

# Multidimensional structured matrices and polynomial systems

Bernard Mourrain INRIA, SAFIR  
2004 route des Lucioles  
B.P. 93, 06902 Sophia Antipolis  
mourrain@sophia.inria.fr

&

Victor Y. Pan  
Mathematics and Computer Science Department  
Lehman College, City University of New York  
Bronx, NY 10468, USA  
VPAN@LCVAX.LEHMAN.CUNY.EDU

(Supported by NSF Grant CCR 9625344 and PSC CUNY Award No. 667340)

February 12, 1997

## Abstract

We apply and extend some well-known and some recent techniques from algebraic residue theory in order to relate to each other two major subjects of algebraic and numerical computing, that is, computations with structured matrices and solving a system of polynomial equations. In the first part of our paper, we extend the Toeplitz and Hankel structures of matrices and some of their known properties to some new classes of structured (quasi-Hankel and quasi-Toeplitz) matrices, naturally associated to systems of multivariate polynomial equations. In the second part of the paper, we prove some relations between these structured matrices, which extend the classical relations of the univariate case.

## 1 Introduction

We apply and extend some well-known and some recent techniques from algebraic residue theory in order to relate to each other two major subjects of algebraic and numerical computing, that is, the computations with structured matrices and solving a system of polynomial equations. We also reveal some hidden correlations between these two subjects via the study of the associated operators of multivariate displacement. The latter operators naturally extend the

univariate displacement operators, which define Toeplitz and/or Hankel structure of matrices (cf. [1]). In our multivariate case, we generalize such a matrix structure and arrive at the new classes of operators and structured matrices, which include operators and matrices associated to the polynomial systems of equations and which we call quasi-Hankel and quasi-Toeplitz operators and matrices since some well-known properties of Toeplitz and Hankel operators and matrices can be extended to them (see section 2). Due to high importance of computations with structured matrices (see e.g. [1]), our study of these matrix classes may be of independent technical interest. In section 3, we recall some basic definitions and facts about algebraic residues, in order to extend classical relations between structured matrices to the multivariate case (section 4).

Next, we will state some definitions.  $R = \mathbb{C}[x_1, \dots, x_n]$  will denote the polynomial ring in variables  $x_1, \dots, x_n$  over the complex field  $\mathbb{C}$ , and  $L = \mathbb{C}[x_1^{\pm 1}, \dots, x_n^{\pm 1}]$  will denote the ring of Laurent's polynomials in the same variables. We will write  $\mathbf{x} = (x_1, \dots, x_n)$  and  $\mathbf{x}^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ . For a vector  $\alpha = (\alpha_1, \dots, \alpha_n)$ , we will write  $|\alpha|$  to denote the 1-norm of this vector,  $|\alpha| = \sum_{i=1}^n |\alpha_i|$ . The total degree of a monomial  $c \mathbf{x}^\alpha$ , with a coefficient  $c$ , is  $|\alpha|$ . The total degree of a polynomial  $\sum_{\alpha} c_{\alpha} \mathbf{x}^{\alpha}$ , with coefficients  $c_{\alpha} \neq 0$ , is the highest total degree of its monomials. We will write  $|S|$  to denote the cardinality of a set  $S$ . *ops* will stand for "arithmetic operations".  $\mathbf{e}_i$  will denote the  $i$ -th unit coordinate vector in  $\mathbb{C}^n$ .

Our study can be immediately extended from the complex field  $\mathbb{C}$  to the case of any number field of constants having characteristic 0. Furthermore, with the exception of the results based on the interpolation techniques of [3] (cf. proposition 2.6), our study can be extended to the case of any field of constants.

## 2 Structured Matrices

In this section, we propose a generalization of the structure of Toeplitz, Hankel and Van der Monde matrices to the case of matrices associated with multivariate polynomials having rows and columns indexed by monomials.

### 2.1 Quasi-Hankel and quasi-Toeplitz matrices, operators, and the associated generating polynomials: definitions and a correlation

DEFINITION 2.1 *Let  $E$  and  $F$  be two subsets of  $\mathbb{Z}^n$  and let  $M = (m_{\alpha, \beta})_{\alpha \in E, \beta \in F}$  be a matrix whose rows and columns are indexed by the elements of  $E$  and  $F$  respectively.*

- *$M$  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  $\alpha + \beta$ . For every  $i = 1, \dots, n$ , we have  $m_{\alpha - \mathbf{e}_i, \beta + \mathbf{e}_i} = m_{\alpha, \beta}$ , provided that  $\alpha, \alpha - \mathbf{e}_i \in E; \beta, \beta + \mathbf{e}_i \in F$ ;*

such a matrix  $M$  is associated with the Laurent polynomial  $H_M(\mathbf{x}) = \sum_{u \in E-F} h_u \mathbf{x}^{-u}$ .

- $M$  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  $\alpha - \beta$ . For every  $i = 1, \dots, n$ , we have  $m_{\alpha + \mathbf{e}_i, \beta + \mathbf{e}_i} = m_{\alpha, \beta}$ , provided that  $\alpha, \alpha + \mathbf{e}_i \in E; \beta, \beta + \mathbf{e}_i \in F$ ; such a matrix  $M$  is associated with the polynomial  $T_M(\mathbf{x}) = \sum_{u \in E+F} t_u \mathbf{x}^u$ .

