-55$$

as part of the basis. This gives only one suspicious factor, namely $ab-1$. This suspicious factor coincides with the resultant of the leading forms of f_1 and f_2.

The consequences for the Bézout matrices extracted from the normal forms obtained via the Gröbner basis and via the Involutive basis are illustrated by the example $f_3 = x + y$. The Bézout matrix constructed using the Gröbner basis with respect to the total degree ordering $x \succ y$ equals (the rows correspond to the normal forms of f_3, xf_3, yf_3, and y^2f_3, and the columns correspond to the power products 1, x, y and y^2):

$$\mathcal{B}=\begin{pmatrix} 0 & 1 & 1 & 0 \\ \frac{17\,a-5\,b-17}{b} & \frac{11\,a-2\,b-11}{b} & \frac{13\,a-3\,b-13}{b} & \frac{a-1}{b} \\ \frac{-17}{b} & \frac{-11}{b} & \frac{-13}{b} & \frac{b-1}{b} \\ \frac{21\,b^2+187\,a+200\,b-408}{ab^2-b} & \frac{-17\,b^2+121\,a+160\,b-264}{ab^2-b} & b_{4,3} & b_{4,4} \end{pmatrix}$$

with $b_{4,3} = \frac{-17\,ab^2+5\,b^3+2\,b^2+143\,a+179\,b-312}{ab^2-b}$ and $b_{4,4} = \frac{3\,b^3-13\,ab^2-5\,b^2+11\,a+28\,b-24}{ab^2-b}$.

Using the Involutive basis we obtain (the rows correspond to the normal forms of f_3, xf_3, yf_3 and xyf_3, and the columns correspond to the power products 1, x, y and xy, as the basis for the residue class ring equals (5.2))

$$\mathcal{B}=\begin{pmatrix} 0 & 1 & 1 & 0 \\ -5 & -2 & -3 & -a+1 \\ -17 & -11 & -13 & -b+1 \\ \frac{55\,a+51\,b-106}{ab-1} & \frac{5\,a+33\,b-38}{ab-1} & \frac{33\,a+34\,b-67}{ab-1} & \frac{11\,a^2+3\,b^2-24\,a-5\,b+15}{ab-1} \end{pmatrix}.$$

The determinant of both matrices is equal to

$$-\frac{289\,a^2-170\,ab+25\,b^2-394\,a+113\,b+137}{ab-1}.$$

We can see that the matrix obtained via Involutive bases has less complex entries. The degree of the numerators of the entries is 2, whereas this is 3 for the matrix obtained via the Gröbner basis. The total number of monomials in the numerators of the matrix entries is 28 instead of 35. Moreover, the number of entries with a non-constant denominator is 4 instead of 12. Note that the denominators b in the matrix obtained via the Gröbner basis do not appear in the determinant result anymore. Via the Involutive basis we already can conclude that b is an extraneous factor.

Appendix A

Maple code Bézout resultant

```
with(linalg): with(grobner):

bezout_eliminate := proc()
local fset, varset, gb, g, m, mdone, mdiff, mselect, cof, mon,
      firsttime, newtime, lasttime,
      i, j, N, res, normalform, den, d, f, df;
global normalforms, total_denom, bezoutmatrix;
option package;
print(`Computing eliminant using construction of Bezout`);

fset := [op(args[1])];
varset := [op(args[2])];
g := args[3];

lasttime := time(): firsttime := lasttime:
gb := gbasis(fset, varset, tdeg):
newtime := time():
print(`Time needed for Groebner basis with respect to tdeg ordering: `
      , newtime - lasttime);
lasttime := newtime:
if (finite(gb, varset) = false) then
  print(`Warning: not a zero-dimensional ideal`);
fi;

m := { 1 }:
mdone := []: normalforms := []:
while (nops(`minus`(m, {op(mdone)})) > 0)
do
  mdiff := `minus`(m, {op(mdone)}):
```

```
  mselect := mdiff[1]:
  normalform := normalf(mselect * g, gb, varset, tdeg):
  normalform := collect(normalform, varset, distributed);
  cof := coeffs(normalform, varset, 'mon'):
  normalforms := [ op(normalforms), normalform ]:
  m := `union`(m, { mon }):
  mdone := [op(mdone), mselect]:
od;
newtime := time():
print(`Time needed for normalforms:`, newtime - lasttime);
lasttime := newtime:

print(`Basis elements: `, mdone);
N := nops(mdone);
print(`Dimension matrix is: `, N);
bezoutmatrix := matrix(N, N):
total_denom := 1:
for i from 1 to N do
  normalform := normalforms [i]:
  den := denom(normalform):
  total_denom := total_denom * den:
  normalform := numer(normalform): # keep matrix fraction-free
  cof := coeffs(collect(normalform, varset, distributed), varset, 'mon'):
  cof := [cof]:
  mon := [mon]:
  for j from 1 to N do
    if (member(mdone[j], mon, 'pos') = true) then
      bezoutmatrix [i, j] := cof[pos]:
    else
      bezoutmatrix [i, j] := 0:
    fi;
  od;
od;
newtime := time():
print(`Time needed for constructing Bezout matrix: `, newtime - lasttime);
lasttime := newtime:

res := eval(det(bezoutmatrix));
newtime := time():
print(`Time needed for evaluating determinant: `, newtime - lasttime);
lasttime := newtime:

print(`Time needed in total: `, lasttime - firsttime);
print(`Denominator was: `, total_denom);

res := simplify(res / total_denom);
end:
```

