ELSEVIER Linear Algebra and its Applications 332–334 (2001) 583–606

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

Bézout and Hankel matrices associated with row reduced matrix polynomials, Barnett-type formulas

Marc Van Barel [a,*,1], Vlastimil Pták [b,2,3], Zdeněk Vavřín [b,2]

[a] *Department of Computer Science, Katholieke Universiteit Leuven, Celestijnenlaan 200 A, B-3001, Leuven (Heverlee), Belgium*

[b] *Institute of Mathematics, Czech Academy of Sciences, Žitná 25, 115 67 Praha, Czech Republic*

Received 2 December 1999; accepted 19 March 2001

Submitted by N.J. Higham

Abstract

Based on the approach introduced by B.D.O. Anderson and E.I. Jury in 1976, the definition of finite Hankel and Bézout matrices corresponding to matrix polynomials is extended to the case where the denominator of the corresponding rational matrix function is not necessarily monic but is row reduced. The matrices introduced keep most of the well-known properties that hold in the monic case. In particular, we derive extensions of formulas giving a connection with polynomials in the companion matrix (usually called Barnett formulas), of the inversion theorem and of formulas concerning alternating products of Hankel and Bézout matrices. © 2001 Elsevier Science Inc. All rights reserved.

* Corresponding author. Fax: +32-0-16-2053-08.

E-mail address: marc.vanbarel@cs.kuleuven.ac.be (M. Van Barel).

[1] This research was partially supported by the Fund for Scientific Research–Flanders (FWO–V), project "Orthogonal systems and their applications", grant #G.0278.97, project "Structured Matrices and their Applications", grant #G.0078.01, by the K.U. Leuven (Bijzonder Onderzoeksfonds), project "SLAP: Structured Linear Algebra Package", grant #OT/00/16, by the Belgian Programme on Interuniversity Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture. The scientific responsibility rests with the authors.

[2] This research has been partially supported by the grant of the Academy of Sciences of the Czech Republic No. A1030003.

[3] During the time this manuscript was written, the second author, Prof. Pták, passed away. Most of the ideas presented in this paper are his work. The other two authors are responsible for all the mistakes when formulating these ideas on paper.

PII: S 0 0 2 4 - 3 7 9 5 (0 1) 0 0 3 2 5 - 1

Keywords: Structured matrices; Bézout matrices; Hankel matrices; Finite and infinite companion matrices; Barnett-type formulas; Row reduced matrix polynomials

1. Introduction/motivation

Bézout and Hankel matrices belong to classes of structured matrices which can be defined based on scalar polynomials. They appear already in very classical works written by such famous authors as Sylvester, Cayley and Hermite (see [20]). The study of their properties and applications continued in fact during the whole 20th century, although the strongest interest has been from the 1970s until the present time. One of the most remarkable papers originating from the first half of the 21st century is Ref. [22]. This survey paper describes different methods to separate roots of polynomials and to solve other associated problems based on Bézout and Hankel matrices (in the language of corresponding quadratic forms). A nice brief historical survey can be found in the paper [19] (see also [32]). Other types of structured matrices connected with polynomials are Vandermonde, Toeplitz, Loewner, Sylvester, companion matrices, etc. Different connections between these classes of matrices were investigated. Stephen Barnett proved formulas relating the Bézout matrix with a polynomial in the companion matrix [3, Theorem. 1.12, p. 46] and, similarly, relating the Hankel matrix with a polynomial in the companion matrix.

Various applications, e.g., in interpolation theory and in control theory, showed that an important direction is the generalization of these classes of matrices to the case of matrix polynomials. These generalized classes keep most of the useful properties and mutual connections.

One of the oldest references where the generalization of Bézout matrices can be implicitly found is Ref. [14]. Not long after this the explicit definition of the Bézout matrix for the general case appeared in [1,16,17]. Another definition based on the Kronecker product can be found in [4]. In [30], Pták and Wimmer gave a Barnett factorization for the Bézoutian defined in this way. But it turned out that the properties of the Bézoutian could be easily generalized from the scalar to the matrix case when the definition of Anderson and Jury was taken.

Several properties and applications were generalized from the scalar to the matrix polynomial case. For example in [5], Bitmead et al. showed how to obtain greatest common right divisors from two matrix polynomials based on the generalized Bézout matrix. Generalizations of classical root-separation results for matrix polynomials can be found in [26]. In [28], the concept of a generalized Bézout matrix for several matrix polynomials was introduced. Based on this definition, a generalization of the Gohberg–Semencul formula for the inverse of a block Toeplitz matrix with nonsquare blocks was given in [13]. Even more abstract notions of Bézoutians were developed in the last two decades, e.g., Bézout matrices based on rational matrix functions, Bézout operators for analytic operator functions, etc. We refer the reader to [15,24,25] and the references therein.

The main goal of the present paper is to generalize Barnett formulas and some other properties to the matrix polynomial case. We could distinguish several levels of generalization depending on the restrictions put on the matrix polynomials. The case usually considered is when the matrix polynomials are monic, i.e., having their leading coefficient equal to the identity matrix. For this case, Barnett-type formulas can be easily obtained, similar to the scalar case.

In this paper, we shall concentrate our attention on structured matrices connected to row reduced matrix polynomials. The motivation for this choice and the reasons why we do not handle the case where the matrix polynomials have no special properties will be explained in Section 3.

The organization of the paper is as follows. In Section 2, we introduce some notation. In Section 3, we define the concept of row reducedness of a matrix polynomial. For a row reduced matrix polynomial we introduce the finite and infinite companion matrix. In the next two sections two other types of structured matrices are defined connected to row reduced matrix polynomials: the finite Hankel matrix and the finite Bézout matrix. These definitions differ slightly from the usual ones. All these matrices are also interpreted as the representation of different operators with respect to certain bases. In Section 6, we characterize Hankel and Bézout matrices by intertwining relations. This result is a direct generalization of a special case of common results of the second author and M. Fiedler, published in [6,7]. The operator interpretation as well as the generating function concept is used to derive the Barnett-type formulas of Section 7. In Section 8, we use these formulas to derive inversion formulas corresponding to our definition of finite Hankel and Bézout matrices. The Barnett-type formulas are also applied in Section 9 to compute mutual products of finite Hankel and Bézout matrices. Also these results extend ideas which arose in collaboration of the second-named author and M. Fiedler (see [6, Theorem 3.2.]). Special cases pointed out in Lemma 32 generalize popular formulas concerning connections between scalar Hankel and Bézout matrices which can be found, e.g., in [18]. For the matrix polynomial case, certain special cases of our product formulas can be found in [33, formulas (1.7) and (1.10)].

2. Notation

In the sequel, the following notation will be used:

- $\mathbf{F}$ denotes an arbitrary (finite or infinite) field.
- $\mathbf{F}^q[z]$ denotes the set of the vector polynomials, i.e., the polynomials with coefficients belonging to $\mathbf{F}^q$.
- Similarly $\mathbf{F}^{p\times q}[z]$ denotes the set of all $p \times q$ matrix polynomials. We will make no distinction between the set of matrix polynomials $\mathbf{F}^{p\times q}[z]$ and the set of polynomial matrices $\mathbf{F}[z]^{p\times q}$.
- $\mathbf{F}[[z^{-1}]]$ denotes the set of all formal power series in z^{-1} of the form $\sum_{k=0}^{\infty} f_k z^{-k}$. Similarly for vector and matrix power series.

- The coefficients of a matrix polynomial $D \in \mathbf{F}^{p\times q}[z]$ are denoted by D_k, i.e., $D(z) = \sum_k D_k z^k$ with $D_k = 0$ for $k > \deg D$.

3. Row reduced matrix polynomials, finite and infinite companion matrices

Our idea to deal with structured matrices corresponding to row reduced matrix polynomials has its origin in our preceding study of generalized companion matrices. In our paper [2], we introduced the concept of an *extended infinite companion matrix* of an arbitrary *nonsingular* matrix polynomial, i.e., a square matrix polynomial whose determinant is a nonzero polynomial. As shown in [2], this $\infty \times \infty$ matrix has only a finite number of nonzero rows. Only in the case of a row reduced matrix polynomial, these nonzero rows can be easily identified.

The *extended infinite companion matrix* is defined as the matrix representation of the remainder operator introduced in [8]. For more information on the division by a square nonsingular matrix polynomial, we refer the reader to [21].

We say that a rational function r is proper if

$$r(z) = \sum_{k=0}^{\infty} r_k z^{-k}, \qquad z \to \infty.$$

We say that a rational function r is strictly proper if

$$r(z) = \sum_{k=1}^{\infty} r_k z^{-k}, \qquad z \to \infty.$$

In these two cases, we will not make a distinction between the rational function and the corresponding formal power series. Vector or matrix rational functions are (strictly) proper if each component is (strictly) proper.

