Int. J. Appl. Math. Comput. Sci., 2001, Vol.11, No.6, 1285–1310

MATRIX QUADRATIC EQUATIONS, COLUMN/ROW REDUCED FACTORIZATIONS AND AN INERTIA THEOREM FOR MATRIX POLYNOMIALS[†]

Irina KARELIN[*], Leonid LERER[*]

It is shown that a certain Bezout operator provides a bijective correspondence between the solutions of the matrix quadratic equation and factorizatons of a certain matrix polynomial $G(\lambda)$ (which is a specification of a Popov-type function) into a product of row and column reduced polynomials. Special attention is paid to the symmetric case, i.e. to the Algebraic Riccati Equation. In particular, it is shown that extremal solutions of such equations correspond to spectral factorizations of $G(\lambda)$. The proof of these results depends heavily on a new inertia theorem for matrix polynomials which is also one of the main results in this paper.

Keywords: matrix quadratic equations, Bezoutians, inertia, column (row) reduced polynomials, factorization, algebraic Riccati equation, extremal solutions

1. Introduction

In this paper we deal with matrix quadratic equations of the form

$$A_1X - XA_2 = -XWX \tag{1}$$

under the standing assumptions that (A_2, W) is controllable and (W, A_1) is observable. It is worth mentioning that if a solution Y_0 of the equation

$$A_1Y - YA_2 = -YWY + Q \tag{2}$$

is known, then any solution of (2) is given by $Y = Y_0 + X$ where X is a solution of (1), A_1 is replaced by $A_1 - Y_0W$, and A_2 is replaced by $A_2 + WY_0$. Among the numerous works devoted to the study of the equation of type (1), (2), and especially of their symmetric versions (known as Algebraic Riccati Equations), we mention only the recent monographs (Ando, 1988; Lancaster and Rodman, 1995) and the paper (Ionescu and Weiss, 1993), where one can find the main elements of the theory and further references.

[†] The research of the second author was partially supported by the Fund for the Promotion of Research at the Technion.

[*] Department of Mathematics, Technion—Israel Institute of Technology, Haifa 32000, Israel, e-mail: {kirina,llerer}@techunix.technion.ac.il

The present paper can be viewed as a continuation and refinement of the papers (Karelin and Lerer, 2001; Lerer, 1989; Lerer and Tismenetsky, 1982). In (Lerer, 1989) we considered (1) with coefficients that are similar to companion matrices and we established that the Bezoutian for a quadruple of matrix polynomials, introduced in (Anderson and Jury, 1976) and thoroughly studied in (Lerer and Tismenetsky, 1982), is an adequate connecting link between the set of all solutions of (1) and the set of all monic divisors of specified degree of a certain matrix polynomial generated by (1). The general case was treated in (Lerer, 1989) by a certain dilation method which seems to be cumbersome. In some recent papers (see, e.g., Gomez and Lerer, 1994; Haimovici and Lerer, 1995; Karelin and Lerer, 2001; Lerer and Rodman, 1996b) more refined notions of Bezout operators were developed (based on the representation of the functions in question in realized form). These Bezout operators were employed in (Karelin and Lerer, 2001; Karelin *et al.*, 2001) to study the connections between the solutions of the general equation (2) and factorization of rational matrix function of the Popov type associated with (2). To be more precise, write

$$G_{C,K}(\lambda) = I - [K \;\; \Phi]\lambda I - \begin{bmatrix} A_2 - \Psi K & 0 \\ Q + CK & A_1 - C\Phi \end{bmatrix}^{-1} \begin{bmatrix} \Psi \\ Q \end{bmatrix}, \tag{3}$$

where K and C are arbitrary (fixed) feedback matrices and $W = \Psi\Phi$ is a rank decomposition. We prove in (Karelin and Lerer, 2001) that a certain Bezout operator for families of rational matrix functions provides a bijective correspondence between the set of all solutions of (2) and the set of factorizations

$$G_{C,K}(\lambda) = G^{(1)}_{C,K}(\lambda)G^{(2)}_{C,K}(\lambda), \tag{4}$$

where $G^{(i)}_{C,K}$ $(i = 1, 2)$ are of a specific form given in terms of realizations:

$$G^{(1)}_{C,K}(\lambda) = I - \Phi(\lambda I - U_C)^{-1}M, \quad G^{(2)}_{C,K}(\lambda) = I - N(\lambda I - V_K)^{-1}\Psi, \tag{5}$$

