Multivariate Polynomials, Duality and Structured Matrices

Bernard Mourrain* Victor Y. Pan †

December 14, 1998

Abstract

We first review the basic properties of the well known classes of Toeplitz, Hankel, Vandermonde, and other related structured matrices and re-examine their correlations to operations with univariate polynomials. Then we define some natural extensions of such classes of matrices based on their correlations to multivariate polynomials. We describe these correlations in terms of the associated operators of multiplication in the polynomial ring and its dual space, which allows us to generalize these structures to the multivariate case. Multivariate Toeplitz, Hankel, and Vandermonde matrices, Bezoutians, algebraic residues and relations among them are studied. Finally, we show some applications of structured matrices, duality and residues to rootfinding problems for a system of multivariate polynomial equations, where the dual space of linear forms, algebraic residues, Bezoutians and other structured matrices play an important role. The developed techniques enable us to obtain a better insight into the major problems of multivariate polynomial computations and to improve substantially the known techniques of the study of these major problems. From the algorithmic point, we yield acceleration by order of magnitude of the known methods for some fundamental problems of solving multivariate polynomial systems of equations.

Keywords: Structured matrices, polynomial equations, ideals, roots, quotient algebra, dual algebra, residues, basis of idempotents, Jacobians

1991 AMS Subject classification: 12E12, 65F15, 68Q40, 47B35, 14Q20, 68Q25

*INRIA, SAGA, 2004 route des Lucioles, B.P. 93, 06902 Sophia Antipolis, mourrain@sophia.inria.fr, (partially supported by European ESPRIT project FRISCO, LTR 21.024)

†Department of Mathematics and Computer Science Lehman College, City University of New York, Bronx, NY 10468, USA, VPAN@LCVAX.LEHMAN.CUNY.EDU (Supported by NSF Grants CCR 9625344 and 9732206 and PSC CUNY Awards 668365 and 669363.)

1 Introduction

The main goal of this paper is to summarize and to develop various techniques in the areas of algebraic residues, dual spaces of linear maps and structured matrices and to demonstrate the power of application of these techniques to algorithmic study of polynomial systems of equations; in particular we accelerate the known solution algorithms by order of magnitude. Let us comment on the structure of our presentation and on some specific new results of this paper.

It is well known that the important classes of Toeplitz, Hankel, Vandermonde, and some other structured matrices have a natural characterization in terms of the associated linear operators of scaling and displacements. We will study some extensions of the classes of such matrices, based on their correlations to the fundamental operations with polynomials, such as polynomial multiplication, multipoint evaluation, interpolation, and rootfinding. We will start with a review of the simpler and well known correlations to operations with univariate polynomials and then will use the patterns of this study as basic samples for our extended study where we involve multivariate polynomials. This will enable us to give a natural introduction to some other large topics, intensively studied and having important applications, such as duality, algebraic residues, and Bezoutians.

The correlations between structured matrices and univariate polynomials have been effectively used for the acceleration of structured matrix computation. We will extend these results to the structured matrices associated with multivariate polynomials and will exploit matrix structure to improve substantially the known methods and algorithms for polynomial systems of equations.

In particular, in section 3.2.3, using these tools enables us to simplify substantially the known derivations of the fundamental upper bounds by Bézout and Bernstein on the number D of the roots of a given polynomial system. In section 3.3, we specify our iterative algorithm outlined in the conference paper [24]. The algorithm quadratically converges right from the start to a selected root of a polynomial system of equations that has D distinct and simple roots, and we approximate such a root by using order of D^2 arithmetic operations (up to a polylogarithmic factor in D). (Hereafter, we will use the abbreviation "ops" for "arithmetic operations". We say "ops" rather than "flops" to cover also rational computations with finite precision.) The algorithm can be applied recursively to compute several roots using order of $D^2\delta$ ops (up to a polylogarithmic factor) to approximate δ roots for $\delta \leq D$. In section 3.4, we devise algorithms, also running in D^2 time (up to a polylog factor), that compute the numbers of distinct roots and distinct real roots of a given polynomial system of equations with real input coefficients. This improves by order of magnitude the known algorithms (not involving structured matrices and algebraic residues), which all require at least order of D^3 time to solve any of the cited computational problems.

Thus, we reached our main technical goal of developing the basic techniques for the improvement of computations with multivariate polynomials by using the associated structured matrices, dual spaces and algebraic residues. We were able to demonstrate the power of such techniques already in the present paper; in our subsequent works we will show how to accentuate this power further (in particular, by removing the assumption that the residue associated with a given polynomial system is known or readily available) and to elaborate and ameliorate the resulting algorithms from numerical and algebraic points of view. Our progress in these directions has been reported in our recent conference papers [4], [25]. In our present paper we have not touched these aspects and only provided an illustrative example for our approach. Some of the presented techniques appeared earlier in less developed form. In particular, some extensions of the well-known classes of structured matrices associated with univariate polynomials were presented in [33], but they only worked in much more restricted cases, and the restrictions do not allow to apply them to solving polynomial systems.

We will use the following order of presentation. Section 1 deals with structured matrices associated with univariate polynomials. In section 2, we present a natural generalization to the multivariate case, and in section 3, we show some applications to the polynomial root-finding problem in the multivariate case. Various advanced techniques that we use have not been collected together so far, so we detail our presentation and give many comments and some illustrative examples, in particular, one basic example is used thoughout sections 2 and 3.

In this section, we will recall the basic results on the classical matrix structure, presenting them from a univariate polynomial point of view. This will give us a sample pattern, which we will use as a springboard for developing similar techniques in the multivariate case. The reader is referred to appendix A, for the summary of the basic definitions, and to appendix B, for the summary of the estimates for the complexity of some fundamental polynomial and matrix computations.

1.1 Toeplitz operators and matrices

Consider a polynomial $t = t_0 + t_1\, x + \cdots + t_{2d}\, x^{2d}$ and the map of multiplication by this polynomial t in the ring $R = \mathbb{C}[x]$ of polynomials in the variable x with coefficients from the complex field $\mathbb{C}$:

$$\begin{array}{rccc} \mathcal{M}_t : & R & \rightarrow & R \\ & p & \mapsto & t\,p. \end{array}$$

The matrix M of this map in the monomial basis (obtained by computing the polynomials $\mathcal{M}_t(1)$, $\mathcal{M}_t(x)$, $\mathcal{M}_t(x^2), \ldots$) has the form

$$\begin{array}{c} 1 \\ \vdots \\ x^d \\ \vdots \\ x^{2d} \\ \vdots \\ \end{array} \left[\begin{array}{cccc} t_0 & & & 0 \\ \vdots & \ddots & & \\ \hline t_d & & t_0 & \cdot \\ \vdots & \ddots & \vdots & \cdot \\ t_{2d} & & t_d & \cdot \\ \hline & \ddots & \vdots & \cdot \\ 0 & & t_{2d} & \cdot \end{array}\right] \left.\begin{array}{c} \\ \\ \\ \end{array}\right\} T \tag{1}$$

The matrix M infinitely continues rightward and downward. Its rows and columns are indexed by the monomials (x^i), and its (i, j)-th entry is the coefficient of x^i in the polynomial $x^j\, t(x)$. The entries of M are invariant in their shift along the diagonal direction. This property characterizes the class of *Toeplitz matrices*:

Definition 1.1.1 *A matrix $T = (t_{i,j})$ is a Toeplitz matrix if for all i, j the entry $t_{i,j}$ depends only on $i - j$, that is, if $t_{i,j} = t_{i+1,j+1}$ for all pairs of (i, j) and $(i + 1, j + 1)$ for which the entries $t_{i,j}$ and $t_{i+1,j+1}$ are defined.*

It is immediately observed that any $h \times k$ Toeplitz matrix T where $\max\{h, k\} \leq d+1$ can be obtained as a submatrix of the matrix M defined in (1). Let $E = \{1, \ldots, x^d\}$ and $F = \{x^d, \ldots, x^{2d}\}$ be two linear subspaces of R and let π_E (resp. π_F) be the projection of R into the vector space generated by E (resp. F). Then the matrix T is just the matrix of the map

$$\mathcal{T}_t = \pi_F \circ \mathcal{M}_t \circ \pi_E.$$

The projections π_E and π_F select the first columns and the middle rows of M, respectively.

Proposition 1.1.2 *A Toeplitz operator (associated with a Toeplitz matrix) is the projection of the multiplication of a fixed polynomial by a polynomial. This is a map from R to R.*

Problem 1.1.1 *Compute the product of an $n \times n$ Toeplitz matrix by a vector.*

We observe that such a product is a subvector of the coefficient vector of the product of two polynomials of R. Therefore, by theorem B.1.1 of appendix B, we may solve problem 1.1.1 in $\mathcal{O}(n \log(n))$ ops.

Hereafter, we will use the abbreviation *f.p.s.* for formal power series and will write W^t to denote the transpose of an operator, matrix or vector W. Similarly to the map $\mathcal{M}_t$, we define the map

$$\begin{array}{rcl} \mathcal{M}_t^{\mathfrak{t}} : \ S & \to & S \\ q(\partial) & \mapsto & t(x) \star q(\partial) = \pi_+(t(\partial^{-1}) q(\partial)), \end{array}$$

where S is the ring of *f.p.s.* in the variable ∂, ∂^i is the differential form: $p \mapsto \frac{1}{i!}p^{(i)}(0)$, and π_+ is the projection of an f.p.s. in ∂ and ∂^{-1} into an f.p.s. in S obtained by deleting all the monomials in ∂^{-1}, that is, π_+ is the projection on the monomials of positive degree in ∂. The matrix of this map is the transpose of the matrix of $\mathcal{M}_t$, where we can extract the transpose of the matrix T:

$$\begin{bmatrix} t_0 & \cdots & t_d & \cdots & t_{2d} & 0 \\ & \ddots & \vdots & \ddots & \vdots & \ddots \\ & & t_0 & \cdots & t_d & \\ & & & \ddots & \vdots & \ddots \\ 0 & & & & t_0 & \end{bmatrix}.$$

1.2 Hankel operators and matrices

Next, consider the multiplication map defined by a f.p.s. $h(\partial) = h_0 + h_1\partial + \cdots + h_{2d}\partial^{2d} + \cdots$ in ∂ as follows: for any polynomial $p \in \mathbb{C}[x]$ we compute the product, $p(\partial^{-1})h(\partial)$, and project it into the space of monomials of positive degree. (Then again, the reader may think of ∂ as a variable and of ∂^{-1} as its reciprocal, and we interpret ∂^i as the linear map $p \mapsto \frac{1}{i!}p^{(i)}(0)$.) Here is the matrix M representing such maps:

$$\begin{array}{c} 1 \\ \vdots \\ \partial^d \\ \vdots \\ \partial^{2d} \\ \cdot \end{array} \left[\begin{array}{cccc} h_0 & & h_d & \cdot \\ \vdots & \iddots & \vdots & \cdot \\ h_d & & h_{2d} & \cdot \\ \hline \vdots & \iddots & \cdot & \cdot \\ h_{2d} & & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array}\right] \begin{array}{l} \left.\begin{array}{c} \\ \\ \\ \end{array}\right\} H \\ \\ \\ \\ \end{array} \tag{2}$$

The matrix M infinitely continues rightward and downward in this case. Its columns are indexed by monomials in x and its rows by monomials in ∂. The (i,j)-th entry of this matrix is the coefficient of ∂^i in $\partial^{-j}h(\partial)$, which explains why its entries are invariant in their shifts into the antidiagonal direction. This property characterizes the class of *Hankel matrices*.

Definition 1.2.1 *A matrix $H = (h_{i,j})$ is a Hankel matrix if its entry $h_{i,j}$ depends only on $i+j$, that is, if $h_{i+1,j-1} = h_{i,j}$ for all pairs (i,j) for which the entries are defined.*

Definition 1.2.2 *The space of linear forms (maps) from R to $\mathbb{C}$, that is, the dual space of the ring of polynomials R, is denoted by $\widehat{R}$. The ring of such maps from $\mathbb{C}[x]$ to $\mathbb{C}[\partial]$ is generalized to the ring of maps from the ring of f.p.s. in x to the ring of f.p.s. in ∂. We denote the latter ring by both S and $\mathbb{C}[[\partial]]$.*

The matrix M is the matrix of the map

$$\begin{array}{rcl} \mathcal{H}_h : \mathbb{C}[x] & \to & \mathbb{C}[[\partial]] \\ p(x) & \mapsto & p(x) \star h(\partial) = \pi_+(p(\partial^{-1}) h(\partial)). \end{array} \tag{3}$$

where π_+ is the projection into the space of monomials of positive degree in ∂.

We immediately observe that any $h \times k$ Hankel matrix H where $\max\{h, k\} \leq d+1$ is a submatrix of the above matrix M, defined in (2), (for some $h \in \mathbb{C}[\partial]$). Let $E = \{1, x, \ldots, x^d\}$, $F = \{1, \partial, \ldots, \partial^d\}$ be the two monomial sets in x and ∂, respectively, and let π_E and π_F be the corresponding projections on the vector spaces generated by these sets. Then the matrix H is the matrix of the following map:

$$\mathcal{H}_h = \pi_F \circ \mathcal{H}_h \circ \pi_E.$$

The projections π_E and π_F select the first columns and the middle rows of the matrix M, respectively.

Proposition 1.2.3 *A Hankel operator (associated with a Hankel matrix) can be defined as the projection of the multiplication of a (projected) polynomial by a fixed Laurent polynomial.*

Problem 1.2.1 *Compute the product of a $(d+1) \times (d+1)$ Hankel matrix by a vector.*

We observe that such a product is a subvector of the coefficient vector of the product of a fixed polynomial $h(\partial)$ by a polynomial in ∂^{-1}. Therefore, by theorem B.1.1 of appendix B, we may solve problem 1.2.1 in $O(d \log(d))$ ops.

1.3 Bezoutians

Next, let us study linear maps from $\mathbb{C}[[\partial]]$ to $\mathbb{C}[x]$. First, consider a polynomial in two variables x and y:

$$\Theta(x, y) = \sum_{i=0, j=0}^{d-1} \theta_{i,j}\, x^i\, y^j.$$

To any element $\Lambda(\partial) \in \mathbb{C}[[\partial]]$, we associate the constant coefficient in ∂ (that is, the ∂-free term) of the product

$$\Theta(x, \partial^{-1})\, \Lambda(\partial).$$

This defines a map $\mathcal{B}$ from $\mathbb{C}[[\partial]]$ to $\mathbb{C}[x]$. We immediately verify that the matrix of this map (which can be obtained by computing the constant coefficients in ∂ of $\Theta(x, \partial^{-1})\, \partial^j : \mathcal{B}(1) = \sum_{i=0}^{d-1} \theta_{i,0} x^i$, $\mathcal{B}(\partial) = \sum_{i=0}^{d-1} \theta_{i,1} x^i, \ldots)$ is precisely the coefficient matrix $[\theta_{i,j}]_{0 \leq i,j \leq d-1}$ of $\Theta(x, y)$.

A fundamental example of such a polynomial is the Bezoutian defined as follows:

Definition 1.3.1 *Let p and q be two polynomials of $\mathbb{C}[x]$. The term Bezoutian of p and q is used for both the bivariate polynomial*

$$\Theta_{q,p}(x,y) = \frac{p(x)\,q(y) - p(y)\,q(x)}{x-y} = \sum_{0\le i,j\le d-1} \theta_{i,j}^{q,p} x^i y^j$$

and the matrix

$$B_{q,p} = \begin{bmatrix} \theta_{0,0}^{q,p} & \cdots & \theta_{0,d-1}^{q,p} \\ \vdots & & \vdots \\ \theta_{d-1,0}^{q,p} & \cdots & \theta_{d-1,d-1}^{q,p} \end{bmatrix}.$$

$\mathcal{B}_{q,p} : \mathbb{C}[[\partial]] \to \mathbb{C}[x]$ denotes the associated map, $\mathcal{B}_{q,p}(\Lambda) \mapsto \partial$-free term of $\Theta_{q,p}(x, \partial^{-1})\Lambda(\partial)$. The image of this map can be expressed as the product

$$[1, x, \ldots, x^{d-1}]\ B_{q,p}[\lambda_0, \ldots, \lambda_{d-1}]^{\mathtt{t}},$$

where $\Lambda(\partial) = \sum_{i=0}^{\infty} \lambda_i \partial^i$.

In particular, if $p = p_0 + p_1\,x + \cdots + p_d\,x^d$, then the polynomial $\Theta_{1,p}$ is of the form

$$\Theta_{1,p}(x,y) = \sum_{i=0}^{d-1} x^i\, \Theta_i^p(y),$$

where $\Theta_i^p(y) = p_{i+1} + p_{i+2}\,y + \cdots + p_d\,y^{d-i-1}$. This polynomial is also called the i-th *Horner polynomial*, for it corresponds to the i-th polynomial, appearing in the so-called Horner rule for polynomial evaluation. It can be also written as

$$\Theta_i^p(y) = \pi_+(y^{-i-1}p(y)), \tag{4}$$

where π_+ is the projection on the set of polynomials in y. We immediately observe that the matrix $B_{1,p}$ associated with $\Theta_{1,p}$ is a triangular Hankel matrix of the form

$$\begin{bmatrix} p_1 & \cdots & p_d \\ \vdots & \iddots & \\ p_d & & 0 \end{bmatrix}. \tag{5}$$

More generally, we have the decomposition

$$\begin{aligned} \Theta_{q,p}(x,y) &= \frac{p(x)\,q(y) - p(y)\,q(x)}{x-y} \\ &= \frac{p(x)-p(y)}{x-y}\,q(y) - \frac{q(x)-q(y)}{x-y}\,p(y) = \Theta_{1,p}(x,y)\,q(y) - \Theta_{1,q}(x,y)\,p(y). \end{aligned}$$

This implies

$$\mathcal{B}_{q,p}(\Lambda) = \mathcal{B}_{1,p}(q \star \Lambda) - \mathcal{B}_{1,q}(p \star \Lambda)$$

for any $\Lambda(\partial) \in \mathbb{C}[[\partial]]$ or, in terms of operators,

$$\mathcal{B}_{q,p} = \mathcal{B}_{1,p} \circ \mathcal{M}_q^{\mathsf{t}} - \mathcal{B}_{1,q} \circ \mathcal{M}_p^{\mathsf{t}}. \tag{6}$$

In term of matrices, this yields the *Barnett formula*,

$$B_{q,p} = \begin{bmatrix} p_1 & \cdots & p_d \\ \vdots & \cdot^{\cdot^{\cdot}} & \\ p_d & & 0 \end{bmatrix} \begin{bmatrix} q_0 & \cdots & q_{d-1} \\ & \ddots & \vdots \\ 0 & & q_0 \end{bmatrix} - \begin{bmatrix} q_1 & \cdots & q_d \\ \vdots & \cdot^{\cdot^{\cdot}} & \\ q_d & & 0 \end{bmatrix} \begin{bmatrix} p_0 & \cdots & p_{d-1} \\ & \ddots & \vdots \\ 0 & & p_0 \end{bmatrix}.$$

1.4 Vandermonde operators and matrices

Consider the linear space R_d of polynomials of degree at most d and $d+1$ distinct points in $\mathbb{C}$: $\Xi = \{\xi_0, \ldots, \xi_d\}$. Also consider the next two bases of R_d:

- the basis of monomials $\langle 1, x, \ldots, x^d \rangle$
- and the basis of Lagrange interpolation polynomials

$$\langle L_i = L_i(x) = \prod_{j \neq i} \frac{x - \xi_l}{\xi_i - \xi_j},\ i = 0, \ldots, d \rangle.$$

Any polynomial $p \in R_d$ can be decomposed in the latter basis as follows:

$$p(x) = \sum_{i=0}^{d} p(\xi_i)\, L_i(x). \tag{7}$$

We deduce from this decomposition that the $(d+1) \times (d+1)$ *matrix of the basis transformation* from $(x^i)_{i=0,\ldots,d}$ to $(L_i(x))_{i=0,1,\ldots,d}$ is the Vandermonde matrix,

$$V(\Xi) = \begin{bmatrix} 1 & \xi_0 & \cdots & \xi_0^d \\ 1 & \xi_1 & \cdots & \xi_1^d \\ \vdots & & & \vdots \\ 1 & \xi_d & \cdots & \xi_d^d \end{bmatrix}.$$

Remark 1 *Many authors use the name "Vandermonde matrix" for $V^{\mathsf{t}}(\Xi)$, the transpose of $V(\Xi)$.*

Problem 1.4.1 *Multiply the matrix $V(\Xi)$ by a vector $p = (p_0, \ldots, p_d)^t$ or, equivalently, evaluate a polynomial $p(x) = \sum_{i=0}^{d} p_i x^i$ on the set of points $\Xi = \{\xi_0, \ldots, \xi_d\}$.*

Clearly, the multiplication of the row vector $(1, \xi_i, \ldots, \xi_i^d)^{\mathsf{t}}$ by the vector $p = (p_0, \ldots, p_d)$ amounts to evaluating the polynomial $p(x) = p_0 + \cdots + p_d\, x^d$ at the point ξ_i. Equivalently, the coefficients $p(\xi_i)$ of $p = p(x)$ in the Lagrange basis can be obtained by means of the evaluation of $p = p(x)$ at the points ξ_i.

Problem 1.4.2 *Solve linear system $V(\Xi)\mathbf{v} = \mathbf{w}$ or, equivalently, interpolate to a polynomial $p(x)$ from its values $w_0, \dots, w_d$ on a set $\Xi = \{\xi_0, \dots, \xi_d\}$.*

The known algorithms solve problems 1.4.1 and 1.4.2 in $O(d \log^2 d)$ ops (see [3], pp. 25-26).