For  $E = [0, \dots, m-1]$  and  $F = [0, \dots, n-1]$ , definition 2.1 turns into the usual definition of Hankel (resp. Toeplitz) matrices [1].

DEFINITION 2.2 Let  $\mathcal{P}_E : L \rightarrow L$  be the projection map such that

$$\mathcal{P}_E(\mathbf{x}^\alpha) = \begin{cases} \mathbf{x}^\alpha & \text{if } \alpha \in E \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

Let

$$\rho_E = Id - \mathcal{P}_E,$$

where  $Id$  denotes the identity operator,  $Id(e) = e$  for all  $e$ . For any element  $Q$  of  $L$ , let  $\mu_Q : L \rightarrow L$  denote the operator of multiplication by  $Q$ . For any matrix  $M = (m_{\alpha, \beta})_{\alpha \in E, \beta \in F}$ , let  $\mathcal{M}$  denote the linear map  $L \rightarrow L$  such that

$$\mathcal{M}(\mathbf{x}^\beta) = \begin{cases} \sum_{\alpha \in E} m_{\alpha, \beta} \mathbf{x}^{-\alpha} & \text{if } \beta \in F \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

The matrix of this linear operator coincides with the matrix  $M$  on  $(\mathbf{x}^{-\alpha}) \times (\mathbf{x}^\beta)$ , for  $\alpha \in E, \beta \in F$ , and is null elsewhere. We will call this operator an  $(E, F)$  quasi-Hankel (resp. an  $(E, F)$  quasi-Toeplitz) operator if the matrix  $M$  is an  $(E, F)$  quasi-Hankel (resp. an  $(E, F)$  quasi-Toeplitz) matrix.

PROPOSITION 2.3 If  $M$  is an  $(E, F)$  quasi-Hankel (resp. an  $(E, F)$  quasi-Toeplitz) matrix, then

$$\mathcal{M} = \mathcal{P}_{-E} \circ \mu_{H_M} \circ \mathcal{P}_F = (\text{resp. } \mathcal{M} = \mathcal{P}_E \circ \mu_{T_M} \circ \mathcal{P}_F). \quad (3)$$

To the end of this section, we will assume that both sets  $E$  and  $F$  contain 0.

## 2.2 Examples

**Quasi-Toeplitz matrices:** Let us be given some multivariate polynomials  $P_0, \dots, P_n$  in the variables  $x_1, \dots, x_n$  and let us consider the matrix associated with the linear map

$$\begin{aligned} \phi : V_0 \times \dots \times V_n &\rightarrow V, \\ (Q_0, \dots, Q_n) &\mapsto \sum_{i=0}^n P_i Q_i. \end{aligned}$$

Here  $V_i$  is the vector space generated by the monomials  $\mathbf{x}^\beta$  for  $\beta \in F_i$ , which is the set of all monomials of the polynomial  $Q_i, i = 0, \dots, n$ , and  $V$  is generated by the monomials  $\mathbf{x}^\alpha$  for  $\alpha \in E$ , where  $E$  denotes the set of the exponents of the monomials of  $(P_i Q_i)_{i=0, \dots, n}$ . Such maps typically appear in the construction of resultant type matrices associated to the system  $\{P_i = 0, i = 0, \dots, n\}$ , of polynomial equations (see [7], [6] or [2] for an effective algorithm).

Let  $M$  denote the matrix of this linear map in the monomial basis of  $V_0 \times \dots \times V_n$  and  $V$ . The rows of this matrix are indexed by the elements of the set  $E$  and the columns by the elements of the set  $F_0 \sqcup \dots \sqcup F_n$ . Let  $x_0$  be a new variable and let us view  $\mathbb{Z}^n$  as the subset of  $\mathbb{Z}^{n+1}$  consisting of elements of the form  $(0, a_1, \dots, a_n)$ . Let  $\mathbf{e}_0$  denote the first canonical vector of  $\mathbb{Z}^{n+1}$ . Then, the elements of the subset

$$F = \{i \mathbf{e}_0 + \beta ; 0 \leq i \leq n, \beta \in F_i\}$$

of  $\mathbb{Z}^{n+1}$  index the columns of  $M$ .

Note that the  $(\alpha, i \mathbf{e}_0 + \beta)$ -th entry of the matrix  $M$  is the coefficient of  $\mathbf{x}^\alpha$  in  $\mathbf{x}^\beta P_i$ . It is also the coefficient of  $\mathbf{x}^{\alpha - \beta}$  in  $P_i$ . Therefore, it depends only on  $\alpha - \beta - i \mathbf{e}_0$ .

REMARK 1 — *Resultant type matrices and their transposes are quasi-Toeplitz matrices.*

REMARK 2 — *The Laurent polynomial associated to a quasi-Toeplitz operator  $\phi$  is just*

$$T_\phi = \sum_{i=0}^n x_0^{-i} P_i.$$

An example of such a structure and an application to the computation of sparse resultants, can also be found in this volume (article of I.Z. Emiris and V.Y Pan).

**Quasi-Hankel matrices:** Let  $\lambda \in \hat{L}$  be a linear form on  $L$  and consider the matrix

$$[\lambda(\mathbf{x}^{\alpha+\beta})]_{\alpha \in E, \beta \in F}.$$

This is an  $(E, F)$  quasi-Hankel matrix. As we will see in the next section, such matrices appear in algebraic residue theory. If  $B$  is a Gorenstein algebra (for instance, if  $B$  is a complete intersection) and has a finite dimension over  $\mathbb{C}$ , then any non-degenerating bilinear form  $q$  can be represented as  $(a, b) \mapsto q(a, b) = \lambda(ab)$  where  $\lambda \in \hat{L}$  is a linear form (see [5]). Furthermore, any Gramm-Schmidt matrix,  $(q(\mathbf{u}_i, \mathbf{u}_j))$ , where  $(\mathbf{u}_i)$  is a basis of  $B$ , is conjugated to a quasi-Hankel matrix.