Any rational function r can be uniquely written as the sum of a polynomial part r_+ and a strictly proper part r_-. The operator Π_- applied to a rational functions r gives this strictly proper part

$$\Pi_- r = r_-, \qquad r = r_+ + r_-.$$

Definition 1 (*Remainder operator*, *remainder space*). For any nonsingular $D \in \mathbf{F}^{q\times q}[z]$, the *remainder operator* $R(D)$ on $\mathbf{F}^q[z]$ (connected with the left division by D) is defined as

$$R(D)p = D\Pi_- D^{-1} p. \tag{1}$$

The *remainder space* $\mathscr{R}(D)$ is defined as

$$\mathscr{R}(D) = \operatorname{Ran}(R(D)).$$

The remainder operator and remainder space, resp. appear in [8] as π_D and K_D, resp.

The extended infinite companion matrix will be the matrix representation of the remainder operator $R(D)$ with respect to the following basis.

Definition 2 (*Monomial basis*). In the space $\mathbf{F}^q[z]$ we choose the basis

$$\mathscr{B} = \{b_t\}_0^\infty \quad \text{where} \quad b_t(z) = e_k z^j \quad \text{if} \quad t = jq + k,\ 0 \leqslant k < q.$$

Based on the remainder operator, we recall now the definition of the infinite companion matrix. For the scalar case, this definition was given in [29].

Definition 3 (*Extended infinite companion matrix*). Assume that D is nonsingular. The $\infty \times \infty$ matrix representation of the operator $R(D)$ with respect to the basis $\mathscr{B}$ will be denoted by $\tilde{C}_\infty(D)$ and will be called the *extended infinite companion matrix* of the matrix polynomial D.

As shown in the paper [2], the extended infinite companion matrix has only a finite number of nonzero rows and its rank equals $n = \deg\det D$. It would be convenient to cut off the zero rows. However, only in the case when D is row reduced we obtain a straightforward way how all zero rows can be omitted.

Definition 4 (*Degree*, *row/column degree*, *highest row/column degree coefficient*). Let $D(z)$ be a $q \times q$ matrix polynomial. If we consider $D(z)$ as an element of $\mathbf{F}[z]^{q\times q}$, $D(z) = [d_{ij}(z)]$, then $\deg D(z) = \max_{ij} \deg d_{ij}(z)$. The *row degrees* are $h_i = \max_j \deg d_{ij}(z)$, $i = 0, \ldots, q-1$. The *highest row degree coefficient* is the $q \times q$ matrix $D_{hrdc} = [a_{ij}]$, where $d_{ij}(z) = a_{ij} z^{h_i} + \mathrm{O}(z^{h_i - 1})$, $z \to \infty$.

The column degrees and the highest column degree coefficient D_{hcdc} are defined analogously.

Definition 5. We say that $D(z)$ is *row reduced* if D_{hrdc} is nonsingular. Similarly, we say that $D(z)$ is *column reduced* if D_{hcdc} is nonsingular.

The definition of column and row reduced matrix polynomials was introduced in algebraic system theory by Wolovich [36]. However, its origin is even much older. The concept of a column reduced matrix polynomial occurs in the book [31, p. 49] as the "normal basis of an integral set".

When the matrix polynomial D is row reduced, the remainder subspace $\mathscr{R}(D)$ is a coordinate subspace with respect to the basis $\mathscr{B}$, i.e., a subspace spanned by n elements of the basis $\mathscr{B}$. This leads us to the following definitions:

Definition 6 (*Index set*). For any nonsingular $D(z)$ with row degrees h_i, we introduce the *index set* $\mathscr{I}(D)$ by

$$\mathscr{I}(D) = \{t \mid t = jq + i, 0 \leqslant j < h_i, 0 \leqslant i < q\}.$$

Lemma 7 (Remainder space for a row reduced D). *Let D be row reduced with row degrees h_i. Then the remainder space $\mathscr{R}(D)$ is spanned by $\{e_i z^j, 0 \leqslant j < h_i\} = \{b_t, t \in \mathscr{I}(D)\}$.*

This lemma occurs in [21] (Lemma 6.3-11 on p. 385).

Definition 8 (*Infinite companion matrix for a row reduced matrix polynomial*). If D is row reduced, then we introduce the *infinite companion matrix* $C_\infty(D)$ as the matrix representation of the operator $R(D) : \mathbf{F}^q[z] \to \mathscr{R}(D)$ with respect to the bases $\mathscr{B} \in \mathbf{F}^q[z]$ and $\{b_t \mid t \in \mathscr{I}(D)\}$ in $\mathscr{R}(D)$. Equivalently, $C_\infty(D) = \tilde{C}_\infty(D)_{(\mathscr{I}(D))}$ indicating the submatrix formed by the rows of $\tilde{C}_\infty(D)$ with indices belonging to the subset $\mathscr{I}(D)$.

We also define a generalization of the notion "finite companion matrix" (or "Frobenius matrix") to row reduced matrix polynomials. We introduce it as the matrix representation of the operator of multiplication by z modulo $D(z)$ on the space $\mathscr{R}(D)$, i.e., the compressed shift, introduced by Fuhrmann in [8, formula (4.1)] as the basic object in his theory of polynomial models. It is necessary to say that our requirement that the matrix polynomial has to be row reduced is dependent on our choice of the monomial basis $\mathscr{B}$. However, another choice of basis would bring another restriction and we believe that our choice is the most natural one.

Definition 9 (*Operator $S(D)$*). For any nonsingular D the operator $S(D) : \mathscr{R}(D) \mapsto \mathscr{R}(D)$ is defined by

$$S(D) = R(D)\mathbf{S}|_{\mathscr{R}(D)},$$

where $\mathbf{S}$ denotes the "shift operator" on the space $\mathbf{F}^q[z]$, given by $\mathbf{S}p(z) = zp(z)$ for any $p(z) \in \mathbf{F}^q[z]$.

If $D(z)$ is row reduced, the matrix of $S(D)$ w.r.t. the basis $\{b_t \mid t \in \mathscr{I}(D)\}$ equals

$$(C_\infty(D)S^q)^{(\mathscr{I}(D))},$$

where S is the "infinite shift matrix", $S = [\delta_{i,j+1}]_{i,j=0}^{\infty}$, and $M^{(\mathscr{I})}$ denotes the submatrix of M formed by columns of M with indices belonging to $\mathscr{I}$.

Definition 10 (*Finite companion matrix for a row reduced matrix polynomial*). The $n \times n$ matrix $C(D)$ defined as

$$C(D) = (C_\infty(D)S^q)^{(\mathscr{I}(D))}$$

is called the *finite companion matrix* of D.

4. Infinite and finite Hankel matrices and their operator interpretation

In this section, we first recall the standard notion of an infinite block Hankel matrix. If this infinite matrix has a finite rank, it can be written as a finite submatrix of the infinite block Hankel matrix multiplied to the left by the transpose of an infinite companion matrix and to the right by an infinite companion matrix. This finite submatrix will be called the finite Hankel matrix. Our definition does not necessarily keep the block Hankel structure but it will have the same properties as the finite block Hankel matrix defined in the classical way when the highest degree coefficients of the denominator matrix polynomials involved are nonsingular. For example, the finite Hankel matrix as defined here corresponding to a left and right coprime matrix fraction description is square and nonsingular (Remark 15).

Definition 11. The block matrix $H = \left(\Omega_{ij}\right)_{i,j=0}^{\infty}$ with blocks Ω_{ij} of dimension $p \times q$ is called *an infinite block Hankel matrix* if

$$\Omega_{ij} = \Omega_{i+j} .$$

For any formal power series $\Omega(z) = \sum \Omega_k z^{-k-1} \in z^{-1}\mathbf{F}^{p\times q}[[z^{-1}]]$, the corresponding Hankel matrix $H(\Omega)$ is defined as

$$H(\Omega) = (\Omega_{i+j})_{i,j=0}^{\infty} .$$

In the sequel we will often need stacking vectors of coefficients of vector polynomials and of strictly proper rational functions. We indicate these stacking vectors by a hat. For example, if a is a vector polynomial, then by $\hat{a}$ we mean the infinite column vector obtained by stacking the coefficients of the vector polynomial a and completing by zeros. Thus the value of a at the point z equals $a(z) = [I_q \; zI_q \; z^2 I_q \; \cdots]\hat{a}$. If b is a vector whose elements are strictly proper rational functions, the stacking vector $\hat{b}$ is defined so that $b(z) = [z^{-1}I_q \; z^{-2}I_q \; \cdots]\hat{b}$. Similarly, if we have a sequence of vectors $\{y_i\}$ having finite length, the stacking vector is denoted by $\hat{y} = [y_0^{\mathrm{T}}, y_1^{\mathrm{T}}, \ldots]^{\mathrm{T}}$.