Appendix B

Maple code Dixon resultant

```
with(linalg):

dixon_eliminate := proc()
local fset, varset, m, v, dm, i, j, f, res, firsttime, lasttime, newtime;
option package;
print(`Computing eliminant using construction of Dixon`);

fset := [ op(args[1]) ]:
varset := [ op(args[2]) ]:
m := nops(fset);

lasttime := time():
firsttime := lasttime:

# create new indeterminates
v := {};
for i from 1 to m-1 do
  v := `union`(v, { cat(`mu_`, i) }):
od;
v := [ op(v) ];

dm := matrix(m, m);
for i from 1 to m do
  f := fset[i]:
  for j from 1 to m do
    dm[i,j] := f:
    if (j < m) then f := subs({varset[j] = v[j] }, f): fi;
  od;
od;

lasttime := time();
```

```
f := det(dm):
newtime := time():
print(`Time needed for raw determinant: `, newtime - lasttime);
lasttime := newtime:

for j from 1 to m-1 do
  f := f / (varset[j] - v[j]):
od;
f := simplify(f):

# extract coefficient matrix with function defined below
lasttime := time();
dm := get_coeff_matrix(f, varset, v):
newtime := time():
print(`Time needed for constructing Dixon matrix: `, newtime - lasttime);

lasttime := newtime:
if (rowdim(dm) = coldim(dm)) then
  res := eval(det(dm));
else
  print(`Dixon matrix is not square: rows = `, rowdim(dm)
        , ` and col = `, coldim(dm));
  res := 0;
fi;
newtime := time():
print(`Time needed for evaluating determinant: `, newtime - lasttime);
print(`Time needed in total: `, newtime - firsttime);
res;
end:


# auxiliary function for extracting a coefficient matrix
get_coeff_matrix := proc()
local f, varset, newvarset, cof, mon, m, ms, N, M, i, j,
      subcof, submon, pos, coeffmatrix;
option package;

f := args[1]:
varset := [ op(args[2]) ]:
newvarset := [ op(args[3]) ]:

# gather all coefficients and monomials (N = number of equations)
f := collect(f, varset, distributed):
cof := [ coeffs(f, varset, 'mon') ]:
mon := [ mon ]:
N := nops(mon):
```

```
# collect all variables of the linear equations
m := {}:
for i from 1 to N do
  subcof := coeffs(collect(cof[i], newvarset, distributed),
                   newvarset, 'submon'):
  m := `union`(m, { submon }):
od;
M := nops(m);
# sorting...
ms := 0:
for i from 1 to M do
  ms := ms + m[i];
od;
ms := sort(ms, newvarset, tdeg);
m := [ op(ms) ];
print(`Number of linear equations:`, N, ` number of variables: `, M);

coeffmatrix := matrix(N, M);

for i from 1 to N do
  subcof := coeffs(collect(cof[i], newvarset, distributed),
                   newvarset, 'submon'):
  subcof := [ subcof ]:
  submon := [ submon ]:
  for j from 1 to M do
    if (member(m[j], submon, 'pos') = true) then
      coeffmatrix[i, j] := subcof[pos]:
    else
      coeffmatrix[i, j] := 0:
    fi;
  od;
od;
eval(coeffmatrix);
end:
```

Bibliography

[1] A.G. Akritas. *Elements of Computer Algebra with Applications.* J.Wiley & Sons, New York, 1989.

