Solving special polynomial systems by using structured matrices and algebraic residues

Bernard Mourrain[1] & Victor Y. Pan[2]

[1] INRIA, SAFIR
2004 route des Lucioles
B.P. 93, 06902 Sophia Antipolis
mourrain@sophia.inria.fr

[2] 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 9020690 and PSC CUNY Awards Nos. 666327)

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 apply some results on computations with matrices of these new classes, together with some techniques from algebraic residues theory, in order to devise an algorithm for approximating a selected solution of a polynomial system of the form

$$\begin{cases} x_1^{d_1} - R_1(x_1, \ldots, x_n) = 0, \\ \vdots \\ x_n^{d_n} - R_n(x_1, \ldots, x_n) = 0, \end{cases}$$

where $\deg(R_i) < d_i$. The complexity of this algorithm is $\mathcal{O}(D^2 \log(D)^c)$, where $D = \prod_{i=1}^n d_i$ is the number of the roots of the system.

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 and extend them in order to apply, in section 4, to the solution of polynomial systems of n equations with n variables. For a special class of such systems (where the i-th equation has the form $P_i = x_i^{d_i} - R_i(x_1, \ldots, x_n)$ and where R_i has a total degree less than d_i), we reduce the solution to computing the associated residues, where we apply some results on computations with structured matrices from section 2. This enables us to compute (under some additional assumptions) a selected solution to the system, by using order of $D^2 \log(D)$ arithmetic operations, where $D = \prod_i^n d_i$. The latter result is a substantial improvement versus the previously known solutions, requiring order of D^3 arithmetic operations. The result may also be of technical interest as the first example where combined application of structured matrices and algebraic residues leads to a substantial improvement of the known methods for solving polynomial systems of equations.

Next, we will state some definitions. $R = \mathbb{C}[x_1, \ldots, x_n]$ will denote the polynomial ring in variables $x_1, \ldots, x_n$ over the complex field $\mathbb{C}$, and $L = \mathbb{C}[x_1^{\pm 1}, \ldots, x_n^{\pm 1}]$ will denote the ring of Laurent's polynomials in the same variables. We will write $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{x}^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. For a vector $\alpha = (\alpha_1, \ldots, \alpha_n)$, we will write $|\alpha|$ to denote the 1-norm of this vector, $|\alpha| = \sum_{i=0}^n |\alpha_i|$. 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_α, is the highest total degree of its monomials, for which $c_\alpha \neq 0$. We will write $\lfloor S \rceil$ 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 26), 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 and Hankel 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 21 *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 are indexed by the elements of E and columns by the elements of F.*

- *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 $m_{\alpha-\mathbf{e}_i,\beta+\mathbf{e}_i} = m_{\alpha,\beta}$ 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 $\alpha - \beta$, that is, if for every $i = 1, \ldots, n$, 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 $E = [0, \cdots, m-1]$ and $F = [0, \ldots, n-1]$, definition 21 turns into the usual definition of Hankel (resp. Toeplitz) matrices [1].

DEFINITION 22 *Let $\mathcal{P}_E : L \to L$ be the projection map such that*

$$\mathcal{P}_E(\mathbf{x}^\alpha) = \mathbf{x}^\alpha$$

if $\alpha \in E$ and $\mathcal{P}_E(\mathbf{x}^\alpha) = 0$ otherwise. 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 $\mu_Q : L \to L$ 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 $L \to L$ such that

$$\mathcal{M}(\mathbf{x}^\beta) = \sum_{\alpha \in E} m_{\alpha,\beta} \mathbf{x}^{-\alpha}$$

if $\beta \in F$ and $\mathcal{M}(\mathbf{x}^{\beta}) = 0$ otherwise. 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 23 *If M is an (E, F) quasi-Hankel (resp. an (E, F) quasi-Toeplitz) matrix, then $\mathcal{M} = \mathcal{P}_{-E} \circ \mu_{H_M} \circ \mathcal{P}_F$ (resp. $\mathcal{M} = \mathcal{P}_E \circ \mu_{T_M} \circ \mathcal{P}_F$).*

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

2.2 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

$$H_M(\mathbf{x})\, V(\mathbf{x}) = \sum_{u \in E-F, \beta \in F} \mathbf{x}^{-u+\beta}\, h_u\, v_\beta$$

$$= \sum_{\alpha = u-\beta \in E-2\,F} \mathbf{x}^{-\alpha} \left(\sum_{\beta \in F} h_{\alpha+\beta}\, v_\beta \right),$$

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 α 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 24 *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 $N = \lfloor E - 2\,F \rceil$ (resp. $\lfloor E + 2F \rceil$) and where $C_{M,N}$ bounds the cost of evaluating the polynomial H_M (resp. T_M) at a fixed set of N points.*

Proof. To obtain the latter complexity estimate, we reduce the problem to computing the product of the two polynomials $H_M(\mathbf{x})$ (resp. $T_M(\mathbf{x})$) and $V(\mathbf{x})$ and then apply a variant of the well-known techniques of evaluation and interpolation (cf. [1]). Namely, we first evaluate the polynomials $H_M(\mathbf{x})$ (resp. $T_M(\mathbf{x})$) and $V(\mathbf{x})$ on a fixed set of N points, then pairwise multiply their values, to obtain the values of the product $H_M(\mathbf{x})\,V(\mathbf{x})$ (resp. $T_M(\mathbf{x})\,V(\mathbf{x})$), and finally, obtain the coefficients of this product by applying the known interpolation techniques for sparse polynomials (see e.g. section 1.9 of [1] or [10]).

In some special cases, we have better complexity estimates.