Lemma 12 (The Hankel operator). *Define the operator* $\mathbf{H}(\Omega) : \mathbf{F}^q[z] \mapsto z^{-1}\mathbf{F}^p[[z^{-1}]]$

$$\mathbf{H}(\Omega) : p \mapsto \Pi_{-}\Omega p, \tag{2}$$

where $\Omega(z) = \sum \Omega_k z^{-k-1} \in z^{-1}\mathbf{F}^{p\times q}[[z^{-1}]]$. *Then*

$$H(\Omega)\hat{p} = \hat{\rho}$$

if and only if the vector polynomial $p(z) \in \mathbf{F}^q[z]$ *corresponding to the stacking vector* $\hat{p}$ *is mapped by* $\mathbf{H}(\Omega)$ *onto the formal power series* $\rho(z) = \sum \rho_k z^{-k-1} \in z^{-1}\mathbf{F}^p[[z^{-1}]]$ *represented by the stacking vector* $\hat{\rho}$:

$$\mathbf{H}(\Omega)p = \rho .$$

Remark 13. It is a well-known fact that rank $(H(\Omega)) < \infty$ if and only if Ω defines a rational function. For the scalar case, this assertion appears in [12, Theorem 8,

p. 207] and is generally attributed to Kronecker. The block case follows immediately because the (block) Hankel matrix contains the scalar Hankel matrices corresponding to all separate entries of Ω.

To motivate our Definition 16 below, we recall a factorization formula for Hankel matrices from our paper [2].

Lemma 14. *Let* $\Omega(z) = \hat{D}_1^{-1}(z)\hat{N}_1(z) = \hat{N}_2(z)\hat{D}_2^{-1}(z)$ *be (not necessarily coprime) left and right matrix fraction descriptions of a strictly proper rational matrix function* Ω *and suppose that* $\hat{D}_1^{\mathrm{T}},\ \hat{D}_2$ *are row reduced. Denote* $m = \deg\det\hat{D}_1,\ n = \deg\det\hat{D}_2$. *Then*

$$H(\Omega) = H = \left[C_\infty(\hat{D}_1^{\mathrm{T}})\right]^{\mathrm{T}} H_{m,n} C_\infty(\hat{D}_2), \tag{3}$$

where $H_{m,n} = H^{(\mathscr{I}(\hat{D}_2))}_{(\mathscr{I}(\hat{D}_1^{\mathrm{T}}))}$ *is an* $m \times n$ *submatrix of* H.

Remark 15. Note that if both matrix fractions are coprime, then $m = n$ and $H_{m,n}$ is a square nonsingular matrix.

We know that in the scalar case a formula similar to (3) holds:

$$H(\omega) = H(h/f) = \left[C_\infty(f)\right]^{\mathrm{T}} H_{n,n} C_\infty(f),$$

where $n = \deg f$ and the matrix $H_{n,n}$ in the middle (which is the leading submatrix of $H(h/f)$ of order n and is nonsingular if h and f are coprime) is the classical scalar finite Hankel matrix. The following definition keeps this analogy.

Definition 16. Given a strictly proper rational matrix function Ω and its left and right matrix fraction descriptions

$$\Omega(z) = \hat{D}_1^{-1}(z)\hat{N}_1(z) = \hat{N}_2(z)\hat{D}_2^{-1}(z)$$

with $\hat{D}_1^{\mathrm{T}},\ \hat{D}_2$ row reduced, we call the matrix

$$H_{m,n}(\Omega) = H(\Omega)^{(\mathscr{I}(\hat{D}_2))}_{(\mathscr{I}(\hat{D}_1^{\mathrm{T}}))}$$

the *finite Hankel matrix* (corresponding to Ω and the choice of $\hat{D}_1,\ \hat{D}_2$).

5. Bézout matrices and their operator interpretation

Like for the Hankel matrix, we shall define the finite Bézout matrix corresponding to row reduced matrix polynomials in a way keeping analogous properties to the scalar case but disturbing the block Bézout structure.

The definition can be explained easier if we start by an infinite extension of the Bézout matrix. If we assume row reducedness, the number of nonzero rows, resp. columns in the Bézout matrix will be equal to $\deg\det D_1$, resp. $\deg\det D_2$ (see (4) for details). Our finite Bézout matrix is obtained by deleting all zero rows and columns. Like in the Hankel case, this matrix keeps important rank properties known from the scalar case. Especially, if the corresponding rational function $\Omega = D_1^{-1}N_1 = N_2D_2^{-1}$ is proper and its left and right matrix fraction representations are coprime, then the Bézout matrix is square of order $\deg\det D_1 = \deg\det D_2$ and nonsingular.

Following the original definition of Anderson and Jury [1], we define the infinite Bézout matrix as follows.

Definition 17. Suppose that D_1, N_2, N_1, and D_2 are matrix polynomials of dimensions $q \times q$, $q \times r$, $q \times r$, and $r \times r$ such that

$$D_1(z)N_2(z) = N_1(z)D_2(z).$$

Then the *infinite Bézout matrix* (but having a finite number of nonzero entries only) is the matrix

$$\tilde{B}(D_1, N_2, N_1, D_2) = (B_{ij})_{i,j=0}^{\infty}$$

with blocks $B_{ij} \in \mathbf{F}^{q\times r}$ such that

$$\sum_{i,j=0}^{\infty} B_{ij}z^iy^j = \frac{D_1(z)N_2(y) - N_1(z)D_2(y)}{z-y} = B(z, y).$$

If the rational function $D_1^{-1}(z)N_1(z) = N_2(z)D_2^{-1}(z)$ is proper, we can derive a finite matrix by cutting off a zero part based on the degrees of the denominator polynomials D_1 and D_2. (In [35], Wimmer drops the properness condition.) Let $\deg D_1 = \delta_1$ and $\deg D_2 = \delta_2$, then in [1] a finite Bézout matrix is defined as

$$(B_{ij})_{i=0,j=0}^{\delta_1-1,\delta_2-1}.$$

However, if D_2^{T}, D_1 are row reduced, *all zero rows and columns* of the infinite Bézout matrix can be deleted.

Hence, we define the *finite Bézout matrix* $B_{l,m}$ as

$$B_{l,m}(D_1, N_2, N_1, D_2) = \tilde{B}(D_1, N_2, N_1, D_2)_{(\mathscr{I}(D_1))}^{(\mathscr{I}(D_2^{\mathrm{T}}))} \tag{4}$$

$(l = \deg\det D_1,\ m = \deg\det D_2)$.

Notation. We write $B(D)$ for $B(D, I, I, D)$.

Next we are going to derive an operator interpretation for the Bézout matrix. Our interpretation will be based on the concept of generating function.

Definition 18. The generating function $B(z, y)$ corresponding to an $\infty \times \infty$ block matrix $B = [b_{i,j}]_{i,j=0}^{\infty}$, where the blocks $b_{i,j}$ are $q \times r$, is the formal power series in two variables z and y defined as

$$B(z, y) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} b_{i,j} z^i y^j.$$

For any generating function, we have the following operator interpretation.

Lemma 19. *For any $q \times r$ generating function $B(z, y)$ the corresponding matrix B with a finite number of nonzero blocks $b_{i,j}$ represents the operator*

$$\mathbf{B} : z^{-1}\mathbf{F}^r[[z^{-1}]] \mapsto \mathbf{F}^q[z] : \quad \mathbf{B}(\omega) = \langle B(z, y)\omega(y)\rangle_{y^{-1}}$$

for any $\omega \in z^{-1}\mathbf{F}^r[[z^{-1}]]$, where $\langle \cdots \rangle_{y^{-1}}$ denotes the coefficient connected to y^{-1} in the given function.

An operator interpretation for the Bézout matrix can be taken as a special case of Lemma 19. In [9, Eq. (4.5)], Fuhrmann gives another operator interpretation for the Bézout matrix as defined by Anderson and Jury in [1]. He defines the map $\mathbf{Z}$: $\mathscr{R}(D_2) \mapsto \mathscr{R}(D_1)$ as

$$\mathbf{Z}p = R(D_1)N_1 p \qquad \text{for } p \in \mathscr{R}(D_2).$$

The block matrix representation of this map $\mathbf{Z}$ with respect to certain bases leads to the Bézout matrix. Based on the intertwining property

$$\mathbf{Z}S(D_2) = S(D_1)\mathbf{Z}$$

several of the following results can be proven in an alternative way. [4] Note that the relations described in [34] can easily be generalized to relations between the Bézout matrix and the Hankel matrix as it is defined in this paper.