Evaluation at a point is an example of a linear form (map), and equation (7) shows that the dual basis of $(L_i)_{i=0,\dots,d}$ (that is, the linear forms (maps) that compute the coefficients of a polynomial p in this basis) is the set of linear forms $(\mathbf{1}_{\xi_i})_{i=0,\dots,d}$ of the *evaluation* at ξ_i: $\mathbf{1}_{\xi_i}(p) = p(\xi_i)$. Such an evaluation will play important role in the following, so we will next define it formally:

Definition 1.4.1 *For any point $\xi \in \mathbb{C}$, let $\mathbf{1}_\xi \in \widehat{R} \subset \widehat{R}_{d-1}$ denote the linear form that corresponds to the evaluation at ξ:*

$$\begin{aligned} \mathbf{1}_\xi : R &\rightarrow \mathbb{C} \\ p &\mapsto p(\xi). \end{aligned}$$

Note that $\widehat{R}$ is the subset of the dual space $\widehat{R}_d$ made by the linear forms on the vector space of polynomials of degree at most d and that the coordinates of the evaluation $\mathbf{1}_\xi \in \widehat{R}_d$ in the dual basis $\langle 1, \partial, \dots, \partial^d \rangle$ of $\widehat{R}_d$ are obtained by computing $\mathbf{1}_\xi(x^i)_{i=0,\dots,d}$. This yields the vector $(1, \xi, \xi^2, \dots, \xi^d)$. In terms of polynomials in ∂, we have

$$\mathbf{1}_\xi = 1 + \xi\,\partial + \cdots + (\xi\,\partial)^d = \frac{1 - (\xi\,\partial)^{d+1}}{1 - \xi\, y}.$$

Thus, the matrix of the basis transformation from the basis $(\mathbf{1}_{\xi_i})_{i=0,\dots,d}$ to the dual basis $\langle 1, \partial, \dots, \partial^d \rangle$ of $\langle 1, x, \dots, x^d \rangle$ is given by

$$V^{\mathtt{t}}(\Xi) = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \xi_0 & \xi_1 & \cdots & \xi_d \\ \vdots & & & \vdots \\ \xi_0^d & \xi_1^d & \cdots & \xi_d^d \end{bmatrix}.$$

Problem 1.4.3 *Multiply $V^{\mathtt{t}}(\Xi)$ by a vector.*

Problem 1.4.4 *Solve the linear system $V^{\mathtt{t}}(\Xi)\mathbf{v} = \mathbf{w}$.*

Problems 1.4.3 and 1.4.4 can be solved in $O(d \log^2 d)$ ops (see [3], pp.141-144). Problem 1.4.3 can be also solved at this cost by reduction to problem 1.4.1 (see theorem B.2.1 of appendix B.2). A slower but technically interesting approach relies on the observation that the multiplication of the matrix $V^{\mathtt{t}}(\Xi)$. by a vector $\Lambda = [\lambda_0, \dots, \lambda_d]$ amounts to the computation, in the monomial basis, of the polynomial

$$\sum_{i=0}^{d} \lambda_i \frac{1 - (\xi_i\,\partial)^{d+1}}{1 - \xi_i\,\partial}.$$

If the interpolation points are the d-th roots of unity, we arrive at a special Vandermonde matrix, sometimes called the Fourier matrix. In this special case, problems 1.4.1-1.4.4 represent forward and inverse discrete Fourier transforms (DFTs) and can be solved by using $O(d \log d)$ ops. The inverse of the Fourier matrix is the transpose of its conjugate (up to the factor d).

1.5 Relations between Bezoutians and Hankel matrices

The Hankel operators correspond to some maps from $\mathbb{C}[x]$ to $\mathbb{C}[[\partial]]$, whereas the Bezoutians define some maps from $\mathbb{C}[[\partial]]$ to $\mathbb{C}[x]$. It is natural to ask if there is a relationship between the maps of these two classes. This is what we are going to examine next. We will use the basic concept of the *ideal* $I = (p)$, generated by $p \in R$, that is, the set of polynomials $\{pq, q \in R\}$.

In order to relate these two classes of operators to each other, we will next describe the elements $h(\partial) \in \mathbb{C}[[\partial]]$ such that $\mathcal{H}_h$ of (3) vanishes on all multiples of a fixed polynomial $p(x) = p_0 + p_1 x + \cdots + p_d x^d$ of degree exactly d (that is, on the ideal generated by p): $\mathcal{H}_h(p\, v) = 0$ for all elements $v \in R$.

Proposition 1.5.1 *The class of f.p.s. $h \in \mathbb{C}[[\partial]]$ such that $\mathcal{H}_h$ vanishes on the ideal (p) generated by a polynomial $p = p_0 + p_1 x + \cdots + p_d x^d$ of degree d ($p_d \neq 0$) coincides with the class of rational functions*

$$h(\partial) = \frac{\partial^{-1} r(\partial)}{p(\partial^{-1})} = h_0 + h_1 \partial + \cdots + h_{d-1} \partial^{d-1} + \cdots, \tag{8}$$

where $r(x) = \sum_{i=0}^{d-1} r_i x^i$ is a polynomial in R_{d-1}.

Proof. First, note that the rational fraction $h(\partial) = \frac{r_0 \partial^{d-1} + r_1 \partial^{d-2} + \cdots + r_{d-1}}{p_d + p_{d-1} \partial + \cdots + p_0 \partial^d}$ is an f.p.s. in ∂, having no terms ∂^{-i} for $i > 0$, since $p_d \neq 0$.

To show that $\mathcal{H}_h$ vanishes on the ideal (p) for $h(\partial)$ of (8), observe that

$$h(\partial) p(\partial^{-1}) v(\partial^{-1}) = \partial^{-1} r(\partial^{-1}) v(\partial^{-1}),$$

for $v \in R$, has only terms with negative powers of ∂ since $r(x)$ and $v(x)$ are polynomials. Therefore, $p(x)v(x) \star h(\partial) = 0$ for any polynomial $v(x) \in R$.

Finally, let $\mathcal{H}_h$ vanish on (p), for an f.p.s. $h = h(\partial)$. This means that

$$\pi_+(p(\partial^{-1})\, h(\partial^{-1})) = 0,$$

that is, $p(\partial^{-1})h(\partial)$ is an f.p.s. in ∂^{-1}, with no constant term: $p(\partial^{-1})h(\partial) = \partial^{-1} r(\partial^{-1})$, where $r(\partial)$ is an f.p.s. from $\mathbb{C}[[\partial]]$. Furthermore, by replacing ∂^{-1} by x, we obtain that $r(x) = x^{-1} p(x) h(x^{-1}) = \pi_+(x^{-1} p(x) h(x^{-1}))$, so that $r(x)$ is a polynomial of degree less than $\deg(p(x)) = d$. $\square$

The proposition implies that the class of the f.p.s. $h \in \mathbb{C}[[\partial]]$ such that $\mathcal{H}_h$ vanishes on (p) is the class of all multiples of the f.p.s. $\tau = \tau(\partial) = \frac{\partial^{-1}}{p(\partial^{-1})} \frac{\partial^{d-1}}{p_d + p_{d-1}\,\partial + \cdots + p_0\,\partial^d}$, called the (algebraic) *residue* of p. (This concept extends the concept of the residue of an analytic function.) We will next give a characterization of this residue that can be easily generalized to the multivariate case.

Proposition 1.5.2 *Let $p = p_0 + p_1 x + \cdots + p_d x^d$ be a fixed polynomial of degree exactly d. Then the residue $\tau = \tau_p(\partial)$ is the unique element of $\mathbb{C}[[\partial]]$ that satisfies:*

1. *τ vanishes on the multiples of p,*
2. *$\mathcal{B}_{1,p}(\tau) = 1$,*

where $\mathcal{B}_{1,p}$ is the map defined in definition 1.3.1.

Proof. Property 1. of τ follows from the definition of τ and proposition 1.5.1. Now, by the definition of $\tau = \tau_p(\partial)$, the element $\tau_p(\partial) = \sum_{i=0}^{\infty} \tau_i \partial^i = \sum_{i=0}^{\infty} \tau(x^i)\partial^i$ of $\mathbb{C}[[\partial]]$ has the form

$$\frac{1}{p_d}\,\partial^{d-1} + \tau_d \partial^d + \cdots,$$

that is, $\tau_0 = \cdots = \tau_{d-2} = 0$, $\tau_{d-1} = \frac{1}{p_d}$, which means that the linear form (map) associated with τ vanishes on $1, x, \ldots, x^{d-2}$ and equals $\frac{1}{p_d}$ on x^{d-1}.

Now we obtain from definition 1.3.1 that

$$\mathcal{B}_{1,p}(\tau) = [1, x, ..., x^{d-1}] B_{1,p} [0, ...0, 1/p_d]^{\mathtt{t}}.$$

As $B_{1,p}$ is of the form (5), we immediately check that

$$B_{1,p}\,[0, \ldots, 0, \frac{1}{p_d}]^{\mathtt{t}} = [1, 0, \ldots, 0]^{\mathtt{t}},$$

which implies property 2. of τ, that is $\mathcal{B}_{1,p}(\tau) = 1$. On the other hand, since $B_{1,p}$ is nonsingular, the latter vector equation implies that $\tau_0 = \cdots = \tau_{d-2} = 0$, $\tau_{d-1} = \frac{1}{p_d}$. Furthermore, as any polynomial q can be decomposed uniquely as $q = p\,u + r$, where $r \in R_{d-1} = \langle 1, x, \ldots, x^{d-1} \rangle$, the linear form τ is entirely determined by the values $\tau_0, \ldots, \tau_{d-1}$ and the fact that it vanishes on the multiples of p. This proves the uniqueness of the element of $\mathbb{C}[[\partial]]$ satisfying properties 1. and 2. and thus completes the proof of the proposition. □

Proposition 1.5.3 *The set $(\Theta_i^p)_{i=0,\ldots,d-1}$ is the dual basis of the monomial basis $(x^i)_{i=0,\ldots,d-1}$ for the inner product associated to τ:*

$$\tau(x^i\,\Theta_j^p(x)) = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ otherwise.} \end{cases} \tag{9}$$

Proof. For $0 \leq i, j \leq d-1$, we have (see (4))

$$\tau(x^i \Theta_j^p(x)) = \tau(x^i \pi_+(x^{-j-1} p(x))) = \tau(x^{i-j-1} p(x)).$$

The last equation holds because $x^i\,(x^{-j-1}p(x) - \pi_+(x^{-j-1}p(x)))$ is in the vector space $R_{-d,d-2}$ and τ vanishes on this vector space. If $i > j$, then $x^{i-j-1}p(x)$ is in the ideal (p) generated by p in R, and τ vanishes on this ideal. On the other hand, if $i < j$, then $x^{i-j-1}p(x)$ is in the vector space $R_{-d,d-2}$, and τ vanishes on this vector space too. For $i = j$, we obtain $\tau(x^{-1}p(x)) = \tau(p_d\, x^{d-1}) = 1$, which proves relations (9). □

We immediately deduce from this result the following corollary.

Corollary 1.5.4 *Let $B_1 = B_{1,p}$ and let $H_1 = H_\tau$ be the Hankel matrix of the map $\mathcal{H}_\tau$ of (3) for $h = \tau$. Then*

$$B_1\, H_1 = H_1\, B_1 = \mathbb{I}_d,$$

where $\mathbb{I}_d$ is the $d \times d$ identity matrix.

Proof. From (9), we deduce that

$$\sum_{j=0}^{d} x^j\, \tau(x^i\ \Theta_j^p(x)) = x^i.$$

On the other hand, the left-hand side of this equation equals $\mathcal{B}_{1,p}(x^i \star \tau)$. Thus, if we compose the two maps $\mathcal{H}_\tau : R_{d-1} \to \mathbb{C}[[\partial]]$ and $\mathcal{B}_{1,p} : \mathbb{C}[[\partial]] \to R_{d-1}$, we obtain that

$$\mathcal{B}_{1,p} \circ \mathcal{H}_\tau(x^i) = \mathcal{B}_{1,p}(x^i \star \tau) = x^i,$$

for $i = 0, \ldots, d-1$. In other words,

$$\mathcal{B}_{1,p} \circ \mathcal{H}_\tau = \mathbb{I}_{R_{d-1}}$$

or, equivalently, $B_1\, H_1 = \mathbb{I}_d$, which shows that the inverse of the Bezoutian B_1 is the Hankel matrix H_1 and vice versa. □

In view of corollary 1.5.4, the Barnett formula of section 1.3 (that is, the matrix version of (6)) gives us an expression for the inverse of a Hankel matrix via the $2d+2$ scalars $p_0, q_0, \ldots, p_d, q_d$. Such an expression extends the *Gohberg-Semencul* celebrated formula, known for the inverses of Toeplitz matrices (cf. [3]).

2 Generalization to the multivariate case

Our next goal, is the extension of the approach and the results of the previous section to the multivariate case. We will start with recalling some definitions and techniques used in [1], [8], [11], [21]-[24], [32]. Then, in sections 2.8, 2.10–2.12, we will develop some new techniques to be used in section 3.

2.1 Polynomial ring

The definitions of the previous section and appendix A can be immediately extended to the n-variate case, for any natural n. In this case, $R = \mathbb{C}[x]$ is replaced by the ring $\mathbb{C}[x_1, \ldots, x_n]$ of multivariate polynomials in $x_1, \ldots, x_n$; x and ∂ are assumed to be vectors, rather than scalars, $\mathbf{x} = (x_1, \ldots, x_n)$ and $\partial = (\partial_1, \ldots, \partial_n)$. We keep denoting by R_d the subspace of all polynomials of degree at most d. Instead of working in the complex space $\mathbb{C}$, we could have allowed the vector spaces over any algebraically closed field $\mathbb{K}$, and then R would denote the space of multivariate polynomials in $\mathbf{x}$, with coefficients from $\mathbb{K}$. Our results of this section would be easily extended, but, to simplify our presentation, we will state them for $\mathbb{K} = \mathbb{C}$. We will let $L = \mathbb{C}[x_1^{\pm 1}, \ldots, x_n^{\pm 1}]$ denote the ring of Laurent's polynomials in the variables $x_1, \ldots, x_n$. For any element p of R, let

$$\begin{array}{rcl} \mathcal{M}_p : R & \rightarrow & R \\ r & \mapsto & p\,r \end{array} \tag{10}$$

denote the operator of multiplication by p in R.

Hereafter, $I = (p_1, \ldots, p_n)$ denotes the ideal of $R = \mathbb{C}[\mathbf{x}]$ generated by the elements $p_1, \ldots, p_n$, that is, the set of polynomial combinations $\sum_i p_i q_i$ of these elements. $\mathcal{A} = R/I$ denotes the quotient ring defined in R by I, and $\equiv$ denotes the equality in $\mathcal{A}$. *We assume that the set of the common zeros of the n polynomials $p_1, \ldots, p_n$ (that is, the set of the roots of the polynomial system $p_1 = \cdots = p_n = 0$) is finite* and denote it $\mathcal{Z} = \mathcal{Z}(I) = \{\zeta_1, \ldots, \zeta_d\}$. This implies that the vector space $\mathcal{A}$ has a finite dimension D, $D \geq d$. (D is the number of the roots counted with their multiplicities.)

2.2 The quotient algebra

Our main objective is the analysis of the structure of $\mathcal{A}$, in particular in order to devise efficient algorithms for computing the zeros in $\mathcal{Z}(I)$.

The first operator that comes naturally in this study is the operator of multiplication by an element of $\mathcal{A}$, based on (10). For any element $a \in \mathcal{A}$, we define the map

$$\begin{array}{rcl} \overline{\mathcal{M}}_a : \mathcal{A} & \rightarrow & \mathcal{A} \\ b & \mapsto & a\,b. \end{array}$$

An important property of this operator is given in the next theorem (see [1], [32], [22]):

Theorem 2.2.1 *The set of the eigenvalues of the linear operator $\overline{\mathcal{M}}_a$ is exactly $\{a(\zeta_1), \ldots, a(\zeta_d)\}$.*

Proof. Let $p(\mathbf{x}) = \prod_{\zeta \in \mathcal{Z}(I)} (a(\mathbf{x}) - a(\zeta))$. This polynomial vanishes on $\mathcal{Z}(I)$, so that (according to the Nullstellensatz, see [9]) there exists $d = d_p \in \mathbb{N}$ such that $p(\mathbf{x})^d \in I$.

Consequently, we have

$$\prod_{\zeta\in\mathcal{Z}(I)} (\overline{\mathcal{M}}_a - a(\zeta)\mathbb{I})^d = 0,$$

where $\mathbb{I}$ is the identity map : $b \to b$, and the minimal polynomial of $\overline{\mathcal{M}}_a$ divides $\prod_{\zeta\in\mathcal{Z}(I)}(T - a(\zeta))^d$, for indeterminate T. This implies that an eigenvalue of $\overline{\mathcal{M}}_a$ is necessarily in the set $\{a(\zeta_1), \ldots, a(\zeta_d)\}$. On the other hand, we will show in theorem 2.4.1, by using duality, that for any $\zeta \in \mathcal{Z}(I)$, $a(\zeta)$ is an eigenvalue of the transpose of $\overline{\mathcal{M}}_a$. □

Example Let $n = 2$,

$$p_1 = {x_1}^2 + 2\,x_2 x_1 - x_1 - 1,\ p_2 = {x_1}^2 + {x_2}^2 - 8\,x_1.$$

We check (by hand computation) that a basis of $\mathcal{A} = \mathbb{C}[x_1, x_2]/(p_1, p_2)$ is $(1, x_1, x_2, x_1\,x_2)$ and that the matrix of multiplication by x_1 in this basis is

$$M_{x_1} = \begin{bmatrix} 0 & 1 & 0 & -\frac{14}{5} \\ 1 & 1 & 0 & -\frac{12}{5} \\ 0 & 0 & 0 & \frac{1}{5} \\ 0 & -2 & 1 & \frac{29}{5} \end{bmatrix}.$$

The eigenvalues of M_{x_1} are the first coordinates of the roots, that is,

$$6.8200982, -0.19395427 + 0.20520688\,\mathbf{i}, -0.19395427 - 0.20520688\,\mathbf{i}, 0.36781361.$$

Theorem 2.2.1 reduces the nonlinear problem of solving a polynomial system of equations to a well known problem of linear algebra. The reduction, however, involves the analysis of the structure of $\mathcal{A}$ and the properties of the operators of multiplication, and this leads us to the study of the dual space, the multivariate Bezoutians, and structured matrices associated with multivariate polynomials. This is needed, in particular, in order to express explicitly the matrices of multiplication associated with the operator $\overline{\mathcal{M}}_a$. (Such matrices are called *multiplication tables.*) The main difficulties stem from the requirement to work modulo the ideal I, and the dual space, Bezoutians, and structured matrices are effective tools for handling this nontrivial problem.

Definition 2.2.2 *Hereafter, $\mathbb{N}$ denotes the set of nonnegative integers, and we fix a subset $E \subset \mathbb{N}^n$, such that $(\mathbf{x}^\alpha)_{\alpha\in E}$ is a basis of $\mathcal{A}$.*

2.3 Dual space

Let $\widehat{R}$ denote the dual of the $\mathbb{C}$-vector space R, that is, the space of linear forms

$$\begin{array}{rcl} \lambda : R & \to & \mathbb{C} \\ p & \mapsto & \lambda(p). \end{array}$$

(R will be the primal space for $\widehat{R}$.) The *evaluation at a fixed point* ζ is a well-known example of such a linear form:

$$\begin{array}{rcl} \mathbf{1}_\zeta \ : R & \to & \mathbb{C} \\ p & \mapsto & p(\zeta). \end{array}$$

Another class of linear forms is obtained by using differential operators. Namely, for any $\mathbf{a} = (a_1, \ldots, a_n) \in \mathbb{N}^n$, consider the map

$$\begin{array}{rcl} \partial^{\mathbf{a}} : R & \to & \mathbb{C} \\ p & \mapsto & \dfrac{1}{\prod_{i=1}^n a_i!} \, (d_{x_1})^{a_1} \cdots (d_{x_n})^{a_n} \, (p)(0), \end{array} \tag{11}$$

where d_{x_i} is the derivative with respect to the variable x_i. We denote this linear form by $\partial^{\mathbf{a}} = (\partial_1)^{a_1} \cdots (\partial_n)^{a_n}$ and for any $(a_1, \ldots, a_n) \in \mathbb{N}^n, (b_1, \ldots, b_n) \in \mathbb{N}^n$ observe that

$$\frac{1}{\prod_{i=1}^n a_i!} \partial^{\mathbf{a}} \left(\prod_{i=1}^n x_i^{b_i} \right) (0) = \begin{cases} 1 \text{ if } \forall i, a_i = b_i, \\ 0 \text{ otherwise.} \end{cases}$$

It immediately follows that $(\partial^{\mathbf{a}})_{\mathbf{a} \in \mathbb{N}^n}$ is the dual basis of the primal monomial basis. By applying Taylor's expansion formula at 0, we decompose any linear form $\Lambda \in \widehat{R}$ as

$$\Lambda = \sum_{\mathbf{a} \in \mathbb{N}^n} \Lambda(\mathbf{x}^{\mathbf{a}}) \, \partial^{\mathbf{a}}.$$

The map $\Lambda \to \sum_{\mathbf{a} \in \mathbb{N}^n} \Lambda(\mathbf{x}^{\mathbf{a}}) \, \partial^{\mathbf{a}}$ defines a one-to-one correspondence between the set of linear forms Λ and the set $\mathbb{C}[[\partial_1, \ldots \partial_n]] = \mathbb{C}[[\partial]] = \{\sum_{\mathbf{a} \in \mathbb{N}^n} \lambda_{\mathbf{a}} \partial_1^{a_1} \cdots \partial_n^{a_n}\}$ of f.p.s. (formal power series) in the variables $\partial_1, \ldots, \partial_n$.

As in the univariate case, *we will identify* $\widehat{R}$ *with* $\mathbb{C}[[\partial_1, \ldots, \partial_n]]$. The evaluation at 0 corresponds to the constant 1, under this definition. It will also be denoted $\mathbf{1}_0 = \partial^0$.

Example

$$(1 + \partial_1^2 \, \partial_2)(1 + 2 \, x_1 \, x_2 + 10 \, x_1^2 \, x_2) = 11.$$

We will continue using this example throughout sections 2 and 3.

Let us next examine the structure of the dual space. We can multiply a linear form by a polynomial (we say that $\widehat{R}$ is an R-module) as follows. For any $p \in R$ and $\lambda \in \widehat{R}$, we define $p \star \Lambda$ as

$$\begin{array}{rcl} p \star \Lambda : R & \to & \mathbb{C} \\ q & \mapsto & \Lambda(p\, q). \end{array}$$

What kind of operation does this multiplication induces on the formal power series representation? For any pair of elements $p \in R$ and $d \in \mathbb{N}$, $d > 1$, we have

$$\begin{array}{rcl} (d_{x_i})^d\,(x_i\,p)(0) & = & (d_{x_i})^{d-1}\,(p + x_i\,d_{x_i}p)\,(0) \\ & = & (d_{x_i})^{d-2}\left(2\,d_{x_i}(p) + x_i\ (d_{x_i})^2\,(p)\right)(0) \\ & = & d\ (d_{x_i})^{d-1}\,(p)(0) + x_i\ (d_{x_i})^d\,(p)(0) \\ & = & d\ (d_{x_i})^{d-1}\,p(0). \end{array}$$

Also we surely have $d_{x_i}(x_i\,p)(0) = d\,p(0)$. Consequently, for any pair of elements $p \in R, \mathtt{d} = (d_1, \ldots, d_n) \in \mathbb{N}^n$, where $d_i \neq 0$ for a fixed i, we obtain that

$$\begin{array}{rcl} x_i \star \partial^{\mathtt{d}}(p) & = & \partial^{\mathtt{d}}(x_i\,p) \\ & = & \partial_1^{d_1} \cdots \partial_{i-1}^{d_{i-1}}\,\partial_i^{d_i-1}\,\partial_{i+1}^{d_{i+1}} \cdots \partial_n^{d_n}(p), \end{array}$$

that is, x_i acts as the *inverse* of ∂_i in $\mathbb{C}[[\partial]]$. This is the reason why in the literature such a representation is referred to as the *inverse systems* (see, for instance, [21]). If $d_i = 0$, then $x_i \star \partial^{\mathtt{d}}(p) = 0$, which allows us to redefine the product $p \star \Lambda$ as follows:

Proposition 2.3.1 *For any $p, q \in R$ and any $\Lambda(\partial) \in \mathbb{C}[[\partial]]$, we have*

$$p \star \Lambda(q) = \Lambda(p\,q) = \pi_+(p(\partial^{-1})\,\Lambda(\partial))(q).$$

Example (continued)
$(x_1 \star (1 + \partial_1^2\,\partial_2))\,(1 + 2\,x_1\,x_2 + 10\,x_1^2\,x_2)$

$$= (1 + \partial_1^2\,\partial_2)(x_1 + 2\,x_1^2 x_2 + 10 x_1^3\,x_2) \tag{12}$$

$$= \partial_1\,\partial_2(1 + 2\,x_1\,x_2 + 10\,x_1^2\,x_2) = 2. \tag{13}$$

For any linear form $\Lambda \in \widehat{R}$, let

$$\begin{array}{rcl} \mathcal{H}_\Lambda : R & \to & \widehat{R} \\ r & \mapsto & r \star \Lambda \end{array}$$

denote the operator of multiplication by Λ, from R to $\widehat{R}$.