Proposition 25 *In the case where $E = F = \{(\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n \ ; \ 0 \leq \alpha_i \leq d_i - 1\}$, 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 $D = \prod_i^n d_i$.*

Proof. At first, as in the proof of proposition 24, we reduce the problem to multiplication of a pair of the associated multivariate polynomials, $U(\mathbf{x})$ and $V(\mathbf{x})$, so that $U\,V$ is of degree at most $3d_i - 3$ in each variable x_i. The latter multiplication can be reduced to multiplication of a pair of univariate polynomials, by means of application of Kronecker's map,

$$x_1 = y,\ x_k = y^{D_k},\ D_k = \prod_{i<k}(3d_i - 2),\ k = 1, \ldots, n,$$

which turns $U(\mathbf{x})\,\mathbf{V}(\mathbf{x})$ into a univariate polynomial of degree at most

$$(3d_n - 3)D_n < \prod_{i=1}^{n}(3d_i - 2) < 3^n D.$$

By applying fast Fourier transform (FFT), we may multiply such polynomials by using order of $3^n D \log(3^n D)$ ops [1].

Remark 1 — *For practical implementation where D is very large, it is better to avoid using Kronecker's map so as to multiply two multivariate polynomials $U(\mathbf{x})$ and $V(\mathbf{x})$ of lower degrees, rather than two univariate polynomials of very high degrees. Then, the number of ops may grow a little, but FFT involves roots of 1 of a lower order.*

Proposition 26 *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*

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

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

Proof. Proceed as in the proof of proposition 24 but use the interpolation technique of [3] for polynomials with bounded total degrees, instead of using the techniques of section 1.9 of [1] or [10].

2.3 Multivariate displacement operators and ranks (definition).

By convention, if $\mathcal{A} : L \to L$ 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 27 *For any subset E of $\mathbb{Z}^n$, we define the two following unit E-displacement matrices (operators):*

$$\mathcal{Z}_i^E = \mathcal{P}_E\, \mu_{x_i} \mathcal{P}_E$$

and

$$\mathcal{Z}_{-i}^E = \mathcal{P}_E\, \mu_{x_i^{-1}} \mathcal{P}_E.$$

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

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

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

Definition 28 *Let E and F denote two subsets of $\mathbb{Z}^n$ and let $\mathcal{A}$ denote a linear operator $L \to L$. Then, the operators*

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

and $T_i^-(\mathcal{A}) = \mathcal{A}-\mathcal{Z}_i^E\mathcal{A}\mathcal{Z}_{-i}^F$ *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.*

2.4 Bounds on displacement ranks of quasi-Hankel and quasi-Toeplitz matrices

Definition 29 *Hereafter, we write*

$$\delta_i(E) = \{\alpha : \alpha \in E\,;\ \alpha + e_i \notin E\}$$

(resp. $\delta_{-i}(E) = \{\alpha : \alpha \in E\,;\ \alpha - e_i \notin E\}$*).*

Proposition 210 *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:*

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

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

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:*

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

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

Proof. The proofs of all the four bounds of this proposition mimic each other, so we will only prove the first bound. According to proposition 23, we have

$$\mathcal{H}_i^+(\mathcal{M}) = \mathcal{P}_{-E}\left(\mu_U - \mu_{x_i^{-1}}\mathcal{P}_{-E}\,\mu_U\mathcal{P}_F\mu_{x_i}\right)\mathcal{P}_F,$$

where U is the polynomial associated with $\mathcal{M}$.

- If $\beta \notin F$ then $\mathcal{H}_i^+(\mathcal{M})(\mathbf{x}^\beta) = 0$.
- If β lies in $\delta_i(F)$, then $\mathcal{Z}_i^F(\mathbf{x}^\beta) = 0$ and $\mathcal{H}_i^+(\mathcal{M})(\mathbf{x}^\beta) = \mathcal{M}(\mathbf{x}^\beta) = \sum_{\alpha \in E} m_{\alpha,\beta}\, \mathbf{x}^\alpha$.
- If β lies in F but not in $\delta_i(F)$, then

$$\mu_{x_i^{-1}} \mathcal{P}_{-E} \mu_U \mathcal{P}_F \mu_{x_i}(\mathbf{x}^\beta) = x_i^{-1} \mathcal{P}_{-E}(U\, \mathbf{x}^\beta\, x_i)$$

and

$$\begin{aligned} \mathcal{H}_i^+(\mathcal{M})(\mathbf{x}^\beta) &= \mathcal{P}_{-E}\left(U\, \mathbf{x}^\beta - x_i^{-1} \mathcal{P}_{-E}(U\, \mathbf{x}^\beta\, x_i)\right) \\ &= \mathcal{P}_{-E}\left(U\, \mathbf{x}^\beta - U\, \mathbf{x}^\beta + x_i^{-1} \rho_{-E}(U\, \mathbf{x}^\beta\, x_i)\right) \\ &= \mathcal{P}_{-E}\left(x_i^{-1} \rho_{-E}(U\, \mathbf{x}^\beta\, x_i)\right) \\ &= \sum_{\alpha \in \delta_i(-E)} m_{\alpha,\beta}\, \mathbf{x}^{-\alpha}. \end{aligned}$$

Therefore, if $\alpha \notin \delta_i(-E)$ and $\beta \notin \delta_i(\mathcal{H}_i^+)$, then $[\mathcal{H}_i^+(\mathcal{M})]_{\alpha,\beta} = 0$. Thus, the rank of $\mathcal{H}_i^+(M)$ is at most $\lfloor \delta_i(-E) \rceil + \lfloor \delta_i(F) \rceil$.