6. Intertwining characterizations of Hankel and Bézout matrices

We shall show that a matrix belongs to the class of all Hankel (Bézout, resp.) matrices if and only if it satisfies certain intertwining relations. This result has various theoretical applications. Among others, it implies the inversion theorem for Hankel and Bézout matrices and the results of Section 9 below.

We start by an obvious fact:

Lemma 20 (Characterization of infinite Hankel matrices). *An infinite matrix H is an infinite Hankel matrix with blocks of dimension $p \times q$ if and only if H satisfies the intertwining relation*

[4] We thank the referee for pointing our attention to this alternative operator interpretation and the possibility to prove some subsequent results based on this interpretation. For the scalar case, we refer the reader to the book [11].

$$(S^{\mathrm{T}})^p H = H S^q,$$

where S is the matrix of the shift operator with respect to the basis $\mathscr{B}$.

Theorem 21 (Characterization of finite Hankel matrices). *Let $\hat{D}_1^{\mathrm{T}},\ \hat{D}_2$ be row reduced, $\deg\det\hat{D}_1 = m,\ \deg\det\hat{D}_2 = n$ and $H_{m,n}$ an arbitrary matrix of dimension $m \times n$. Then the following assertions are equivalent*:
(i) *There exist $\hat{N}_1,\ \hat{N}_2$ such that*

$$\hat{D}_1^{-1}\hat{N}_1 = \hat{N}_2\hat{D}_2^{-1} \qquad (\textit{strictly proper})$$

and $H_{m,n}$ is the corresponding finite Hankel matrix.
(ii) *$H_{m,n}$ satisfies the intertwining relation*

$$C^{\mathrm{T}}(\hat{D}_1^{\mathrm{T}})H_{m,n} = H_{m,n}C(\hat{D}_2). \tag{5}$$

Proof. (i) → (ii) Using Lemma 14 we write the infinite Hankel matrix $H(\hat{D}_1^{-1}\hat{N}_1) = H(\hat{N}_2\hat{D}_2^{-1})$ in the form (3). We apply the intertwining relation concerning the infinite companion matrix,

$$C_\infty(D)S^q = C(D)C_\infty(D)\,, \tag{6}$$

given in [2]. We also use Lemma 20 to get

$$\begin{aligned}
&\left[C_\infty(\hat{D}_1^{\mathrm{T}})\right]^{\mathrm{T}} C(\hat{D}_1^{\mathrm{T}})H_{m,n}C_\infty(\hat{D}_2) = \left(S^{\mathrm{T}}\right)^p \left[C_\infty(\hat{D}_1^{\mathrm{T}})\right]^{\mathrm{T}} H_{m,n}C_\infty(\hat{D}_2)\\
&\quad = \left(S^{\mathrm{T}}\right)^p H = HS^q = \left[C_\infty(\hat{D}_1^{\mathrm{T}})\right]^{\mathrm{T}} H_{m,n}C_\infty(\hat{D}_2)S^q\\
&\quad = \left[C_\infty(\hat{D}_1^{\mathrm{T}})\right]^{\mathrm{T}} H_{m,n}C(\hat{D}_2)C_\infty(\hat{D}_2).
\end{aligned}$$

Since the infinite companion matrices $\left[C_\infty(\hat{D}_1^{\mathrm{T}})\right]^{\mathrm{T}},\ C_\infty(\hat{D}_2)$ are full rank matrices, the equality (5) follows.

(ii) → (i) Starting from (5) and proceeding in the opposite direction, we easily verify the relation

$$\left[C_\infty(\hat{D}_1^{\mathrm{T}})\right]^{\mathrm{T}} H_{m,n}C_\infty(\hat{D}_2)S^q = \left(S^{\mathrm{T}}\right)^p \left[C_\infty(\hat{D}_1^{\mathrm{T}})\right]^{\mathrm{T}} H_{m,n}C_\infty(\hat{D}_2)\,.$$

It follows that

$$\left[C_\infty(\hat{D}_1^{\mathrm{T}})\right]^{\mathrm{T}} H_{m,n}C_\infty(\hat{D}_2) = H \tag{7}$$

is an infinite Hankel matrix with blocks of dimension $p \times q$ which has a finite rank. We conclude that $H = H(\Omega)$ where Ω is rational. Suppose that $\Omega = N_0D_0^{-1}$ where $N_0,\ D_0$ are right coprime. Then, with respect to the operator interpretation (see, e.g., [21, Lemma 6.6-1, p. 471]),

$$\operatorname{Ker}\mathbf{H}\,(\Omega) = D_0\mathbf{F}^q[z]\,.$$

On the other hand, (7) shows that

$$\operatorname{Ker}\mathbf{H}(\Omega) \supset \operatorname{Ker}\mathscr{R}(\hat{D}_2) = \hat{D}_2\mathbf{F}^q[z].$$

The inclusion $D_0\mathbf{F}^q[z] \supset \hat{D}_2\mathbf{F}^q[z]$ is equivalent with the relation $\hat{D}_2 = D_0M$ for some nonsingular matrix polynomial M so that $\Omega = (N_0M)\hat{D}_2^{-1}$ and we put $\hat{N}_2 = N_0M$. The left matrix fraction for Ω can be found similarly. $\square$

It remains to state a similar assertion for Bézout matrices.

Theorem 22 (Characterization of finite Bézout matrices). *Let* D_2^{T}, D_1 *be row reduced*, $\deg\det D_1 = l$, $\deg\det D_2 = m$ *and let* $B_{l,m}$ *be an arbitrary matrix of dimension* $l \times m$. *Then the following two assertions are equivalent*:
(i) *There are* N_1, N_2 *such that the rational functions* $D_1^{-1}N_1$, $N_2D_2^{-1}$ *are proper and equal to each other and*

$$B_{l,m} = B_{l,m}(D_1, N_2, N_1, D_2).$$

(ii) *The matrix* $B_{l,m}$ *satisfies the intertwining relation*

$$C(D_1)B_{l,m} = B_{l,m}C^{\mathrm{T}}(D_2^{\mathrm{T}}).$$

Proof. The generating function $L_1(z, y)$ of $C(D_1)B_{l,m}$ has the following form:

$$L_1(z, y) = zB_{l,m}(z, y) - D_1(z)Q_1(y).$$

Similarly, the generating function $L_2(z, y)$ of $B_{l,m}C^{\mathrm{T}}(D_2^{\mathrm{T}})$ has the form

$$L_2(z, y) = yB_{l,m}(z, y) - Q_2^{\mathrm{T}}(z)D_2(y).$$

($\Rightarrow$) Proving the intertwining relation is equivalent to showing the equality of the generating functions of the left-hand and right-hand sides. The matrix polynomial $Q_1(y)$ can be determined as follows. By definition of the Bézout matrix, we derive

$$L_1(z, y) = z\left(\frac{D_1(z)N_2(y) - N_1(z)D_2(y)}{z - y}\right) - D_1(z)Q_1(y).$$

Hence,

$$(z - y)L_1(z, y) = z(D_1(z)N_2(y) - N_1(z)D_2(y)) - (z - y)D_1(z)Q_1(y).$$

By multiplying both sides to the left by $D_1^{-1}(z)$, we get

$$\begin{aligned}&(z - y)D_1^{-1}(z)L_1(z, y)\\ &\quad = z(N_2(y) - D_1^{-1}(z)N_1(z)D_2(y)) - (z - y)Q_1(y).\end{aligned} \tag{8}$$

Because

$$D_1^{-1}(z)L_1(z, y) = O_-(z^{-1}), \qquad D_1^{-1}(z)N_1(z) = A + O_-(z^{-1})$$

with A a constant matrix, taking the coefficient of z of both sides of (8) gives us

$$0 = N_2(y) - AD_2(y) - Q_1(y)$$

or

$$N_2(y) = AD_2(y) + Q_1(y).$$

We get a similar relation for the matrix polynomial $Q_2^{\mathrm{T}}(z)$:

$$N_1(z) = D_1(z)A + Q_2^{\mathrm{T}}(z).$$

We have to prove now that $L_1(z, y) = L_2(z, y)$.

$$\begin{aligned}
L_1(z, y) &= zB_{l,m}(z, y) - D_1(z)\,(N_2(y) - AD_2(y)) \\
&= yB_{l,m}(z, y) + (z - y)B_{l,m}(z, y) - D_1(z)N_2(y) + D_1(z)AD_2(y) \\
&= yB_{l,m}(z, y) + D_1(z)N_2(y) - N_1(z)D_2(y) - D_1(z)N_2(y) \\
&\qquad + D_1(z)AD_2(y) \\
&= yB_{l,m}(z, y) + (D_1(z)A - N_1(z))D_2(y) \\
&= yB_{l,m}(z, y) - Q_2^{\mathrm{T}}(z)D_2(y) \\
&= L_2(z, y).
\end{aligned}$$

This proves the first part of the theorem.