### 2.3 Multiplication of quasi-Hankel and quasi-Toeplitz matrices by vectors.

Multiplication of an  $(E, F)$  quasi-Hankel matrix by a vector  $\mathbf{v} = [v_\beta] \in \mathbb{C}^F$  can be reduced to (Laurent) polynomial multiplication in the following way. Let  $M = (m_{\alpha, \beta})_{\alpha \in E, \beta \in F}$  denote an  $(E, F)$  quasi-Hankel matrix, let  $H_M(\mathbf{x}) = \sum_{u \in E-F} h_u \mathbf{x}^{-u}$  denote the associated Laurent polynomial, and let  $V(\mathbf{x}) = \sum_{\beta \in F} v_\beta \mathbf{x}^\beta$ . Then we have

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

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

$$\sum_{\beta \in F} h_{\alpha+\beta} v_\beta = \sum_{\beta \in F} m_{\alpha, \beta} v_\beta,$$

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

A similar argument reduces multiplication of an  $(E, F)$  quasi-Toeplitz matrix by a vector to multiplication of a pair of Laurent's polynomials.

The stated reductions enable us to deduce the following result:

**PROPOSITION 2.4** *An  $(E, F)$  quasi-Hankel (resp. an  $(E, F)$  quasi-Toeplitz) matrix  $M$  can be multiplied by a vector in  $\mathcal{O}(N \log^2 N + C_{M,N})$  ops, where  $N = \lfloor E - 2F \rfloor$  (resp.  $\lfloor E + 2F \rfloor$ ) and where  $C_{M,N}$  bounds the cost of evaluating the polynomial  $H_M$  (resp.  $T_M$ ) on a fixed set of  $N$  points.*

*Proof.* See [8]. □

In some special cases, we have better complexity estimates.

**PROPOSITION 2.5** *In the case where  $E = F = \{(\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n ; 0 \leq \alpha_i \leq d_i - 1\}$ , an  $(E, F)$  quasi-Hankel (resp. an  $(E, F)$  quasi-Toeplitz) matrix can be multiplied by a vector in  $\mathcal{O}(3^n D \log(3^n D))$  ops, where*

$$D = \prod_i^n d_i. \tag{4}$$

*Proof.* See [8]. □

**PROPOSITION 2.6** *In the case where  $E = \{\alpha \in \mathbb{N}^n, |\alpha| \leq k\}$ ,  $F = \{\beta \in \mathbb{N}^n, |\beta| \leq l\}$  and where the computations are over a field of constants containing the field of rational numbers, an  $(E, F)$  quasi-Toeplitz matrix can be multiplied by a vector in  $\mathcal{O}(\sigma \log^3 \sigma)$  ops, where*

$$\sigma = \sigma_{2l+k, n} = \binom{2l+k+n}{n} = \mathcal{O}\left(\left(e \frac{2l+k+n}{n}\right)^n / \sqrt{n}\right), e = 2.718\dots$$

*(The latter equation is implied by Stirling's formula.)*

*Proof.* See [8]. □

## 2.4 Multivariate displacement operators and ranks: definition

By convention, if  $\mathcal{A} : L \rightarrow L$  is a linear operator,  $A$  will denote its matrix in a (sub)basis of  $L$ . These operators may have infinite dimension, but in our case, we will only study the finite dimension case. The rank of the operator  $\mathcal{A}$  is the rank of the matrix  $A$ . For all  $\alpha, \beta \in \mathbb{Z}^n$ ,  $[\mathcal{A}]_{\alpha, \beta} = A_{\alpha, \beta}$  is the coefficient of  $\mathbf{x}^\alpha$  in  $\mathcal{A}(\mathbf{x}^\beta)$ .

**DEFINITION 2.7** *For every subset  $E$  of  $\mathbb{Z}^n$  and for every  $i, i = 1, \dots, n$ , we define the two following unit  $E$ -displacement matrices (operators):*

$$Z_i^E = \mathcal{P}_E \mu_{x_i} \mathcal{P}_E \tag{5}$$

and

$$Z_{-i}^E = \mathcal{P}_E \mu_{x_i^{-1}} \mathcal{P}_E. \tag{6}$$

In particular, for  $E = [0, \dots, n-1]$  and  $i = 1$ , we arrive at the well-known displacement matrix

$$Z_1^E = \begin{pmatrix} 0 & \dots & \dots & 0 \\ 1 & \ddots & & \vdots \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}$$

and its transpose,  $Z_{-1}^E$  (cf. [1]).