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

2.5 Examples

Quasi-Toeplitz matrices: Let us be given some multivariate polynomials $P_0, \ldots, P_n$ and let us consider the matrix associated with the linear map

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

where 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 $Q_i, i = 0, \ldots, n$, where E denotes the set of the exponents of the monomials of $(P_i Q_i)_{i=0,\ldots,n}$, and where V is generated by the monomials $\mathbf{x}^\alpha$ for $\alpha \in E$. Such maps typically appear in the construction of resultant type matrices associated to the system $\{P_i = 0, \ i = 0, \ldots, n\}$, of polynomial equations [8], [2].

Let M denote the matrix of this linear map in the monomial basis of $V_0 \times \cdots \times V_n$ 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 $F_0 \sqcup \cdots \sqcup F_n$. 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 $(0, a_1, \ldots, a_n)$. Let e_0 denote the first canonical vector of $\mathbb{Z}^{n+1}$. Then, the elements of the subset

$$F = \{i\, e_0 + \alpha \ ; \ 0 \leq i \leq n, \, \alpha \in F_{i+1}\}$$

index the rows of M.

Note that the $(\alpha, i\, e_0 + \beta)$-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 $\alpha - \beta - i\, e_0$.

Remark 2 — *Resultant type matrices and their transposes are quasi-Toeplitz matrices.*

Remark 3 — *The polynomial associated to the Toeplitz operator ϕ is just*

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

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 $(a, b) \mapsto q(a, b) = \lambda(a\, b)$ where $\lambda \in \hat{L}$ is a linear form (see [6]). 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.

3 Algebraic residues

In this section, we will recall some basic definitions from algebraic residue theory, referring the reader to [5], [6] for further details.

3.1 Definitions and basic facts

Let $R = \mathbb{C}[\mathbf{x}] = \mathbb{C}[x_1, \ldots, x_n]$ 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 vectors $\mathbf{y} = (y_1, \ldots, y_n)$ and write $\mathbf{x}^{(0)} = \mathbf{x}$, $\mathbf{x}^{(1)} = (y_1, x_2, \ldots, x_n)$, $\ldots$,

$\mathbf{x}^{(n)} = \mathbf{y}$. 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 $P_0, \ldots, P_n \in R$, let us construct the following polynomial in $\mathbf{x}$ and $\mathbf{y}$:

$$\Phi(P_0, P_1, \ldots, P_n) = \begin{vmatrix} P_0(\mathbf{x}) & \theta_1(P_0) & \cdots & \theta_n(P_0) \\ \vdots & \vdots & & \vdots \\ P_n(\mathbf{x}) & \theta_1(P_n) & \cdots & \theta_n(P_n) \end{vmatrix}, \tag{1}$$

and let us write $\Delta_{\mathbf{P}} = \Phi(1, P_1, \ldots, P_n) \in \mathbb{C}[\mathbf{x}, \mathbf{y}]$. Now, we can define the residue of $\mathbf{P} = (P_1, \ldots, P_n)$ as a unique linear form τ in the set of linear forms on R such that

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

Hereafter, I will denote the ideal generated by the polynomials $P_1, \ldots, P_n$; $B = R/I$ 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, then we have the following property:

$$\Delta_{\mathbf{P}} \equiv \sum_{\alpha \in E} \mathbf{x}^\alpha \, \mathbf{w}_\alpha(\mathbf{y}) \equiv \sum_{\alpha \in E} \mathbf{w}_\alpha(\mathbf{x}) \, \mathbf{y}^\alpha \mod (\mathbf{P}(\mathbf{x}), \mathbf{P}(\mathbf{y})).$$

Here, $(\mathbf{w}_\alpha)$ is the dual basis of $(\mathbf{x}^\alpha)$ for τ:

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

$\delta_{\alpha,\beta}$ is 1 if $\alpha = \beta$ and 0 otherwise. Thus, for any $b \in B$, we have the relations

$$b \equiv \sum_{\alpha \in E} \tau(b \, \mathbf{x}^\alpha) \, \mathbf{w}_\alpha \equiv \sum_{\alpha \in E} \tau(b \, \mathbf{w}_\alpha) \, \mathbf{x}^\alpha. \tag{2}$$

Moreover, for any polynomial $Q \in R$, we have

$$\Delta_{\mathbf{P}}(\mathbf{x}, \mathbf{y}) \, Q(\mathbf{x}) \equiv \Delta_{\mathbf{P}}(\mathbf{x}, \mathbf{y}) \, Q(\mathbf{y}) \mod (\mathbf{P}(\mathbf{x}), \mathbf{P}(\mathbf{y})). \tag{3}$$

Consequently, for any pair of distinct roots, ζ and η, of the polynomial system $\mathbf{P} = \mathbf{0}$, we have

$$\Delta_{\mathbf{P}}(\zeta, \eta) = 0. \tag{4}$$

3.2 Computation of the residues associated to systems of polynomial equations

We consider a system of polynomial equations of the special form

$$\begin{cases} P_1 = x_1^{d_1} - R_1(x_1, \ldots, x_n) = 0, \\ \vdots \\ P_n = x_n^{d_n} - R_n(x_1, \ldots, x_n) = 0, \end{cases}$$

with $\deg(R_i) < d_i$. Then, the vector space B has dimension $D = \prod_{i=1}^{n} d_i$, and a basis of B is the set of monomials $\mathbf{x}^\alpha$ with α in the set

$$E = \{(\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n \ ; \ 0 \leq \alpha_i \leq d_i - 1\}.$$