($\Leftarrow$) We start from the fact that

$$L_1(z, y) = L_2(z, y).$$

Hence,

$$(z - y)B_{l,m}(z, y) = D_1(z)Q_1(y) - Q_2^{\mathrm{T}}(z)D_2(y)$$

or

$$B_{l,m}(z, y) = \frac{D_1(z)Q_1(y) - Q_2^{\mathrm{T}}(z)D_2(y)}{z - y}.$$

Note that the row degrees of $Q_1^{\mathrm{T}}(y)$ are smaller than the corresponding row degrees of $D_2^{\mathrm{T}}(y)$. Similarly, the row degrees of $Q_2^{\mathrm{T}}(z)$ are smaller than the corresponding row degrees of $D_1(z)$. Hence, we can take $N_1 = Q_2^{\mathrm{T}}$ and $N_2 = Q_1$. This proves the other part of the theorem. □

Theorems 21 and 22 give us a tool to show that, under certain assumptions, the alternating product of Hankel and Bézout matrices starting and ending with a Hankel matrix (with a Bézout matrix resp.) is again a Hankel matrix (a Bézout matrix). On the other hand, Barnett formulas give another tool to prove the same and even a quantitative version of this assertion (describing connections between the corresponding polynomials). This motivates us to include first Barnett formulas and then a section concerning the products.

7. Barnett formulae

Like the classical scalar-case Barnett formulas, the following theorem relates a generalized Hankel matrix based on a numerator and a denominator matrix polynomial and the Hankel matrix only based on the denominator matrix polynomial, e.g., $H(\hat{D}_1^{-1}\hat{N}_1)$ and $H(\hat{D}_1^{-1})$. The sum expressions in the formulas reduce to polynomials in the finite companion matrix if the scalar case is considered.

Theorem 23. *Let* $\hat{D}_1,\ \hat{N}_1,\ \hat{D}_2,\ \hat{N}_2$ *be matrix polynomials*, $\hat{D}_1^{\mathrm{T}},\ \hat{D}_2$ *square and row reduced*, $\deg\det\hat{D}_1 = m,\ \deg\det\hat{D}_2 = n$, *and let*

$$\hat{D}_1^{-1}\hat{N}_1 = \hat{N}_2\hat{D}_2^{-1} \qquad (\textit{strictly proper}).$$

Then:

$$H(\hat{D}_1^{-1}\hat{N}_1) = H(\hat{D}_1^{-1})\sum_k \tilde{C}^k(\hat{D}_1)\tilde{C}_\infty(\hat{D}_1)\mathrm{diag}\,((\hat{N}_1)_k), \tag{9}$$

$$H_{m,n}(\hat{D}_1^{-1}\hat{N}_1) = H_{m,m}(\hat{D}_1^{-1})\sum_k C^k(\hat{D}_1)C_\infty(\hat{D}_1)\left[\mathrm{diag}\,((\hat{N}_1)_k)\right]^{(\mathscr{I}(\hat{D}_2))}, \tag{10}$$

$$H(\hat{N}_2\hat{D}_2^{-1}) = \sum_k \mathrm{diag}\,((\hat{N}_2)_k)\left[\tilde{C}_\infty(\hat{D}_2^{\mathrm{T}})\right]^{\mathrm{T}}\left[\tilde{C}^k(\hat{D}_2^{\mathrm{T}})\right]^{\mathrm{T}} H(\hat{D}_2^{-1})\,, \tag{11}$$

$$\begin{aligned} H_{m,n}(\hat{N}_2\hat{D}_2^{-1}) = \sum_k \left[\mathrm{diag}\,((\hat{N}_2)_k)\right]_{(\mathscr{I}(\hat{D}_1^{\mathrm{T}})}\left[C_\infty(\hat{D}_2^{\mathrm{T}})\right]^{\mathrm{T}}\left[C^k(\hat{D}_2^{\mathrm{T}})\right]^{\mathrm{T}} \\ \times H_{n,n}(\hat{D}_2^{-1})\,. \end{aligned} \tag{12}$$

Proof. It is sufficient to prove the first formula (9). Formula (10) follows just by taking corresponding submatrices and omitting the zero part of the matrix $\tilde{C}(\hat{D}_1)$.

Formulas (11) and (12) are obtained as an application of (9) and (10) to transposed matrices.

To prove (9), we apply the operator represented by the r.h.s. to a vector polynomial $p(z)$ (which is described by a stacking vector $\hat{p}$). Due to (6),

$$\sum_k \tilde{C}^k(\hat{D}_1)\tilde{C}_\infty(\hat{D}_1)\mathrm{diag}\,((\hat{N}_1)_k)\hat{p} = \tilde{C}_\infty(\hat{D}_1)\sum_k \left(S^q\right)^k \mathrm{diag}\,((\hat{N}_1)_k)\hat{p}$$

which evidently represents the remainder r under division of $\hat{N}_1 p$ by $\hat{D}_1$:

$$\hat{N}_1 p = \hat{D}_1 q + r\,.$$

Due to Lemma 12,

$$\mathbf{H}\,(\hat{D}_1^{-1})r = \Pi_-\hat{D}_1^{-1}r = \hat{D}_1^{-1}r = \hat{D}_1^{-1}\hat{N}_1 p - q = \Pi_-\hat{D}_1^{-1}\hat{N}_1 p.$$

Using Lemma 12 once more, we see that this corresponds to $H(\hat{D}_1^{-1}\hat{N}_1)\hat{p}$. □

Now, a similar theorem on Bézout matrices follows. It gives the connections between a general Bézout matrix based on four matrix polynomials and the Bézout matrix only based on one of the denominator matrix polynomials.

Theorem 24. *Let* D_1, N_1, D_2, N_2 *be matrix polynomials*, D_2^{T}, D_1 *square and row reduced*, $\deg\det D_1 = l,\ \deg\det D_2 = m$, *and let*

$$D_1^{-1}N_1 = N_2 D_2^{-1}.$$

Then:

$$\tilde{B}(D_1, N_2, N_1, D_2) = \sum_k \tilde{C}^k(D_1)\tilde{C}_\infty(D_1)\mathrm{diag}\ ((N_1)_k)\ \tilde{B}(D_2), \tag{13}$$

$$\begin{aligned} &B_{l,m}(D_1, N_2, N_1, D_2) \\ &\quad = \sum_k C^k(D_1)C_\infty(D_1)\left[\mathrm{diag}\ ((N_1)_k)\right]^{(\mathscr{I}(D_2^{\mathrm{T}}))} B_{m,m}(D_2), \end{aligned} \tag{14}$$

$$\begin{aligned} &\tilde{B}(D_1, N_2, N_1, D_2) \\ &\quad = \tilde{B}(D_1)\sum_k \mathrm{diag}\ ((N_2)_k)\left[\tilde{C}_\infty(D_2^{\mathrm{T}})\right]^{\mathrm{T}}\left[\tilde{C}^k(D_2^{\mathrm{T}})\right]^{\mathrm{T}}, \end{aligned} \tag{15}$$

$$\begin{aligned} &B_{l,m}(D_1, N_2, N_1, D_2) \\ &\quad = B_{l,l}(D_1)\sum_k \left[\mathrm{diag}\ ((N_2)_k)\right]_{(\mathscr{I}(D_1))}\left[C_\infty(D_2^{\mathrm{T}})\right]^{\mathrm{T}}\left[C^k(D_2^{\mathrm{T}})\right]^{\mathrm{T}}. \end{aligned} \tag{16}$$

Proof. Again, it is sufficient to prove the first formula (13). The others are derived similarly as in the Hankel case.

The proof of (13) will be based on a formula for the generating function. We use the equality $N_1(z)D_2(z) = D_1(z)N_2(z)$:

$$\begin{aligned} \frac{D_1(z)N_2(y) - N_1(z)D_2(y)}{z-y} &= D_1(z)\frac{N_2(y) - N_2(z)}{z-y} \\ &\quad + N_1(z)\frac{D_2(z) - D_2(y)}{z-y}. \end{aligned}$$

Considering z as a variable and y as a parameter, we conclude that the right-hand side is the remainder in division of

$$N_1(z)\frac{D_2(z) - D_2(y)}{z-y}$$

by $D_1(z)$. Since

$$\frac{D_2(z) - D_2(y)}{z-y}$$

is the generating function of $\tilde{B}(D_2)$, we obtain the matrix equality

$$\begin{aligned}\tilde{B}(D_1, N_2, N_1, D_2) &= \tilde{C}_\infty(D_1) \sum_k \left(S^q\right)^k \operatorname{diag}\ ((N_1)_k)\ \tilde{B}(D_2) \\ &= \sum_k \tilde{C}^k(D_1)\tilde{C}_\infty(D_1)\operatorname{diag}\ ((N_1)_k)\ \tilde{B}(D_2)\end{aligned}$$

(we used (6) again) and (13) is proved. □

8. Inversion of Hankel and Bézout matrices

It is a well-known fact that the inverse of any scalar nonsingular Hankel matrix is a Bézout matrix and vice versa. We refer to the paper of Lander [23]. It is also well known that the same fact can be generalized to the case of block matrices (the corresponding matrix polynomials satisfy some monicness conditions [10,13,14,27,33]). It is natural to put the question whether an analogous property holds also for our generalized definition, related to row reduced matrix polynomials.