2.4 The dual of the quotient algebra

Now, let $\widehat{\mathcal{A}}$ be the dual space of $\mathcal{A}$. It is possible to identify the set $\widehat{\mathcal{A}}$ with the elements of $\widehat{R}$ that vanish on I. Thus, the set $\widehat{\mathcal{A}}$ will be also denoted $I^{\perp}$. Now, for any element $a \in \mathcal{A}$, we can describe the transposed operator $\overline{\mathcal{M}}_a^{\mathsf{t}}$:

$$\begin{array}{rcl} \overline{\mathcal{M}}_a^{\mathsf{t}} : \widehat{\mathcal{A}} & \rightarrow & \widehat{\mathcal{A}} \\ \Lambda & \mapsto & a \star \Lambda = \Lambda \circ \overline{\mathcal{M}}_a. \end{array}$$

The matrix associated to this operator is the transpose of the matrix associated to the matrix $\overline{\mathcal{M}}_a$.

We have already described the eigenvalues of this operator in theorem 2.2.1 and will give now a description of its eigenvectors (see [22], [32]):

Theorem 2.4.1 *The common eigenvectors of the operators $\overline{\mathcal{M}}_a^{\mathsf{t}}$, for $a \in \mathcal{A}$, are (up to a scalar factor) the evaluations $\mathbf{1}_{\zeta_1}, \ldots, \mathbf{1}_{\zeta_d}$, where $\mathbf{1}_{\zeta} : p \rightarrow p(\zeta)$.*

Proof. For any pair of polynomials $a, b \in R$ and any $\zeta_i \in \mathcal{Z}(I)$, we have

$$\overline{\mathcal{M}}_a^{\mathsf{t}}(\mathbf{1}_{\zeta_i})(b) = \mathbf{1}_{\zeta_i}(a\, b) = a(\zeta_i)\, \mathbf{1}_{\zeta_i}(b),$$

that is, $\overline{\mathcal{M}}_a^{\mathsf{t}}(\mathbf{1}_{\zeta_i}) = a(\zeta_i)\, \mathbf{1}_{\zeta_i}$. Moreover, $\mathbf{1}_{\zeta_i}$ is in $\widehat{\mathcal{A}}$, because ζ_i is a common root of the polynomials in I. Therefore, for any $a \in R$, $\mathbf{1}_{\zeta_i}$ is an eigenvector of $\overline{\mathcal{M}}_a^{\mathsf{t}}$ associated with the eigenvalue $a(\zeta_i)$.

Conversely, let us prove that the common eigenvectors of $(\overline{\mathcal{M}}_{x_i}^{\mathsf{t}})_{i=1,\ldots,n}$ are (up to a scalar factor) exactly $\mathbf{1}_{\zeta_1}, \ldots, \mathbf{1}_{\zeta_d}$. Let $\Lambda \in \widehat{\mathcal{A}}$ be a non-zero common eigenvector of $(\overline{\mathcal{M}}_{x_i}^{\mathsf{t}})_{i=1,\ldots,n}$ for the eigenvalues $(\gamma_i)_{i=1,\ldots,n}$: $x_i \star \Lambda - \gamma_i\, \Lambda = 0$. Then, for any monomial $\mathbf{x}^{\alpha}$ of R, we have

$$x_i \star \Lambda(\mathbf{x}^{\alpha}) = \Lambda(x_i\, \mathbf{x}^{\alpha}) = \gamma_i \Lambda(\mathbf{x}^{\alpha}).$$

By induction, this implies that $\Lambda(\mathbf{x}^{\alpha}) = \gamma^{\alpha}\, \Lambda(1)$. In other words, $\Lambda = \Lambda(1)\, \mathbf{1}_{\gamma}$, where $\gamma = (\gamma_1, \ldots, \gamma_n) \in \mathbb{C}^n$ and $\mathbf{1}_{\gamma} \in \widehat{R}$ is the evaluation at γ. As $\Lambda \in \widehat{\mathcal{A}} \equiv I^{\top}$, we have $\Lambda(p) = \Lambda(1)\, \mathbf{1}_{\gamma}(p) = \Lambda(1)\, p(\gamma) = 0$, for any $p \in I$, which implies that $\gamma \in \mathcal{Z}(I)$. $\square$

Both theorems 2.2.1 and 2.4.1 reduce the solution of a polynomial system to matrix eigenproblem, but theorem 2.4.1 has an advantage compared to theorem 2.2.1: *Each eigenvector of an operator $\overline{\mathcal{M}}_a^{\mathsf{t}}$ defines all the coordinates of a root* (whereas each eigenvalue of $\mathcal{M}_a$ defines only one coordinate or the inner product of the vector of a root by a fixed vector defined by $a \in \mathcal{A}$). Indeed, the evaluations $\mathbf{1}_{\zeta_i}$ at the roots $\zeta_i \in \mathcal{Z}(I)$ are eigenvectors of $\overline{\mathcal{M}}_a^{\mathsf{t}}$. From these evaluations $\mathbf{1}_{\zeta_i}$, we can recover the coordinates $\zeta_{i,j} = \mathbf{1}_{\zeta_i}(x_j)$ of the root $\mathbf{1}_{\zeta_i}$. We will specify this computation in section 3.1.

2.5 Quasi-Toeplitz and quasi-Hankel matrices

Definition 2.5.1 *Let E and F be two finite subsets of $\mathbb{N}^n$ and let $M = (m_{\alpha,\beta})_{\alpha \in E, \beta \in F}$ be a matrix whose rows are indexed by the elements of E and columns by the elements of F. To the end of this definition, let $\mathbf{i}$ denote the i-th coordinate vector of $\mathbb{N}^n$, $i = 1, \ldots, n$.*

- *M is an (E, F)* quasi-Toeplitz *matrix iff, for all $\alpha \in E, \beta \in F$, the entries $m_{\alpha,\beta} = t_{\alpha-\beta}$ depend only on $\alpha - \beta$, that is, if for every $i = 1, \ldots, n$, we have $m_{\alpha+\mathbf{i},\beta+\mathbf{i}} = m_{\alpha,\beta}$, provided that $\alpha, \alpha + \mathbf{i} \in E; \beta, \beta + \mathbf{i} \in F$; such a matrix M is associated with the polynomial $T_M(\mathbf{x}) = \sum_{\mathbf{u} \in E+F} t_{\mathbf{u}} \mathbf{x}^{\mathbf{u}}$.*

- *M is an (E, F)* quasi-Hankel *matrix iff, for all $\alpha \in E, \beta \in F$, the entries $m_{\alpha,\beta} = h_{\alpha+\beta}$ depend only on $\alpha + \beta$, that is, if for every $i = 1, \ldots, n$, we have $m_{\alpha-\mathbf{i},\beta+\mathbf{i}} = m_{\alpha,\beta}$ provided that $\alpha, \alpha - \mathbf{i} \in E; \beta, \beta + \mathbf{i} \in F$; such a matrix M is associated with the Laurent polynomial $H_M(\partial) = \sum_{\mathbf{u} \in E-F} h_{\mathbf{u}} \partial^{\mathbf{u}}$.*

If we work with the Laurent polynomials, these definitions can be immediately extended to subsets E, F of $\mathbb{Z}^n$, $\mathbb{Z}$ denoting the set of all integers.

For $E = [0, \ldots, h-1]$ and $F = [0, \ldots, k-1]$, definition 2.5.1 turns into the usual definition of $h \times k$ Hankel (resp. Toeplitz) matrices (see sections 1.1 and 1.2). For E and F forming rectangles in $\mathbb{N}^n$, the quasi-Toeplitz matrices appeared in [33] under the name of multilevel Toeplitz matrices. For our study of the multivariate polynomial systems the latter class is not sufficiently general, and we need our definition 2.5.1 due to [23] (cf. also [24]).

The two classes of quasi-Toeplitz and quasi-Hankel matrices can be immediately transformed into each other by their pre- or post-multiplicaiton by the reflection matrix, having its antidiagonal filled with ones and having all other entries zero.

Definition 2.5.2 *Let $\pi_E : L \to L$ be the projection map such that*

$$\pi_E(\mathbf{x}^\alpha) = \mathbf{x}^\alpha$$

if $\alpha \in E$ and $\pi_E(\mathbf{x}^\alpha) = 0$ otherwise. We also let $\pi_E : \mathbb{C}[[\partial]] \to \mathbb{C}[[\partial]]$ denote the projection map such that $\pi_E(\partial^\alpha) = \partial^\alpha$ is $\alpha \in E$ and $\pi_E(\partial^\alpha) = 0$ otherwise.

We can describe the quasi-Toeplitz and quasi-Hankel operators in terms of polynomial multiplication (see [24], [23]).

Proposition 2.5.3 *The matrix M is an (E, F) quasi-Toeplitz (resp. an (E, F) quasi-Hankel) matrix, if and only if it is the matrix of the operator $\pi_E \circ \mathcal{M}_{T_M} \circ \pi_F$ (resp. $\pi_E \circ \mathcal{H}_{H_M} \circ \pi_F$).*

Proof. We will give a proof only for an (E, F) quasi-Toeplitz matrix $M = (M_{\alpha,\beta})_{\alpha \in E, \beta \in F}$. (The proof is similar for a quasi-Hankel matrix.) The associated polynomial is

$T_M(\mathbf{x}) = \sum_{\mathbf{u}\in E+F} t_{\mathbf{u}} \mathbf{x}^u$. For any vector $\mathbf{v} = [v_\beta] \in \mathbb{C}^F$, let $v(\mathbf{x})$ denote the polynomial $v(\mathbf{x}) = \sum_{\beta\in F} v_\beta \mathbf{x}^\beta$. Then

$$\begin{aligned} T_M(\mathbf{x})\, v(\mathbf{x}) &= \sum_{\mathbf{u}\in E+F, \beta\in F} \mathbf{x}^{\mathbf{u}+\beta}\, t_{\mathbf{u}}\, v_\beta \\ &= \sum_{\alpha=\mathbf{u}+\beta\in E+2\,F} \mathbf{x}^\alpha \left(\sum_{\beta\in F} t_{\alpha-\beta}\, v_\beta \right), \end{aligned}$$

where we assume that $v_\beta = 0$ if $\mathbf{u} \notin E + F$, $t_{\mathbf{u}} = 0$ if $u \notin E + F$. Therefore, for $\alpha \in E$, the coefficient of $\mathbf{x}^\alpha$ equals

$$\sum_{\beta\in F} t_{\alpha-\beta}\, v_\beta = \sum_{\beta\in F} M_{\alpha,\beta}\, v_\beta,$$

which is precisely the coefficient α of $M\mathbf{v}$. □

Due to proposition 2.5.3, multiplication of an (E, F) quasi-Toeplitz (resp. quasi-Hankel) matrix by a vector $\mathbf{v} = [v_\beta] \in \mathbb{C}^F$ reduces to (Laurent's) polynomial multiplication.

Algorithm 2.5.1 (Multiplication of the (E, F) quasi-Toeplitz (resp. quasi-Hankel) matrix $M = (M_{\alpha,\beta})_{\alpha\in E,\beta\in F}$ by a vector $\mathbf{v} = [v_\beta] \in \mathbb{C}^F$ *). Multiply the two polynomials,* $T_M = \sum_{\mathbf{u}\in E+F} t_{\mathbf{u}}\, \mathbf{x}^{\mathbf{u}}$ *(resp.* $H_M(\partial) = \sum_{\mathbf{u}\in E-F} h_{\mathbf{u}} \partial^{\mathbf{u}}$*) by* $v(\mathbf{x}) = \sum_{\beta\in F} v_\beta \mathbf{x}^\beta$ *(resp.* $v(\partial^{-1}) = \sum_{\beta\in F} v_\beta \partial^{-\beta}$*), and output the projection of the product on* $\mathbf{x}^E$ *(resp.* ∂^E*).*

Hereafter, $C_{PolMult}(E, F)$ denotes the number of ops required to multiply a polynomial with a support in E by a polynomial with a support in F. (We will estimate $C_{PolMult}(E, F)$ in appendix B.) Algorithm 2.5.1 can be performed by using $C_{PolMult}(E + F, F)$, resp. $C_{PolMult}(E - F, -F)$, ops. According to the estimates of the appendix B, this means $\mathcal{O}(N \log^2 N + C_{M,N})$ ops, where $N = \lfloor E - 2\,F \rceil$ (resp. $\lfloor E + 2F \rceil$) and where $C_{M,N}$ bounds the cost of evaluating the polynomial H_M (resp. T_M) on a fixed set of N points.

The displacement rank analysis (cf. [3]) can be also generalized to the multivariate case. Instead of the well-known displacement matrices

$$Z = \begin{pmatrix} 0 \cdots\cdots\cdots\cdots & 0 \\ 1 \ \ddots & \vdots \\ 0 \ \ddots \ \ddots & \vdots \\ \vdots \ \ \ddots \ \ddots \ \ddots & \vdots \\ 0 \cdots\cdots 0 \ \ 1 & 0 \end{pmatrix}$$

and $Z^{\mathbf{t}}$, we use the following operators (one per variable):

$$\mathcal{Z}_i^E = \pi_E\, \mathcal{M}_{x_i} \pi_E \tag{12}$$

and

$$\mathcal{Z}_{-i}^{E} = \pi_E \, \mathcal{M}_{x_i^{-1}} \pi_E, \tag{13}$$

respectively. The displacement rank (that is, the rank of the matrix obtained by applying the displacement operator of the matrix) is bounded by the sum in i of the sizes of the boundary of E and F in the direction i (see [24], [23]).

Example Let the sets E and F correspond to the set of the monomials in x_1, x_2 graphically represented as follows:

$$\begin{array}{ccccc} \circ & \circ & & & \\ \circ & \circ & \circ & \circ & \circ \end{array}$$

Then the displacement rank is less than $2 \times 2 = 4$ in the direction x_1 and is less than $2 \times 5 = 10$ in the direction x_2.

In other words, the flatter the sets E and F in a fixed direction, the smaller the displacement rank in this direction.

If $E = F = \{(\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n \; ; \; 0 \le \alpha_i \le d_i - 1\}$, the displacement rank for the operator associated to $\mathcal{Z}_i$ is $\le 2 \prod_{j \ne i} d_j$.

2.6 Multivariate Bezoutians

In this section and in the next one, we will recall some basic definitions from the theories of Bezoutians and algebraic residues (compare the special univariate cases of sections 1.3 and 1.5), referring the reader to [8], [11] for further details and to section 3 for some applications.

In addition to the vector of variables $\mathbf{x}$, consider another vector of varialbes, $\mathbf{y} = (y_1, \ldots, y_n)$, and write $\mathbf{x}^{(0)} = \mathbf{x}$, $\mathbf{x}^{(1)} = (y_1, x_2, \ldots, x_n)$, $\ldots$, $\mathbf{x}^{(n)} = \mathbf{y}$. For a polynomial $q \in R$, define $\theta_i(q) = \frac{q(\mathbf{x}^{(i)}) - q(\mathbf{x}^{(i-1)})}{y_i - x_i}$, the *discrete differentiation* of q. For a sequence of $n+1$ polynomials $q, p_1, \ldots, p_n \in R$, construct the following polynomial in $\mathbf{x}$ and $\mathbf{y}$:

$$\Theta_{\mathbf{p}}(q) = \Theta_{q,\mathbf{p}} = \det \begin{pmatrix} q(\mathbf{x}) & \theta_1(q) & \cdots & \theta_n(q) \\ \vdots & \vdots & & \vdots \\ p_n(\mathbf{x}) & \theta_1(p_n) & \cdots & \theta_n(p_n) \end{pmatrix} = \sum_{\alpha,\beta} \theta_{\alpha,\beta}^{q,\mathbf{p}} \, \mathbf{x}^{\alpha} \, \mathbf{y}^{\beta}, \tag{14}$$

where $\det(A)$ denotes the determinant of a matrix A, $\mathbf{p} = (p_1, \ldots, p_n)$, and α and β vary in fixed ranges. This polynomial of $\mathbb{C}[\mathbf{x}, \mathbf{y}]$ is called the *Bezoutian* of $q, p_1, \ldots, p_n$. It defines a map $\mathcal{B}_{q,\mathbf{p}}$:

$$\begin{array}{rcl} \mathcal{B}_{q,\mathbf{p}} : \widehat{R} & \to & R \\ \Lambda & \mapsto & \sum_{\alpha,\beta} \theta_{\alpha,\beta}^{q,\mathbf{p}} \, \mathbf{x}^{\alpha} \, \Lambda(\mathbf{y}^{\beta}). \end{array}$$

Using the representation of Λ as a formal power series in $\partial_1, \ldots, \partial_n$, we obtain the value of $\mathcal{B}_{q,\mathbf{p}}(\Lambda(\partial))$ by computing the term free of $\partial_1, \ldots, \partial_n$ in the product

$$\Theta_{q,\mathbf{p}}(\mathbf{x}, \partial^{-1})\Lambda(\partial).$$

This construction extends the construction of section 1.3 to the multivariate case. The matrix of the map $\mathcal{B}_{q,\mathbf{p}}$ in the monomial basis is the matrix of the coefficients $[\theta^q_{\alpha,\beta}]$.

If $(\mathbf{x}^\alpha)_{\alpha\in E}$ is a basis of $\mathcal{A}$, then for any q in R, the polynomial $\Theta_{\mathbf{p}}(q)$ can be rewritten as

$$\Theta_{\mathbf{p}}(q) \equiv \sum_{\alpha,\beta\in E} B^{q,\mathbf{p}}_{\alpha,\beta}\, \mathbf{x}^\alpha \mathbf{y}^\beta. \tag{15}$$

This polynomial is obtained from (14) by reducing $\Theta_{q,\mathbf{p}}$ modulo I.

To simplify the notation, we will write $B^q_{\alpha,\beta}$, dropping the superscript $\mathbf{p}$ for a fixed ideal $(\mathbf{p})$.

Example (continued) We have

$$\begin{aligned}
\Theta_{\mathbf{p}}(1) &= x_1x_2 + 2\,{x_2}^2 + (-2\,y_1 + y_2)\,x_1 + (y_1 + 2\,y_2 - 1)\,x_2 \\
&\quad -2\,{y_1}^2 + y_1y_2 + 16\,y_1 - y_2 \\
&\equiv 5\,x_1x_2 + (y_2 - 2\,y_1 + 14)\,x_1 + (2\,y_2 + y_1 - 1)\,x_2 \\
&\quad +5\,y_1y_2 - y_2 + 14\,y_1 - 4.
\end{aligned} \tag{16}$$

Definition 2.6.1 *The matrix*

$$B_{q,\mathbf{p}} = [B^{q,\mathbf{p}}_{\alpha,\beta}]_{\alpha,\beta\in E}, \tag{17}$$

associated to the polynomial $\Theta_{\mathbf{p}}(q)$ of (15), is called the Bezoutian matrix *or the* Bezoutian *of $q, \mathbf{p}$. This is the matrix of the map*

$$\begin{aligned}
\overline{\mathcal{B}}_{q,\mathbf{p}} : \widehat{\mathcal{A}} &\rightarrow \mathcal{A} \\
\Lambda &\mapsto \sum_{\alpha,\beta\in E} B^q_{\alpha,\beta}\, \mathbf{x}^\alpha \Lambda(\mathbf{y}^\beta)
\end{aligned}$$

in the monomial basis $(\mathbf{x}^\alpha)_{\alpha\in E}$ and its dual. When $\mathbf{p}$ is fixed, we will write B_q and $\overline{\mathcal{B}}_q$ instead of $B_{q,\mathbf{p}}$ and $\overline{\mathcal{B}}_{q,\mathbf{p}}$.

Example (continued) In the basis $(1, x_1, x_2, x_1\,x_2)$ of $\mathcal{A} = \mathbb{C}[x_1, x_2]/(p_1, p_2)$ we have

$$B_1 = \begin{bmatrix} -4 & 14 & -1 & 5 \\ 14 & -2 & 1 & 0 \\ -1 & 1 & 2 & 0 \\ 5 & 0 & 0 & 0 \end{bmatrix}.$$

The rows of this matrix are filled with the coefficients of the monomials in x_1, x_2 in (16). It is a symmetric matrix, which is a property of the Bezoutians.

2.7 Bezoutians and algebraic residues

We will next define the residue and recall some fundamental properties of the multivariate Bezoutians and residues, to end with some correlations between primal and dual multiplication tables in the next section.

Definition 2.7.1 *The residue of* $\mathbf{p} = (p_1, \ldots, p_n)$ *is the unique linear form* τ *in the set of linear forms on* R *such that*

1. τ *vanishes on* $(\mathbf{p})$,
2. $\mathcal{B}_{1,\mathbf{p}}(\tau) - 1 \in (\mathbf{p})$.

This definition extends the characterization of the residue of proposition 1.5.2, given in the univariate case; we now consider all polynomials modulo the ideal $(\mathbf{p})$, in particular, $\mathcal{B}_{\mathbf{p}}(q)$ is modulo $(\mathbf{p})$. This is not a constructive definition; we prove the existence of τ but give no general recipe for computing τ yet.

Consider the decomposition $\Theta_{1,\mathbf{p}} \equiv \sum_{\alpha,\beta\in E} B^1_{\alpha,\beta}\, \mathbf{x}^\alpha\, \mathbf{y}^\beta$ and let us write $\mathbf{w}_\alpha(\mathbf{y}) = \sum_{\beta\in E} B^1_{\alpha,\beta}\mathbf{y}^\beta$, so that

$$\Theta_{1,\mathbf{p}} \equiv \sum_{\alpha\in E} \mathbf{x}^\alpha\, \mathbf{w}_\alpha(\mathbf{y}).$$

Then we have the following property:

Proposition 2.7.2 *The set* $(\mathbf{w}_\alpha)_{\alpha\in E}$ *is the dual basis of* $(\mathbf{x}^\alpha)$ *for* τ*:*

$$\tau(\mathbf{x}^\alpha\, \mathbf{w}_\beta) = \begin{cases} 1 \text{ if } \alpha = \beta \\ 0 \text{ otherwise.} \end{cases}$$

Example (continued) The residue is defined on $(1, x_1, x_2, x_1\, x_2)$ by

$$\tau(1) = \tau(x_1) = \tau(x_2) = 0, \quad \tau(x_1\, x_2) = \frac{1}{5}.$$

It vanishes on all multiples of p_1, p_2. According to (16), the dual basis of $(1, x_1, x_2, x_1\, x_2)$ is given by

$$\mathbf{w}_1 = 5\, y_1 y_2 - y_2 + 14\, y_1 - 4,\;\; \mathbf{w}_{x_1} = y_2 - 2\, y_1 + 14,\;\; \mathbf{w}_{x_2} = 2\, y_2 + y_1 - 1,\;\; \mathbf{w}_{x_1 x_2} = 5.$$

Again, we are going to study the properties of the dual basis but do not give yet any algorithm for actually computing this basis. According to proposition 2.7.2, for any $a \in \mathcal{A}$, we have the relations

$$a \equiv \sum_{\alpha\in E} \tau(a\, \mathbf{x}^\alpha)\, \mathbf{w}_\alpha \equiv \sum_{\alpha\in E} \tau(a\, \mathbf{w}_\alpha)\, \mathbf{x}^\alpha. \tag{18}$$

We also have the following simple but fundamental property ([8], [11]):

$$\Theta_{1,\mathbf{p}} \equiv \sum_{\alpha \in E} \mathbf{x}^{\alpha}\, \mathbf{w}_{\alpha}(\mathbf{y}) \equiv \sum_{\alpha \in E} \mathbf{w}_{\alpha}(\mathbf{x})\, \mathbf{y}^{\alpha} \bmod (\mathbf{p}(\mathbf{x}), \mathbf{p}(\mathbf{y})), \tag{19}$$

which shows that B_1 is a symmetric matrix.