$U_C = A_1 - C\Phi$, $V_K = A_2 - \Psi K$, and M, N are some matrices. Clearly, if one takes C and K in (3) so that the matrices U_C and V_K are nilpotent (which is possible due to the controllability of (A_2, Ψ) and the observability of (Φ, A_1)), then $G_{C,K}(\lambda)$ becomes a matrix polynomial and in the particular case of (1) one easily sees that, in fact, $G_{C,K}(\lambda)$ can be written as a product of two matrix polynomials:

$$G_{C,K}(\lambda) = L_1(\lambda)L_2(\lambda). \tag{6}$$

In this paper we show that for a specific choice of C and K the polynomials $L_1(\lambda)$ and $L_2(\lambda)$ in (6) are row and column reduced, respectively, and the requirement (5) on the factors in (4) can be expressed also in a more convenient form in terms of row and column reduced polynomials. This approach provides more insight into the structure of hermitian solutions of the Algebraic Riccati Equation

$$A^*X + XA = XWX \quad (W \geq 0). \tag{7}$$

In this case the associated matrix polynomial $G(\lambda)$ is non-negative definite on the real axis and there is a bijective correspondence between the set of hermitian solutions of (7) and the set of all symmetric factorizations of $G(\lambda)$. Moreover, we show that the minimal and maximal solutions of (7) correspond to left and right spectral factorizations of $G(\lambda)$, respectively.

Remark that in the framework of this paper the above-mentioned bijective correspondence between the solutions of (1) or (7) and the set of appropriate factorizations into a product of matrix polynomials is provided by the Bezoutian of a quadruple of matrix polynomials given in realized form as introduced in (Haimovici and Lerer, 1995). Using this Bezoutian one can describe completely the common zero structure of the underlying polynomials (see, e.g., Haimovici and Lerer, 2001; Lerer and Rodman, 1996b). Based on this fact we obtain conditions for the existence of an invertible solution for linear matrix equations of the Sylvester type, a problem which has been addressed in several publications (see, e.g., (Gohberg *et al.*, 1984; Hearon, 1977; Lerer, 1989; Lerer and Ran, 1996; Lerer and Rodman, 1996c; 1999) and references therein).

Moreover, the above-mentioned property of the Bezoutian alows us to prove in this paper a new inertia theorem of the Hermite-Fujiwara type for matrix polynomials. The previously known result of this type established in (Lerer and Tismenetsky, 1982) describes the localization of the zeroes of matrix polynomials in the extended complex plane $\mathbb{C}\bigcup\{\infty\}$ (i.e. taking into account possible zeroes at infinity). The inertia theorem in the present paper provides results on zero localization in the finite complex plane in terms of the Bezoutian mentioned above. Some results of this type were obtained previously in (Dym, 1991; Dym and Young, 1990) in other terms and by different methods. Note that this inertia theorem is heavily used in the study of extremal solutions of (6).

This paper is organized as follows. After preliminary Section 2 we study in Section 3 the connection between the solutions of (1) and factorizations of an associated matrix polynomial $G(\lambda)$ into a product of row and column reduced polynomials. In Section 4 we prove the inertia theorem which is used in Section 5 to study the extremal solutions of the Algebraic Riccati Equation (7).

2. Preliminaries

First we remind some well-known notions and results concerning the spectral theory of matrix polynomials. More detailed information can be found, e.g., in (Gohberg *et al.*, 1981; 1982; Lerer and Rodman, 1996a).

In this work we deal with *right admissible pairs* (Φ, V) of order p, where V is an $p\times p$ matrix and Φ is an $n\times p$ matrix. A pair of matrices (A,Ψ) is a *left admissible pair of order* p if A and Ψ are $p\times p$ and $p\times n$, respectively. It is clear that the pair (Φ, V) is right admissible if and only if the pair (V^T, Φ^T) is left admissible. (Here the superscript 'T' means matrix transposition.)

A pair of matrices (A,Ψ) is called *controllable* if for some integer $l\geq 1$ the matrix $\operatorname{row}[A^{j-1}B]_{j=1}^{l} := [B, AB, \dots, A^{l-1}B]$ is right invertible. A pair of matrices (Φ, V) is called *observable* if $\operatorname{col}[\Phi V^{j-1}]_{j=1}^{n} := [\operatorname{row}[(V^T)^{j-1}\Phi^T]_{j=1}^{l}]^T$ is left invertible, i.e. (V^T,Φ^T) is controllable.