[2] M. Alonso, E. Becker, M.-F. Roy, and T. Wörmann. Zeroes, Multiplicities and Idempotents for Zerodimensional Systems. In *MEGA '94*, 1994.

[3] P. Aubry, D. Lazard, and M. Moreno Maza. On the Theories of Triangular Sets. *Journal of Symbolic Computation*, 28(1):105–124, 1999.

[4] W. Auzinger and H.J. Stetter. An Elimination Algorithm for the Computation of all Zeros of a System of Multivariate Polynomial Equations. In *Conference in Numerical Analysis*, volume 86 of *ISNM*, pages 11–30. Birkhäuser, 1988.

[5] E.H. Bareiss. Sylvester's Identity and Multistep Integer-preserving Gaussian elimination. *Mathematics of Computation*, 22:565–578, 1968.

[6] S. Barnett. Greatest common divisor of two polynomials. *Linear Algebra and its Applications*, 3:7–9, 1970.

[7] E. Becker, M.G. Marinari, T. Mora, and C. Traverso. The shape of the Shape Lemma. In *Proceedings ISSAC '94*, pages 129–133. ACM, 1994.

[8] E. Becker and T. Wörmann. Radical computations of zero-dimensional ideals and real root counting. In *IMACS Symposium SC-1993*, pages 197–202, 1993.

[9] T. Becker and V. Weispfenning. *Gröbner Bases - A Computational Approach to Commutative Algebra.* Springer-Verlag, 1993.

[10] C.A. Berenstein and D.C. Struppa. Small degree solutions for the polynomial Bézout equation. *Linear Algebra and Its Applications*, 98:41–56, 1988.

[11] C.A. Berenstein and A. Yger. Effective Bézout identities in $Q[z_1, \ldots, z_n]$. *Acta Mathematica*, 166:69–120, 1991.

[12] É. Bézout. *Théorie générale des Équations Algébriques.* P.-D. Pierres, Paris, 1779.

[13] P. Bikker. Non-Admissible Orderings and Involutive Bases. Technical Report 97-16, RISC-Linz, 1997.

[14] P. Bikker and A.Yu. Uteshev. On the Bézout construction of the Resultant. *Journal of Symbolic Computation*, 28(1):45–88, 1999.

[15] G. Björck and R. Fröberg. A Faster Way to Count the Solutions of Inhomogeneous Systems of Algebraic Equations, with Applications to Cyclic n-roots. *Journal of Symbolic Computation*, 12(1):329–336, 1991.

[16] W. Boege, R. Gebauer, and H. Kredel. Some Examples for Solving Systems of Algebraic Equations by Calculating Groebner Bases. *Journal of Symbolic Computation*, 1(1):83–98, 1986.

[17] B. Buchberger. *Ein Algorithmus zum Auffinden der Basiselemente des Restklassenrings nach einem nulldimensionalen Polynomideal.* Ph.D. Thesis, University of Innsbruck, 1965.

[18] B. Buchberger. Gröbner Bases: An Algorithmic Method in Polynomial Ideal Theory. In N.K. Bose, editor, *Recent Trends in Multidimensional Systems Theory*, pages 184–232. D. Riedel Publishing Company, 1985.

[19] B. Buchberger and F. Winkler. A Criterion for Detecting Unnecessary Reductions in the Construction of Gröbner Bases. In W. Ng., editor, *Proceedings EUROSAM 79*, volume 72 of *Lecture Notes in Computer Science*, pages 3–21. Marseille, 1979.

[20] J. Canny. Generalized Characteristic Polynomials. In P. Gianni, editor, *Proceedings ISSAC '88*, volume 358 of *Lecture Notes in Computer Science*, pages 293–299. Springer-Verlag, 1988.

[21] J. Canny and I. Emiris. An Efficient Algorithm for the Sparse Mixed Resultant. In G. Cohen, T. Mora, and O. Moreno, editors, *Proceedings of AAECC*, volume 263 of *Lecture Notes in Computer Science*, pages 89–104. Springer-Verlag, 1993.

[22] A. Cayley. On the Theory of Elimination. *Cambridge and Dublin Math. Journal*, 3:116–120, 1848.

[23] A. Cayley. Note sur la Méthode d'Élimination de Bézout. *J. Reine Angew. Math.*, 53:336–376, 1857.