DEFINITION 2.8 Let  $E$  and  $F$  denote two subsets of  $\mathbb{Z}^n$  and let  $\mathcal{A}$  denote a linear operator  $L \rightarrow L$ . Then, the operators

$$\mathcal{H}_i^+(\mathcal{A}) = \mathcal{A} - \mathcal{Z}_{-i}^{-E} \mathcal{A} \mathcal{Z}_i^F, \quad \mathcal{H}_i^-(\mathcal{A}) = \mathcal{A} - \mathcal{Z}_i^{-E} \mathcal{A} \mathcal{Z}_{-i}^F, \quad (7)$$

$$\mathcal{T}_i^+(\mathcal{A}) = \mathcal{A} - \mathcal{Z}_{-i}^E \mathcal{A} \mathcal{Z}_i^F, \quad \mathcal{T}_i^-(\mathcal{A}) = \mathcal{A} - \mathcal{Z}_i^E \mathcal{A} \mathcal{Z}_{-i}^F \quad (8)$$

will be called the  $(-, +, -E, F, i)$ ,  $(+, -, -E, F, i)$ ,  $(-, +, E, F, i)$ , and  $(+, -, E, F, i)$  displacements of  $\mathcal{A}$ , respectively. The ranks of these displacements will be called the  $(-, +, -E, F, i)$ ,  $(+, -, -E, F, i)$ ,  $(-, +, E, F, i)$ , and  $(+, -, E, F, i)$  displacement ranks of  $\mathcal{A}$ , resp., and will be denoted  $r_{-,+,-E,F,i}(\mathcal{A})$ ,  $r_{+,-,-E,F,i}(\mathcal{A})$ ,  $r_{-,+,E,F,i}(\mathcal{A})$ , and  $r_{+,-,E,F,i}(\mathcal{A})$ , resp. The operators transforming  $\mathcal{A}$  into the above displacements will be called the  $(-, +, -E, F, i)$ ,  $(+, -, -E, F, i)$ ,  $(-, +, E, F, i)$ , and  $(+, -, E, F, i)$  displacement operators, resp.

## 2.5 Bounds on displacement ranks of quasi-Hankel and quasi-Toeplitz matrices

DEFINITION 2.9 Hereafter, we write

$$\delta_i(E) = \{\alpha : \alpha \in E; \alpha + \mathbf{e}_i \notin E\}$$

(resp.  $\delta_{-i}(E) = \{\alpha : \alpha \in E; \alpha - \mathbf{e}_i \notin E\}$ ).

PROPOSITION 2.10 For an  $(E, F)$  quasi-Hankel operator  $\mathcal{M}$ , we have the following bounds on its  $(-, +, E, F, i)$  and  $(+, -, E, F, i)$  displacement ranks:

$$r_{-,+,-E,F,i}(\mathcal{M}) \leq \lfloor \delta_i(-E) \rfloor + \lfloor \delta_i(F) \rfloor,$$

$$r_{+,-,-E,F,i}(\mathcal{M}) \leq \lfloor \delta_{-i}(-E) \rfloor + \lfloor \delta_{-i}(F) \rfloor.$$

For an  $(E, F)$  quasi-Toeplitz operator  $\mathcal{M}$ , we have the following bounds on its  $(-, +, E, F, i)$  and  $(+, -, E, F, i)$  displacement ranks:

$$r_{-,+,E,F,i}(\mathcal{M}) \leq \lfloor \delta_{-i}(E) \rfloor + \lfloor \delta_i(F) \rfloor,$$

$$r_{+,-,E,F,i}(\mathcal{M}) \leq \lfloor \delta_i(E) \rfloor + \lfloor \delta_{-i}(F) \rfloor.$$

*Proof.* See [8]. □

In the particular case, where  $E = F = \{(\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n; 0 \leq \alpha_i \leq d_i - 1\}$ , the displacement rank of  $\mathcal{H}_i^+(\mathcal{M})$  is bounded by  $2 \frac{D}{d_i} = 2 \prod_{j \neq i} d_j$ .

## 2.6 Generalized Van der Monde matrices: definition

DEFINITION 2.11 For two sets of exponents  $E = \{\alpha_1, \dots, \alpha_D\}$  and points  $\mathcal{Z} = \{\zeta_1, \dots, \zeta_D\}$ , let the matrix

$$V_E(\mathcal{Z}) = (\zeta_j^{\alpha_i})_{1 \leq i, j \leq D}$$

be called the generalized Van der Monde matrix of basis  $\mathcal{Z}$  and exponents  $E$ .

## 3 Algebraic residues: definitions and basic facts

In this section, we will recall some basic definitions from algebraic residue theory, referring the reader to [4], [5] for further details. Let  $R = \mathbb{C}[\mathbf{x}] = \mathbb{C}[x_1, \dots, x_n]$  be the algebra of polynomials in  $x_i$  over the field  $\mathbb{C}$ . In addition to the vector (set) of variables  $\mathbf{x}$ , we consider the vector  $\mathbf{y} = (y_1, \dots, y_n)$  and write  $\mathbf{x}^{(0)} = \mathbf{x}$ ,  $\mathbf{x}^{(1)} = (y_1, x_2, \dots, x_n)$ ,  $\dots$ ,  $\mathbf{x}^{(n)} = \mathbf{y}$ . We define  $\theta_i(P) = \frac{P(\mathbf{x}^{(i)}) - P(\mathbf{x}^{(i-1)})}{y_i - x_i}$ , the *discrete differentiation* of  $P$ . For any sequence of  $n+1$  polynomials  $P_0, \dots, P_n \in R$ , let us construct the following polynomial in  $\mathbf{x}$  and  $\mathbf{y}$ :

$$\Theta_{\mathbf{P}}(P_0) = \Theta(P_0, P_1, \dots, P_n) = \begin{vmatrix} P_0(\mathbf{x}) & \theta_1(P_0) & \cdots & \theta_n(P_0) \\ \vdots & \vdots & & \vdots \\ P_n(\mathbf{x}) & \theta_1(P_n) & \cdots & \theta_n(P_n) \end{vmatrix}, \quad (9)$$

where  $|A|$  denotes the determinant of  $A$ , and  $\mathbf{P} = (P_1, \dots, P_n)$ . Let us write  $\Theta_{\mathbf{P}} = \Theta_{\mathbf{P}}(1) = \Theta(1, P_1, \dots, P_n) \in \mathbb{C}[\mathbf{x}, \mathbf{y}]$ . Now, we define the residue of  $\mathbf{P} = (P_1, \dots, P_n)$  as a unique linear form  $\tau$  in the set of linear forms on  $R$  such that