Moreover, we recall from [8], [11] that for any polynomial $q \in R$ we have

$$\Theta_{q,\mathbf{p}} = \Theta_{1,\mathbf{p}}(\mathbf{x}, \mathbf{y})\, q(\mathbf{x}) \equiv \Theta_{1,\mathbf{p}}(\mathbf{x}, \mathbf{y})\, q(\mathbf{y}) \mod (\mathbf{p}(\mathbf{x}), \mathbf{p}(\mathbf{y})). \tag{20}$$

In particular, we set $q(\mathbf{x}) = x_i$ for $i = 1, \ldots, n$, and for any pair, ζ and η, of distinct roots of the polynomial system $\mathbf{p} = \mathbf{0}$, we set $\mathbf{x} = \zeta$, $\mathbf{y} = \eta \in \mathcal{Z}(I)$ and deduce that

$$\Theta_{1,\mathbf{p}}(\zeta, \eta) = 0. \tag{21}$$

If $\zeta = \eta$, then $\Theta_{1,\mathbf{p}}(\zeta, \eta) = J(\zeta)$, where $J = J_{\mathbf{p}} = (\partial p_i / \partial x_j)$ is the Jacobian of $\mathbf{p}$.

2.8 Bezoutians and multiplication tables in primal and dual bases

The notion of dual basis (for τ), defined in the previous section, should not be confused with the following notion of dual basis in the dual space $\widehat{\mathcal{A}}$:

Definition 2.8.1 *Given a basis $(b_i)_{i=1,\ldots,D}$ of $\mathcal{A}$, let $(\widehat{b_i})_{i=1,\ldots,D}$ denote the dual basis of (b_i), that is, the basis set of linear forms in $\widehat{R}$ that compute the coefficients of any $a \in \mathcal{A}$ in the primal basis.*

The next proposition relates the matrix $B_a = B_{a,\mathbf{p}}$ of the map $\overline{\mathcal{B}}_a$ of definition 2.6.1 with $q = a$ (cf. (17)), to the transformations between the primal bases $(\mathbf{x}^{\alpha})$ and $(\mathbf{w}_{\alpha})$ and their dual bases $(\widehat{\mathbf{x}^{\alpha}})$ and $(\widehat{\mathbf{w}_{\alpha}})$, respectively.

Proposition 2.8.2 *The matrix of the map $\overline{\mathcal{B}}_a$ of definition 2.6.1,*

1. *from the basis $(\widehat{\mathbf{x}^{\alpha}})$ of $\widehat{\mathcal{A}}$ to the basis $(\mathbf{x}^{\alpha})$ of $\mathcal{A}$ is $B_a = (\tau(a\, \mathbf{w}_{\alpha}\, \mathbf{w}_{\beta}))$,*
2. *from the basis $(\widehat{\mathbf{w}_{\alpha}})$ to the basis $(\mathbf{w}_{\alpha})$ is $H_a = (\tau(a\, \mathbf{x}^{\alpha}\, \mathbf{x}^{\beta}))$.*

Proof. According to proposition 2.7.2, the coordinates of $\overline{\mathcal{B}}_a(\widehat{\mathbf{x}^{\beta}})$ in the basis $(\mathbf{x}^{\alpha})_{\alpha \in E}$ are given by

$$\tau(\overline{\mathcal{B}}_a(\widehat{\mathbf{x}^{\beta}})\, \mathbf{w}_{\alpha}).$$

The identities (20) and (19) imply that $\Theta_{\mathbf{p}}(a) \equiv a(\mathbf{x})\Theta_{\mathbf{p}}(1)$, and $\overline{\mathcal{B}}_a(\widehat{\mathbf{x}^{\beta}}) \equiv a\, \overline{\mathcal{B}}_1(\widehat{\mathbf{x}^{\beta}}) \equiv a\, \mathbf{w}_{\beta}$. Therefore,

$$\tau(\overline{\mathcal{B}}_a(\widehat{\mathbf{x}^{\beta}})\, \mathbf{w}_{\alpha}) = \tau(a\, \overline{\mathcal{B}}_1(\widehat{\mathbf{x}^{\beta}})\, \mathbf{w}_{\alpha}) = \tau(a\, \mathbf{w}_{\alpha}\, \mathbf{w}_{\beta}).$$

In other words, we have $B^a_{\alpha,\beta} = \tau(\overline{\mathcal{B}}_a(\widehat{\mathbf{x}^\beta})\,\mathbf{w}_\alpha)$. This proves the first part of the proposition.

The coordinates of $\overline{\mathcal{B}}_a(\widehat{\mathbf{w}_\beta})$ in the basis $(\mathbf{w}_\alpha)_{\alpha\in E}$ are given by

$$\tau(\overline{\mathcal{B}}_a(\widehat{\mathbf{w}_\beta})\,\mathbf{x}^\alpha).$$

According to identities (20) and (19), we also have

$$\tau(\overline{\mathcal{B}}_a(\widehat{\mathbf{w}_\beta})\,\mathbf{x}^\alpha) = \tau(a\,\overline{\mathcal{B}}_1(\widehat{\mathbf{w}_\beta})\,\mathbf{x}^\alpha) = \tau(a\,\mathbf{x}^\alpha\,\mathbf{x}^\beta),$$

which proves the second part of the proposition. □

Now, we deduce some simple correlations between multiplication tables in the bases $(\mathbf{x}^\alpha)$ and $(\mathbf{w}_\alpha)$.

Definition 2.8.3 *For any a in $\mathcal{A}$, let $M_a = (M^a_{\alpha,\beta})$ denote the matrix of the map $\overline{\mathcal{M}}_a$ in the basis $(\mathbf{x}^\alpha)$ and let $N_a = (N^a_{\alpha,\beta})_{\alpha,\beta\in E}$ denote its matrix in the basis $(\mathbf{w}_\alpha)$.*

Proposition 2.8.4 *The matrix N_a of multiplication by a in $\mathcal{A}$, in the basis $(\mathbf{w}_\alpha)$, is the transpose M_a^{t} of the matrix M_a of multiplication by a in $\mathcal{A}$, in the basis $(\mathbf{x}^\alpha)$.*

Proof. For any $\alpha \in E$, we have

$$b\,\mathbf{x}^\beta \equiv \sum_{\gamma\in E} M^a_{\gamma,\beta}\,\mathbf{x}^\gamma,\;\; b\,\mathbf{w}_\beta \equiv \sum_{\gamma\in E} N^a_{\gamma,\beta}\,\mathbf{w}_\gamma,$$

and

$$\begin{aligned} M^a_{\alpha,\beta} &= \tau(b\,\mathbf{x}^\beta\,\mathbf{w}_\alpha),\\ N^a_{\alpha,\beta} &= \tau(a\,\mathbf{x}^\alpha\,\mathbf{w}_\beta). \end{aligned}$$

Therefore, $N_a = M_a^{\mathsf{t}}$. □

The proposition also implies that the matrix of the transposed map $\overline{\mathcal{M}}_a^{\mathsf{t}}$ in the dual basis $(\widehat{\mathbf{x}^\alpha})$ of $(\mathbf{w}_\alpha)$ is M_a.

2.9 Multivariate (stipped) Vandermonde matrices

Vandermonde matrices can be immediately generalized to the multivariate case, in the following way.

Definition 2.9.1 *For a set $(\mathbf{x}^\alpha)_{\alpha\in E}$ of D monomials and a set $\xi = (\xi_1, \ldots, \xi_D)$ of D points of $\mathbb{C}^n$, define the Vandermonde matrix of ξ on E by*

$$V_E(\xi) = [\xi_i^\alpha]_{i=1,\ldots,D;\ \alpha\in E}.$$

The rows of this matrix are the vectors $[\mathbf{x}^\alpha]_{\alpha\in E}$ of monomials evaluated at points ξ_i (for $i = 1, \ldots, D$).

$V_E(\xi)$ is the matrix of the coefficients (of $(\partial^\alpha)_{\alpha\in E}$) in the f.p.s. representing the evaluations $\mathbf{1}_{\xi_i}$ at the points ξ_i. Multivariate Vandermonde matrices are also called *stripped* because the powers ξ_i^α do not generally fill the entire D-dimensional box of integers.

Algorithm 2.9.1 *(*Multiplication of a multivariate Vandermonde matrix $V_E(\xi)$ by a vector $\mathbf{v}$ and the solution in $\mathbf{v}$ of a linear system $V_E(\xi)\mathbf{v} = \mathbf{w}$, for given ξ, E and $\mathbf{w}$). *Perform multipoint evaluation at the node-points ξ_i, $i = 1, \ldots, D$, of the associated multivariate polynomial with the coefficient vector $\mathbf{w}$ (resp. perform the converse operation of multivariate polynomial interpolation).*

See [6] and [12], for a record (asymptotic) bounds on the number of ops involved in algorithm 2.9.1.

2.10 Relations between Hankel and Bezoutian matrices

Motivated by applications of matrix computations to the solution of polynomial systems, we are particularly interested in studying *multiplication tables* (see theorems 2.2.1, 2.4.1).

Definition 2.10.1 *For any Λ in $\widehat{\mathcal{A}}$, let H_Λ denote the quasi-Hankel matrix of residues,*

$$H_\Lambda = (\Lambda(\mathbf{x}^{\alpha+\beta}))_{\alpha,\beta\in E}.$$

For any element a in $\mathcal{A}$, we will also write $H_a = H_{a\star\tau}$, where τ is the residue of $\mathbf{p}$.

Let us extend corollary 1.5.4, by relating the multivariate Bezoutian B_1 with the quasi-Hankel matrix of residues H_1.

Theorem 2.10.2 *The inverse of H_1 is B_1.*

Proof. By definition, $\mathbf{w}_\alpha(\mathbf{x}) = \sum_{\beta\in E} B^1_{\alpha,\beta}\, \mathbf{x}^\beta$. Therefore, by using proposition 2.7.2, we obtain that

$$\tau(\mathbf{w}_\alpha\, \mathbf{x}^\beta) = \sum_{\gamma\in E} B^1_{\alpha,\gamma}\, \tau(\mathbf{x}^{\gamma+\beta})$$

equals 1 if $\alpha = \beta$ and is 0 otherwise. This is precisely the coefficient (α, β) of the matrix $B_1\, H_1$. Thus, we have

$$B_1\, H_1 = \mathbb{I}_D,$$

where $\mathbb{I}_D$ is the $D \times D$ identity matrix. $\square$

Example (continued) We have

$$\tau(1) = \tau(x_1) = \tau(x_2) = 0,$$
$$\tau(x_1x_2) = \frac{1}{5}, \tau(x_1^2) = -\frac{2}{5}, \tau(x_2^2) = \frac{2}{5}, \tau(x_1^2x_2) = \frac{29}{25}, \tau(x_1^2x_2^2) = -\frac{12}{25}, \tau(x_1^2x_2^2) = -\frac{398}{125},$$

and

$$H_1 = \begin{bmatrix} 0 & 0 & 0 & \frac{1}{5} \\ 0 & -\frac{2}{5} & \frac{1}{5} & \frac{29}{25} \\ 0 & \frac{1}{5} & \frac{2}{5} & -\frac{12}{25} \\ \frac{1}{5} & \frac{29}{25} & -\frac{12}{25} & -\frac{398}{125} \end{bmatrix}.$$

Here is the polynomial associated to this quasi-Hankel matrix:

$$P = \frac{1}{5}\partial_1\partial_2 - \frac{2}{5}\partial_1^2 + \frac{2}{5}\partial_2^2 + \frac{29}{25}\partial_1^2\partial_2 - \frac{12}{25}\partial_1\partial_2^2 - \frac{398}{125}\partial_1^2\partial_2^2.$$

To illustrate the multiplication property, we compute for instance the coordinates of the vector $[1, 0, -1, 0]^T H_1$, which are the coefficients of $1, \partial_1, \partial_2, \partial_1\partial_2$ in the next product:

$$(1 - \partial_2^{-1})\, P =$$
$$2\,\partial_1{}^2\partial_2^{-1} - \partial_1 - 2\,\partial_2 - \frac{39}{5}\,\partial_1{}^2 + \frac{17}{5}\,\partial_2\partial_1 + 2\,\partial_2{}^2 + \frac{543}{25}\,\partial_2\partial_1{}^2 - \frac{12}{5}\,\partial_2{}^2\partial_1 - \frac{398}{25}\,\partial_2{}^2\partial_1{}^2,$$

which yields the vector $[0, -1, -2, \frac{17}{5}]$. We may verify that H_1 is the inverse of the Bezoutian B_1 of the example of section 2.6:

$$\begin{bmatrix} -4 & 14 & -1 & 5 \\ 14 & -2 & 1 & 0 \\ -1 & 1 & 2 & 0 \\ 5 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & \frac{1}{5} \\ 0 & -\frac{2}{5} & \frac{1}{5} & \frac{29}{25} \\ 0 & \frac{1}{5} & \frac{2}{5} & -\frac{12}{25} \\ \frac{1}{5} & \frac{29}{25} & -\frac{12}{25} & -\frac{398}{125} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The matrices B_1 and H_1 express the transformation from the basis $(\mathbf{x}^\alpha)$ to the dual basis $(\mathbf{w}_\alpha)_{\alpha\in E}$:

Proposition 2.10.3 *For any $a \in \mathcal{A}$, if $\mathbf{v}$ is the coordinate vector of a in the monomial basis $(\mathbf{x}^\alpha)_{\alpha\in E}$ and $\mathbf{w}$ is the coordinate vector of a in the dual basis $(\mathbf{w}_\alpha)_{\alpha\in E}$, then we have*

$$\mathbf{v} = B_1\,\mathbf{w}, \quad \mathbf{w} = H_1\,\mathbf{v}.$$

Let us relate the matrices above to multiplication tables (compare section 2.8).

Proposition 2.10.4 *For any linear form $\Lambda \in \widehat{\mathcal{A}}$ and any $a \in \mathcal{A}$, we have*

$$H_{a\star\Lambda} \;=\; M_a^{\mathsf{t}} H_\Lambda = H_\Lambda\, M_a, \tag{22}$$

where M_a is the matrix of definition 2.8.3. In particular, we have

$$H_a \;=\; H_1\, M_a = M_a^{\mathsf{t}}\, H_1. \tag{23}$$

Proof. For any pair $a, b \in R$, we define the operator

$$\begin{aligned}\mathcal{H}_{a\star\Lambda}(b) &= b \star (a \star \Lambda) = a\, b \star \Lambda = \mathcal{H}_\Lambda(a\, b)\\ &= a \star (b \star \Lambda) = a \star \mathcal{H}_\Lambda(b).\end{aligned}$$

Therefore, the operator $\mathcal{H}_{a\star\Lambda}$ can be decomposed as follows:

$$\mathcal{H}_{a\star\Lambda} = \mathcal{H}_\Lambda \circ \mathcal{M}_a = \mathcal{M}_a^{\mathsf{t}} \circ \mathcal{H}_\Lambda.$$

In terms of matrices, this yields equations (22). □

A similar relation is also valid for the Bezoutian matrices (see definition 2.6.1):

Theorem 2.10.5 *For any $a \in \mathcal{A}$, we have*

$$B_a = B_1\, M_a^{\mathsf{t}} = M_a\, B_1. \tag{24}$$

Proof. According to (20), in terms of operators (see definition 2.6.1 with $a = q$) we have $\forall \Lambda \in \widehat{\mathcal{A}}$ that

$$\begin{aligned}\overline{\mathcal{B}}_a(\Lambda) &= \sum_{\alpha,\beta\in E} B^a_{\alpha,\beta}\, \mathbf{x}^\alpha\, \Lambda(\mathbf{y}^\beta)\\ &= a(\mathbf{x}) \sum_{\alpha,\beta\in E} B^1_{\alpha,\beta} \mathbf{x}^\alpha\, \Lambda(\mathbf{y}^\beta) = a(\mathbf{x})\, \overline{\mathcal{B}}_1(\Lambda)\\ &= \sum_{\alpha,\beta\in E} B^1_{\alpha,\beta}\, \mathbf{x}^\alpha\, \Lambda(a(\mathbf{y})\, \mathbf{y}^\beta) = \overline{\mathcal{B}}_1(a \star \Lambda).\end{aligned}$$

Thus, we can decompose the map $\overline{\mathcal{B}}_a$ as follows:

$$\overline{\mathcal{B}}_a = \overline{\mathcal{M}}_a \circ \overline{\mathcal{B}}_1 = \overline{\mathcal{B}}_1 \circ \overline{\mathcal{M}}_a^{\mathsf{t}}.$$

In terms of matrices, this implies (24). □

According to proposition 2.10.3, the latter theorem can be also reformulated as follows: *For any pair of elements a and b of $\mathcal{A}$, let $\mathbf{v}$ be the coordinate vector of b in $(\mathbf{w}_\alpha)_{\alpha\in E}$. Then the coordinate vector of $a\, b$ in the monomial basis $(\mathbf{x}^\alpha)_{\alpha\in E}$ is $B_a\mathbf{v}$.*

We will use relations (24) and (23) in section 3, in order to transform the eigenproblem of sections 2.2 and 2.3 into a generalized *structured* eigenproblem (see in particular our demonstration in section 3.1.2).

Proposition 2.10.6 *If $a\,b \equiv 1$ in $\mathcal{A}$, then*

$$B_a\, H_b = B_b\, H_a = \mathbb{I}_D.$$

Proof. According to (24), we have

$$B_a H_b = B_1\, M_a^{\mathtt{t}}\, M_b^{\mathtt{t}}\, H_1 B_1 H_1 = \mathbb{I}_D,$$

for $M_a\, M_b = M_{a\,b} = \mathbb{I}_D$. Similarly, we deduce that $B_b\, H_a = \mathbb{I}_D$. □

Proposition 2.10.7 *For any $a \in \mathcal{A}$, we have the relations*

- $B_a = B_1\, H_a\, B_1$,
- $H_a = H_1\, B_a\, H_1$.

Proof. According to (24) and (23) and proposition 2.10.2, we have

$$B_a = B_1 M_a^{\mathtt{t}} \ \text{ and } \ M_a^{\mathtt{t}} = H_a H_1^{-1} = H_a B_1,$$

which implies the first relation of this proposition. The other relation is obtained by inverting the first one and applying proposition 2.10.6. □

2.11 Some relations between the values of the Jacobian at the roots and structured matrices, in the case of simple roots

Let us assume that *the roots $\zeta \in \mathcal{Z}$ are simple.* Then $J(\zeta_i) \neq 0$, where $J = J_{\mathbf{p}} = \left(\frac{\partial p_i}{\partial x_j}\right)$ is the Jacobian of $\mathbf{p} = (p_1, \ldots, p_n)$.

Let $V_E(\mathcal{Z})$ be the Vandermonde matrix, defined in section 2.9. We recall that for any vector $\mathbf{v} = [v_\alpha]_{\alpha \in E}$, the product $V_E(\mathcal{Z})\,\mathbf{v}$ is the vector $[v(\zeta_1), \ldots, v(\zeta_D)]$ of the evaluations of the polynomial $v(\mathbf{x}) = \sum_\alpha v_\alpha\, \mathbf{x}^\alpha$ at the roots $\zeta_i \in \mathcal{Z}(I)$.

Proposition 2.11.1 *For any polynomial $a \in R$, we have*

$$B_a = V_E(\mathcal{Z})^{-1}\, \mathrm{diag}\,(a(\zeta_1)\, J(\zeta_1), \ldots, a(\zeta_D)\, J(\zeta_D))\, V_E(\mathcal{Z})^{-t},$$

where $\mathrm{diag}\,(l_1, \ldots, l_D)$ *represents the $D \times D$ diagonal matrix, with the diagonal entries $l_1, \ldots, l_D$.*

Proof. As the rows of $V_E(\mathcal{Z})$ are given by the values of the monomial vector $[\mathbf{x}^\alpha]$ at the roots $\zeta_i \in \mathcal{Z}(I)$, the matrix $V_E(\mathcal{Z})\, B_a\, V_E^{\mathbf{t}}(\mathcal{Z})$ is the matrix

$$[\Theta_{a,\mathbf{p}}(\zeta_i, \zeta_j)]_{i,j=1,\ldots,D}.$$

According to equations (20), (21), we have $\Theta_{a,\mathbf{p}}(\zeta, \eta) = \Theta_{a,\mathbf{p}}(\zeta, \eta) = 0$ if $\zeta \neq \eta$.

If $\eta = \zeta$, then, by construction, $\Theta_{a,\mathbf{p}}(a)(\zeta, \zeta) = a(\zeta)\, J(\zeta)$. Consequently, $(\Theta_{1,\mathbf{p}}(\zeta_i, \zeta_j))$ is the diagonal matrix

$$\operatorname{diag}\left(a(\zeta_1)\, J(\zeta_1), \ldots, a(\zeta_D)\, J(\zeta_D)\right).$$

□

Corollary 2.11.2 *If the roots of the system* $\mathbf{p} = 0$ *are simple, then*

$$H_1 = V_E(\mathcal{Z})^t \operatorname{diag}\left(\frac{1}{J(\zeta_1)}, \ldots, \frac{1}{J(\zeta_D)}\right) V_E(\mathcal{Z}).$$

Proof. We have $B_1 = V_E(\mathcal{Z})^{-1} \operatorname{diag}\left(J(\zeta_1), \ldots, J(\zeta_D)\right) V_E(\mathcal{Z})^{-t}$, according to proposition 2.11.1, and we deduce from theorem 2.10.2 that

$$H_1 = B_1^{-1} = V_E(\mathcal{Z})^t \operatorname{diag}\left(\frac{1}{J(\zeta_1)}, \ldots, \frac{1}{J(\zeta_D)}\right) V_E(\mathcal{Z}).$$

□

If we substitute these relations into (24), we obtain the following property:

Corollary 2.11.3 *If the roots of the system* $\mathbf{p} = \mathbf{0}$ *are simple, then*

$$M_a = V_E^{-1}(\mathcal{Z}) \operatorname{diag}\left(a(\zeta_1), \ldots, a(\zeta_d)\right) V_E(\mathcal{Z}). \tag{25}$$

According to theorem 2.10.5, we have $H_a = H_1\, M_a$, which yields:

Corollary 2.11.4 *If the roots of the system* $\mathbf{p} = 0$ *are simple, then*

$$H_a = V_E(\mathcal{Z})^t \operatorname{diag}\left(\frac{a(\zeta_1)}{J(\zeta_1)}, \ldots, \frac{a(\zeta_D)}{J(\zeta_D)}\right) V_E(\mathcal{Z}). \tag{26}$$

2.12 Relations between Bezoutians and idempotents

As in section 2.11, we assume that *the roots* $\zeta \in \mathcal{Z}$ *are simple* and denote by J the Jacobian of $\mathbf{p}$. For any $\zeta \in \mathcal{Z}$, $J(\zeta) \neq 0$.