Let $L(\lambda) = \sum_{j=0}^{l} l_j\lambda^j$ be an $n\times n$ matrix polynomial, where l_j are $n\times n$ matrices. $\lambda_0\in\mathbb{C}$ is called an *eigenvalue* of the polynomial $L(\lambda)$ if $\det L(\lambda_0) = 0$.

The set of the eigenvalues of $L(\lambda)$ is denoted by $\sigma(L)$. We consider only polynomials which are *regular*, i.e. $\sigma(L) \neq \mathbb{C}$. A vector $\phi_0 \neq 0$ is a *right eigenvector* of the polynomial $L(\lambda)$ corresponding to an eigenvalue λ_0 if $L(\lambda_0)\phi_0 = 0$. If ϕ_0 is an eigenvector of $L(\lambda)$ corresponding to λ_0, then the sequence of n-dimensional vectors $\phi_0, \phi_1, \ldots, \phi_m$ for which the equalities $\sum_{j=0}^{k} \frac{1}{j!} L^{(j)}(\lambda_0)\phi_{k-j} = 0$ $(k = 0, 1, \ldots, m-1)$ hold true is called the *right Jordan chain* of length $m+1$ for the polynomial $L(\lambda)$ corresponding to λ_0. Its leading vector $\phi_0 (\neq 0)$ is a right eigenvector and the subsequence $\phi_1, \ldots, \phi_m$ is called the sequence of *generalized eigenvectors*. By constructing a certain canonical set of Jordan chains corresponding to all the eigenvalues of $L(\lambda)$ one can encode the zero information about $L(\lambda)$, as a right admissible pair (Φ, V) $(V \in \mathbb{C}^{p\times p},\ \Phi \in \mathbb{C}^{n\times p},\ p = \deg\det L(\lambda))$ called the *right finite null pair* of $L(\lambda)$ (see, e.g., (Gohberg *et al.*, 1982; Lerer and Rodman, 1996a). One can see that an admissible pair (Φ, V) of order p is a right null pair of a regular polynomial $L(\lambda)$ if $p = \deg\det L(\lambda)$, $\operatorname{rank}\operatorname{col}(\Phi V^{j-1})_{j=1}^{l} = p$ and $L(\Phi, V) = \sum_{j=0}^{l} l_j \Phi V^j = 0$. It can be proved that $L(\lambda)$ is uniquely defined by its right finite null pair up to multiplication from the left by a matrix polynomial with constant determinant. In an obvious way one defines the 'left' counterparts of the above notions.

We also need the definitions of divisors and common divisors of matrix polynomials. We say that a polynomial $D(\lambda)$ is a *right divisor* of $L(\lambda)$ if there exists a matrix polynomial $Q(\lambda)$ such that $L(\lambda) = Q(\lambda)D(\lambda)$. A polynomial $D(\lambda)$ is called a *right common divisor* of the polynomials $L_1(\lambda)$ and $L_2(\lambda)$ if $D(\lambda)$ is a divisor of both the polynomials $L_1(\lambda)$ and $L_2(\lambda)$. A right common divisor $D_0(\lambda)$ is the *greatest right common divisor* of the polynomials $L_1(\lambda)$ and $L_2(\lambda)$ if any right common divisor of $L_1(\lambda)$ and $L_2(\lambda)$ is also a right divisor of $D_0(\lambda)$. The dual concepts concerning *left common divisors* are defined similarly. Matrix polynomials $L_1(\lambda)$ and $L_2(\lambda)$ are called *right (left) coprime* if I is their greatest right (left) common divisor. For more details from the divisibility theory of matrix polynomials see, e.g., (Gohberg *et al.*, 1981; 1982; Lerer and Rodman, 1996a).

We shall use the representation of polynomials in realized form:

$$L(\lambda) = L(0) + \lambda C(I - \lambda A)^{-1}B. \tag{8}$$

The following result reflects the connection between finite null pairs of matrix polynomials and their realizations (Bart *et al.*, 1979).