In this section, we compute the vector

$$[\tau(\mathbf{x}^\alpha)]_{\nu < |\alpha| \leq \mu},$$

where $\nu = \sum_{i=1}^{n} d_i - n$ and $\mu > \nu$. Let $T_l = \{\alpha \in \mathbb{N}^n$ such that $|\alpha| \leq l\}$ and $T^{(l)} = \{\alpha \in \mathbb{N}^n$ such that $|\alpha| = l\}$. The set T_l contains $\sigma_l := \sigma_{l,n} := \lfloor T_l \rceil = \binom{l+n}{n}$ elements. Due to Stirling's formula, $\sigma_l = \lfloor T_l \rceil$ (the cardinality of T_l) is asymptotically equivalent to

$$\mathcal{O}((e\frac{l+n}{n})^n/\sqrt{n}), \ e = 2.718\ldots.$$

Let $\mathcal{N}$ denote the set of all monomials $\mathbf{x}^\alpha$ for $\alpha \in T_\mu \backslash T_\nu = \{\alpha \in T_\mu; \, \alpha \notin T_\nu\}$. An element of $\mathcal{N}$ is divisible by one of the monomials $x_i^{d_i}$. Therefore, $\mathcal{N}$ can be partitioned into n subsets as follows:

$$\mathcal{N} = x_1^{d_1}\mathcal{N}_1 \sqcup \cdots \sqcup x_n^{d_n}\mathcal{N}_n.$$

(For instance, $x_1^{d_1}\,\mathcal{N}_1$ is the subset of elements of $\mathcal{N}$ divisible by $x_1^{d_1}$, $x_2^{d_2}\,\mathcal{N}_2$ is the subset of elements of $\mathcal{N}$ divisible by $x_2^{d_2}$ and not by $x_1^{d_1}$, …). Let V_i (resp. V) denote the vector space generated by the monomials in $\mathcal{N}_i$ (resp. $\mathcal{N}$). As in section 2.5, consider the map

$$\begin{aligned} \phi \ : \ & V_1 \times \cdots \times V_n \rightarrow V, \\ & (Q_1, \ldots, Q_n) \mapsto \sum_{i=1}^{n} P_i\, Q_i. \end{aligned}$$

For any set S in R, let $S^{(l)}$ (resp. $S^{(\leq l)}$) denote the subset of S formed by the elements that are homogeneous of a degree l (resp. of a degree at most l).

Due to the structure of the polynomials of the system $\mathbf{P} = \mathbf{0}$, the image of $(V^{(l-d_1)} \times \cdots \times V_n^{(l-d_n)})$ is in $V^{(\leq l)}$, and the matrix of this map [in the monomial basis of $(V^{(l-d_1)} \times \cdots \times V_n^{(l-d_n)})$ and $V^{(\leq l)}$] has the following form:

$$\left(\begin{array}{c} U_l \\ \\ \\ I_{\sigma_l} \end{array}\right) \begin{array}{l} \left.\vphantom{\begin{array}{c} a \\ b \end{array}}\right\} V^{(\leq l-1)} \\ \left.\vphantom{\begin{array}{c} a \\ b \end{array}}\right\} V^{(l)} \end{array},$$

where U_l is a $(T_{l-1}, T^{(l)})$ quasi-Toeplitz matrix, $\sigma_l = \lfloor T_l \rceil$, and I_{σ_l} is the $\sigma_l \times \sigma_l$ identity matrix. Thus, the complete matrix of ϕ in the monomial basis is a block upper triangular matrix of the size $\lfloor T_\mu \rceil \times (\lfloor T_\mu \rceil - \lfloor T_\nu \rceil)$.

Let M be the submatrix of this matrix whose indices of rows and columns are in $\mathcal{N} = T_\mu \backslash T_\nu$ and let $\mathbf{v} = [\mathbf{v}_\alpha]_{\alpha \in N}$ be the row corresponding to the monomial $x_1^{d_1-1} \cdots x_n^{d_n-1}$ in the matrix of ϕ. Then, we can compute the residues based on the next result:

Proposition 31 *(see [6]). The following equations hold:*

$$\mathbf{t} = [\tau(\mathbf{x}^\alpha)]_{\alpha \in \mathcal{N}} = -\mathbf{v}\, M^{-1}.$$

The matrix M is upper triangular, of size $(\lfloor T_\mu \rceil - \lfloor T_\nu \rceil) \times (\lfloor T_\mu \rceil - \lfloor T_\nu \rceil)$. Therefore, the vector $\mathbf{t}$ can be computed in $\mathcal{O}((\lfloor T_\mu \rceil - \lfloor T_\nu \rceil)^2)$ ops.

By using the quasi-Toeplitz structure of the matrix M, we can obtain a better complexity bound.

Proposition 32 *The vector of the residues, $[\tau(\mathbf{x}^\alpha)]_{\alpha \in N}$, can be computed in $\mathcal{O}((\mu - \nu)\,\sigma \log^3 \sigma)$ ops, where*

$$\sigma = \sigma_{3\mu,n} = \binom{3\,\mu + n}{n} = \mathcal{O}((e\frac{3\,\mu + n}{n})^n / \sqrt{n}),\ e = 2.718\ldots.$$

Proof. Computation of $[\tau(\mathbf{x}^\alpha)]_{\alpha \in N}$ is essentially reduced to solving the following linear system of equations:

$$[\mathbf{t}_{\nu+1}, \ldots, \mathbf{t}_\mu] = [\mathbf{v}_{\nu+1}, \ldots, \mathbf{v}_\mu] \begin{bmatrix} I_{\sigma_{\nu+1}} & U_{\nu+1,\nu+1} & \cdots & U_{\nu,\mu} \\ 0 & I_{\sigma_{\nu+2}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & U_{\mu-1,\mu} \\ 0 & \cdots & 0 & I_{\sigma_\mu} \end{bmatrix}. \quad (5)$$