Proposition 2.12.1 *If the roots of the system* $\mathbf{p} = \mathbf{0}$ *are simple, then the vectors*

$$\mathbf{e}_\zeta = \frac{1}{J(\zeta)} \Theta_{1,\mathbf{p}}(\mathbf{x}, \zeta),\ \ \zeta \in Z,$$

form a basis, consisting of orthogonal idempotents of $\mathcal{A}$*, whose sum equals* 1*, that is,* $\mathbf{e}_\zeta^2 \equiv \mathbf{e}_\zeta$*,* $\mathbf{e}_\zeta \mathbf{e}_\eta \equiv 0$ *if* $\zeta \neq \eta$*, and* $\sum_{\zeta \in \mathcal{Z}(I)} \mathbf{e}_\zeta \equiv 1$.

Proof. According to equation (20), for any $q \in R$ and any $\zeta \in \mathcal{Z}(I)$, we have

$$\Theta_{1,\mathbf{p}}(\mathbf{x},\zeta)\, q(\mathbf{x}) \equiv \Theta_{1,\mathbf{p}}(\mathbf{x},\zeta)\, q(\zeta)$$

in the quotient ring B. Therefore,

$$\Theta_{1,\mathbf{p}}(\mathbf{x},\zeta)\, \Theta_{1,\mathbf{p}}(\mathbf{x},\zeta) \equiv J(\zeta)\Theta_{1,\mathbf{p}}(\mathbf{x},\zeta),$$

and $\mathbf{e}_\zeta = \frac{1}{J(\zeta)}\, \Theta_{1,\mathbf{p}}(\mathbf{x},\zeta) \equiv \mathbf{e}_\zeta^2$ is an idempotent ($J(\zeta) \neq 0$ assuming all roots of the system $\mathbf{p} = \mathbf{0}$ are simple). Moreover, according to (21), we have

$$\Theta_{1,\mathbf{p}}(\mathbf{x},\zeta)\, \Theta_{1,\mathbf{p}}(\mathbf{x},\eta) \equiv \Theta_{1,\mathbf{p}}(\mathbf{x},\zeta)\Theta_{1,\mathbf{p}}(\zeta,\eta) \equiv 0,$$

for any pair of distinct roots $\zeta, \eta \in \mathcal{Z}(I)$, which shows that $\mathbf{e}_\zeta\, \mathbf{e}_\eta \equiv 0$ unless $\zeta = \eta$. We obtain from the definition of the residue τ and from the Euler-Jacobi identity (cf. [11]) that

$$\begin{aligned} \Theta_{1,\mathbf{p}}(\tau) &\equiv 1 \text{ (by definition 2.7.1)} \\ &\equiv \sum_{\zeta \in Z} \frac{1}{J(\zeta)}\, \Theta_{1,\mathbf{p}}(\mathbf{x},\zeta) \equiv \sum_{\zeta \in Z} \mathbf{e}_\zeta \text{ (by the Euler-Jacobi identity).} \end{aligned}$$

This shows that the sum of the idempotents equals 1 in $\mathcal{A}$, and thus they form a basis of $\mathcal{A}$ (which is of dimension D). □

Now let us recover the root ζ from the idempotent $\mathbf{e}_\zeta$. By definition, we have

$$\mathbf{e}_\zeta = \frac{1}{J(\zeta)}\Theta_{1,\mathbf{p}}(\mathbf{x},\zeta) = \frac{1}{J(\zeta)} \sum_{\alpha \in E} \mathbf{x}^\alpha \left(\sum_\beta B^1_{\alpha,\beta}\, \zeta^\beta \right),$$

so that the coordinate vector $[\mathbf{e}_\zeta]$ of $\mathbf{e}_\zeta$ in the basis $(\mathbf{x}^\alpha)_{\alpha\in E}$ is

$$[\mathbf{e}_\zeta] = \frac{1}{J(\zeta)} B_1\, [\zeta^\alpha]_{\alpha \in E}.$$

Equivalently, we have

$$[\zeta^\alpha]_{\alpha\in E} = J(\zeta)\, H_1\, [\mathbf{e}_\zeta]. \tag{27}$$

Corollary 2.12.2 *The coordinates of* $\mathbf{e}_\zeta$ *in the dual basis* $(\mathbf{w}_\alpha)$ *are* $\frac{1}{J(\zeta)}[\zeta^\alpha]$.

Algorithm 2.12.1 *(*Recovery of the root vector ζ from the idempotent vector $\mathbf{e}_\zeta$*). Multiply* $\mathbf{e}_\zeta$ *by the quasi-Hankel matrix* H_1 *and output the ratios of the coordinates of the resulting vector.*

Let us estimate the computational cost of performing the algorithm. If $\mathbf{v} = H_1[\mathbf{e}_\zeta] = \frac{1}{J(\zeta)}[\zeta^\alpha]_{\alpha\in E} = [v_1, v_{x_1}, \ldots, v_{x_n}, v_{x_1^2}, \ldots]$, then the i-th coordinate of ζ is

$$\zeta_i = \frac{v_{x_i}}{v_1}.$$

Thus, the roots can be computed from the idempotent $\mathbf{e}_\zeta$ in at most $C_{PolMult}(E, 2E)$ ops, by using algorithm 2.5.1 applied for $F = 2E$.

3 Applications

In this section, we exploit the properties of and the relations among multivariate polynomial roots, dual space of linear maps, residue and structured matrices in order to devise algorithms for solving polynomial systems of equations. First we focus on structured generalized eigenproblem, involving quasi-Hankel and Bezoutian matrices. Then we consider quasi-Toeplitz matrices that generalize the Sylvester matrices. They are used for computing a basis $(\mathbf{x}^\alpha)_{\alpha\in E}$ of $\mathcal{A}$, the multiplication tables, and the first coefficients of the dual basis of $(\mathbf{x}^\alpha)_{\alpha\in E}$, for *generic input*. Using the machinery of the previous section enables us to yield better insight into the subject and simplify substantially the proofs of some known fundamental results. Finally, we focus on iterative methods converging to idempotents (and based on using quasi-Hankel matrices) and on application of structured matrices to counting distinct roots and real roots of a polynomial system. In this part, we improve dramatically the known computational complexity estimates, though the algorithms are proposed in preliminary form and require further elaboration for their implementation.

3.1 Reduction of solving a polynomial system to matrix eigenproblems

Let us restate theorem 2.2.1 and 2.4.1 in terms of matrices rather than their associated operators. For a fixed element $a \in \mathcal{A}$, we consider the operator of multiplication by a :

$$\begin{array}{rcl}\overline{\mathcal{M}}_a : \mathcal{A} & \to & \mathcal{A} \\ b & \mapsto & a\,b,\end{array}$$

whose matrix in the monomial basis $(\mathbf{x}^\alpha)_{\alpha\in E}$ is denoted M_a. The transposed operator from $\widehat{\mathcal{A}}$ to $\widehat{\mathcal{A}}$ is defined by the map:

$$\begin{array}{rcl}\overline{\mathcal{M}}_a^{\mathtt{t}} : \widehat{\mathcal{A}} & \to & \widehat{\mathcal{A}} \\ \Lambda & \mapsto & a\star\Lambda = \Lambda\circ\overline{\mathcal{M}}_a,\end{array}$$

and its matrix in the dual basis is $M_a^{\mathtt{t}}$. We have the following theorem, whose first two parts restate theorems 2.2.1 and 2.4.1 in terms of matrices (see [1], [22]):

Theorem 3.1.1

1. *The eigenvalues of the matrices of the linear operators M_a and $M_a^{\mathtt{t}}$ are $\{a(\zeta_1), \ldots, a(\zeta_d)\}$.*

2. *The common eigenvectors of the matrices $(M_{x_i}^{\mathtt{t}})_{i=1,\ldots,n}$ are (up to a scalar) $[\zeta_i^\alpha]_{\alpha\in E}$.*

3. *If $n = m$, then the common eigenvectors of the matrices $(M_{x_i})_{i=1,\ldots,n}$ are (up to a scalar factor) $J(\mathbf{x})\,\mathbf{e}_1, \ldots, J(\mathbf{x})\,\mathbf{e}_d$, where $J(\mathbf{x})$ is the Jacobian of $p_1, \ldots, p_n$, and $\mathbf{e}_i$ are the idempotents associated with the roots.*

Part 1 amounts to theorem 2.2.1. Part 2 is deduced from theorem 2.4.1: the coordinates of the evaluation $\mathbf{1}_{\zeta_i}$ at the root ζ_i in the dual basis of $(\mathbf{x}^\alpha)_{\alpha\in E}$ are precisely $[\zeta_i^\alpha]_{\alpha\in E}$. The third part is proved in [11].

As a consequence of theorem 3.1.1, we may compute easily the roots ζ_i from the eigenvectors of $M_{x_i}^{\mathtt{t}}$, as in algorithm 2.12.1:

Proposition 3.1.2 *If $(\mathbf{x}^\alpha)_{\alpha\in E} = (1, x_1, \ldots, x_n, \ldots)$ contains the monomials $1, x_1, \ldots, x_n$ and if $\mathbf{v} = [v_\alpha]_{\alpha\in E} = (v_1, v_{x_1}, \ldots, v_{x_n}, \ldots,)$ is a common vector of $(M_{x_i})_{i=1,\ldots,n}$, then*

$$\zeta = (\frac{v_{x_1}}{v_1}, \ldots, \frac{v_{x_n}}{v_1})$$

is a root of $\mathbf{p} = \mathbf{0}$.

Algorithm 3.1.1 (Computation of the roots of the polynomial system $\mathbf{p} = \mathbf{0}$). *Compute and output the roots as the scaled common eigenvectors of the matrices $M_a^{\mathtt{t}}$ for $a \in R$.*

Example (continued) Here is the normalized matrix V of eigenvectors (with eight digit accuracy) of the matrix $M_{x_1}^{\mathtt{t}}$:

$$\begin{bmatrix} 1.0 & 1.0 & 1.0 & 1.0 \\ 6.8200982 & -0.19395427 + 0.20520688\,\mathbf{i} & -0.19395427 - 0.20520688\,\mathbf{i} & 0.36781361 \\ -2.8367388 & -0.61937124 - 1.3895199\,\mathbf{i} & -0.61937124 + 1.3895199\,\mathbf{i} & 1.6754769 \\ -19.346814 & 0.40526841 + 0.14240419\,\mathbf{i} & 0.40526841 - 0.14240419\,\mathbf{i} & 0.61626304 \end{bmatrix}.$$

The columns of this matrix are the vectors $[\zeta_i^\alpha]_{\alpha\in E}$ for $\zeta_i \in \mathcal{Z}(I)$. Thus, we immediately deduce that the four roots of $p_1(x_1, x_2) = p_2(x_1, x_2) = 0$ are given by the next table:

ζ_1	ζ_2	ζ_3	ζ_4
6.8200982	$-0.19395427 + 0.20520688\,\mathbf{i}$	$-0.19395427 - 0.20520688\,\mathbf{i}$	0.36781361
-2.8367388	$-0.61937124 - 1.3895199\,\mathbf{i}$	$-0.61937124 + 1.3895199\,\mathbf{i}$	1.6754769

We immediately check that $V_{2,i}V_{3,i} = V_{4,i}$ for $i = 1, 2, 3, 4$. The structure of these multiplication matrices is not easy to analyze directly, however. Thus, we will multiply the matrices $M_a^{\mathtt{t}}$ by two fixed invertible matrices A and B in order to transform the problem into an equivalent generalized eigenproblem, $(AM_a^{\mathtt{t}}B - \lambda\, A\, B)\mathbf{v} = \mathbf{0}$, where the structure can be explicitly exploited. We will give some examples of such a reduction involving structured matrices.

3.1.1 Reduction by using Hankel matrices

According to (22), for any $\Lambda \in \widehat{\mathcal{A}}$ and any $a \in R$, we have

$$H_{a\star\Lambda} = M_a^{\mathtt{t}} H_\Lambda,$$

so that solving the eigenproblem $(H_{a\star\Lambda} - \lambda H_\Lambda)\mathbf{v} = \mathbf{0}$ yields the eigenvector $H_\Lambda\,\mathbf{v}$ of $M_a^{\mathtt{t}}$. Let us next exploit this matrix equation assuming that we have a normal form algorithm $\mathtt{Nf}$ that projects R onto $\langle \mathbf{x}^\alpha \rangle_{\alpha\in E}$ along I, that is, computes the unique element of $\langle \mathbf{x}^\alpha \rangle_{\alpha\in E}$ in the same class modulo I.

Algorithm 3.1.2 (Solution of a polynomial system via the solution of a generalized eigenvector problem defined by using Hankel matrices). *Fix two exponents $\alpha_0, \alpha_1 \in E$. Then proceed as follows:*

1. *For all monomials $\mathbf{x}^{\alpha+\beta}$ with $\alpha, \beta \in E$, compute in the normal form $\mathtt{Nf}(\mathbf{x}^{\alpha+\beta})$ of $\mathbf{x}^{\alpha+\beta}$:*
   - *the coefficient of $\mathbf{x}^{\alpha_0}$, which we denote $\sigma_0(\mathbf{x}^{\alpha+\beta})$,*
   - *the coefficient of $\mathbf{x}^{\alpha_1}$, which we denote $\sigma_1(\mathbf{x}^{\alpha+\beta})$.*
2. *Construct the two quasi-Hankel matrices:*
   - $H_{\sigma_0} = (\sigma_0(\mathbf{x}^{\alpha+\beta}))_{\alpha,\beta\in E}$,
   - $H_{\sigma_1} = (\sigma_1(\mathbf{x}^{\alpha+\beta}))_{\alpha,\beta\in E}$.
3. *Solve the polynomial system $\mathbf{p}$=$\mathbf{0}$ via the solution of the generalized eigenvector problem:*
$$(H_{\sigma_1} - \lambda H_{\sigma_0})\,\mathbf{v} = 0. \tag{28}$$

Let us specify stage 3. The coefficient of $\mathbf{x}^\alpha$ in $\mathcal{A}$ (for any $\alpha \in E$) is output by the linear form $p \to \tau(\mathbf{w}_\alpha\, p) = \mathbf{w}_\alpha \star \tau(p)$. Thus, we have

$$H_{\sigma_i} = M^{\mathtt{t}}_{\mathbf{w}_{\alpha_i}} H_1,$$

for $i = 0, 1$. Therefore, if $\mathbf{v}$ is a generalized eigenvector of (28), then $\tilde{\mathbf{v}} = H_1\mathbf{v}$ is a generalized eigenvector of $(M^{\mathtt{t}}_{\mathbf{w}_{\alpha_1}} - \lambda M^{\mathtt{t}}_{\mathbf{w}_{\alpha_0}})\,\tilde{\mathbf{v}} = 0$, and its associated eigenvalue is equal to $\frac{\mathbf{w}_{\alpha_1}(\zeta)}{\mathbf{w}_{\alpha_0}(\zeta)}$ (if $\mathbf{w}_{\alpha_0}(\zeta) \neq 0$) for one of the roots $\zeta \in \mathcal{Z}(I)$.

According to theorem 3.1.1, the common eigenvectors of $M^{\mathtt{t}}_{\mathbf{w}_{\alpha_1}} - \lambda M^{\mathtt{t}}_{\mathbf{w}_{\alpha_0}}$ for all pairs $\alpha_0, \alpha_1 \in E$ are the multiples of the vectors $[\zeta^\alpha]_{\alpha\in E}$ for $\zeta \in \mathcal{Z}(I)$. The roots ζ are easily computed from these vectors, by using algorithm 3.1.1.

Example (continued) Suppose that we have computed the following normal forms in the basis $(1, x_1, x_2, x_1\,x_2)$ of $\mathcal{A} = \mathbb{C}[x_1, x_2]/(p_1, p_2)$:

$$\begin{aligned}
&\mathtt{Nf}(1) = 1, \mathtt{Nf}(x_1) = x_1, \mathtt{Nf}(x_2) = x_2, \mathtt{Nf}(x_1x_2) = x_1x_2,\\
&\quad \mathtt{Nf}({x_1}^2) = 1 + x_1 - 2\,x_1x_2, \mathtt{Nf}({x_2}^2) = -1 + 7\,x_1 + 2\,x_1x_2,\\
&\quad \mathtt{Nf}(x_2{x_1}^2) = -\frac{14}{5} - \frac{12\,x_1}{5} + \frac{x_2}{5} + \frac{29\,x_1x_2}{5},\\
&\quad \mathtt{Nf}(x_1{x_2}^2) = \frac{7}{5} + \frac{6\,x_1}{5} + \frac{2\,x_2}{5} - \frac{12\,x_1x_2}{5},\\
&\quad \mathtt{Nf}({x_1}^2{x_2}^2) = \frac{198}{25} + \frac{209\,x_1}{25} - \frac{12\,x_2}{25} - \frac{398\,x_1x_2}{25}.
\end{aligned}$$

We choose the monomial $\mathbf{x}^{\alpha_0} = x_1 x_2$ and $\mathbf{x}^{\alpha_1} = x_2$, which yields the following matrices:

$$H_{\sigma_0} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & -2 & 1 & \frac{29}{5} \\ 0 & 1 & 2 & -\frac{12}{5} \\ 1 & \frac{29}{5} & -\frac{12}{5} & -\frac{398}{25} \end{bmatrix}, \quad H_{\sigma_1} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \frac{1}{5} \\ 1 & 0 & 0 & \frac{2}{5} \\ 0 & \frac{1}{5} & \frac{2}{5} & -\frac{12}{25} \end{bmatrix},$$

and we obtain

$$H_{\sigma_1} H_{\sigma_0}^{-1} = \begin{bmatrix} -\frac{1}{5} & \frac{1}{5} & \frac{2}{5} & 0 \\ \frac{1}{5} & 0 & 0 & 0 \\ -\frac{2}{5} & \frac{14}{5} & -\frac{1}{5} & 1 \\ 0 & 0 & \frac{1}{5} & 0 \end{bmatrix}.$$

We have $\sigma_0 = \mathbf{w}_{\alpha_0} \star \tau = (2\,x_2 + x_1 - 1) \star \tau$ and $\sigma_{\alpha_1} = \mathbf{w}_{\alpha_1} \star \tau = 5\,\tau$. Therefore, $H_{\sigma_0} = 5\,H_1$, and $H_{\sigma_1} = H_{2\,x_2+x_1-1}$, so that

$$H_{\sigma_1} H_{\sigma_0}^{-1} = M^{\mathsf{t}}_{\frac{1}{5}(2\,x_2+x_1-1)}.$$

Indeed, the first row of the latter matrix represents the polynomial $\frac{1}{5}(2\,x_2 + x_1 - 1)$, the second row is $x_1 \times \frac{1}{5}(2\,x_2 + x_1 - 1)$, which is reduced to $\frac{1}{5}$ in $\mathcal{A}$. This implies that $x_1^{-1} \equiv (2\,x_2 + x_1 - 1)$.

3.1.2 Reduction by using Bezoutian matrices

The relations (23) on Bezoutians imply that

$$B_a = B_1\, M^{\mathsf{t}}_a.$$

As in algorithm 3.1.2, assume that we have a normal form algorithm that computes an element in $\mathcal{A}$ reduced modulo I.

Algorithm 3.1.3 (Solution of a polynomial system via the solution of a generalized eigenvector problem defined by using Bezoutian matrices).

1. *Compute the polynomials $\Theta_{1,\mathbf{p}}$ and $\Theta_{x_1,\mathbf{p}}$ and their normal forms in $\mathbf{x}$ and $\mathbf{y}$.*
2. *Compute the matrices B_1 and B_{x_1} associated with these normal forms.*
3. *Solve the polynomial system $\mathbf{p}$=$\mathbf{0}$ via the solution of the generalized eigenvector problem*

$$(B_{x_1} - \lambda\, B_1)\,\mathbf{v} = \mathbf{0}.$$

The generalized eigenvector of the pencil (B_{x_1}, B_1) (computed at stage 3) yields immediately the eigenvectors $[\zeta_i^{\alpha}]_{\alpha \in E}$, and then scaling immediately gives us the coordinates of the roots ζ_i (cf. algorithm 3.1.1).

Example (continued) The Bezoutian of x_1 is

$$B_{x_1} = \begin{bmatrix} 0 & -2 & 1 & 0 \\ -2 & 12 & 0 & 5 \\ 1 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \end{bmatrix} \text{ and } B_1^{-1} B_{x_1} = M_{x_1}^{\mathtt{t}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & -2 \\ 0 & 0 & 0 & 1 \\ -\frac{14}{5} & -\frac{12}{5} & \frac{1}{5} & \frac{29}{5} \end{bmatrix}.$$

The first row of this matrix represents multiplication by x_1 in $\mathcal{A}$, the second represents multiplication by x_1^2, and so on.

3.2 Computation of multiplication matrices and the dual space

3.2.1 Sylvester's matrices

As a basic pattern, we will first revisit the construction of the well-known Sylvester matrix in the univariate case.

Given two univariate polynomials, $p_0 = p_{0,0} + \cdots + p_{0,d_0}\, x^{d_0}$ of degree d_0 and $p_1 = p_{1,0} + \ldots + p_{1,d_1}\, x^{d_1}$ of degree d_1, we will define the multiplication by p_0 modulo p_1 by the map:

$$\begin{array}{rcl} \overline{\mathcal{M}}_{p_0} : \mathcal{A} & \to & \mathcal{A} \\ a & \mapsto & a\, p_0, \end{array}$$

in the basis $\langle 1, \ldots, x^{d_1-1} \rangle$ of $\mathcal{A} = \mathbb{C}[x]/(p_1)$. The matrix of this map is defined via the Sylvester matrix S of p_0 and p_1, that is, the matrix of the coefficients of the polynomials

$$p_0, x\, p_0, \ldots, x^{d_1-1}\, p_0, p_1, x\, p_1, \ldots, x^{d_0-1}\, p_1$$

in the monomial basis. The matrix S takes the following form:

$$\begin{array}{c} \overbrace{\begin{array}{cccccc} p_0 & \cdots & x^{d_1-1}p_0 & p_1 & \cdots & x^{d_0-1}p_1 \end{array}}^{d_0+d_1} \\ \left[\begin{array}{ccc|ccc} p_{0,0} & & 0 & p_{1,0} & & p_{1,1-d_0} \\ \vdots & \ddots & & \vdots & \ddots & \\ p_{0,d_1-1} & \cdots & p_{0,0} & p_{1,d_1-1} & \cdots & p_{1,d_1-d_0} \\ \hline p_{0,d_1} & \cdots & p_{0,1} & p_{1,d_1} & \cdots & p_{1,d_1-d_0+1} \\ \vdots & & \vdots & & \ddots & \vdots \\ p_{0,d_0+d_1-1} & \cdots & p_{0,d_0} & 0 & & p_{1,d_1} \end{array}\right] \end{array} \begin{array}{c} \\ \left.\begin{array}{c} 1 \\ x \\ . \\ x^{d_1-1} \\ . \\ . \\ x^{d_0+d_1-1} \end{array}\right\} d_0 + d_1 \end{array} \qquad (29)$$

under the convention that $p_{0,i} = 0$ if $i > d_0$, $p_{1,j} = 0$ if $j < 0$. Let $\mathcal{V}_0, \mathcal{V}_1$, and $\mathcal{V}$ denote the vector spaces generated by the monomials $\{1, \ldots, x^{d_1-1}\}$, $\{1, \ldots, x^{d_0-1}\}$, and $\{1, \ldots, x^{d_0+d_1-1}\}$, respectively. Then the Sylvester matrix is the matrix of the map

$$\begin{array}{rcl} \mathcal{S} : \mathcal{V}_0 \times \mathcal{V}_1 & \rightarrow & \mathcal{V} \\ (q_0, q_1) & \mapsto & p_0\, q_0 + p_1\, q_1, \end{array}$$

in the corresponding monomial basis. The determinant of this $(d_0 + d_1) \times (d_0 + d_1)$ matrix is the *resultant* of p_0 and p_1.