In fact, intertwining characterizations themselves, given in Section 6, show that the positive answer is true. We give a quantitative result (describing the connection of the corresponding polynomials), based on the Barnett-type formulas of the previous section.

Let us start by a simple lemma:

Lemma 25.
(i) *For any square nonsingular* D,

$$\tilde{C}_\infty(D) = \tilde{B}(D)H(D^{-1}).$$

(ii) *For* D *row reduced*, $\deg\det D = m$,

$$B_{m,m}(D)H_{m,m}(D^{-1}) = I_m.$$

This is in fact Theorem 8 of [2].

The next theorem will be important both for the inversion formulas of this section and for the product formulas given in the next section.

Theorem 26. *Let* D_1, D_2^{T}, D_3 *be row reduced*, *suppose that* $D_1^{-1}N_1 = N_2D_2^{-1}$ *and* $D_2^{-1}\hat{N}_2 = \hat{N}_3D_3^{-1}$. *Then the matrix*

$$P = B_{l,m}(D_1, N_2, N_1, D_2)H_{m,n}(D_2^{-1}\hat{N}_2 = \hat{N}_3D_3^{-1})$$

satisfies the intertwining relation

$$C(D_1)P = PC(D_3)\,. \tag{17}$$

Moreover,

$$P = \sum_k C^k(D_1)C_\infty(D_1)\left[\operatorname{diag}\ ((N)_k)\right]^{(\mathscr{I}(D_3))},$$

where $N(z) = N_1(z)\hat{N}_2(z)$.

Remark 27. This means in fact that P represents consecutive multiplication by $\hat{N}_2$ and by N_1 and then taking remainder under division by D_1.

Proof. Relation (17) is an obvious consequence of intertwining characterizations of Hankel and Bézout matrices.

To prove the expression for P, we substitute first from the corresponding Barnett formulas to get $B_{m,m}(D_2)H_{m,m}(D_2^{-1}) = I_m$ (see Lemma 25) in the middle. Then apply the matrix

$$P = \sum_k C^k(D_1)C_\infty(D_1)\left[\operatorname{diag}((N_1))_k\right]^{(\mathscr{I}(D_2^{\mathrm{T}}))} \times \sum_j C^j(D_2)C_\infty(D_2)\left[\operatorname{diag}((\hat{N}_2))_j\right]^{(\mathscr{I}(D_3))}$$

to a vector polynomial $p \in \mathscr{R}(D_3)$:

$$\hat{N}_2 p = D_2 q' + r'$$

$$N_1 r' = D_1 q + r$$

$$N_1 \hat{N}_2 p = D_1 N_2 q' + D_1 q + r$$

and $D_1^{-1} r$ is strictly proper. From this, both assertions of the theorem are immediately clear. □

The next theorem concerning the rank of the Bézout matrix was already proven by Anderson and Jury [1, Theorem 2.1]. See also [9, Corol. 4.5].

Theorem 28 (Rank of the Bézout matrix). *Let* $D_1,\ D_2^{\mathrm{T}}$ *be square and nonsingular,* $\deg\det D_1 = l,\ \deg\det D_2 = m$ *and let the functions* $D_1^{-1}(z)N_1(z) = N_2(z)D_2^{-1}(z)$ *be proper. Then we can compute the rank of the Bézout matrix*:

$$\operatorname{rank} \tilde{B}(D_1, N_2, N_1, D_2) = \mu,$$

where μ *is the degree of the determinant of the denominator matrix polynomial of any (right or left) coprime fraction defining the rational function* $\Omega = N_2 D_2^{-1} = D_1^{-1} N_1$.

Proof. If $D_1 = P_1 D_1',\ N_1 = P_1 N_1',\ D_2 = D_2' P_2,\ N_2 = N_2' P_2$ such that $D_1',\ N_1'$ are left coprime, $D_2',\ N_2'$ are right coprime, then

$$B(z, y) = P_1(z)B'(z, y)P_2(y).$$

In matrix representation,

$$\tilde{B}(D_1, N_2, N_1, D_2) = M(P_1)\tilde{B}(D_1', N_2', N_1', D_2')\,(M(P_2))^{\mathrm{T}}$$

and, cutting out the zero parts at the right, we get

$$\begin{aligned}&\tilde{B}(D_1, N_2, N_1, D_2)\\&\quad = M(P_1)^{(\mathscr{I}(D_1'))} B_{\mu,\mu}(D_1', N_2', N_1', D_2') \left(M(P_2^{\mathrm{T}})^{\mathrm{T}}\right)_{(\mathscr{I}(D_2'^{\mathrm{T}}))},\end{aligned}$$

where $M(P_1)$ represents the multiplication by the polynomial matrix $P_1(z)$ to the left of a polynomial vector. Similarly for $M(P_2^{\mathrm{T}})^{\mathrm{T}}$. The Bézout matrix in the r.h.s. of the second equality is square of order μ. Thus the rank of $\tilde{B}(D_1, N_2, N_1, D_2)$ is at most μ. The equality will be a consequence of the following Theorem 29. $\square$

Theorem 29 (Inversion of Bézout matrices). *Suppose D_1, D_2^{T} row reduced and N_1, N_2 such that the matrix fractions $D_1^{-1}N_1$ and $N_2D_2^{-1}$ are coprime, proper and define the same rational matrix function. Then the Bézout matrix $B_{m,m}(D_1, N_2, N_1, D_2)$ with $m = \det D_1 = \det D_2$ is square nonsingular and its inverse is a Hankel matrix. More specifically,*

$$B_{m,m}(D_1, N_2, N_1, D_2)^{-1} = H_{m,m}(D_2^{-1}\hat{N}_2 = \hat{N}_1 D_1^{-1}),$$

where $\hat{N}_1$ is the unique solution of

$$\hat{N}_1 N_2 + Q_1 D_2 = I \tag{18}$$

(*for unknown matrix polynomials $\hat{N}_1$, Q_1*) *s.t.*

$$\hat{N}_1 D_1^{-1} \tag{19}$$

is strictly proper and $\hat{N}_2$ is the unique solution of

$$N_1 \hat{N}_2 + D_1 Q_2 = I \tag{20}$$

(*for unknown $\hat{N}_2$, Q_2*) *s.t. $D_2^{-1}\hat{N}_2$ is strictly proper.*

Proof. We prove first existence and unicity of $\hat{N}_1$.

Since N_2, D_2 are right coprime and N_1 and D_1 are left coprime, there exist matrix polynomials A, B, E and F s.t.

$$U = \begin{bmatrix} N_2 & A \\ D_2 & B \end{bmatrix}$$

is unimodular and U^{-1} is equal to

$$U^{-1} = \begin{bmatrix} E & F \\ D_1 & -N_1 \end{bmatrix}.$$

Hence, all solutions of

$$XN_2 + YD_2 = I$$

can be written as

$$X = E + VD_1, \qquad Y = F - VN_1$$

where V can be any matrix polynomial. Writing $E = QD_1 + R$ s.t. RD_1^{-1} is strictly proper, we obtain the unique solution of (18) by putting V equal to $-Q$.

Similarly existence and uniqueness of $\hat{N}_2$ can be proved.

We now apply Remark 27 to the product

$$P = B_{m,m}(D_1, N_2, N_1, D_2)H_{m,m}(D_2^{-1}\hat{N}_2 = \hat{N}_1 D_1^{-1}) .$$

We conclude that P represents multiplication by

$$N_1\hat{N}_2 = I - D_1 Q_2$$

on the space $\mathscr{R}(D_1)$, followed by taking the remainder under left division by D_1. This proves that $P = I_m$ ($m = \deg\det D_1$) and the proof of Theorem 29 is completed. □

Theorem 30. *Suppose D_2^{T}, D_1 row reduced and $\hat{N}_1$, $\hat{N}_2$ such that the matrix fractions $D_2^{-1}\hat{N}_2$ and $\hat{N}_1 D_1^{-1}$ are coprime, strictly proper and define the same rational matrix function Ω. Then the finite Hankel matrix corresponding to Ω and the choice of denominators D_1, D_2 is square nonsingular and its inverse is the Bézout matrix $B_{m,m}(D_1, N_2, N_1, D_2)$ with $m = \det D_1 = \det D_2$ and where*

$$\hat{N}_2 N_1 + D_2 Q_1 = I , \tag{21}$$

$$N_2 \hat{N}_1 + Q_2 D_1 = I \tag{22}$$

such that both $D_1^{-1}N_1$ and $N_2 D_2^{-1}$ are strictly proper.