1.  $\tau$  vanishes on  $(\mathbf{P})$ ,
2.  $\Theta_{\mathbf{P}}(\tau) - 1 \in (\mathbf{P})$ ,

where  $\Theta_{\mathbf{P}}(\tau)$  denotes the polynomial in  $\mathbf{x}$ , that we obtain when we apply  $\lambda$  on the monomials in  $\mathbf{y}$  of  $\Theta_{\mathbf{P}}$ . Hereafter,  $I$  will denote the ideal generated by the polynomials  $P_1, \dots, P_n$ ;  $B = R/I$  will denote the quotient ring defined in  $R$  by  $I$ , and  $\equiv$  will denote an equality in  $B$ .

If  $(\mathbf{x}^\alpha)_{\alpha \in E}$  is a basis of  $B$ , let  $(\mathbf{w}_\alpha)$  be the dual basis of  $(\mathbf{x}^\alpha)$  for  $\tau$ :

$$\tau(\mathbf{x}^\alpha \mathbf{w}_\beta) = \delta_{\alpha, \beta},$$

$\delta_{\alpha, \beta}$  is 1 if  $\alpha = \beta$  and 0 otherwise. Let us write  $\mathbf{w}_\alpha(\mathbf{x}) = \sum_{\beta \in E} w_{\beta, \alpha} \mathbf{x}^\beta$  and  $\Omega_1 = (w_{\alpha, \beta})_{\alpha, \beta \in E}$ .

Then we have the following property ([4], [5]):

$$\Theta_{\mathbf{P}} \equiv \sum_{\alpha \in E} \mathbf{x}^\alpha \mathbf{w}_\alpha(\mathbf{y}) \equiv \sum_{\alpha \in E} \mathbf{w}_\alpha(\mathbf{x}) \mathbf{y}^\alpha \pmod{(\mathbf{P}(\mathbf{x}), \mathbf{P}(\mathbf{y}))}. \quad (10)$$



Thus, for any  $b \in B$ , we have the relations

$$b \equiv \sum_{\alpha \in E} \tau(b \mathbf{x}^\alpha) \mathbf{w}_\alpha \equiv \sum_{\alpha \in E} \tau(b \mathbf{w}_\alpha) \mathbf{x}^\alpha. \quad (11)$$

For any  $b$  in  $B$ , we let

$$\Theta_{\mathbf{P}}(b) = \Theta_{\mathbf{P}}(b)(\mathbf{x}, \mathbf{y}) = \sum_{\alpha, \beta \in E} u_{\alpha, \beta}^{(b)} \mathbf{x}^\alpha \mathbf{y}^\beta \quad (12)$$

denote the decomposition of  $\Theta_{\mathbf{P}}(b)$  in the basis  $(\mathbf{x}^\alpha)$  of  $B$ .

Moreover, we recall from [4], [5], that for any polynomial  $Q \in R$ , we have

$$\Theta_{\mathbf{P}}(1)(\mathbf{x}, \mathbf{y}) Q(\mathbf{x}) \equiv \Theta_{\mathbf{P}}(1)(\mathbf{x}, \mathbf{y}) Q(\mathbf{y}) \pmod{(\mathbf{P}(\mathbf{x}), \mathbf{P}(\mathbf{y}))}. \quad (13)$$

In particular, we set  $Q(\mathbf{x}) = x_i$   $i = 1, \dots, n$ , and for any pair  $\zeta$  and  $\eta$  of distinct roots of the polynomial system  $\mathbf{P} = \mathbf{0}$ , we set  $\mathbf{x} = \zeta$ ,  $\mathbf{y} = \eta$  and deduce that

$$\Theta_{\mathbf{P}}(\zeta, \eta) = 0. \quad (14)$$

## 4 Relations among structured matrices

For any  $b$  in  $B$ , let  $\mu_b$  denote the operator of multiplication by  $b$  in  $B$  and let  $M_b = (m_{\alpha, \beta}^{(b)})_{\alpha, \beta \in E}$  denote its matrix in the basis  $(\mathbf{x}^\alpha)$ . Let  $N_b = (n_{\alpha, \beta}^{(b)})_{\alpha, \beta \in E}$  denote its matrix in the basis  $(\mathbf{w}_\alpha)$ .