For computing the matrix M_{p_0} of the multiplication by p_0 modulo p_1, we have to reduce the polynomials $p_0, x\, p_0, \ldots, x^{d_1-1}\, p_0$ modulo p_1, by subtracting some multiples of p_1, to obtain linear combinations of the monomial basis $(1, \ldots, x^{d_1-1})$ of $\mathcal{A}$. The partition of the Sylvester matrix into four blocks as in (29),

$$S = \begin{bmatrix} U & V \\ Z & W \end{bmatrix},$$

enables us to interpret these operations in terms of matrix operations and thus to analyze the structure of the matrix of multiplication. The block $P_0 = \begin{bmatrix} U \\ Z \end{bmatrix}$ represents the multiples of p_0, and the block $P_1 = \begin{bmatrix} V \\ W \end{bmatrix}$ represents the multiples of p_1. Therefore, reducing the multiples of p_0 by p_1 consists in subtracting some linear combinations of the columns of P_1 from the columns of P_0 so that Z is replaced by a zero block. These operations on the columns of the Sylvester matrix are given explicitly by the following formula:

$$\begin{bmatrix} U & V \\ Z & W \end{bmatrix} \begin{bmatrix} \mathbb{I}_{d_1} & 0 \\ -W^{-1}Z & \mathbb{I}_{d_0} \end{bmatrix} = \begin{bmatrix} U - V\, W^{-1} Z & V \\ 0 & W \end{bmatrix},$$

and we have the following property:

Proposition 3.2.1 *The matrix M_{p_0} of multiplication by p_0 modulo p_1 in the monomial basis $\langle 1, x, \ldots, x^{d_1-1}\rangle$ is the Schur complement of W in S:*

$$M_{p_0} = U - V\, W^{-1}\, Z.$$

Note that the blocks U, V, W, and Z have Toeplitz structure, U and W are triangular, and if $d_0 \leq d_1$ (resp. $d_0 \geq d_1$), then so is Z (resp. V) also. Thus, we have the following algorithm:

Algorithm 3.2.1 (Multiplication by a polynomial modulo a polynomial, in the univariate case). *Given three polynomials p_0, p_1 and a of degrees d_0, d_1 and less than d_1, respectively, compute the coefficient vector of the polynomial ap_0 mod p_1 as the matrix-by-vector product:*

$$M_{p_0}\mathbf{a} = (U - VW^{-1}Z)\mathbf{a},$$

where $\mathbf{a}$ is the coefficient vector of the polynomial a and U, V, W and Z are the above blocks of the Sylvester matrix S of (29).

The computation reduces to multiplication of the Toeplitz matrices Z of size $d_0 \times d_1$ and U of size $d_1 \times d_1$ by the vector $\mathbf{a}$, solving the triangular Toeplitz system

$$W\mathbf{q} = Z\mathbf{a}$$

of d_0 equations, multiplying the Toeplitz matrix V by the solution $\mathbf{q}$ of this system, and subtracting the vectors $V\mathbf{q}$ from $U\mathbf{a}$.

With application of the algorithms of section B.1, one may perform algorithm 3.2.1 in $\mathcal{O}(d \log d)$ ops, where $d = \max(d_0, d_1)$.

We are going to extend this approach to the multivariate case. Let us mention some of the main difficulties that are peculiar to the multivariate case but do not occur in the univariate case:

- We lose the notion of the leading monomial of the highest degree.
- We have no natural monomial basis for representing the quotient modulo a set of polynomials.
- When we homogenize the polynomials, we may introduce spurious solutions (at *infinity*) to a polynomial system of equations.

For the latter reasons and many others, we need to restrict our study to the cases where we may describe easily the structure of the matrices. These are the generic cases of two types that we are going to describe.

3.2.2 The generic multivariate case

In order to generalize the Sylvester matrix construction to the multivariate case, we consider $n+1$ polynomials $p_0, \ldots, p_n$ and $n+1$ vector spaces $\mathcal{V}_0, \ldots, \mathcal{V}_n$ generated by the monomials $\mathbf{x}^{F_i} = \{\mathbf{x}^\alpha, \alpha \in F_i\}$, where F_i is the set of the exponents,

$$F_i = \{\beta_{i,1}, \beta_{i,2}, \ldots\}.$$

Let $\mathcal{V}$ be a vector space containing all the monomials of the polynomials $p_i\, \mathbf{x}^{\beta_i}$, for $\beta_i \in F_i$, so that we can define the following map:

$$\begin{aligned} \mathcal{S} : \mathcal{V}_0 \times \cdots \times \mathcal{V}_n &\rightarrow \mathcal{V} \\ (q_0, \ldots, q_n) &\mapsto \sum_{i=0}^{n} p_i\, q_i. \end{aligned} \tag{30}$$

Let the matrix of the map $\mathcal{S}$ in the monomial basis of $\mathcal{V}_0 \times \cdots \times \mathcal{V}_n$ and $\mathcal{V}$ be also denoted by S and take the following form:

$$\begin{array}{ccc} & & \begin{array}{cccc} \overbrace{\hphantom{\mathbf{x}^{\beta_{0,1}}p_0 \ \cdots}}^{\mathcal{V}_0} & \overbrace{\hphantom{\mathbf{x}^{\beta_{1,1}}p_1 \ \cdots}}^{\mathcal{V}_1} & \hphantom{\cdots\cdots\cdots} & \overbrace{\hphantom{\mathbf{x}^{\beta_{n,1}}p_n \ \cdots}}^{\mathcal{V}_n} \end{array} \\ \mathcal{V} \left\{ \begin{array}{c} \mathbf{x}^{\alpha_1} \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{x}^{\alpha_N} \end{array} \right. & & \left[\begin{array}{c|c|c|c} \begin{array}{cc} \cdot & \\ \cdot & \\ \mathbf{x}^{\beta_{0,1}}p_0 & \cdots \\ \cdot & \\ \cdot & \end{array} & \begin{array}{cc} \cdot & \\ \cdot & \\ \mathbf{x}^{\beta_{1,1}}p_1 & \cdots \\ \cdot & \\ \cdot & \end{array} & \cdots\cdots\cdots & \begin{array}{cc} \cdot & \\ \cdot & \\ \mathbf{x}^{\beta_{n,1}}p_n & \cdots \\ \cdot & \\ \cdot & \end{array} \end{array} \right]. \end{array} \tag{31}$$

Let us decompose such a matrix S into blocks $S = [S_0, \ldots, S_n]$, where S_i involves only the coefficients of p_i. The matrix S_i is a submatrix of the matrix of multiplication by p_i, defined in section 2.5. More precisely, S_i is the matrix of the map

$$\pi_F \circ \mathcal{M}_{p_i} \circ \pi_{F_i}.$$

Thus, it is a *quasi-Toeplitz* matrix.

Algorithm 3.2.2 *(* Multiplication of the matrix S of (31) by a vector). *For every j, compute the products $\mathbf{x}^{\beta_{i,j}} p_i\, q_i$ for all i and sum them together in i. Output the sum for every j.*

The complexity of this algorithm is bounded by $C_{PolMult}(F_0, F) + \cdots + C_{PolMult}(F_n, F)$ (due to using the algorithms of section B.1).

It is possible to consider the global matrix S as a quasi-Toeplitz matrix by adding a new variable x_0. The sum $\sum_{i=0}^{n} p_i\, q_i$ can be computed from the product of $p = \sum_i p_i\, x_0^i$ by $\sum_{i=0}^{n} q_i\, x_0^{n-i}$. Indeed, this sum is the coefficient of x_0^n in the product. Let F' and F'' be the sets of the exponents of the monomials in $x_0^n\, \mathbf{x}^F$ and $\cup_{i=0}^{n} \mathbf{x}_0^{n-i} \mathbf{x}^{E_i}$, respectively. Then the matrix S is the matrix of the operator

$$\pi_{F'} \circ \mathcal{M}_p \circ \pi_{F''}.$$

Remark 2 *We can extend easily the construction of the map $\mathcal{S}$ to the case where the number of polynomials $p_0, \ldots, p_m$ is greater than $n+1$ ($m \geq n$).*

Operators of this type have been extensively used in the literature, in order, for instance, to define resultants (see [20], [34], [15]). Let us recall that *the vanishing of the resultant is the necessary and sufficient condition on the coefficients of the polynomials $p_0, \ldots, p_n$, under which these polynomials have a common root (in a projective variety X).* Two main examples appear in the literature:

- The classical case corresponds to $X = \mathbb{P}^n$, the projective space of dimension n. In this case, the polynomials $p_0, \ldots, p_n$ of degree $d_0, \ldots, d_n$ are homogenized, and the vanishing of the resultant gives a necessary and sufficient condition on their coefficients under which the homogenized polynomials have a common zero in $\mathbb{P}^n$. This case is referred to as *Macaulay case* (see [20]).

- In the second case, the variety $X = \mathcal{T}$ is a *toric variety*, and the map $\mathcal{S}$ is used to define the toric resultant of the polynomials $p_0, \ldots, p_n$. The polynomials can also be homogenized in a toric sense, and the resultant gives a necessary and sufficient condition on their coefficients under which the toric-homogenized polynomials have a common zero in the toric variety $\mathcal{T}$ (see [15]). We refer to this case as the *toric case.*

Let us describe more carefully the monomials with exponents in F_i used in the construction of $\mathcal{S}$.

The Macaulay case Let us fix integers $d_0, \ldots, d_n$, and $\nu = d_0 + \cdots + d_n - n$. For any $d \in \mathbb{N}$, let R_d denote the set of polynomials of degree not greater than d. Let $p_0, \ldots, p_n$ be polynomials of degrees $d_0, \ldots, d_n$, respectively. To construct the map $\mathcal{S}$ that yields the resultant of these polynomials, we follow Macaulay's work and choose $\mathcal{V}_i = R_{\nu - d_i}$, $\mathcal{V} = R_\nu$, so that we define the map

$$\begin{array}{rcl}\mathcal{S} : R_{\nu-d_0} \times \cdots \times R_{\nu-d_n} & \to & R_\nu \\ (q_0, \ldots, q_n) & \mapsto & \displaystyle\sum_{i=0}^{n} p_i\, q_i.\end{array}$$

The toric case In this case, we replace the constraints on the degree of the polynomials by the constraints on the support of the polynomials p_i (that is, the set of the exponents of the monomials with non-zero coefficients in p_i). Let $C_0, \ldots, C_n$ be *polytopes* in $\mathbb{Z}^N$ and let $p_0, \ldots, p_n \in L$ be Laurent's polynomials, whose supports are in $C_0, \ldots, C_n$, respectively. In order to construct the map $\mathcal{S}$ that yields the toric resultant, we fix (at random) a direction $\delta \in \mathbb{Q}^n$. For any polytope C, let C^δ denote the polytope obtained from C, by removing its facets whose normals have positive inner products with δ (see [5],[27]). For $F_i = (\sum_{j \neq i} C_j)^\delta$ and $F = (\sum_j C_j)^\delta$, we define the map

$$\begin{array}{rcl}\mathcal{S} : \langle \mathbf{x}^{F_0} \rangle \times \cdots \times \langle \mathbf{x}^{F_n} \rangle & \to & \langle \mathbf{x}^F \rangle \\ (q_0, \ldots, q_n) & \mapsto & \displaystyle\sum_{i=0}^{n} p_i\, q_i.\end{array}$$

Many other examples of this type can be obtained by means of convenient choices of the vector spaces $\mathcal{V}_0, \ldots, \mathcal{V}_n$, and $\mathcal{V}$. We are going to examine the properties of these maps in the *generic cases.*

Definition 3.2.2 *A property is generically true in the Macaulay case (or in the toric case), if this property is true for an (algebraically) open subset of the set of all possible values of the coefficients, according to the constraints on the degree (or on the support) of the polynomials.*

Given polynomials $p_1, \ldots, p_n$, we will compute from the matrix S:

- a basis $(\mathbf{x}^\alpha)_{\alpha \in E}$ of the quotient $\mathcal{A} = R/(p_1, \ldots, p_n)$,
- the table of multiplication by a polynomial p_0 in $\mathcal{A}$, from the matrix S (note that the matrix S of $\mathcal{S}$ is not a square matrix anymore),
- the dual basis of the monomial basis $(\mathbf{x}^\alpha)_{\alpha \in E}$ of $\mathcal{A}$.

These constructions will be valid for *generic* values of the coefficients of $p_1, \ldots, p_n$ but may fail for specific values. A more sophisticated method, described in [22], circumvents this difficulty by the compression of pencils of matrices.

3.2.3 A basis of $\mathcal{A}$

First, we will define a subset E_0 of exponents such that $\mathbf{x}^{E_0}$ is generically a basis of $\mathcal{A} = R/(p_1, \ldots, p_n)$. For that purpose, we choose $p_0 = u_0 + u_1\, x_1 + \cdots + u_n\, x_n$ (or $p_0 = u_0 + u_1\, x_1 + \cdots + u_n\, x_n + u_{-1}\, x_1^{-1} + \cdots u_{-n}\, x_n^{-1}$ in the toric case), where u_i are parameters. We also choose subsets $E_i \subset F_i$ for $i = 0, \ldots, n$, such that

(a) $|E_0| + \cdots + |E_n| = |F|$

(b) and the matrix of the map

$$\begin{array}{rcl} \tilde{\mathcal{S}} : \langle \mathbf{x}^{E_0} \rangle \times \cdots \times \langle \mathbf{x}^{E_n} \rangle & \rightarrow & \langle \mathbf{x}^{F} \rangle \\ (q_0, \ldots, q_n) & \mapsto & \displaystyle\sum_{i=0}^{n} p_i\, q_i \end{array}$$

takes the form

$$\tilde{S} = \begin{array}{c} \\ \left\{\begin{array}{c} E_0 \\ F' \end{array}\right. \end{array} \begin{array}{c} \begin{array}{cc} \overbrace{\;E_0\;} & \overbrace{E_1 \cdots E_n} \end{array} \\ \left[\begin{array}{c|c} U & V \\ \hline Z & W \end{array}\right] \end{array}, \tag{32}$$

where W is generically invertible.

In order to prove this generic property, it is sufficient to specify the coefficients of polynomials p_i, for which the property is satisfied.

Theorem 3.2.3 *If condition (a) and (b) are satisfied, then for generic values of the coefficients of $p_1, \ldots, p_n$, $(\mathbf{x}^\alpha)_{\alpha \in E}$ is a generating set of $\mathcal{A}$, and we have*

$$\dim_{\mathbb{C}}(\mathcal{A}) \leq |E_0|.$$

Proof. As W is *generically* invertible, the same process as in section 3.2.1 enables us to reduce modulo $(\mathbf{p})$ the elements $\mathbf{x}^\alpha\, p_0$ for $\alpha \in E_0$, in $\langle \mathbf{x}^\alpha \rangle_{\alpha \in E_0}$. As this is valid for any value of the parameter u_i, we can reduce in $\langle \mathbf{x}^\alpha \rangle_{\alpha \in E_0}$ modulo $(\mathbf{p})$ the monomial $\mathbf{x}^\alpha\, x_i$ (resp. $\mathbf{x}^\alpha\, x_i^{-1}$ in the toric case), for any variable x_i and any $\alpha \in E_0$. By induction, for any polynomial p in R (or for any Laurent polynomial in L in the toric case) and any $\alpha \in E_0$, we can reduce modulo $(\mathbf{p})$ the polynomial $\mathbf{x}^\alpha\, p$ in $\langle \mathbf{x}^\alpha \rangle_{\alpha \in E_0}$.

Therefore, as $1 \in (\mathbf{x}^\alpha)_{\alpha \in E_0}$ in the Macaulay case (or because any Laurent's monomial $p \in L$ is of the form $p = p' \mathbf{x}^\alpha$ with $\alpha \in E_0$ and $p' \in L$ in the toric case), we can reduce modulo $(\mathbf{p})$ any polynomial or Laurent polynomial p in $\langle \mathbf{x}^\alpha \rangle_{\alpha \in E_0}$ (where $p \in R$, in the Macaulay case, or $p \in L$, in the toric case). This proves that $(\mathbf{x}^\alpha)_{\alpha \in E_0}$ is a generating set of $\mathcal{A} = R/(p_1, \ldots, p_n)$ ($\mathcal{A} = L/(p_1, \ldots, p_n)$ in the toric case). $\square$

Let us show more details on how we choose the subsets E_i in the Macaulay case and in the toric case.

Macaulay case Let us choose E_i such that the matrix $\tilde{S}$ becomes the identity matrix (see [20]), when we replace the polynomial p_i by $x_i^{d_i}$. We can choose, for instance,

$$
\begin{aligned}
E_0 &= \{(\alpha_1, \ldots, \alpha_n);\ 0 \leq \alpha_i \leq d_i - 1, i = 1, \ldots, n\}, \\
E_1 &= \{\alpha = (\alpha_1, \ldots, \alpha_n); |\alpha| \leq \nu - d_1;\ 0 \leq \alpha_i \leq d_i - 1, i = 2, \ldots, n\}, \\
&\vdots \\
E_n &= \{\alpha = (\alpha_1, \ldots, \alpha_n); |\alpha| \leq \nu - d_n\},
\end{aligned}
$$

where $|\alpha| = |\alpha_1| + \cdots + |\alpha_n|$.

According to theorem 3.2.3, $(\mathbf{x}^\alpha)_{\alpha \in E_0}$ is *generically* a generating set of $\mathcal{A}$, and

$$\dim_{\mathbb{C}}(\mathcal{A}) \leq |E_0| = \prod_{i=1}^{n} d_i,$$

which gives BEZOUT THEOREM.

Toric case In the toric case, the polynomial p_i is replaced by $p_i^{\mathbf{t}} = \sum_\alpha a_{i,\alpha}\, t^{w_\alpha} \mathbf{x}^\alpha$ (where t is a new variable and $w_\alpha \in \mathbb{Q}_+$). The subsets of the exponents E_i are chosen so that the corresponding matrix $S(t) = (s_{i,j}(t))$ satisfies

$$deg_t(s_{i,i}(t)) < deg_t(s_{i,j}(t)) \text{ for } i \neq j$$

(see [15],[5] for more details). The set E_0 is the set of the exponents in the *mixed cells* of a regular triangulation of $C_1 \oplus \cdots \oplus C_n$, so that, by construction, $|E_0|$ is the mixed volume of $C_1, \ldots, C_n$. This yields BERNSTEIN THEOREM (part 1) (see [2], [19]).

Part 2 of Bernstein theorem shows that *generically* the number of common zeros of the system $p_1 = \ldots = p_n = 0$ is at least $|E_0|$. Thus, we deduce that $\dim_{\mathbb{K}}(\mathcal{A}) \geq |E_0|$, and we have the following theorem:

Theorem 3.2.4 *For generic values of the coefficients of $p_1, \ldots, p_n$, $(\mathbf{x}^\alpha)_{\alpha \in E_0}$ is a basis of $\mathcal{A}$, in both Macaulay and toric case.*

Note that we gave simpler proofs than in the articles [13], [28].

3.2.4 Matrices of multiplication in $\mathcal{A}$

In this section, we still let $\mathcal{S}$ denote the map (30), constructed with using the fixed polynomials $p_1, \ldots, p_n$ and vector spaces $\mathcal{V}_1, \ldots, \mathcal{V}_n, \mathcal{V}$ and with various choices of polynomial p_0 and vector space $\mathcal{V}_0 = \langle \mathbf{x}^\alpha \rangle_{\alpha \in E_0}$. The set of monomials $(\mathbf{x}^\alpha)_{\alpha \in E_0}$, defined in the previous section, is a basis of $\mathcal{A}$.

For any polynomial p_0, we can also construct the table of multiplication by p_0, starting with a submatrix of $\mathcal{S}$. Namely, we choose any subset $E_i' \subset F_i$, $i = 1, \ldots, n$, such that simultaneously

(a') $|E'_1| + \cdots + |E'_n| = |F| - |E_0|$,

(b') and the corresponding columns in the matrix of $\mathcal{S}$ are *linearly independent*.

Generically, this is always possible, which we can show by giving a specific example. Decomposing again the matrix of the map

$$\begin{array}{rcl} \tilde{\mathcal{S}} : \langle \mathbf{x}^{E_0} \rangle \times \langle \mathbf{x}^{E'_1} \rangle \times \cdots \times \langle \mathbf{x}^{E'_n} \rangle & \rightarrow & \langle \mathbf{x}^F \rangle \\ (q_0, \ldots, q_n) & \mapsto & \sum_{i=0}^{n} p_i\, q_i \end{array}$$

in the form (32), we obtain the following property:

Theorem 3.2.5 *For generic values of the coefficients of $p_1, \ldots, p_n$, the matrix of multiplication by p_0 in $\mathcal{A}$ is given by*

$$M_{p_0} = U - V\, W^{-1}\, Z.$$

Proof. First, we will show that W is invertible. Otherwise, there exists a vector $\mathbf{v} \neq \mathbf{0}$ in the kernel of W. Then we have

$$\left[\begin{array}{c} V \\ W \end{array} \right] \mathbf{v} = \left[\begin{array}{c} \mathbf{w} \\ \mathbf{0} \end{array} \right],$$

and $\mathbf{w}$ is not $\mathbf{0}$, because the columns $\left[\begin{array}{c} V \\ W \end{array} \right]$ of the matrix S are linearly independent (condition (b')). This implies that there is a non-zero polynomial of the form $w(x) = \sum_{i=1}^{n} p_i q_i$ in $\langle \mathbf{x}^\alpha \rangle_{\alpha \in E_0}$, which contradicts the fact that $(\mathbf{x}^\alpha)_{\alpha \in E_0}$ is a basis of $\mathcal{A}$. Consequently, W is invertible and, by the same argument as in section 3.2.1, $U - VW^{-1}Z$ is the matrix M_{p_0} of multiplication by p_0 in the basis $(\mathbf{x}^\alpha)_\alpha$ of $\mathcal{A}$. $\square$

Example (continued) Let $p_0 = x_1$, $\mathbf{x}^{E_0} = (1, x_1, x_2, x_1\, x_2)$, $\mathbf{x}^{E_1} = \mathbf{x}^{E_2} = (1, x_1, x_2)$, and

$$\mathbf{x}^F = (1, x_1, x_2, x_1 x_2, {x_1}^2, {x_2}^2, {x_1}^3, {x_2}^3, {x_1}^2 x_2, x_1 {x_2}^2).$$

Then we have

$$\tilde{S} = \left[\begin{array}{cccc|cccccc} 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & -1 & 0 & -8 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 2 & 0 & -1 & 0 & 0 & -8 \\ \hline 0 & 1 & 0 & 0 & 1 & -1 & 0 & 1 & -8 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 2 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 1 & 0 \end{array} \right].$$

We may verify that

$$U - V\,W^{-1}\,Z = \begin{bmatrix} 0 & 1 & 0 & -\frac{14}{5} \\ 1 & 1 & 0 & -\frac{12}{5} \\ 0 & 0 & 0 & \frac{1}{5} \\ 0 & -2 & 1 & \frac{29}{5} \end{bmatrix}$$

is the matrix of multiplication M_{x_1} of the example of section 2.8.