We immediately obtain from the latter system that $\mathbf{t}_{\nu+1} = \mathbf{v}_{\nu+1}$. Substituting this part of the solution vector into the system yields a subsystem in $[\mathbf{t}_{\nu+2}, \dots, \mathbf{t}_\mu]$, having similar form. We will iterate this process in order to compute the entire vector $\mathbf{v}\ M^{-1}$, by using $\mu - \nu$ iteration steps. At the k-th iteration step, we multiply the vector $\mathbf{t}_{\nu+k}$ by the quasi-Toeplitz matrix $[U_{\nu+k,\nu+k}, \dots, U_{\nu+k,\mu}]$ and also perform $\lfloor T_\mu \rceil - \lfloor T_{k+\nu} \rceil$ vector subtractions. By applying proposition 26, we multiply the vector $\mathbf{t}_{\nu+k}$ by the matrix $[U_{\nu+k,\nu+k},\ \dots,\ U_{\nu+k,\mu}]$ by using order of

$$\binom{2\,\nu + 2\,k + \mu + n}{n} \log^3 (\binom{2\,\nu + 2\,k + \mu + n}{n})$$

ops, which dominates the computational cost of performing the k-th iteration step. Since $k \leq \mu - \nu$, the entire computation of the vector $(\mathbf{v}_{\nu+\mathbf{1}}, \cdots, \mathbf{v}_\mu)$ only requires

$$\mathcal{O}((\mu - \nu)\, \sigma \log^3 \sigma)$$

ops, where $\sigma = \sigma_{3\,\mu,n} = \binom{3\,\mu + n}{n}$.

Hereafter, we will write $H_0 = [\tau(\mathbf{x}^{\alpha+\beta})]_{\alpha,\beta \in E}$ and $H_i = [\tau(x_i \mathbf{x}^{\alpha+\beta})]_{\alpha,\beta \in E}$. Clearly, H_i are quasi-Hankel matrices for all i. Since $(\mathbf{x}^\alpha)_{\alpha \in \mathbf{E}}$ is a basis in the quotient ring B, we will call H_0 a *basis residue matrix* of B.

Proposition 33 *The matrices $H_i, i = 0, \dots, n$ can be computed by using*

$$\mathcal{O}\left(n^{3.5}\,(6\,e)^n \vartheta^{n+1} \log^3(\vartheta)\right)$$

ops, where $\vartheta = \frac{\sum_i d_i}{n}$.

Proof. In order to compute H_i, we need to compute $\tau(\mathbf{x}^\alpha)$ for all α such that $|\alpha| \leq \mu = 2 \sum_{i=1}^n d_i - 2\,n + 1$. By using Stirling formula, we obtain that

$$\sigma_{3\,\mu,n} = \mathcal{O}\left((6\,e \frac{\sum_i d_i}{n})^n / \sqrt{n} \right),\ e = 2.718\dots.$$

Substitution into proposition (32) gives us the desired complexity.

If all the degrees d_i are equal to each other, that is, if $d_i = d,\ i = 1, \cdots, n$, then, the complexity bound of the latter proposition turns into

$$\mathcal{O}(6^n\ e^n\ d^{n+1}\ n^{3.5} \log^3 d) = \mathcal{O}(\frac{1}{n^3}\, 6^n\ e^n\ D^{1+\frac{1}{n}} \log(D))$$

ops, which is asymptotically better than D^2.

3.3 Linear solver with the basis residue matrix $\boldsymbol{H_0}$

Proposition 34 *The inverse of* W_0 *is the quasi-Hankel matrix* H_0.

Proof. We have the equations

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

where $\delta_{\alpha,\beta} = 1$ if $\alpha = \beta$ and 0 otherwise, and we have

$$H_0\, W_0 = I_D,$$

where I_D is the $D \times D$ identity matrix. Therefore, the inverse of H_0 is the matrix of the coefficients of the dual basis of $(\mathbf{x}^{\alpha})$ in this basis.

Proposition 35 *The solution vector* $\mathbf{z}$ *to the system* $H_0\,\mathbf{z} = \mathbf{w}$ *can be computed in* $\mathcal{O}(3^n\, D^2 \log(3^n D))$ *ops.*

Proof. We first compute the values $s_k = \mathbf{v}^T\, \widetilde{H}_0^k\, \mathbf{u}$, where $\mathbf{u}$ and $\mathbf{v}$ are two random vectors, $\widetilde{H}_0 = T H_0$, T is a random square Toeplitz matrix, and $k = 0, \ldots, 2\,D+1$. We reduce multiplication of the matrix $\widetilde{H}_0$ by a vector to multiplication of H_0 by a vector, followed by multiplication of T by the resulting vector. By applying proposition 25, we multiply H_0 by a vector by using $\mathcal{O}(3^n D \log(3^n D))$ ops. Multiplication of the Toeplitz matrix T by a vector takes $\mathcal{O}(D \log D)$ ops (cf. [1]). Therefore, we may evaluate $\widetilde{H}_0^k\, \mathbf{u}$, for $k = 0, \ldots, 2D+1$, by using $\mathcal{O}(3^n D^2 \log(3^n D))$ ops. Then, the values $s_0, \ldots, s_{2\,D+1}$ can be computed by using $\mathcal{O}(D^2)$ ops.