Proposition 1. *Let $L(\lambda)$ be a matrix polynomial with an invertible coefficient $L(0)$ and let (8) be its minimal realization. Write $A^{\times} = A - BL(0)^{-1}C$. Then λ_0 is an eigenvalue of $A^{\times}$ if and only if λ_0^{-1} is an eigenvalue of $L(\lambda)$ and the partial multiplicities of λ_0 as an eigenvalue of $A^{\times}$ coincide with the partial multiplicities of λ_0^{-1} as an eigenvalue of $L(\lambda)$. Moreover, $(L(0)^{-1}C, (A^{\times})^{-1})$ is a right finite null pair of $L(\lambda)$, and $((A^{\times})^{-1}, BL(0)^{-1})$ is a left finite null pair of $L(\lambda)$.*

Now we shall introduce the notion of the Bezoutian of a quadruple of matrix polynomials following the work (Haimovici and Lerer, 1995).

Let $L(\lambda)$ and $M(\lambda)$ be two $n \times n$ matrix polynomials. One can always find (Lerer and Tismenetsky, 1982) two complementary polynomials $L_1(\lambda)$ and $M_{1(\lambda)}$ satisfying

$$L_1(\lambda)L(\lambda) = M_1(\lambda)M(\lambda). \tag{9}$$

Write a joint observable realization of $L_1(\lambda)$ and $M_1(\lambda)$:

$$\big[L_1(\lambda)\ M_1(\lambda)\big] = \big[L_1(0)\ M_1(0)\big] + \lambda\Phi(I-\lambda U)^{-1}[C_L\ C_M], \tag{10}$$

and a joint controllable realization of $L(\lambda)$ and $M(\lambda)$

$$\begin{bmatrix} L(\lambda) \\ M(\lambda) \end{bmatrix} = \begin{bmatrix} L(0) \\ M(0) \end{bmatrix} + \lambda \begin{bmatrix} K_L \\ K_M \end{bmatrix} (I-\lambda V)^{-1}\Psi. \tag{11}$$

Theorem 1. (Haimovici and Lerer, 1995) *There exists a unique matrix $\mathbb{B}$ such that the equality*

$$\frac{L_1(\lambda)L(\mu) - M_1(\lambda)M(\mu)}{\lambda-\mu} = \Phi(I-\lambda U)^{-1}\mathbb{B}(I-\mu V)^{-1}\Psi \tag{12}$$

holds true for all (λ,μ) for which both the sides make sense.

The matrix $\mathbb{B}$ from Theorem 1 is called the *Bezoutian* of the quadruple of the polynomials $(L_1, M_1; L, M)$ associated with (9) and realizations (10) and (11).

Assume now that $L(0) = M(0) = D$ and $L_1(0) = M_1(0) = E$ are invertible matrices, and introduce the matrices $V_L^\times = V - \Psi D^{-1}K_L$, $V_M^\times = V - \Psi D^{-1}K_M$, $U_L^\times = U - C_L E^{-1}\Phi$ and $U_M^\times = U - C_M E^{-1}\Phi$. It was shown in (Haimovici and Lerer, 2001) (see also (Lerer and Rodman, 1996c; 1999)) that for functions which are analytic at the origin the Bezoutian of $(L_1, M_1; L, M)$ associated with (9) and realizations (10) and (11) satisfies the following equations:

Proposition 2.

$$U_M^\times\mathbb{B} = \mathbb{B}V_L^\times, \quad \mathbb{B}\Psi = (C_L - C_M)E, \quad \Phi\mathbb{B} = (K_M - K_L)D,$$

$$U_L^\times\mathbb{B} - \mathbb{B}V_L^\times = -(C_L - C_M)(K_L - K_M).$$

We also need a description of the kernel of the Bezoutian $\mathbb{B}$ for the quadruple of the polynomials $(L_1, M_1; L, M)$ associated with the equality (9) and the realizations (10) and (11). The theorem below (Haimovici and Lerer, 2001; Lerer and Rodman, 1996b) will play an important role in the sequel.

Theorem 2. *Let $\mathbb{B}$ be the Bezoutian of the quadruple of the polynomials $(L_1, M_1; L, M)$ associated with the equality (9), the observable realization (11) and the controllable realization (10). Then the kernel of $\mathbb{B}$ is the maximal subspace contained in $\operatorname{Ker}(K_L - K_M)$ which is invariant with respect to the operator $V_L^\times$. If, in addition, $\operatorname{Ker}\Psi = \{0\}$, then $\operatorname{Ker}\mathbb{B}$ is the maximal subspace which is invariant with respect to $V_L^\times$ such that on this space the operators $V_L^\times$ and $V_M^\times$ coincide.*

If the realizations (10) and (11) are minimal, then λ_0 *is a common eigenvalue of* $L(\lambda)$ *and* $M(\lambda)$ *if and only if* λ_0^{-1} *is an eigenvalue of* $V_L^{\times}|_{\operatorname{Ker}\mathbb{B}}$ *and the common multiplicity of* λ_0 *as a common eigenvalue of* $L(\lambda)$ *and* $M(\lambda)$ *equals the multiplicity of* λ_0^{-1} *as an eigenvalue of* $V_L^{\times}|_{\operatorname{Ker}\mathbb{B}}$.