[24] S. Collart, M. Kalkbrener, and D. Mall. Converting bases with the Gröbner walk. *Journal of Symbolic Computation*, 24:465–470, 1997.

[25] G.E. Collins. Subresultants and Polynomial Remainder Sequences. *Journal of the ACM*, 14(1):128–142, 1967.

[26] G.E. Collins. The Calculation of Multivariate Polynomial Resultants. *Journal of the ACM*, 18(4):515–532, 1971.

[27] D. Cox, J. Little, and D. O'Shea. *Ideals, Varieties, and Algorithms.* Springer-Verlag, 1992.

[28] D. Cox, J. Little, and D. O'Shea. *Using Algebraic Geometry.* Springer-Verlag, 1998.

[29] C. D'Andrea and A. Dickenstein. Explicit Formulas for the Multivariate Resultant. In *Special issue MEGA 2000*, 2000.

[30] A.L. Dixon. The eliminant of three quantics in two independent variables. *Proceedings of The London Mathematical Society*, 6:468–478, 1908.

[31] M. Elkadi and B. Mourrain. A New Algorithm for the Geometric Decomposition of a Variety. In *Proceedings ISSAC '99*, pages 9–16. ACM, 1999.

[32] I.Z. Emiris and J. Canny. A Practical Method for the Sparse Resultant. In *Proceedings ISSAC '93*, pages 183–192. ACM, 1993.

[33] I.Z. Emiris and B. Mourrain. Matrices in Elimination Theory. *Journal of Symbolic Computation*, 28(1):3–44, 1999.

[34] D.K. Faddeev. *Lectures in Higher Algebra (in Russian).* Nauka, Moscow, 1984.

[35] J.C. Faugère. A new efficient algorithm for computing Gröbner bases (F_4). In *MEGA '98*, 1998.

[36] J.C. Faugère, P. Gianni, D. Lazard, and T. Mora. Efficient Computation of Zero-dimensional Gröbner Bases by Change of Ordering. *Journal of Symbolic Computation*, 16(4):329–344, 1993.

[37] F.R. Gantmacher. *The Theory of Matrices.* Chelsea, New York, 1959.

[38] P.A. Garcia-Sánchez. Gröbner and Involutive Bases for Zero-dimensional Ideals. *SIGSAM Bulletin*, 29.2:12–15, 1995.

[39] K. Gatermann. Symbolic solution of polynomial equation systems with symmetry. In *Proceedings ISSAC '90*, pages 112–119. ACM, 1990.

[40] V.P. Gerdt. Involutive Division Technique: Some Generalizations and Optimizations. *St. Petersburg Math. Seminars*, 1999.

[41] V.P. Gerdt and Yu.A. Blinkov. Involutive Bases of Polynomial Ideals. *Mathematics and Computers in Simulation*, 45(5–6):519–542, 1998.

[42] P. Gianni. Properties of Gröbner basis under Specialization. In J.H. Davenport, editor, *Eurocal 87*, volume 378 of *Lecture Notes in Computer Science*, pages 293–297. Springer-Verlag, 1987.

[43] A. Giovini, T. Mora, G. Nielsi, L. Robbiano, and C. Traverso. 'One sugar cube, please' or Selection strategies in the Buchberger algorithm. In *Proceedings ISSAC '91*, pages 49–54. ACM, 1991.

[44] L. González-Vega. An Elementary proof of Barnett's theorem about the greatest common divisor of several univariate polynomials. *Linear Algebra and Its Applications*, 247:185–202, 1996.

[45] L. González-Vega and M-J. González-López. Newton Identities in the multivariate case: Pham Systems. In B. Buchberger and F. Winkler, editors, *Gröbner Bases and Applications*, volume 215 of *Lecture Notes Series*, pages 351–366. Cambridge University Press, London Mathematical Society, 1998.

[46] L. González-Vega, G. Trujillo, and J.R. Sendra. Efficient manipulation of the trace matrix via Hankel matrices. In *EACA '95*, pages 97–107, 1995.

[47] G. Hagel. *Algorithmische Strategien zur Berechnung der u-Resultante.* Ph.D. Thesis, Eberhard Karls Universität Tübingen, 1994.

[48] C.M. Hoffmann. Algebraic and Numeric Techniques for Offsets and Blends. In W. Dahmen, M. Gasca, and C. Michelli, editors, *Computations of Curves and Surfaces.* Kluwer Academic Publishers, 1990.