Due to our random choice of Toeplitz matrix T, the characteristic and minimal polynomials of $\widetilde{H}_0$ coincide with each other, $c_{\widetilde{H}_0}(x) = m_{\widetilde{H}_0}(x)$, with a sufficiently high probability (cf. [7] and regularization 2.13.3 on page 206 of [1]). Now, assuming that they do coincide and having the values $s_0, \ldots, s_{2\,D+1}$ available, we will follow [1] and [7] and will compute the coefficients of the characteristic polynomial $c_{\widetilde{H}_0}(\lambda) = \sum_{i=0}^{D} c_i\, \lambda^i$ of $\widetilde{H}_0$. This computation is reduced to the solution (by using $\mathcal{O}(D \log^2 D)$ ops, cf. [1]) of a Toeplitz linear system of D equations. The latter system is non-singular, with a high probability, due to the equality, with a high probability, of the minimal polynomial of $\widetilde{H}_0$ to the characteristic polynomial of $\widetilde{H}_0$ and to the random choice of the vectors $\mathbf{u}$ and $\mathbf{v}$ (cf. [1] or [7]). Having the coefficients c_i, $i = 0, \ldots, D$, of the characteristic polynomial of the matrix $\widetilde{H}_0$ available and applying the Cayley-Hamilton theorem, we obtain the expression

$$\widetilde{H}_0^{-1} = -\sum_{i=1}^{D} \frac{c_i}{c_0} \widetilde{H}_0^{i-1},$$

where $c_0 = (-1)^D \det(\widetilde{H}_0)$ [$c_0 \neq 0$, with a high probability, since so are $\det(T)$ and $\det(H_0)$]. Thus, computing $H_0^{-1}\mathbf{w}$ requires $D-1$ multiplications of $\widetilde{H}_0$ by vectors and D additions of vectors. According to proposition 25, these operations can be performed in $\mathcal{O}(3^n D^2 \log(3^n D))$ ops.

3.4 Multiplication in $B = R/I$

Let us write $\mathbf{w}_\alpha = \sum_{\beta \in E} w_{\beta,\alpha} \mathbf{x}^\alpha$ and $W_0 = (w_{\alpha,\beta})_{\alpha,\beta \in E}$.

Let μ_i denote the operator of multiplication by x_i in B and let $M_i = (m^{(i)}_{\alpha,\beta})_{\alpha,\beta \in E}$ denote its matrix in the basis $(\mathbf{x}^\alpha)$. Then, we have

$$x_i \, \mathbf{x}^\alpha \equiv \sum_{\gamma \in E} m^{(i)}_{\gamma,\alpha} \, \mathbf{x}^\gamma$$

and

$$\tau(x_i \, \mathbf{x}^{\alpha+\beta}) = \sum_{\gamma \in E} \tau(\mathbf{x}^{\beta+\gamma}) \, m^{(i)}_{\gamma,\alpha}.$$

In other words, we have

$$H_i = H_0 \, M_i. \tag{6}$$

Note that the matrix $(n^{(i)}_{\alpha,\beta})$ of multiplication by x_i in the basis $(\mathbf{w}_\alpha)$ is M_i^T, for we have

$$m^{(i)}_{\alpha,\beta} = \langle x_i \mathbf{x}^\beta | \mathbf{w}_\alpha \rangle = \tau(x_i \, \mathbf{x}^\beta \, \mathbf{w}_\alpha)$$

and

$$n^{(i)}_{\alpha,\beta} = \langle x_i \mathbf{w}_\beta | \mathbf{x}^\alpha \rangle = \tau(x_i \, \mathbf{w}_\beta \, \mathbf{x}^\alpha).$$

Proposition 36 *After a precomputation of*

$$\mathcal{O}\left(n^{3.5} \, (9\,e)^n \, \vartheta^{n+1} \log^3(\vartheta)\right) \ \textit{ops}$$

where $\vartheta = \frac{\sum_i d_i}{n}$, *two elements of* B *can be multiplied in*

$$\mathcal{O}\left(3^n D^2 \log(3^n D)) + 4^n D \log^2(4^n D)\right)$$

ops.

Proof. We want to compute $f\,g$ in B where

$$f := \textstyle\sum_{\alpha\in E} f_\alpha\, \mathbf{x}^\alpha,$$
$$g := \textstyle\sum_{\alpha\in E} g_\alpha\, \mathbf{x}^\alpha.$$

According to the equation (2), we have

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

Let $S := \sum_{u\in 3\,E} \tau(\mathbf{x}^u)\mathbf{x}^{-u} = \sum_{u\in 3\,E} \tau_u \mathbf{x}^{-u}$, with the convention that $\tau_u = 0$ if $u \notin 3\,E$. According to (32), this polynomial can be computed within

$$\mathcal{O}\left(n^{3.5}\,(9\,e)^n\,\vartheta^{n+1}\log^3(\vartheta)\right)$$

ops. This computation has to be done once. Then,

$$\begin{aligned} S\,g &= \sum_{u\in 3\,E,\gamma\in E} \mathbf{x}^{-u+\gamma}\,\tau_u\,g_\gamma \\ &= \sum_{v=u-\gamma\in 3\,E-E} \mathbf{x}^{-v} \sum_{\gamma\in E,\ v+\gamma\in 3\,E} \tau_{v+\gamma}\,g_\gamma. \end{aligned}$$

The support of this polynomial is in $3\,E - E$. Therefore, by applying the known interpolation techniques for sparse polynomials ([1], [10]), the product of these two polynomials can be computed in $\mathcal{O}(4^n D \log^2(4^n D))$ ops. Similarly,

$$\begin{aligned} S\,f\,g &= \sum_{v\in 3\,E,\beta\in E} \mathbf{x}^{-v+\beta} \left(\sum_{\gamma\in E, v+\gamma\in 3\,E-E} \tau_{v+\gamma}\,g_\gamma \right) f_\beta \\ &= \sum_{\alpha=v-\beta\in 3\,E-2\,E} \mathbf{x}^{-\alpha} \sum_{\beta\in E,\gamma\in E} \tau_{\alpha+\beta+\gamma}\,f_\beta\,g_\gamma. \end{aligned}$$

Therefore, the coefficients of $\mathbf{x}^{-\alpha}$ in $S\,f\,g$ for $\alpha \in E$ are precisely the coefficients of $f\,g$ in the dual basis $(\mathbf{w}_\alpha)$ of B. Note that these coefficients only involve the coefficients of $\mathbf{x}^{-v}$ in $S\,g$ for which $v \in 2\,E$. Therefore, the cost of such a computation is $\mathcal{O}(3^n D \log^2(3^n D))$ ops.