3.2.5 The dual basis

It is possible to construct the dual basis $(\sigma_\alpha)_{\alpha\in E_0}$ of $(\mathbf{x}^\alpha)_{\alpha\in E}$, from the matrix S. Let

$$\sigma_\alpha = \sum_{\beta\in\mathbb{N}^n} \sigma_{\alpha,\beta}\,\partial^\beta$$

be the f.p.s. representing σ_α in $\mathbb{C}[[\partial]]$. Then we have the following property:

Proposition 3.2.6 *The coefficients $[\sigma_{\alpha,\beta}]_{\alpha\in E_0,\beta\in F}$ of $(\partial^\beta)_{\beta\in F}$ in the dual basis $(\sigma_\alpha)_{\alpha\in E_0}$ are given by the matrix*

$$\left[\ \mathbb{I}_D \,\middle|\, -VW^{-1}\ \right].$$

Proof. Let $[\sigma_\alpha] = [\sigma_{\alpha.\beta}]_{\beta\in F}$ denote the vector of the first coordinates of σ_α and let Σ denote the matrix $\Sigma = [\sigma_{\alpha,\beta}]_{\alpha\in E_0,\beta\in F}$. As $E_0 \subset F$, we represent this matrix as a 1×2 block matrix $\Sigma = [\Sigma'|\Sigma'']$, where $\Sigma' = [\sigma_{\alpha,\beta}]_{\alpha,\beta\in E_0}$ and $\Sigma'' = [\sigma_{\alpha,\beta}]_{\alpha\in E_0,\beta\in F-E_0}$. The linear forms σ_α vanish on the multiples of $p_1,\ldots,p_n$, which implies that

$$[\Sigma'|\Sigma'']\begin{bmatrix} V \\ W \end{bmatrix} = 0$$

or, equivalently,

$$\Sigma'\,V + \Sigma''\,W = 0. \tag{33}$$

The set $(\sigma_\alpha)_{\alpha\in E_0}$ is the dual basis of $(\mathbf{x}^\alpha)_{\alpha\in E_0}$, which means that, for any $\alpha,\beta\in E_0$, $\sigma_\alpha(\mathbf{x}^\beta) = \sigma_{\alpha,\beta}$ equals 1 if $\alpha=\beta$ and 0 otherwise. In other words, $\Sigma' = \mathbb{I}_D$ is the identity matrix and, according to (33) we have

$$\Sigma'' = -V\,W^{-1}.$$

□

Algorithm 3.2.3 (Computation of the normal form of a multivariate polynomial). *For any polynomial $p\in\langle\mathbf{x}^\beta\rangle_{\beta\in F}$, compute its normal form by multiplying the matrix $[\mathbb{I}_D| - VW^{-1}]$ by the coordinate vector of p.*

Proposition 3.2.7 *Algorithm 3.2.3 can be performed by using $C_{LinSolve}(W)+C_{PolMult}(E_0, F)+D$ ops, where $C_{LinSolve}(W)$ denotes the arithmetic complexity of solving a linear system of equations with the coefficient matrix W.*

Proof. The normal form of a polynomial $p = \sum_{\beta\in F} p_\beta\, \mathbf{x}^\beta$ is by definition

$$\sum_{\alpha\in E_0} \sigma_\alpha(p)\, \mathbf{x}^\alpha.$$

The coefficients $\sigma_\alpha(p) = \sum_{\beta\in F} \sigma_{\alpha,\beta}\, \partial^\alpha(p) = \sum_{\beta\in F} \sigma_{\alpha,\beta}\, p_\beta$ are obtained by multiplication of $\Sigma = [\mathbb{I}_D| - V\, W^{-1}]$ by the vector $[p_\beta]_{\beta\in F}$. $\square$

Example (continued) Let us be given the matrix

$$[\mathbb{I}_4| - V\, W^{-1}] = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & -1 & \frac{33}{5} & \frac{14}{5} & -\frac{14}{5} & \frac{7}{5} \\ 0 & 1 & 0 & 0 & 1 & 7 & \frac{34}{5} & \frac{12}{5} & -\frac{12}{5} & \frac{6}{5} \\ 0 & 0 & 1 & 0 & 0 & 0 & -\frac{2}{5} & -\frac{1}{5} & \frac{1}{5} & \frac{2}{5} \\ 0 & 0 & 0 & 1 & -2 & 2 & -\frac{68}{5} & \frac{11}{5} & \frac{29}{5} & -\frac{12}{5} \end{bmatrix}.$$

The normal form of $x_1x_2^2$ is defined by the last column of this matrix:

$$\mathtt{Nf}(x_1x_2^2) = \frac{7}{5} + \frac{6}{5}\, x_1 + \frac{2}{5}\, x_2 - \frac{12}{5}\, x_1x_2,$$

as found in the example of section 3.1.1. The linear form $\sigma_{x_1x_2}$ (in the last row of this matrix) turns into

$$\sigma_{x_1x_2} = \partial_1\partial_2 - 2\,\partial_1{}^2 + 2\,\partial_2{}^2 - \frac{68}{5}\,\partial_1^3 + \frac{11}{5}\,\partial_2{}^3 + \frac{29}{5}\,\partial_1{}^2\partial_2 - \frac{12}{5}\,\partial_1\partial_2{}^2 + \cdots.$$

3.3 Iterative methods in $\mathcal{A}$

Efficient iterative methods for solving the system $\mathbf{p} = \mathbf{0}$ rely on fast multiplication in $\mathcal{A}$. We will next present some examples where the structure of the matrices can be exploited.

3.3.1 Fast multiplication in $\mathcal{A}$

For a large class of polynomial ideals, specified, for instance, in [24], we may effectively compute the residue. Hereafter, we assume that *the residue τ is known*, and we will use it for computing efficiently the product of two elements in $\mathcal{A}$.

For any element $f \in \mathcal{A}$, let $[f]$ denote the coordinate vector of f in the basis $(\mathbf{x}^\alpha)_{\alpha\in E}$. Let us write $\mathbf{w}_\alpha(\mathbf{x}) = \sum_{\beta\in E} B^1_{\alpha,\beta}\mathbf{x}^\alpha$ to denote the dual basis of $(\mathbf{x}^\alpha)_{\alpha\in E}$ and $B_1 = (B^1_{\alpha,\beta})_{\alpha,\beta\in E}$ to denote the Bezoutian of 1.

We want to compute the product $[f\,g]$ in $\mathcal{A}$ where

$$\begin{array}{l} f := \sum_{\alpha\in E} f_\alpha\, \mathbf{x}^\alpha, \\ g := \sum_{\alpha\in E} g_\alpha\, \mathbf{x}^\alpha. \end{array}$$

We may first compute the polynomial $f\,g$ and then reduce it to a linear combination of the elements of the monomial basis $\langle \mathbf{x}^\alpha \rangle$ in order to obtain $[f\,g]$. We may also proceed directly by using the projection formula:

$$\begin{aligned} f\,g &= \sum_{\alpha\in E} \tau(f\,g\,\mathbf{x}^\alpha)\mathbf{w}_\alpha \\ &= \sum_{\alpha\in E} f\,g\star\tau(\mathbf{x}^\alpha)\,\mathbf{w}_\alpha. \end{aligned}$$

In this case, we have to compute the coefficients of the linear form $f\,g\star\tau$ and then shift from the basis $(\mathbf{w}_\alpha)_{\alpha\in E}$ to the monomial basis $(\mathbf{x}^\alpha)_{\alpha\in E}$. By using relations (23), we may also proceed in an equivalent way, based on the formula

$$[f\,g] = M_g[f] = H_1^{-1}\,H_{f\star\tau}[g] = B_1\,H_{f\star\tau}[g]. \tag{34}$$

As we want to compute the coefficients $f\,g\star\tau(\mathbf{x}^\alpha) = \tau(f\,g\,\mathbf{x}^\alpha)$ for $\alpha\in E$, we need to know the value of τ for the monomials $\mathbf{x}^{\alpha+\beta+\gamma}$ for $\alpha,\beta,\gamma\in E$. Let $\tilde\tau := \sum_{u\in 3\,E}\tau(\mathbf{x}^u)\,\partial^u$ denote the leading part of the series τ associated with the residue τ. We first compute

$$\begin{aligned} g\star\tilde\tau &= \pi_+(g(\partial^{-1})\,\tilde\tau(\partial)) \\ &= \pi_+((\sum_{\alpha\in E} g_\alpha\,\partial^{-\alpha})(\sum_{u\in 3\,E}\tau_u\partial^u)) \end{aligned}$$

and then

$$\begin{aligned} f\,g\star\tilde\tau &= f\star(g\star\tilde\tau) = \pi_+(f(\partial^{-1})\,g(\partial^{-1})\,\tilde\tau(\partial)) \\ &= \pi_+\left(\left(\sum_{\alpha\in E} f_\alpha\partial^{-\alpha}\right)\left(\sum_{\beta\in E} g_\beta\partial^{-\beta}\right)\left(\sum_{u\in 3\,E}\tau_u\,\partial^u\right)\right). \end{aligned}$$

The coefficients λ_α of ∂^α in $f\,g\star\tilde\tau$ for $\alpha\in E$ are precisely the coefficients of $f\,g$ in the dual basis $(\mathbf{w}_\alpha)_{\alpha\in E}$ of $\mathcal{A}$. Summarizing, we obtain the following algorithm:

Algorithm 3.3.1 (Multiplication by a polynomial modulo an ideal in a monomial basis). *To obtain the coefficients of $f\,g$ in the basis $(\mathbf{x}^\alpha)$:*

- *Compute the coefficient vector $\Lambda = [\lambda_\alpha]_{\alpha\in E}$ of ∂^α for $\alpha\in E$, by multiplying the Laurent polynomial $f(\partial^{-1})g(\partial^{-1})$ by $\tilde\tau(\partial)$.*
- *Multiply the vector $\Lambda = [\lambda_\alpha]_{\alpha\in E}$ by the matrix $B_1 = H_1^{-1}$, that is, solve the linear system of equations $H_1\,\mathbf{v} = \Lambda$. Output the vector $\mathbf{v}$.*

3.3.2 Fast inversion in $\mathcal{A}$

Similar techniques can be used to compute the inverse (reciprocal) of an invertible element $f \in \mathcal{A}$. By relation (34), for $g = f^{-1}$, we have

$$[1] = H_1^{-1} H_{f\star\tau}[f^{-1}] \text{ or, equivalently, } H_1[1] = H_{f\star\tau}[f^{-1}],$$

where $H_1[1]$ is the coefficient vector of $(\partial^{\alpha})_{\alpha\in E}$ in $\tilde{\tau}$. This yields the following algorithm:

Algorithm 3.3.2 (Inverse of a polynomial modulo an ideal in a monomial basis). *To obtain the coefficients of f^{-1} in the basis $(\mathbf{x}^{\alpha})$:*

- *Compute the coefficient vector $\mathbf{u} = [\lambda_{\alpha}]_{\alpha\in E}$ of ∂^{α} for $\alpha \in E$, in $\tilde{\tau}$.*
- *Compute the coefficients of $\partial^{\alpha+\beta}$ for $\alpha, \beta \in E$ in the Laurent polynomial $f(\partial^{-1})\,\tilde{\tau}(\partial)$, to dedine the matrix $H_{f\star\tau}$.*
- *Solve the linear system $H_{f\star\tau}\,\mathbf{v} = \mathbf{u}$. Output the vector $\mathbf{v}$.*

3.3.3 Computing selected roots of a polynomial system

As before, let $\mathcal{Z}$ denote the set of all common roots of the system $\mathbf{p} = \mathbf{0}$. Then again, we assume here that *the roots are simple.*

By decomposing any element h of $\mathcal{A}$ in the basis of idempotents $\mathbf{e}_{\zeta}$ (see section 2.12), we obtain that

$$h(\mathbf{x}) = \sum_{\zeta\in Z} h(\mathbf{x})\,\mathbf{e}_{\zeta} \equiv \sum_{\zeta\in Z} h(\zeta)\,\mathbf{e}_{\zeta}.$$

The second equation follows since $\mathbf{e}_{\zeta}h(x) \equiv \mathbf{e}_{\zeta}h(\zeta)$. Squaring h in the quotient ring $\mathcal{A}$ gives us that

$$h^2 \equiv \sum_{\zeta\in Z} h(\zeta)^2\,\mathbf{e}_{\zeta}.$$

Here and hereafter, for any element $a \in \mathcal{A}$, $[a]$ denotes the vector of the coefficients of a in the basis $(\mathbf{x}^{\alpha})_{\alpha\in E}$. In particular, $[1] = (1, 0, \cdots, 0)^t$ if the basis starts with the monomial 1. Let $||\cdot||$ denote a norm in $\mathbb{C}^D$, say, the Euclidean or Hermitian norm,

$$||\mathbf{v}|| = (\mathbf{v}, \mathbf{v}) = (\sum_{i=1}^{D} |v_i|^2)^{1/2},\ \mathbf{v} = (v_i),\ i = 1, \ldots, D.$$

By minor abuse of notation, for any element $a \in \mathcal{A}$, we will let $||a||$ denote $||[a]||$. Let $h \in R$ and assume that there is a unique root $\zeta \in Z$, for which the absolute value of $h(\zeta)$ is maximum, so that

$$|h(\zeta)/h(\eta)| - 1 \geq \rho, \tag{35}$$

for some fixed positive ρ and for any $\eta \in \mathcal{Z}$ distinct from ζ. (Since all the roots in $\mathcal{Z}$ are assumed to be distinct, we may, in principle, ensure the latter relation with a high probability, by means of a random linear substitution of the vector of the variables $\mathbf{x}$. For the uniform distribution of all roots in a box in $\mathbb{C}^n$, we easily yield $1/\rho = O(D)$.) Then, by iteratively computing and normalizing the squares, we obtain

$$h_0 = h,\ h_{i+1} \equiv h_i^2/||h_i^2||,\ i = 0, 1, \ldots, k-1,$$

and arrive at the bounds

$$\epsilon_k := ||\frac{h_k}{||h_k||} - \frac{\mathbf{e}_\zeta}{||\mathbf{e}_\zeta||}|| \leq \frac{c}{(1+\rho)^{2^k}}.$$

Therefore, we ensure the bound $\epsilon_k \leq 2^{-b}$ in $k = k(\rho, b) = \mathcal{O}(\log(b/\rho))$ recursive steps for any positive b, which shows that, right from the start, the process quadratically converges to a multiple of the idempotent $\mathbf{e}_\zeta$.

Proposition 3.3.1 *In the case of simple roots and for $h \in R$ such that $|h(\zeta)| > |h(\eta)|$ for any $\eta \neq \zeta \in \mathcal{Z}(I)$, the latter process of repeated squaring and normalization in $\mathcal{A}$, always converges quadratically right from the start to a multiple of the idempotent $\mathbf{e}_\zeta$.*

We refer the reader to [31] and [7] for some preceding works on a similar approach in the univariate case.

Algorithm 2.12.1 computes the root vector ζ as the product of H_1 by the idempotent vector $\mathbf{e}_\zeta$. Therefore, the transition from $\mathbf{e}_\zeta$ to the root ζ of the system $\mathbf{p} = \mathbf{0}$ can be performed in $C_{LinSolve}(H_1)$ ops.

Thus, we have the following algorithm:

Algorithm 3.3.3 (Computation of the simple root that maximizes the modulus of a fixed polynomial). *Assume that all the roots in $\mathcal{Z}(I)$ are simple and $h \in R$ is such that there exists $\zeta \in \mathcal{Z}(I)$, with $|h(\zeta)| > |h(\eta)|$ for any $\eta \neq \zeta \in \mathcal{Z}(I)$.*

- *Set $u_0 := h$ and fix a positive tolerance value $\epsilon = 2^{-b}$.*
- *Recursively, for $k = 0, 1, \cdots, N-1$, apply algorithm 3.3.1 to compute $v_{k+1} \equiv u_k^2$ and $u_{k+1} = \frac{v_k}{||v_k||}$ in $\mathcal{A}$, until the norm $||u_{k+1} - u_k||$ becomes smaller than ϵ,*
- *Compute and output the product of the matrix H_1 by u_N.*

The algorithm outputs a multiple of the vector $[\zeta^\alpha]_{\alpha \in E}$, whose scaling gives us the root ζ for which $|h(\zeta)|$ is maximal (compare algorithm 3.1.1). The overall cost of the approximation of the root within an error norm 2^{-b} is bounded by $\mathcal{O}(D^2 \log(D) \log(b/\rho))$ ops.

Algorithm 3.3.3 can be applied recursively to approximate several roots of the polynomial system $\mathbf{p} = \mathbf{0}$. In each application, the basic element $h(\mathbf{x})$ of $\mathcal{A}$ should be updated; we propose the following policy:

$$h(\mathbf{x}) = \prod_\zeta h_\zeta(\mathbf{x})$$

where the product is over all previously computed approximations ζ to the roots of the system $\mathbf{p} = \mathbf{0}$ and $h_\zeta(\zeta) = 0$ for all ζ.

One may choose $h_\zeta(\mathbf{x})$ to be linear functions in $\mathbf{x}$, $h_\zeta(x) = \sum_{i=1}^n h_{i,\zeta} x_i - h_{0,\zeta}$ with fixed or random coefficients $h_{1,\zeta}, \ldots, h_{n,\zeta}$ and with $h_{0,\zeta} = \sum_{i=1}^n h_{i,\zeta}\zeta_i$. Under this choice, assumption (35) is satisfied for all computed approximations ζ to the roots, in the case of generic input, and we yield the following result.

Theorem 3.3.2 *In δ recursive applications of algorithm 3.3.1, δ distinct roots of the generic polynomial system $\mathbf{p} = \mathbf{0}$ can be approximated at the cost of performing $\mathcal{O}(a\delta D^2 \log(D))$ ops, where a is the average number of iterations per root, $a = \mathcal{O}(\log(\frac{b}{\rho}))$ for the approximation within the error 2^{-b} of roots ζ satisfying (35). In particular, $a = \mathcal{O}(\log D)$ if $\frac{b}{\rho} = D^{\mathcal{O}(1)}$.*

A distinct extension to approximation of several roots is shown in [25].

3.3.4 Computing the closest root

Suppose that we seek a root of the system $\mathbf{p} = \mathbf{0}$ whose coordinate x_1 is the closest to a given value $u \in \mathbb{C}$. Let us assume that u is not a projection of any root of the system $\mathbf{p} = \mathbf{0}$, so that $x_1 - u$ has reciprocal in $\mathcal{A}$. Let $\rho_1(\mathbf{x})$ denote such a reciprocal. We have $\rho_1(\mathbf{x})(x_1 - u) \equiv 1$ and $\rho_1(\zeta) = \frac{1}{\zeta_1 - u}$. Therefore, a root whose coordinate x_1 is the closest to u_1 is a root for which $|\rho_1(\zeta)|$ is the largest. Consequently, iterative squaring of $\rho_1 = \rho_1(\mathbf{x})$ shall converge to this root.

The polynomial ρ_1 can be computed in the following way. Let $\overline{\mathcal{M}}_{x_1-u}$ denote the multiplication by $x_1 - u$ in $\mathcal{A}$. Then $\rho_1 = (\overline{\mathcal{M}}_{x_1-u})^{-1}(1)$, and according to the matrix equation (24), we have

$$[\rho_1] = H_1\,(H_{x_1} - u\,H_1)^{-1}[\mathbf{1}].$$

$[\rho_1]$ defined by the latter equation can be computed within $C_{LinSolve}(H_{x_1-u}) + C_{PolMult}(-2E, E)$ ops, by using the black box algorithms of the appendix.

One may compute several roots of the polynomial system by applying the latter computation (successively or concurrently) to several initial values u.

Example (continued) We illustrate this approach by first computing the root for which x_1 is maximal. We start with $u_0 = x_1$. After 4 iterations, we obtain

$$u_4 = 7.6055995 + 7.7975926x_1 - 0.46159096x_2 - 15.740471x_1x_2.$$

By multiplying the coefficient vector of this polynomial by H_1 and dividing by the first coordinate, we obtain

$$[\zeta_1^\alpha]_{\alpha \in E} = [1., 6.820095, -2.836734, -19.34680],$$

where $\zeta_1 = (6.820095, -2.836734)$.

If we start with

$$u_0 \equiv (x_1 - \frac{1}{2})^{-1} \equiv -\frac{78}{35} - \frac{228}{35} x_1 - \frac{32}{35} x_2 - \frac{16}{7} x_1 x_2,$$

the algorithm should converge to the root with the coordinate x_1 closest to $\frac{1}{2}$. Indeed, after 4 iterations, we obtain

$$u_4 = 0.15292071 + 0.89409187 x_1 + 0.16270766 x_2 + 0.29923055 x_1 x_2,$$

and after multiplication by H_1 and normalization, we arrive at

$$[\zeta_4^\alpha]_{\alpha \in E} = [1., 0.3678148, 1.675476, 0.6162664],$$

where $\zeta_4 = (0.3678148, 1.675476)$ is the root with the coordinate x_1 closest to $\frac{1}{2}$.

3.3.5 Splitting the set of roots

In addition to the repeated squaring iteration of algorithm 3.3.3, we will also consider iteration associated to a slight modification of the so-called *Joukovski* map (see [17],[7]): $z \mapsto \frac{1}{2}(z + \frac{1}{z})$ and its variant $z \mapsto \frac{1}{2}(z - \frac{1}{z})$.

The two attractive fixed points of this map are 1 and -1; for its variant, they turn into $\mathbf{i}$ and $-\mathbf{i}$, $\mathbf{i}$ denoting here $\sqrt{-1}$.

Algorithm 3.3.4 *(*Sign iteration*). Fix $u_0 = h \in \langle \mathbf{x}^\alpha \rangle_{\alpha \in E}$. Then recusively compute the polynomials $u_{n+1} \equiv \frac{1}{2}(u_n - \frac{1}{u_n}) \in \mathcal{A}$, $n = 0, 1, \ldots$.*

Each iteration step of the algorithm can be performed within $C_{LinSolve}(H_{u_n}) + C_{PolMult}(-3E, E)$ ops, by using algorithm 3.3.2 and the black box algorithms of the appendix.

Hereafter, $\Re(h)$ and $\Im(h)$ denote the real and imaginary parts of a complex h, respectively.