Proof. Existence and unicity of N_1, N_2 can be proved in the same way as in Theorem 29. If (22) holds then

$$\begin{aligned} &N_2 D_2^{-1} D_2 \hat{N}_1 + Q_2 D_1 = I , \\ &D_1^{-1} N_1 \hat{N}_2 D_1 + Q_2 D_1 = I , \\ &N_1 \hat{N}_2 + D_1 Q_2 = I . \end{aligned}$$

Then it remains to use Theorem 26 to prove that

$$B(D_1, N_2, N_1, D_2)H(D_2^{-1}\hat{N}_2 = \hat{N}_1 D_1^{-1}) = I$$

(in the same way as in Theorem 29). □

9. Products of Hankel and Bézout matrices

In this section, we show that under certain conditions the alternating product of Hankel and Bézout matrices leads to a finite Hankel or Bézout matrix.

Theorem 31 (Alternating products of Hankel and Bézout matrices). *Let* $D_0, D_1, \dots, D_{2l+2},\ N_1, N_2, \dots, N_{2l+2},\ \hat{N}_0, \hat{N}_1, \dots, \hat{N}_{2l+1}$ *be matrix polynomials such that all following assumptions hold*:

(a) $D_0^{\mathrm{T}},\ D_1,\ D_2^{\mathrm{T}},\ D_3, \dots, D_{2l+1},\ D_{2l+2}^{\mathrm{T}}$ *are square and row reduced. Denote* $\deg\det D_k = m_k,\ k = 0, 1, \dots, 2l+2$.

(b) $N_i,\ \hat{N}_i$ *are not necessarily square. Their dimensions are determined by the requirement that the conditions* (c) *have sense.*

(c) *The functions* $D_{2i-1}^{-1}N_{2i-1},\ N_{2i}D_{2i}^{-1}$ *are proper and equal to each other*, $i = 1, \dots, l+1$ *and* $D_{2i}^{-1}\hat{N}_{2i},\ \hat{N}_{2i+1}D_{2i+1}^{-1}$ *are strictly proper and equal to each other*, $i = 0, \dots, l$.
Denote

$$B_{2i-1} = B_{m_{2i-1},m_{2i}}(D_{2i-1}, N_{2i}, N_{2i-1}, D_{2i}),$$
$$H_{2i} = H_{m_{2i},m_{2i+1}}(D_{2i}^{-1}\hat{N}_{2i} = \hat{N}_{2i+1}D_{2i+1}^{-1}).$$

Then:

(i) *The product*

$$P = B_1H_2B_3H_4 \dots B_{2l-1}H_{2l}$$

satisfies the intertwining relation

$$C(D_1)P = PC(D_{2l+1})\,.$$

The matrix P represents the operator of subsequent multiplication by $\hat{N}_{2i}$ *and* N_{2i-1}, $i = l, l-1, \dots, 2, 1$ *and then division by* D_1, *i.e.*,

$$P = \sum_k C^k(D_1)C_\infty(D_1)\left[\mathrm{diag}\ ((N)_k)\right]^{(\mathscr{I}(D_{2l+1}))}\,,$$

where

$$\sum (N)_k z^k = N(z) = N_1(z)\hat{N}_2(z)N_3(z)\hat{N}_4(z) \dots N_{2l-1}(z)\hat{N}_{2l}(z)\,.$$

(ii) *The product*

$$P = H_0B_1H_2B_3 \dots H_{2l-1}B_{2l-1}H_{2l}$$

is the finite Hankel matrix of the form

$$P = H_{m_0,m_{2l+1}}(D_0^{-1}N_L = N_RD_{2l+1}^{-1}),$$

where

$$N_L = \hat{N}_0N_1\hat{N}_2N_3 \cdots N_{2l-1}\hat{N}_{2l}\,,$$
$$N_R = \hat{N}_1N_2\hat{N}_3N_4 \cdots N_{2l}\hat{N}_{2l+1}.$$

(iii) *The product*

$$P = B_1H_2B_3H_4 \dots B_{2l-1}H_{2l}B_{2l+1}$$

is the finite Bézout matrix of the form $B_{m_1,m_{2l+2}}(D_1, N_R, N_L, D_{2l+2})$, *where*

$$\begin{aligned} N_L &= N_1\hat{N}_2N_3\hat{N}_4\cdots N_{2l-1}\hat{N}_{2l}N_{2l+1} \\ N_R &= N_2\hat{N}_3N_4\hat{N}_5\cdots N_{2l}\hat{N}_{2l+1}N_{2l+2}\,. \end{aligned}$$

Proof. First we prove (i) by induction on l.

The first induction step: Consider the assertion (i) for two factors, i.e., for $l = 1$. Then, it is equivalent to Theorem 26.

The second induction step: Suppose that the assertion (i) holds for $2l$ factors. We want to prove that (i) holds for $2l + 2$ factors.

First, we want to prove that

$$P_{l+1} = P_l P_{l\to l+1}$$

represents the operator of subsequent multiplication by $\hat{N}_2i$ and N_{2i-1}, for $i = l + 1, l, \ldots, 1$ and then division by D_1. Because $P_{l\to l+1} = B_{2l+1}H_{2l+2}$ satisfies the conditions of Theorem 26, we know that $P_{l\to l+1}$ represents the operator of multiplication by $N_{2l+1}\hat{N}_{2l+2}$ and then division by D_{2l+1}. Similar to the proof of Theorem 26, it is clear that the division by D_{2l+1} of $P_{l\to l+1}$ can be postponed after the subsequent multiplications by $\hat{N}_2i$ and N_{2i-1} for $i = l, l - 1, \ldots, 1$ corresponding to the operator P_l because

$$N_1\hat{N}_2N_3\cdots N_{2l-1}\hat{N}_{2l}D_{2l+1} = D_1N_2\hat{N}_3\cdots N_{2l}\hat{N}_{2l+1}.$$

Hence, the first part is proven.

To prove the intertwining relation, we again use Theorem 26, to obtain

$$\begin{aligned} C(D_1)P_{l+1} &= C(D_1)P_lP_{l\to l+1} \\ &= P_lC(D_{2l+1})P_{l\to l+1} \\ &= P_lP_{l\to l+1}C(D_{2l+3}) \\ &= P_{l+1}C(D_{2l+3}). \end{aligned}$$

This proves part (i) of the theorem.

Now we prove part (ii). Using (i), the product

$$P = H_0B_1\cdots B_{2l-1}H_{2l}$$

equals to

$$\begin{aligned} H_{m_0,m_1}(D_0^{-1}\hat{N}_0 = \hat{N}_1D_1^{-1})\sum_k C^k(D_1)C_\infty(D_1)\left[\operatorname{diag}((N)_k)\right]^{(\mathscr{I}(D_{2l+1}))} \\ = H_{m_0,m_0}(D_0^{-1})\sum_k C^k(D_0)C_\infty(D_0)\left[\operatorname{diag}((\hat{N}_0)_k)\right]^{(\mathscr{I}(D_1))} \\ \times\sum_k C^k(D_1)C_\infty(D_1)\left[\operatorname{diag}((N)_k)\right]^{(\mathscr{I}(D_{2l+1}))}. \end{aligned}$$

The same reasoning as in the proof of Theorem 26 can be applied to show that

$$P = H_{m_0,m_0}(D_0^{-1}) \sum_k C^k(D_0) C_\infty(D_0) \left[\operatorname{diag}((N_L)_k)\right]^{(\mathscr{I}(D_{2l+1}))}$$

and, using the corresponding generalized Barnett formula,

$$P = H_{m_0,m_{2l+1}}(D_0^{-1} N_L = N_R D_{2l+1}^{-1})$$

where

$$N_L = \hat{N}_0 N_1 \hat{N}_2 N_3 \ldots \hat{N}_{2l} N_{2l+1} .$$

If we use the same arguments for the matrix P^{T}, we prove the analogous expression for N_R.

The proof of (iii) is similar and we omit it. □

Some particular cases of the general formulas given in Theorem 31 are of special interest. Similar formulas for the scalar case are well known and can be found, e.g., in [18].