In order to obtain the coefficients of $f\,g$ in the basis $(\mathbf{x}^\alpha)$, we multiply the vector

$$\mathbf{t} = [\sum_{\beta\in E,\gamma\in E} \tau_{\alpha+\beta+\gamma}\,f_\beta\,g_\gamma]_{\alpha\in E}$$

by the matrix $W_0 = [\mathbf{w}_\alpha]_{\alpha \in E} = H_0^{-1}$, that is, we solve the linear system of equations $H_0\,\mathbf{s} = \mathbf{t}$. According to proposition (35), this computation can be done within $\mathcal{O}(3^n\,D^2 \log(3^n D))$ ops.

4 Application to solving polynomial systems of equations

4.1 Computing selected roots of a polynomial system.

Let Z denote the set of all common roots of the system $\mathbf{P} = \mathbf{0}$. We will assume that *they are all distinct*. Let J be the Jacobian of $\mathbf{P}$. For any $\zeta \in Z$, we have $J(\zeta) \neq 0$.

PROPOSITION 41 *If the roots of* $\mathbf{P}$ *are simple, then*

$$e_\zeta = \frac{1}{J(\zeta)} \mathit{\Delta}_{\mathbf{P}}(\mathbf{x}, \zeta),\ \ \zeta \in Z,$$

is a linear basis of orthogonal idempotent of B, of sum 1.

Proof. According to the equation (3), for any $Q \in R$ and for any $\zeta \in Z$, we have

$$\mathit{\Delta}_{\mathbf{P}}(\mathbf{x}, \zeta)\, Q(\mathbf{x}) \equiv \mathit{\Delta}_{\mathbf{P}}(\mathbf{x}, \zeta)\, Q(\zeta)$$

in the quotient ring B. Therefore,

$$\mathit{\Delta}_{\mathbf{P}}(\mathbf{x}, \zeta)\, \mathit{\Delta}_{\mathbf{P}}(\mathbf{x}, \zeta) \equiv J(\zeta) \mathit{\Delta}_{\mathbf{P}}(\mathbf{x}, \zeta),$$

and $e_\zeta = \frac{1}{J(\zeta)}\, \mathit{\Delta}_{\mathbf{P}}(\mathbf{x}, \zeta)$ is an idempotent ($J(\zeta) \neq 0$, assuming all roots of the system $\mathbf{P} = \mathbf{0}$ are distinct). Moreover, according to (4), we have

$$\mathit{\Delta}_{\mathbf{P}}(\mathbf{x}, \zeta)\, \mathit{\Delta}_{\mathbf{P}}(\mathbf{x}, \eta) \equiv \mathit{\Delta}_{\mathbf{P}}(\mathbf{x}, \zeta) \mathit{\Delta}_{\mathbf{P}}(\zeta, \eta) \equiv 0,$$

for any pair of distinct roots $\zeta, \eta \in Z$, which shows that $e_\zeta\, e_\eta \equiv 0$ unless $\zeta = \eta$. Now, we recall from the definition of the residue τ and from the the Euler-Jacobi identity (cf. [6]) that

$$\begin{aligned} \mathit{\Delta}_{\mathbf{P}}(\tau) &= 1 \text{ (by definition)} \\ &= \sum_{\zeta \in Z} \frac{1}{J(\zeta)}\, \mathit{\Delta}_{\mathbf{P}}(\mathbf{x}, \zeta) = \sum_{\zeta \in Z} e_\zeta \text{ (by the Euler}-\text{Jacobi identity).} \end{aligned}$$

By decomposing any element h of B in the basis e_ζ, we obtain that

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

Here, the second equation follows since $e_\zeta h(x) \equiv e_\zeta h(\zeta)$. Squaring h in the quotient ring B gives us

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

Here and hereafter, for any element $b \in B$, $[b]$ denotes the vector of the coefficients of b in the basis $(\mathbf{x}^\alpha)_{\alpha \in E}$. In particular, $[1] = (1, 0, \cdots, 0)$ if the basis starts with the monomial 1. Let $|| \cdot ||$ denote a norm in $\mathbb{C}^D$ [say, for the Euclidean 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 abuse of notation, for any element $b \in B$, $||b||$ will denote $||[b]||$. Let $h \in R$ and assume that there is a unique root $\zeta \in Z$, for which the norm of $h(\zeta)$ is maximum, so that

$$|h(\zeta)|/|h(\eta)| - 1 \geq \rho, \tag{7}$$

for some fixed positive ρ and for any $\eta \in Z$ distinct from ζ. (Since all the roots in Z are assumed to be distinct, we may, in principle, ensure the latter relation with a high probability, by means of random linear substitution of the vector of the variables $\mathbf{x}$.) Then, by iteratively computing and normalizing the squares,

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

so that we have

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

and $\epsilon_k \leq 2^{-b}$ in $k = k(\rho, b) = \mathcal{O}(\log(b/\rho))$ recursive steps. Therefore, squaring and normalizing in B, we will make our process converge to a multiple of the element e_ζ.