3. The 'Homogeneous' Matrix Quadratic Equation and Factorizations of Matrix Polynomials

In this section we shall be concerned with relations between the solutions of the matrix quadratic equation (1) and factorizations of some matrix polynomials. Some different aspects of such relations are studied in (Lerer and Ran, 1996).

Consider the matrix quadratic equation

$$A_1X - XA_2 = -XWX, \tag{13}$$

where X is a $q \times p$ complex matrix, and A_1, A_2, W are $q \times q$, $p \times p$, $p \times q$ complex matrices, respectively. It will be assumed throughout that the pair (W, A_1) is observable and the pair (A_2, W) is controllable. Also, without loss of generality, we may (and will) assume that A_1 and A_2 are invertible. Write a rank decomposition of the matrix W: $W = \Psi\Phi$, where Ψ and Φ are $p \times n$ and $n \times q$ matrices, respectively. (Here $n = \operatorname{rank} W$.) Since (A_2, Ψ) is controllable, we can choose an $n \times p$ matrix K such that $\sigma(A_2 - \Psi K) = \{0\}$. Write $V = A_2 - \Psi K$. Similarly, there is a matrix C such that $U = A_1 - C\Phi$ is nilpotent. Clearly, the pair (Φ, U) is observable and the pair (V, Ψ) is controllable. Since V and U are nilpotent matrices, the functions

$$\begin{aligned} L_1(\lambda) &= E - \lambda E\Phi(I - \lambda U)^{-1}C, \\ L_2(\lambda) &= D - \lambda K(I - \lambda V)^{-1}\Psi D \end{aligned} \tag{14}$$

are polynomials. (Here we fix some invertible matrices D and E.) Define

$$G_{C,K}(\lambda) = L_1(\lambda)L_2(\lambda).$$

In this section we establish a bijective correspondence between the solutions of (13) and some factorizations of $G_{C,K}(\lambda)$. To be more precise, we consider a factorization of $G_{C,K}(\lambda)$

$$G(\lambda) = R_1(\lambda)R_2(\lambda) \tag{15}$$

with factors of the form

$$\begin{aligned} R_1(\lambda) &= E - \lambda E\Phi(I - \lambda U)^{-1}C_R, \\ R_2(\lambda) &= D - \lambda K_R(I - \lambda V)^{-1}\Psi D, \end{aligned} \tag{16}$$

where C_R and K_R are some matrices of appropriate sizes.

Theorem 3. *Let $R_1(\lambda)$ and $R_2(\lambda)$ be two $n\times n$ matrix polynomials of the form (16), respectively, satisfying (15). Then the Bezoutian of the quadruple $(L_1, R_1; L_2, R_2)$ associated with the equality (15) and the realizations (14), (16) is a solution of (13).*

Conversely, if X is a solution of (13), define the polynomials

$$\begin{aligned} R_1(\lambda) &= E - \lambda E\Phi(I-\lambda U)^{-1}(C+X\Psi), \\ R_2(\lambda) &= D - \lambda(K-\Phi X)(I-\lambda V)^{-1}\Psi D. \end{aligned} \tag{17}$$

Then the polynomials $R_1(\lambda), R_2(\lambda)$ satisfy (15). Moreover, the solution X of (13) coincides with the Bezoutian of the quadruple $(L_1, R_1; L_2, R_2)$ associated with the equality (13) and the realizations (14), (17). The above correspondence between the set of the solutions of (13) and factorizations of $G_{C,K}(\lambda)$ of the type (16) is bijective.

The above theorem in the case of $D = E = I$ can be obtained as a corollary from the results of (Karelin and Lerer, 2001). For the sake of completeness, we present an independent proof in the case of arbitrary invertible D and E.