Lemma 32. *Let* D_1, D_2^{T} *be row reduced, with* $\deg\det D_1 = l$, $\deg\det D_2 = m$, *and let* N_1, N_2 *be s.t.* $N_1 D_2 = D_1 N_2$. *Then*

$$\begin{aligned} &H_{l,l}(D_1^{-1}) B_{l,m}(D_1, N_2, N_1, D_2) H_{m,m}(D_2^{-1}) \\ &\quad = H_{l,m}(D_1^{-1} N_1 = N_2 D_2^{-1}). \end{aligned} \tag{23}$$

If, moreover, $m = l$ *and* $D_1^{-1} N_1$, $N_2 D_2^{-1}$ *are coprime, then the matrices on both sides of* (23) *are square nonsingular and the equality can be rewritten as*:

1. $B_{l,l}(D_1)^{-1} B_{l,l}(D_1, N_2, N_1, D_2) B_{l,l}(D_2)^{-1} = B_{l,l}(D_2, \hat{N}_1, \hat{N}_2, D_1)^{-1}$,

where $\hat{N}_1$, $\hat{N}_2$ *are such that*

$$\begin{aligned} N_2 \hat{N}_1 + D_2 Q_1 &= I, \\ \hat{N}_2 N_1 + Q_2 D_1 &= I. \end{aligned}$$

2. $H_{l,l}(D_1^{-1}) H_{l,l}(D_1^{-1} \hat{N}_1 = \hat{N}_2 D_2^{-1})^{-1} H_{l,l}(D_2^{-1}) = H_{l,l}(D_1^{-1} N_1 = N_2 D_2^{-1})$.

Let D_1^{T}, D_2 *be row reduced*, $\deg\det D_1 = l$, $\deg\det D_2 = m$, *then*

$$B_{l,l}(D_1) H_{l,m}(D_1^{-1} N_1 = N_2 D_2^{-1}) B_{m,m}(D_2) = B_{l,m}(D_1, N_2, N_1, D_2). \tag{24}$$

If, moreover, $m = l$, *then the matrices on both sides of* (24) *are nonsingular and the equality can be rewritten as*:

1. $B_{l,l}(D_1) B_{l,l}(D_1, \hat{N}_2, \hat{N}_1, D_2)^{-1} B_{l,l}(D_2) = B_{l,l}(D_1, N_2, N_1, D_2)$,
2. $H_{l,l}(D_1^{-1})^{-1} H_{l,l}(D_1^{-1} N_1 = N_2 D_2^{-1}) H_{l,l}(D_2^{-1})^{-1} = H_{l,l}(D_1^{-1} \hat{N}_1 = \hat{N}_2 D_2^{-1})^{-1}$.

References

[1] B.O.D. Anderson, E.I. Jury, Generalized bezoutian and sylvester matrices in multivariable linear control, IEEE Trans. Automat. Control 21 (1976) 551–556.

[2] M. Van Barel, V. Pták, Z. Vavřín, Extending the notions of companion and infinite companion to matrix polynomials, Linear Algebra Appl. 290 (1999) 61–94.
[3] S. Barnett, Polynomials and Linear Control Systems, Pure and Applied Mathematics, vol. 77, Marcel Dekker, New York, 1983.
[4] S. Barnett, P. Lancaster, Some properties of the bezoutian for polynomial matrices, Linear Multilinear Algebra 9 (1980) 99–110.
[5] R.R. Bitmead, S.-Y. Kung, B.D.O. Anderson, T. Kailath, Greatest common divisors via generalized sylvester and bezout matrices, IEEE Trans. Automat. Control AC-23 (1978) 1043–1047.
[6] M. Fiedler, V. Pták, Bézoutians and intertwining matrices, Linear Algebra Appl. 86 (1987) 43–51.
[7] M. Fiedler, V. Pták, Intertwining and testing matrices corresponding to a polynomial, Linear Algebra Appl. 86 (1987) 53–74.
[8] P.A. Fuhrmann, Algebraic system theory: An analyst's point of view, J. Franklin Institute 301 (6) (1976) 521–540.
[9] P.A. Fuhrmann, On symmetric rational transfer functions, Linear Algebra Appl. 50 (1983) 167–250.
[10] P.A. Fuhrmann, Block hankel matrix inversion—the polynomial approach, in: H. Bart, I. Gohberg, M.A. Kaashoek (Eds.), Operator Theory and Systems, vol. OT19, Proc. Workshop, Amsterdam, June 4–7, 1985, Birkhäuser, Basel, 1986.
[11] P.A. Fuhrmann, A Polynomial Approach to Linear Algebra, Springer, Berlin, 1996.
[12] F.R. Gantmacher, The Theory of Matrices, vol. II, Chelsea, New York, 1959.
[13] I. Gohberg, T. Shalom, On bezoutians of nonsquare matrix polynomials and inversion of matrices with nonsquare blocks, Linear Algebra Appl. 137/138 (1990) 249–323.
[14] I.C. Gohberg, G. Heinig, Inversion of finite Toeplitz matrices composed of elements of a noncommutative algebra, Rev. Roum. Math. Pure Appl. 19 (5) (1974) 623–663.
[15] I. Haimovici, L. Lerer, Bezout operators for analytic operator functions. I. A general concept of Bezout operator, Integral Equations and Operator Theory 21 (1995) 33–70 (In Russian).
[16] G. Heinig, Über Block-Hankelmatrizen und den Begriff der Resultante für Matrixpolynome, Wiss. Ze. Tech. Hochsch. Karl-Marx-Stadt 19 (4) (1977) 513–519 (In German).
[17] G. Heinig, Resultante, Bezoutiante und Spektralverteilungsprobleme fuer Operatorpolynome, Math. Nachr. 91 (1979) 23–43 (In German).
[18] G. Heinig, K. Rost, Algebraic Methods for Toeplitz-like Matrices and Operators, Birkhäuser, Basel, 1984.
[19] U. Helmke, P.A. Fuhrmann, Bezoutians, Linear Algebra Appl. 122–124 (1989) 1039–1097.
[20] C. Hermite, Sur le nombre des racines d'une équation algébraique comprise entre des limites donnés, J. Reine Angew. Math. 52 (1856) 39–51.
[21] T. Kailath, Linear Systems, Prentice-Hall, Englewood, Cliffs, 1980.
[22] M.G. Krein, M.A. Naimark, The method of symmetric and hermitian forms in the theory of the separation of roots of algebraic equations, Linear Multilinear Algebra 10 (1981) 265–308.
[23] F.I. Lander, The bezoutian and the inversion of hankel and toeplitz matrices (English translation of the Russian original from 1936), Matem. Issled. 9 (2) (1974) 69–87.
[24] L. Lerer, L. Rodman, Bezoutians of rational matrix functions, matrix equations and factorizations, Linear Algebra Appl. 302–303 (1999) 105–133.
[25] L. Lerer, L. Rodman, M. Tismenetsky, Bezoutian and schur–cohn problem for operator polynomials, J. Math. Anal. Appl. 103 (1984) 83–102.
[26] L. Lerer, M. Tismenetsky, The bezoutian and the eigenvalue-separation problem for matrix polynomials, Integral Equations Operator Theory 5 (1982) 386–445.
[27] L. Lerer, M. Tismenetsky, Generalized Bezoutian and the inversion problem for block matrices, I. General scheme, Integral Equations Operator Theory 9 (1986) 790–819.
[28] L. Lerer, M. Tismenetsky, Generalized bézoutian and matrix equations, Linear Algebra Appl. 99 (1988) 123–160.
[29] V. Pták, Spectral radius, norms and iterates and the critical exponent, Linear Algebra Appl. 1 (1968) 245–260.

[30] V. Pták, H.K. Wimmer, On the bézoutian for polynomial matrices, Linear Algebra Appl. 71 (1985) 267–272.
[31] J.H.M. Wedderburn, Lectures on Matrices, AMS, New York, 1934.
[32] H.K. Wimmer, On the history of the bezoutian and the resultant matrix, Linear Algebra Appl. 128 (1990) 27–34.
[33] H.K. Wimmer, Bezoutians of polynomial matrices and their generalized inverses, Linear Algebra Appl. 122/123/124 (1989) 475–487.
[34] H.K. Wimmer, Generalized bezoutians and block hankel matrices, Linear Algebra Appl. 117 (1989) 105–113.
[35] H.K. Wimmer, The kernel of the bezoutian of polynomial matrices, Systems Control Lett. 12 (1989) 287–290.
[36] W.A. Wolovich, The determination of state-space representations for linear multivariable systems, in: Proc. Second IFAC Symp. on Multivariable Technical Control Systems, Duesseldorf, October 1971.