From the previous section, we have the bound of $\mathcal{O}(3^n\, D \log(3^n\, D) + D^2 \, \log\, D)$ ops on the computational cost of squaring in B, which means that

$$\mathcal{O}\left(3^n D^2 \log(3^n D) + 4^n D \log^2(4^n D)\right)$$

ops suffice in order to approximate the element $e_\zeta/||e_\zeta||$ *within the error norm bound* 2^{-b}, assuming the equation (7).

We refer the reader to [9] and [4] for preceding works on a similar approach in the univariate case.

4.2 Transition from e_ζ to ζ

By definition, we have

$$e_\zeta = \frac{1}{J(\zeta)} \Delta_{\mathbf{P}}(\mathbf{x}, \zeta) = \frac{1}{J(\zeta)} \sum_{\alpha \in E} \mathbf{w}_\alpha(\mathbf{x})\, \zeta^\alpha.$$

This can be rewritten as

$$[e_\zeta] = \frac{1}{J(\zeta)} W_0\, [\zeta^\alpha]_{\alpha \in E}.$$

Here, W_0 denotes the matrix $(\mathbf{w}_\alpha)_{\alpha \in E}$ in the monomial basis $(\mathbf{x}^\alpha)_{\alpha \in E}$. Then, we obtain that

$$[\zeta^\alpha]_{\alpha \in E} = J(\zeta)\, H_0\, [e_\zeta]. \tag{8}$$

Let $f = H_0 e_\zeta$. Then according to (8), we have $f_{\mathbf{0}} = 1/J(\zeta)$ and the i$^{\text{th}}$ coordinate of ζ is $\frac{f_{x_i}}{f_0}$. According to proposition (35), we arrive at the following result:

Proposition 42 *The transition from e_ζ to the root ζ of the system $\mathbf{P} = \mathbf{0}$ can be performed by using $\mathcal{O}(3^n\, D^2 \log(3^n\, D))$ ops.*

4.3 The closest root

Suppose that we seek a root of the system $\mathbf{P} = \mathbf{0}$ for which 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 $\mathbf{P} = \mathbf{0}$ and that $x_1 - u$ has reciprocal in B. 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 for which x_1 is the closest to u_1 is a root for which $|\rho_1(\zeta)|$ is the largest. Consequently, iterative squaring of $\rho_1 = \rho_1(\zeta)$ shall converge to this root.

The polynomial ρ_1 can be computed in the following way. Let μ_1 denote multiplication by x_1 in B. Then $\rho_1 = (\mu_1 - u)^{-1}[1]$, and according to the matrix equation (6), we have

$$[\rho_1] = H_0\, (H_1 - u\, H_0)^{-1}[1],$$

which can be computed within $\mathcal{O}(3^n\, D^2 \log(3^n\, D))$ ops.

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

5 Conclusion

In this paper, we extend the structure of Toeplitz and Hankel matrices to a new class of structured matrices and operators, which includes matrices associated to the polynomial systems of equations and which we call quasi-Hankel and quasi-Toeplitz matrices and operators. Exploiting the fact that multiplication of such matrices by vectors is "fast", we devise an algorithm, based on residues computations, for approximating a selected solution of a polynomial system of the form $P_i = x_i^{d_i} - R_i(x_1, \ldots, x_n)$ where R_i has a total degree less than d_i.

The key ingredients of this algorithm are a) the inversion of basis residue matrices b) fast multiplication of two elements of the quotient ring $B = R/I$, and c) an iterative method for approximating a single root of the system. The complexity analysis shows that our algorithm gives us a substantial improvement of the known methods for solving polynomial systems of equations. Though we consider a special class of polynomial systems, we expect to extend such an approach to general systems $R_i = 0, i = 1, \ldots, n$, by means of the known techniques of homotopic deformation and the local continuity of the residue.

References

1. D. Bini and V. Y. Pan. *Polynomial and matrix computations, Vol 1 : Fundamental Algorithms.* Birkaüser, 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, Springer's Lecture Notes in Computer Science, vol. 263*, pages 89–104, 1993.
3. J. Canny, E. Kaltofen, and Y. Lakshman. Solving systems of non-linear polynomial equations faster, *Proc. ACM-SIGZAM Intern. Symp. on Symb. and Alg. Comput. (ISSAC'89)*, pages 121-128, ACM Press, New York, 1989.
4. J. P. Cardinal. On two iterative methods for approximating the roots of a polynomial. In J. Renegar, M. Shub, and S. Smale, editors, *Proc. AMS-SIAM Summer Seminar on Math. of Numerical Analysis*, Park City, Utah, July 1995. *Lectures in Applied Math.*, 1996.
5. 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, July 1995. *Lectures in Applied Math.*, 1996.
6. M. Elkadi and B. Mourrain. Approche Effective des Résidus Algébriques. Rapport de Recherche 2884, INRIA, 1996.
7. E. Kaltofen and V.Y. Pan. Processor efficient parallel solution of linear systems over an abstract field. In *Proc. 3rd Ann. ACM Symp. on Parallel Algorithms and Architectures*, pages 180–191. ACM Press, New York, 1991.

8. F. S. Macaulay. Some formulae in elimination. *Proc. London Math. Soc.*, 1(33):3–27, 1902.

9. J. Sebastiao e Sylva. Sur une méthode d'approximation semblable á celle de Graeffe. *Portugal. Math.*, 2:271-279, 1941.

10. R. Zippel. Interpolating polynomials from their values. *J. Symbolic Computation*, 9:375–403, 1990.