**PROPOSITION 4.1** *The matrix  $N_b$  of multiplication by  $b$  in  $B$ , in the basis  $(\mathbf{w}_\alpha)$ , is the transpose  $M_b^t$  of the matrix  $M_b$  of multiplication by  $b$  in  $B$ , in the basis  $(\mathbf{x}^\alpha)$ .*

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

$$b \mathbf{x}^\alpha \equiv \sum_{\gamma \in E} m_{\gamma, \alpha}^{(b)} \mathbf{x}^\gamma, \quad b \mathbf{w}_\alpha \equiv \sum_{\gamma \in E} n_{\gamma, \alpha}^{(b)} \mathbf{w}_\gamma,$$

and

$$\begin{aligned} m_{\alpha, \beta}^{(b)} &= \langle b \mathbf{x}^\beta | \mathbf{w}_\alpha \rangle = \tau(b \mathbf{x}^\beta \mathbf{w}_\alpha), \\ n_{\alpha, \beta}^{(b)} &= \langle b \mathbf{w}_\beta | \mathbf{x}^\alpha \rangle = \tau(b \mathbf{x}^\alpha \mathbf{w}_\beta) = m_{\beta, \alpha}^{(b)}. \end{aligned}$$

Therefore,  $N_b = M_b^t$ .  $\square$

Let  $(\widehat{\mathbf{x}}^\alpha)$  (resp.  $(\widehat{\mathbf{w}}_\alpha)$ ) be the dual basis of  $(\mathbf{x}^\alpha)$  (resp.  $(\mathbf{w}_\alpha)$ ). Let  $\mu_b^*$  denotes the multiplication by  $b$  in  $\widehat{B}$ :  $\mu_b^*(\lambda) = b \cdot \lambda$ . The map  $\mu_b^*$  is the transpose of the map  $\mu_b$ , and its matrix in the basis  $(\widehat{\mathbf{x}}^\alpha)$  is  $M_b^t$ . Its matrix in the basis  $(\widehat{\mathbf{w}}_\alpha)$  is  $M_b$ .

DEFINITION 4.2 For any  $b$  in  $B$ , let  $\Xi_b$  denote the quasi-Hankel matrix of residues

$$\Xi_b = (\tau(b \mathbf{x}^\alpha \mathbf{x}^\beta))_{\alpha, \beta \in E}.$$

Let  $\Omega_b$  denote the matrix of residues

$$\Omega_b = (\tau(b \mathbf{w}_\alpha \mathbf{w}_\beta))_{\alpha, \beta \in E}.$$

For any  $b \in B$ , let  $\psi_b : \widehat{B} \rightarrow B$  be the following map :

$$\lambda \mapsto \psi_b(\lambda) = \sum_{\alpha, \beta \in E} u_{\alpha, \beta}^{(b)} \lambda(\mathbf{x}^\beta) \mathbf{x}^\alpha,$$

where the coefficients  $u_{\alpha, \beta}^{(b)}$  are defined in (12). This map is independent of the decomposition.

PROPOSITION 4.3 The matrix of the map  $\psi_b$

1. from the basis  $(\widehat{\mathbf{x}}^\alpha)$  of  $\widehat{B}$  to the basis  $(\mathbf{x}^\alpha)$  of  $B$  is  $\Omega_b = (\tau(b \mathbf{w}_\alpha \mathbf{w}_\beta))$ ,
2. from the basis  $(\widehat{\mathbf{w}}_\alpha)$  to the basis  $(\mathbf{w}_\alpha)$  is  $\Xi_b = (\tau(b \mathbf{x}^\alpha \mathbf{x}^\beta))$ .

*Proof.* The image of  $\widehat{\mathbf{x}}^\beta$  by this map is

$$\psi_b(\widehat{\mathbf{x}}^\beta) = \sum_{\alpha} u_{\alpha, \beta}^{(b)} \mathbf{x}^\alpha,$$

whose coordinates in the basis  $(\mathbf{x}^\alpha)_{\alpha \in E}$  are  $[u_{\alpha, \beta}^{(b)}]_{\alpha \in E}$ . Therefore, the matrix of the map  $\psi_b$  from  $(\widehat{\mathbf{x}}^\alpha)$  to  $(\mathbf{x}^\alpha)$  is the matrix  $(u_{\alpha, \beta}^{(b)})$ . The coefficients  $u_{\alpha, \beta}^{(b)}$  can be computed by the formula

$$u_{\alpha, \beta}^{(b)} = \tau(\psi_b(\widehat{\mathbf{x}}^\beta) \mathbf{w}_\alpha).$$

By using the identity (13), we obtain  $\Theta_{\mathbf{P}}(b) \equiv b(\mathbf{x})\Theta_{\mathbf{P}}(1)$ , so that  $\forall \lambda \in \widehat{B}$ ,  $\psi_b(\lambda) = b \psi_1(\lambda)$  and

$$\tau(\psi_b(\widehat{\mathbf{x}}^\beta) \mathbf{w}_\alpha) = \tau(b \psi_1(\widehat{\mathbf{x}}^\beta) \mathbf{w}_\alpha).$$

According to the identity (10), we have

$$\psi_1(\widehat{\mathbf{x}}^\beta) = \mathbf{w}_\beta, \tag{15}$$

which implies that

$$u_{\alpha, \beta}^{(b)} = \tau(b \mathbf{w}_\alpha \mathbf{w}_\beta).$$

This proves part 1 of proposition.