[49] C.M. Hoffmann. Implicit Curves and Surfaces in CAGD. *IEEE Computer Graphics and Applications*, 13(1):79–88, 1993.

[50] M. Kalkbrener. Solving Systems of Algebraic Equations by Using Gröbner Bases. In J.H. Davenport, editor, *Eurocal 87*, volume 378 of *Lecture Notes in Computer Science*, pages 282–292. Springer-Verlag, 1987.

[51] M. Kalkbrener. On the Stability of Gröbner Bases Under Specializations. *Journal of Symbolic Computation*, 24:51–58, 1997.

[52] H. Kapferer. Über Resultanten und Resultanten-Systeme. *Sitzungsber. Bayer. Akad.*, pages 179–200, 1929.

[53] D. Kapur. Automated Geometric Reasoning: Dixon Resultants, Gröbner Bases and Characteristic Sets. In D. Wang, editor, *Automated Deduction in Geometry*, volume 1360 of *LNAI*, pages 1–36. Springer-Verlag, 1997.

[54] D. Kapur and Y.N. Lakshman. Elimination Methods: an Introduction. In B.R. Donald, D. Kapur, and J.L. Mundy, editors, *Symbolic and Numerical Computation for Artificial Intelligence*, pages 45–87. Academic Press, 1992.

[55] D. Kapur and T. Saxena. Extraneous Factors in the Dixon Resultant Formulation. In *Proceedings ISSAC '97*, pages 141–148. ACM, 1997.

[56] H. Kobayashi, T. Fujise, and A. Furukawa. Solving systems of algebraic equations by a general elimination method. *Journal of Symbolic Computation*, 5:303–320, 1988.

[57] M.G. Krein and M.A. Naimark. The Method of Symmetric and Hermitian Forms in the Theory of the Separation of the Roots of Algebraic Equations. *Linear and Multilinear Algebra*, 10:265–308, 1981.

[58] H. Laurent. L'Élimination. In *Scientia, Phys.-Mathématique*, volume 7. Gauthier-Villars, Paris, 1900.

[59] D. Lazard. Résolution des Systèmes d'Équations Algébriques. *Theoretical Comput. Sci.*, 15:77–110, 1981.

[60] F.S. Macaulay. On some Formulae in Elimination. *Proceedings of The London Mathematical Society*, 35:3–27, 1903.

[61] F.S. Macaulay. *The Algebraic Theory of Modular Systems.* Cambridge University Press, 1916.

[62] D. Manocha and J.F. Canny. Multipolynomial Resultant Algorithms. *Journal of Symbolic Computation*, 15:99–122, 1993.

[63] H.M. Moeller and H.J. Stetter. Multivariate polynomial equations with multiple zeros solved by matrix eigenproblems. *Numer. Math.*, 70:311–329, 1995.

[64] S. Moritsugu and K. Kuriyama. On Multiple Zeros of Systems of Algebraic Equations. In *Proceedings ISSAC '99*, pages 23–30. ACM, 1999.

[65] E. Netto. *Vorlesungen über Algebra*, volume 2. Teubner, Leipzig, 1900.

[66] M. Noro and T. Takeshima. Risa/Asir - A Computer Algebra System. In *Proceedings ISSAC '92*, pages 387–396. ACM, 1992.

[67] M. Noro and K. Yokoyama. New methods for the change-of-ordering in Gröbner basis computation. *Research Report ISIS-RR-95-8E*, 1995.

[68] M. Noro and K. Yokoyama. A Modular Method to Compute the Rational Univariate Representation of Zero-dimensional Ideals. *Journal of Symbolic Computation*, 28(1):243–263, 1999.

[69] A.M. Ostrowski. Über ein algebraisches Übertragungsprinzip. In *Abh. Mathematisches Seminar*, volume 1. Universität Hamburg, 1922.

[70] A.M. Ostrowski. *Solution of Equations and Systems of Equations.* Academic Press, New York and London, 1960.

[71] P. Pedersen. Calculating multidimensional symmetric functions using Jacobi's formula. In H.F. Mattson, T. Mora, and T.R.N. Rao, editors, *Proceedings of the AAECC-9*, volume 539 of *Lecture Notes in Computer Science*, pages 304–317. Springer-Verlag, 1991.