Proposition 3.3.3 *Assume that for any root $\zeta \in \mathcal{Z}$, $\Re(h(\zeta)) \neq 0$. Then the sequence (u_n) converges quadratically to $\sigma = \sum_{\Im(h(\zeta))>0} \mathbf{e}_\zeta - \sum_{\Im(h(\zeta))<0} \mathbf{e}_\zeta$, and we have*

$$\|u_n - \sigma\| \leq K \rho^{2^n}$$

(for some constant K), where

$$\rho^+ = max_{\Im(h(\zeta))>0, \zeta \in \mathcal{Z}(I)} \left| \frac{h(\zeta) - \mathbf{i}}{h(\zeta) + \mathbf{i}} \right|,$$

$$\rho^- = max_{\Im(h(\zeta))<0, \zeta \in \mathcal{Z}(I)} \left| \frac{h(\zeta) + \mathbf{i}}{h(\zeta) - \mathbf{i}} \right|,$$

and $\rho = \max\{\rho^+, \rho^-\}$.

Proof. We apply the classical convergence analysis of the Joukovski map (see [17]) to the matrices of multiplication by u_n in $\mathcal{A}$, whose eigenvalues are $\{u_n(\zeta), \zeta \in \mathcal{Z}(I)\}$. $\square$

This iteration can be applied to count the number of roots in a half-space, based on the following proposition:

Proposition 3.3.4 *The rank of the matrix H_{σ^+} is the number of roots such that $\Im(h(\zeta)) > 0$ (where the roots are counted with their multiplicity).*

Proof. As H_1 is invertible, the rank of $H_{\sigma^+} = H_1 M_{\sigma^+}$ is the rank of M_{σ^+}, that is, the dimension of $\sigma^+\mathcal{A}$, which equals $\sum_{\Im(h(\zeta))>0} \mathbf{e}_\zeta \mathcal{A}$. The dimension of $\mathcal{A}_\zeta = \mathbf{e}_\zeta \mathcal{A}$ is the mutiplicity of ζ, which yields the proposition. $\square$

The ranks can be computed efficiently by the algorithm supporting theorem B.3.2 (of appendix B).

Note that we may replace the iteration of algorithm 3.3.4 by the iteration $u_{n+1} \equiv \frac{1}{2}(u_n + \frac{1}{u_n}) \in \mathcal{A}$ and then replace $\Im$ by $\Re$ in the previous two propositions.

By $4n - 4m$ successive applications of the above splitting procedures, we can compute efficiently the number of roots in a fixed box in $\mathbb{C}^n$ bounded by $2n - 2m$ pairs of parallel hyperplanes, $0 \leq m < n$. See [25] for more advanced applications of these techniques, which in particular enable us to restrict the approximation of the roots by algorithm 3.3.3 to a fixed box in $\mathbb{C}^n$.

3.4 Traces and real roots

In this section, we will keep assuming that the residue τ is known, will suppose that the coefficients of the polynomials p_i are real, and will study the problem of computing the numbers of distinct roots and of real roots. We will next define a special element of $\widehat{R}$, called *the trace.*

Definition 3.4.1 *The linear form* Tr *is defined over any fixed field $\mathbb{K}$ by*

$$\begin{aligned} \mathrm{Tr} : R &\rightarrow \mathbb{K} \\ p &\mapsto trace(\overline{\mathcal{M}}_p), \end{aligned}$$

where $trace(\overline{\mathcal{M}}_p)$ is the usual trace of the linear operator $\overline{\mathcal{M}}_p$.

By using this linear form, we define the *quasi-Hankel* matrix

$$H_{\mathrm{Tr}} = [\mathrm{Tr}(\mathbf{x}^{\alpha+\beta})]_{\alpha,\beta \in E}.$$

In order to compute H_{Tr}, assuming that we know the table of the multiplications by x_i in $\mathcal{A}$ ($i = 1, \ldots, n$), we may compute the values of $\mathbf{x}^\gamma$ (for $\gamma = \alpha + \beta$ and $\alpha, \beta \in E$) by induction, for we have $\mathbf{x}^\gamma = x_i \mathbf{x}^{\gamma'}$ with $|\gamma'| < |\gamma|$ and $\mathrm{Tr}(1) = D = \dim_{\mathbb{R}}(\mathcal{A})$. By using the linearity of the trace, we compute all the coefficients of H_{Tr} (see, for instance, [30]). Alternatively, we may apply the following theorem (see [11]):

Theorem 3.4.2 *Let $J \in R$ be the Jacobian of the polynomials $p_1, \ldots, p_n$. Then*

$$\mathrm{Tr} = J \star \tau.$$

Example (continued) According to the example of section 2.8, we have

$$\mathrm{Tr}(x_1) = 1 + \frac{29}{5} = \frac{34}{5}$$

and also

$$\tau(x_1\, J) = \tau(-16 - 16\, x_1 + 4\, x_2 + 34\, x_1 x_2) = \frac{34}{5}.$$

Algorithm 3.4.1 *(*Application of the trace to a monomial set). *Compute and output $H_{\mathrm{Tr}} = [\mathrm{Tr}(\mathbf{x}^{\alpha+\beta})]_{\alpha,\beta\in E}$ as the product of the linear form*

$$\tilde{\tau} = \sum_{\alpha\in 3\, E} \tau_\alpha\, \partial^\alpha$$

by $J(\partial^{-1})$.

The number of ops involved in this algorithm is bounded by $C_{PolMult}(3\, E,\, -E)$.

Once the matrix H_{Tr} is computed, we apply the following theorem, due to Hermite (see [18], [26], [10]):

Proposition 3.4.3 (Hermite). *Let J be the Jacobian of $\mathbf{p} = (p_1, \ldots, p_n)$ and let B_J be the Bezoutian of J. Then*

- *the rank of H_{Tr} or B_J is the number of distinct roots,*
- *the signature of H_{Tr} or B_J is the number of real roots.*

Algorithm 3.4.2 *(*Computation of the numbers of distinct roots and real roots). *For a polynomial system $p_1 = \cdots = p_n = 0$, define the matrix H_{Tr}, then compute the numbers of the distinct roots and the real roots of the system by applying proposition 3.4.3 and the algorithm supporting theorem B.3.2 (of appendix B).*

Example (continued) The normal form of the Jacobian J is

$$J = -8 + 40\, x_1 - 2\, x_2 + 20\, x_1 x_2.$$

Note that $\tau(J) = \frac{1}{5} \times 20 = 4$ is the dimension of $\mathcal{A}$. The matrix H_{Tr} is given by

$$H_{\mathrm{Tr}} = \begin{bmatrix} 4 & \frac{34}{5} & -\frac{12}{5} & -\frac{448}{25} \\ \frac{34}{5} & \frac{1166}{25} & -\frac{448}{25} & -\frac{16492}{125} \\ -\frac{12}{5} & -\frac{448}{25} & \frac{194}{25} & \frac{6976}{125} \\ -\frac{448}{25} & -\frac{16492}{125} & \frac{6976}{125} & \frac{234354}{625} \end{bmatrix}.$$

The matrix of the Bezoutian B_J is

$$B_J = \begin{bmatrix} -4 & -50 & 52 & -40 \\ -50 & 602 & -36 & 200 \\ 52 & -36 & 6 & -10 \\ -40 & 200 & -10 & 100 \end{bmatrix}.$$

The rank and the signature of both matrices are 4 and 2, respectively. The number of distinct roots is 4, and the number of distinct real roots is 2.

The overall randomized cost of computing the numbers of distinct roots and of real roots is $\mathcal{O}(D^2)$ up to a polylogarithmic factor.

4 Conclusion

We have deduced the results of sections 3.3 and 3.4 assuming that the residue τ associated with the ideal $I = (p_1, \cdots, p_n)$ is known (or readily available). This somewhat restricts the class of polynomial systems to which the results apply. A major research challenge is an extension of these results to a more general class of polynomial systems of equations having a finite number of solutions. Substantial progress in this direction based on further extension of the techniques of this paper combined with some other new techniques has been reported in [25]. In [4] some further elaboration of the presented approach towards some practical problems of multivariate polynomial rootfinding and optimization was shown. Another research challenge (still widely open) is to extend the results of section 3.3 to approximating all the D roots of the system at the cost $\mathcal{O}(D^2)$ (up to a polylogarithmic factor).

Our goal, throughout this paper, was to demonstrate the power of the application of dual spaces, algebraic residues and the generalization of the structure of Toeplitz and Hankel matrices to the solution of a polynomial system in the multivariate case. In order to be able to yield the latter generalization, we reinterpreted the associated operators in terms of operations in the polynomial ring and in its dual. Multivariate Bezoutians and residues come naturally under these studies, and the algebraic interpretation of the associated operators yielded the relations between these matrices. Section 3 was devoted to applications of the developed techniques to yield order of magnitude improvement of the known algorithms for some fundamental problems of multivariate polynomial rootfinding. We hope that this work and our subsequent progress will motivate new interest in this recently open and challenging area.

References

[1] W. Auzinger and H. Stetter, *An elimination algorithm for the computation of all zeros of a system of multivariate polynomial equations*, in Proc Intern. Conf. on Numerical Math., vol. 86 of Int. Series of Numerical Math., Birkhäuser, 1988, pp. 11–30.

[2] D. Bernstein, *The number of roots of a system of equations*, Funct. Anal. and Appl., 9 (1975), pp. 183–185.

[3] D. Bini and V.Y. Pan, *Polynomial and Matrix Computations, Vol. 1: Fundamental Algorithms*, Birkhäuser, Boston, 1994.

[4] D. Bondifalat, B. Mourrain, and V. Y. Pan, *Controlled iterative methods for solving polynomial systems,* Proc. Annual ACM-SIGSAM Intern. Symp. on Symb. and Algebr. Comp. (ISSAC'98), ACM Press, New York, 1998, pp. 252-259.

[5] J. Canny and I. Emiris, *An efficient algorithm for the sparse mixed resultant*, in Proc. Intern. Symp. Applied Algebra, Algebraic Algor. and Error-Corr. Codes (Puerto Rico), G. Cohen, T. Mora, and O. Moreno, eds., vol. 263 of Lect. Notes in Comp. Science, Springer Verlag, 1993, pp. 89–104.

[6] J. Canny, E. Kaltofen, and Y. Lakshman, *Solving systems of non-linear polynomial equations faster*, in Proc. of the Annual ACM-SIGSAM Int. Symp. on Symb. and Alg. Comp. (ISSAC'89), ACM Press, New York, 1989, pp. 121–128.

[7] J. Cardinal, *On two iterative methods for approximating the roots of a polynomial*, in Proc. AMS-SIAM Summer Seminar on Math. of Numerical Analysis, (Park City, Utah, 1995), J. Renegar, M. Shub, and S. Smale, eds., vol. 32 of Lectures in Applied Math., American Math. Soc. Press, Providence, Rhode Island, 1996, pp. 165–188.

[8] J. Cardinal and B. Mourrain, *Algebraic approach of residues and applications*, in Proc. AMS-SIAM Summer Seminar on Math. of Numerical Analysis, (Park City, Utah, 1995), J. Renegar, M. Shub, and S. Smale, eds., vol. 32 of Lectures in Applied Math., American Math. Soc. Press, Providence, Rhode Island, 1996, pp. 189–210.

[9] D. Cox, J. Little, and D. O'Shea, *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*, Undergraduate Texts in Mathematics, Springer, 1992.

[10] M. Demazure, *Charles Hermite, déjà ...* , Notes informelles de calcul formel 8, Centre de Math., Ecole Polytechnique (France), 1987.

[11] M. Elkadi and B. Mourrain, *Approche effective des résidues algébriques*, Rapport de Recherche 2884, INRIA, Sophia Antipolis, 1996.

[12] I. Z. Emiris and V. Y. Pan, *The Structure of Sparse Resultant Matrices*, in Proc. Annual ACM-SIGSAM Intern. Symp. on Symbolic and Algebraic Computation (ISSAC'97), ACM Press, New York, 1997, pp. 189-196.

[13] I. Emiris and A. Rege, *Monomial bases and polynomial system solving*, in Proc. Annual ACM-SIGSAM Intern. Symp. on Symbolic and Algebraic Computation (ISSAC'94), ACM Press, New York, 1994, pp. 114–122.

[14] P. Fuhrmann, *A Polynomial Approach to Linear Algebra*, Springer-Verlag, 1996.

[15] I. Gelfand, M. Kapranov, and A. Zelevinsky, *Discriminants, Resultants and Multidimensional Determinants*, Birkhäuser, Boston-Basel-Berlin, 1994.

[16] G. H. Golub and C. F. Van Loan, *Matrix Computations* (third edition), John Hopkins Univ. Press, Baltimore, Maryland, 1996.

[17] P. Henrici. *Applied and Computational Complex Analysis*, volume I. Wiley, 1988.

[18] C. Hermite, *Remarques sur le théorème de Sturm*, C. R. Acad. Sci. de Paris, 36 (1853), pp. 52–54.

[19] A. Kushnirenko, *A Newton polyhedron and the number of solutions of a system of k equations in k unknowns*, Uspekhi Matem. Nauk, 30 (1975), pp. 266–267.

[20] F. Macaulay, *Some formulae in elimination*, Proc. London Math. Soc., 1 (1902), pp. 3–27.

[21] F. Macaulay, *The Algebraic Theory of Modular Systems*, Cambridge Univ. Press, 1916.

[22] B. Mourrain, *Computing isolated polynomial roots by matrix methods*, J. Symb. Comp., December, 1998.

[23] B. Mourrain and V. Y. Pan, *Multidimensional structured matrices and polynomial systems*, Calcolo, 33 (1996), pp. 389-401.

[24] ——, *Solving special polynomial systems by using structured matrices and algebraic residues*, in Proc. of the workshop on Foundations of Computational Mathematics (Rio de Janeiro, 1997), F. Cucker and M. Shub, eds., Springer, 1997, pp. 287–304.

[25] B. Mourrain and V. Y. Pan, *Asymptotic acceleration of solving multivariate polynomial systems of equations, in* Proc. 30th Ann. ACM Symp. on Theory of Computing, ACM Press, New York, 1998, pp. 488-496.

[26] P. S. Pedersen, M.-F. Roy, and A. Szpirglas, *Counting real zeros in the multivariate case*, in Effective Methods in Algebraic Geometry (MEGA'92), A. Galligo and F. Eyssette, eds., Progress in Math., Nice (France), Birkhäuser, 1993, pp. 203–223.

[27] P. S. Pedersen and B. Sturmfels, *Product formulas for resultants and Chow forms*, Math. Zeitschrift, 214 (1993), pp. 377–396.

[28] ——, *Mixed monomial basis*, in Effective Methods in Algebraic Geometry (MEGA'94), vol. 143 of Progress in Math., Santander (Spain), Birkhäuser, 1994, pp. 285–306.

[29] P. Penfield Jr., R. Spencer, and S. Duinker, *Tellegen's Theorem and Electrical Networks*, M.I.T. Press, Cambridge, Massachussetts, 1970.

[30] F. Rouillier, *Systèmes polynomiaux*, PhD thesis, Univ. de Rennes, 1996.

[31] J. Sebastiao e Silva, *Sur une méthode d'approximation semblable à celle de Graeffe*, Portugal Math. J., 2 (1941), pp. 271–279.

[32] H. J. Stetter, *Eigenproblems are at the heart of polynomial system solving*, SIGSAM Bulletin, 30, 4, ACM Press, New York (1996), pp. 22–25.

[33] E. E. Tyrtyshnikov, *A unifed approach to old and new theorems on distribution and clustering,* Linear Algebra and Its Applications, 232 (1998) pp. 1-43.

[34] B. Van der Waerden, *Modern Algebra, Vol. II*, Frederick Ungar Publishing Co, 1948.

A Polynomials, Laurent's polynomials, and dual spaces (univariate case). Basic Definitions

Consider univariate polynomials $p = p(x) = \sum_{i=0}^{d} p_i x^i \in R = \mathbb{C}[x]$, represented by vectors of their complex coefficients $\mathbf{p} = (p_0, \ldots, p_d)$. Let the subspace R_d denote the vector space (of dimension $d+1$) of polynomials in R of degree at most d.

A fixed polynomial $p(x)$ of R generates the ideal $I = (p(x))$ in R, formed by all polynomial multiples $q(x)$ of $p(x)$. Let $\mathcal{A} = R/I$ denote the quotient ring of polynomials reduced modulo $p(x)$ (that is, modulo the ideal I). If $p(x)$ is of degree d, then $\mathcal{A}$ is isomorphic to R_{d-1}, as a *vector space.*

By introducing the reciprocal x^{-1}, we arrive at the ring of Laurent's polynomials $\mathbb{C}[x, x^{-1}] = L$ and let $L_{-c,d}$ denote the subspace of Laurent's polynomials of the form $\sum_{i=-c}^{d} \lambda_i x^i$.

A polynomial $p \in R_d$ can be represented by the vector of its $d+1$ coefficients or, equivalently, by the values $p(0), p'(0), \ldots, \frac{1}{d\,!} p^{(d)}(0)$. In other words, a primal basis of R_d is $\langle 1, x, \ldots, x^d \rangle$, and *its dual basis* (that is, the set of linear forms (maps) that compute the coefficients of p in the primal basis) is the set of linear forms

$$\langle p \mapsto \frac{1}{i\,!} p^{(i)}(0) \rangle_{i=0,\ldots,d}.$$

We introduce a new variable ∂ and let ∂^i denote the i$^{\text{th}}$ element, $p \mapsto \frac{1}{i\,!} p^{(i)}(0)$, of this *dual basis.* Thus, a linear form on R_d, that is, an element Λ of the dual space $\widehat{R}_d$ of R_d, is represented by a polynomial

$$\Lambda = \sum_{i=0}^{d} \lambda_i \, \partial^i.$$

For any $p \in R_d$, we have $\Lambda(p) = \sum_{i=0}^{d} \lambda_i \frac{1}{i\,!} \frac{d^{(i)}}{dx^i}(p)(0)$ and $\lambda_i = \Lambda(x^i)$.

Next, consider linear forms $\Lambda \in \widehat{R}$ on the primal space R. The restrictions of the linear forms to $R_d \subset R$ are the elements of $\widehat{R}_d$, which can be represented by polynomials in ∂ of degree at most d. This is valid for any d; therefore, an element $\Lambda \in \widehat{R}$ is a formal power series (f.p.s.) in ∂:

$$\Lambda = \sum_{i=0}^{\infty} \Lambda(x^i) \partial^i.$$

Such a ring of f.p.s. in the variable ∂ is denoted $S = \mathbb{C}[[\partial]]$.

For any $p(x) \in \mathbb{C}[x]$ and $\Lambda(\partial) \in \mathbb{C}[[\partial]]$, we define

$$p(x) \star \Lambda(\partial) = \pi_+(p(\partial^{-1})\Lambda(\partial)),$$

where $\pi_+ : \mathbb{C}[\partial^{-1}][[\partial]] \to \mathbb{C}[[\partial]]$ is the projection on the monomials having non-negative exponents in ∂.

Example

$$(1+x^2)\star(\partial^3+3\,\partial-2)=\partial^3+4\,\partial-2.$$

Contrary to [14], we introduce a new variable ∂ for the "inverse" of x, which we consider as an element of the dual space.

B Algorithms and complexity for some linear algebra computations

We will recall the known estimates for the computational cost of performing some basic algorithms used in this paper.

B.1 Polynomial multiplication

In sections 1 and 2, we reduced multiplication of various structured matrices by vectors to polynomial multiplication. Now, let us recall the known arithmetic complexity bounds for the latter operation. As before, let $C_{PolMult}(E,F)$ denote the number of arithmetic operations required for the multiplication of a polynomial with support in E by a polynomial with support in F.

Theorem B.1.1 *([3], p. 13) Let $E_d=[0,\ldots,d]\subset\mathbb{N}$. Then*

$$C_{PolMult}(E_d,E_d)=\mathcal{O}(d\log(d)).$$

Theorem B.1.2 *([3], p. 64) Let $E_{\mathbf{d}}=\{(\alpha_1,\ldots,\alpha_n)\ ;\ 0\le\alpha_i\le d_i-1\}$. Then we have*

$$C_{PolMult}(E_{\mathbf{d}},E_{\mathbf{d}})=\mathcal{O}(M\log(M)),$$

where $d=max(d_1,\ldots,d_n)$, $M=c^n$, and $c=2\,d+1$.

Theorem B.1.3 *[6], [12]. Let $E_{d,n}$ be a fixed set of exponents having total degree at most d in n variables and let S be a given set containing the set $E_{d,n}+E_{d,n}$. Then*

$$C_{PolMult}(E_{d,n},E_{d,n})=\mathcal{O}(C_{PolMult}(E_T,E_T)\log T),$$

where $T=\binom{n+d}{n}$ is the maximum possible number of distinct monomials of degree at most d in n variables. Furthermore, $C_{PolMult}(E_{d,n},E_{d,n})=\mathcal{O}(s(\log^2(s)+n)+n\sqrt{d})$ where s is the cardinality of the set S.

Remark Theorems B.1.1 and B.1.2 can be extended to the computations over any ring of constants (rather than over the complex field) at the expense of increasing their complexity bounds by factors at most $\log\log(d)$ and $\log\log(c)$, respectively. Theorem B.1.3 applies over any field of constants having characteristic 0.

B.2 Tellegen's theorem on duality of multiplication of a matrix and its transpose by a vector

Theorem B.2.1 *[29]. Let W be a square $n \times n$ matrix with no zero rows or columns. Let C_W ops suffice to compute the product $W\mathbf{v}$ for a vector $\mathbf{v}$. Then C_W ops also suffice to compute the product $W^t\mathbf{v} = (\mathbf{v}^t W)^t$.*

B.3 Solving a linear system of equations and the computation of the rank and the signature of a matrix

Application of the conjugate gradient algorithm ([16], sect. 10.2) gives us the following result:

Theorem B.3.1 *Let W be a nonsingular $N \times N$ matrix. Performing $2N$ multiplications of W and W^t by vectors and $\mathcal{O}(N^2)$ other arithmetic operations suffice to compute the solution $\mathbf{v}$ to a linear system $W\mathbf{v} = \mathbf{w}$.*

In section 3.4, we needed an algorithm for computing the rank and the signature of an $N \times N$ real symmetric (and quasi-Hankel) matrix W. Such an algorithm may start with tridiagonalizing the matrix. In exact arithmetic, this can be done by means of the Lanczos algorithm, whose randomized computational cost is dominated by the cost of performing $\mathcal{O}(N)$ multiplications of W by vectors and $\mathcal{O}(N^2)$ other ops (cf. [3], p.118, or [16], ch. 9]). Then both the rank and the signature of W can be computed at the cost $\mathcal{O}(N)$, by using the three-term recurrence relations for the characteristic polynomials of the leading principal submatrices of W (cf. [16], p. 439). We arrive at the following result.

Theorem B.3.2 *Let W be an N-by-N real symmetric matrix. Then application of Lanczos randomized algorithm (which uses $\mathcal{O}(N)$ multiplications of W by vectors and $\mathcal{O}(N^2)$ other ops) and performing $O(N)$ additional ops suffice to compute the rank and the signature of W.*

If W is a structured (resp. and real symmetric) matrix, whose multiplication by a vector is expressed in terms of polynomial multiplication, one may combine theorems B.1.1-B.1.3 and B.3.2 in order to express the arithmetic cost of the solution of the linear system $W\mathbf{v} = \mathbf{w}$ and the randomized arithmetic cost of computing the rank (resp. and signature) of W in terms of the size of W.