Now, express  $\Theta_{\mathbf{P}}(b)$  in the basis  $(\mathbf{w}_\alpha)$ :

$$\Theta_{\mathbf{P}}(b) = \sum_{\alpha, \beta \in E} v_{\alpha, \beta}^{(b)} \mathbf{w}_\alpha(\mathbf{x}) \mathbf{w}_\beta(\mathbf{y}).$$

By interchanging the roles of  $(\mathbf{x}^\alpha)$  and  $(\mathbf{w}_\alpha)$ , we prove part 2 of the proposition, that is, we prove that the matrix of the map  $\psi_b$  from  $(\widehat{\mathbf{w}}_\alpha)$  to  $(\mathbf{w}_\alpha)$  is  $(v_{\alpha, \beta}^{(b)})$ , where

$$v_{\alpha, \beta}^{(b)} = \tau(b \mathbf{x}^\alpha \mathbf{x}^\beta).$$

□

Note that the matrices  $(u_{\alpha, \beta}^{(b)})$ ,  $(v_{\alpha, \beta}^{(b)})$  also correspond to the matrices of the map

$$\begin{aligned} \phi_b : \widehat{B} \times \widehat{B} &\rightarrow \mathbf{C} \\ (\lambda, \mu) &\mapsto \sum_{\alpha, \beta \in E} u_{\alpha, \beta}^{(b)} \lambda(\mathbf{x}^\beta) \mu(\mathbf{x}^\alpha) \end{aligned}$$

in the bases  $(\widehat{\mathbf{x}}^\alpha) \times (\widehat{\mathbf{x}}^\alpha)$  and  $(\widehat{\mathbf{w}}^\alpha) \times (\widehat{\mathbf{w}}^\alpha)$  of  $\widehat{B} \times \widehat{B}$ , respectively.

The matrix  $\Omega_1$  is just the matrix of the coefficients of  $(\mathbf{w}_\alpha)$  in the basis  $(\mathbf{x}^\alpha)$  of  $B$ . We relate it to the quasi-Hankel matrix of residues  $\Xi_1$ :

**THEOREM 4.4** *The inverse of  $\Xi_1$  is  $\Omega_1$ .*

*Proof.* According to (10), we have the equations

$$\tau(\mathbf{x}^\alpha \mathbf{w}_\beta) = \sum_{\beta \in E} \tau(\mathbf{x}^\alpha \mathbf{x}^\beta) \mathbf{w}_{\beta, \alpha} = \delta_{\alpha, \beta},$$

where  $\delta_{\alpha, \beta}$  equals 1 if  $\alpha = \beta$  and is 0 otherwise. In term of matrices, this is equivalent to the equation

$$\Xi_1 \Omega_1 = I,$$

where  $\mathbb{I}_D$  is the identity matrix, of size  $D$ . □

**THEOREM 4.5** *For any  $b \in B$ , we have*

$$\Omega_b = \Omega_1 M_b^t = M_b \Omega_1, \tag{16}$$

$$\Xi_b = \Xi_1 M_b = M_b^t \Xi_1. \tag{17}$$

*Proof.* According to (13),  $\forall \lambda \in \widehat{B}$ ,  $\psi_b(\lambda) = b \psi_1(\lambda) = \psi(b \cdot \lambda)$ . Thus we can decompose the map  $\psi_b$  as

$$\psi_b = \mu_b \circ \psi_1 = \psi_1 \circ \mu_b^*.$$

If we choose  $(\widehat{\mathbf{x}}^\alpha)$  and  $(\mathbf{x}^\alpha)$  respectively, as the bases of  $\widehat{B}$  and  $B$ ,  $(\widehat{\mathbf{x}}^\alpha)$  and  $(\mathbf{x}^\alpha)$  we obtain the relations (16). If we choose the bases  $(\widehat{\mathbf{w}}_\alpha)$  and  $(\mathbf{w}_\alpha)$ , we obtain the relations (17).  $\square$

This also proves that the matrix of the map  $\psi_b$

1. from the basis  $(\widehat{\mathbf{w}}_\alpha)$  to the basis  $(\mathbf{x}^\alpha)$  is  $M_b^t$ ,
2. from the basis  $(\widehat{\mathbf{x}}^\alpha)$  to the basis  $(\mathbf{w}_\alpha)$  is  $M_b$ .

Indeed, to go from the basis  $(\mathbf{w}_\alpha)$  to the basis  $(\mathbf{x}^\alpha)$ , we multiply by the matrix  $\Omega_1$ .

PROPOSITION 4.6 *If  $ab \equiv 1$  in  $B$ , then*

$$\Omega_a \Xi_b = \Omega_b \Xi_a = \mathbb{I}_D.$$

*Proof.* According to (16), we have

$$\Omega_a \Xi_b = \Omega_1 M_a^t M_b^t \Xi_1 = \Omega_1 \Xi_1 = \mathbb{I}_D,$$

for  $M_a M_b = M_{ab} = \mathbb{I}_D$ . Similarly, we also have  $\Omega_b \Xi_a = \mathbb{I}_D$ .  $\square$

Let  $\mathcal{Z} = \{\zeta_1, \dots, \zeta_D\}$  denote the set of all common roots of the  $n$  polynomials  $P_1, \dots, P_n$ . Let us assume that *the roots  $\zeta \in \mathcal{Z}$  are simple*. Then  $J_{\mathbf{P}}(\zeta_i) \neq 0$ , where  $J_{\mathbf{P}}$  is the Jacobian of  $\mathbf{P}$ .