[72] P. Pedersen. Multivariate Sturm theory. In H.F. Mattson, T. Mora, and T.R.N. Rao, editors, *Proceedings of the AAECC-9*, volume 539 of *Lecture Notes in Computer Science*, pages 318–332. Springer-Verlag, 1991.

[73] P. Pedersen and B. Sturmfels. Product formulas for resultants and Chow forms. *Mathematische Zeitschrift*, 214:377–396, 1993.

[74] O. Perron. *Algebra - Die Grundlagen*. Walter de Gruyter & Co., Berlin-Leipzig, 1932.

[75] Tran Quoc-Nam. *A Hybrid Symbolic-Numerical Approach in Computer Aided Geometric Design (CAGD) and Visualization*. Ph.D. Thesis, Johannes Kepler University Linz, 1996.

[76] F. Rouillier. Solving Zero-Dimensional Systems Through the Rational Univariate Representation. In *AAECC*, volume 9, pages 433–461. Springer-Verlag, 1999.

[77] T. Saxena. *Efficient Variable Elimination Using Resultants*. Ph.D. Thesis, State University of New York, 1997.

[78] G. Scheja and U. Storch. *Lehrbuch der Algebra*. Teubner, 1988.

[79] L. Schläfli. Über die Resultante eines Systemes mehrerer algebraischer Gleichungen (1852). In *Gesammelte Mathematische Abhandlungen*, volume 2, pages 9–112. Birkhäuser, Basel, 1953.

[80] J.-A. Serret. *Cours d'Algèbre Supérieure*, volume 1. Gauthier-Villars, Paris, 1928.

[81] H.J. Stetter. Matrix Eigenproblems are at the Heart of Polynomial System Solving. *SIGSAM Bulletin*, 30.4:22–25, 1996.

[82] B. Sturmfels. *Algorithms in Invariant Theory*. Springer-Verlag, 1993.

[83] B. Sturmfels and A. Zelevinsky. Multigraded Resultants of Sylvester Type. *Journal of Algebra*, 163:115–127, 1994.

[84] J.J. Sylvester. On a general method of determining by mere inspection the derivations from two equations of any degree. *Philosophical Magazine*, 16:132–135, 1840.

[85] W. Trinks. On B. Buchberger's Method for Solving Systems of Algebraic Equations. *Journal of Number Theory*, 10:475–488, 1978.

[86] A.Yu. Uteshev and T.M. Cherkasov. Localization of Solutions to Systems of Algebraic Equations and Inequalities: The Hermite's method. *Doklady Math. (English translation)*, 53.2:227–229, 1996.

[87] A.Yu. Uteshev and T.M. Cherkasov. The Search for the Maximum of a Polynomial. *Journal of Symbolic Computation*, 25:587–618, 1998.

[88] A.Yu. Uteshev and S.G. Shulyak. Hermite's Method of Separation of Solutions of Systems of Algebraic Equations and Its Applications. *Linear Algebra and Its Applications*, 177:49–88, 1992.

[89] B.L. van der Waerden. *Modern Algebra*, volume II. F. Ungar Publishing Co., New York, 1950.

[90] B.L. van der Waerden. *Algebra, Erster Teil.* Springer-Verlag, 1966.

[91] B. Wall. *Symbolic Computation with Algebraic Sets.* Ph.D. Thesis, Johannes Kepler University Linz, 1993.

[92] D. Wang. Solving Polynomial Equations: Characteristic Sets and Triangular Systems. In *IMACS Symposium SC-1993*, pages 3–9, 1993.

[93] D. Wang. *Elimination Methods.* Springer-Verlag, Wien, 2000.

[94] V. Weispfenning. Comprehensive Gröbner Bases. *Journal of Symbolic Computation*, 14:1–29, 1992.

[95] F. Winkler. A p-adic Approach to the Computation of Gröbner Bases. *Journal of Symbolic Computation*, 6:287–304, 1988.

[96] W.T. Wu. Basic Principles of Mechanical Theorem Proving in Elementary Geometries. *Journal of System Sciences and Mathematical Sciences*, 4:207–235, 1984.

[97] K. Yokoyama, M. Noro, and T. Takeshima. Solutions of Systems of Algebraic Equations and Linear Maps on Residue Class Rings. *Journal of Symbolic Computation*, 14:399–417, 1992.

[98] A.Yu. Zharkov and A.Yu. Blinkov. Involution Approach to Solving Systems of Algebraic Equations. In *IMACS Symposium SC-1993*, pages 11–16, 1993.