Let  $V_E(\mathcal{Z})$  be the generalized Van der Monde matrix defined in 2.11. Let  $\delta_{\zeta_i} : R \rightarrow \mathbb{C}$  denotes the evaluation at  $\zeta$ :

$$\delta_{\zeta_i}(p) = p(\zeta_i).$$

We immediately check that the columns of the matrix  $V_E(\mathcal{Z})$ , are nothing else but the coefficients of the evaluations  $\delta_{\zeta_i}$  in the dual basis of  $(\mathbf{x}^\alpha)_{\alpha \in E}$ .

PROPOSITION 4.7 — *For any polynomial  $a \in R$ , we have*

$$V_E(\mathcal{Z})^t \Omega_a V_E(\mathcal{Z}) = \text{diag}(a(\zeta_1) J_{\mathbf{P}}(\zeta_1), \dots, a(\zeta_D) J_{\mathbf{P}}(\zeta_D)).$$

*Proof.* As  $V_E(\mathcal{Z})$  represents the coefficients of the evaluation  $\delta_{\zeta_i}$  in the dual basis  $(\widehat{\mathbf{x}}^\alpha)_{\alpha, \beta}$ , the matrix  $V_E(\mathcal{Z})^t \Omega_a V_E(\mathcal{Z})$  is the matrix of the map

$$\begin{aligned} \widehat{B} \times \widehat{B} &\rightarrow \mathbf{C} \\ (\lambda, \mu) &\mapsto \lambda(\psi_a(\mu)) \end{aligned}$$

in the basis  $(\delta_{\zeta_i})_{i=1, \dots, D}$  of  $\widehat{B}$ . In other words, the entry  $(i, j)$  of this matrix is

$$\delta_{\zeta_i}(\psi_a(\delta_{\zeta_j})) = \Theta_{\mathbf{P}}(a)(\zeta_i, \zeta_j).$$

According to the equation (14), if  $\zeta \neq \eta$ , we have  $\Theta_{\mathbf{P}}(a)(\zeta, \eta) = \Theta_{\mathbf{P}}(\zeta, \eta) = 0$ .

If  $\eta = \zeta$ , then, by definition,  $\Theta_{\mathbf{P}}(a)(\zeta, \zeta) = a(\zeta) J_{\mathbf{P}}(\zeta)$ . Consequently,  $(\Theta_{\mathbf{P}}(a)(\zeta_i, \zeta_j))$  is the diagonal matrix

$$\text{diag}(a(\zeta_1) J_{\mathbf{P}}(\zeta_1), \dots, a(\zeta_D) J_{\mathbf{P}}(\zeta_D)).$$

□

**COROLLARY 4.8** *If the roots of the system  $\mathbf{P} = 0$  are simple, then*

$$\Xi_1 = V_E(\mathcal{Z}) \text{diag}\left(\frac{1}{J_{\mathbf{P}}(\zeta_1)}, \dots, \frac{1}{J_{\mathbf{P}}(\zeta_D)}\right) V_E(\mathcal{Z})^t.$$

*Proof.* We have  $\Omega_1 = V_E(\mathcal{Z})^{-t} \text{diag}(J_{\mathbf{P}}(\zeta_1), \dots, J_{\mathbf{P}}(\zeta_D)) V_E(\mathcal{Z})^{-1}$ , according to proposition (4.7), and we deduce from theorem (4.4) that

$$\Xi_1 = \Omega_1^{-1} = V_E(\mathcal{Z}) \text{diag}\left(\frac{1}{J_{\mathbf{P}}(\zeta_1)}, \dots, \frac{1}{J_{\mathbf{P}}(\zeta_D)}\right) V_E(\mathcal{Z})^t.$$

□

## References

- [1] D. Bini and V. Pan. *Polynomial and matrix computations, Vol 1 ; Fundamental Algorithms*. Progress in Theoretical Computer Science. Birkhäuser, Boston, 1994.
- [2] J. Canny and I. Emiris. An efficient algorithm for the sparse mixed resultant. In G. Cohen, T. Mora, and O. Moreno, editors, *Proc. Intern. Symp. Applied Algebra, Algebraic Algor. and Error-Corr. Codes (Puerto Rico)*, volume 263 of *LNCS*, pages 89–104. Springer Verlag, 1993.
- [3] J. Canny, E. Kaltofen, and Y. Lakshman. Solving Systems of Non-linear Polynomial Equations Faster. In *ISSAC'89*, pages 121–128. ACM Press, 1989.
- [4] J.P. Cardinal and B. Mourrain. Algebraic approach of residues and applications. In J. Renegar, M. Shub, and S. Smale, editors, *Proc. AMS-SIAM Summer Seminar on Math. of Numerical Analysis, (Park City, Utah'95)*, volume 32 of *Lectures in Applied Math.*, pages 189–210. Am. Math. Soc. Press, 1996.
- [5] M. Elkadi and B. Mourrain. Approche Effective des Résidus Algébriques. Rapport de Recherche 2884, INRIA, 1996.

- [6] I.M. Gelfand, M.M. Kapranov, and A.V. Zelevinsky. Discriminants of polynomials in several variables and triangulations of Newton polytopes. *Leningrad Math. J.*, 2(3):449–505, 1991. (Translated from *Algebra i Analiz*, 2, 1990, pp. 1–62).
- [7] F.S. Macaulay. Some formulae in elimination. *Proc. London Math. Soc.*, 1(33):3–27, 1902.
- [8] B. Mourrain and V. Y. Pan. Solving special polynomial systems by using structured matrices and algebraic residues. In F. Cucker, editor, *Foundations of Computational Mathematics*. (To appear in LNCS, Springer-Verlag, 1997).