Proof . We prove first the converse part of the theorem. Assume that X is a solution of (13). Set $\Lambda(\lambda) = \lambda E\Phi(I-\lambda U)^{-1}$, $\Delta(\mu) = \mu(I-\mu V)^{-1}\Psi D$ and multiply (13) by $\Lambda(\lambda)$ from the left and by $\Delta(\mu)$ from the right:

$$\Lambda(\lambda)(A_1X - XA_2 + XWX)\Delta(\mu) = 0. \tag{18}$$

Compute

$$\begin{aligned} A_2\Delta(\mu) = (V+\Psi K)\Delta(\mu) &= -(I-\mu V)(I-\mu V)^{-1}\Psi D \\ &\quad + (I-\mu V)^{-1}\Psi D + \Psi K\Delta(\mu) = -\Psi L_2(\mu) + \Delta(\mu)\mu^{-1}. \end{aligned}$$

Similarly, we get

$$\Lambda(\lambda)A_1 = -L_1(\lambda)\Phi + \Lambda(\lambda)\lambda^{-1}.$$

Thus (18) becomes

$$\begin{aligned} 0 &= -L_1(\lambda)X_2(\mu) + \mu X(\lambda,\mu) + X_1(\lambda)L_2(\mu) - \lambda X(\lambda,\mu) \\ &\quad + X_1(\lambda)X_2(\mu) - \big(L_1(\lambda) - X_1(\lambda)\big)\big(L_2(\mu) + X_2(\mu)\big) \\ &\quad + L_1(\lambda)L_2(\mu) - (\lambda-\mu)X(\lambda,\mu), \end{aligned} \tag{19}$$

where $X_1(\lambda) = \Lambda(\lambda)X\Psi$, $X_2(\mu) = \Phi X\Delta(\mu)$, $X(\lambda,\mu) = \Lambda(\lambda)X\Delta(\mu)/\lambda\mu$. Setting $\lambda = \mu$ in (19) and

$$R_1(\lambda) = L_1(\lambda) - X_1(\lambda) = E - \lambda E\Phi(I-\lambda U)^{-1}(C+X\Psi), \tag{20}$$

$$R_2(\lambda) = L_2(\lambda) + X_2(\lambda) = D - \lambda(K-\Phi X)(I-\lambda V)^{-1}\Psi D, \tag{21}$$

we obtain the factorization (15). From (19) one sees that X is the Bezoutian of $(L_1, R_1; L_2, R_2)$ associated with the factorization (15) and the realizations (14), (20), (21).

Conversely, let $R_1(\lambda)$ and $R_2(\lambda)$ be polynomials of the form (16), satisfying (15). We claim that the Bezoutian $\mathbb{B}$ of the quadruple $(L_1, R_1; L_2, R_2)$ associated with (15) and the realizations (14), (16) is a solution of (13). Indeed, from (19) we have

$$\Lambda(\lambda)(A_1\mathbb{B}-\mathbb{B}A_2)\Delta(\mu)=(\mu-\lambda)\mathbb{B}(\lambda,\mu)+\mathbb{B}_1(\lambda)L_2(\mu)-L_1(\lambda)\mathbb{B}_2(\mu), \quad (22)$$

where

$$\mathbb{B}(\lambda,\mu) = \frac{\Lambda(\lambda)\mathbb{B}\Delta(\mu)}{\lambda\mu}, \quad \mathbb{B}_1(\lambda) = \Lambda(\lambda)\mathbb{B}\Psi, \quad \mathbb{B}_2(\mu) = \Phi\mathbb{B}\Delta(\mu).$$

Since $\mathbb{B}$ is the Bezoutian of the above polynomials,

$$(\mu-\lambda)\mathbb{B}(\lambda,\mu) = -L_1(\lambda)L_2(\mu) + R_1(\lambda)R_2(\mu).$$

Substituting this expressions into (22), we obtain

$$\begin{aligned}\Lambda(\lambda)(A_1\mathbb{B}-\mathbb{B}A_2)\Delta(\mu) &= -L_1(\lambda)L_2(\mu) + R_1(\lambda)R_2(\mu)\\ &\quad + \mathbb{B}_1(\lambda)L_2(\mu) - L_1(\lambda)\mathbb{B}_2(\mu). \end{aligned} \quad (23)$$

Setting $\lambda = 0$ in (23), we obtain $R_2(\mu) = L_2(\mu) + \mathbb{B}_2(\mu)$. Similarly, setting $\mu = 0$, we observe that $R_1(\lambda) = L_1(\lambda) - \mathbb{B}_1(\lambda)$. Substituting these expressions into (23), we see that $\Lambda(\lambda)(A_1\mathbb{B} - \mathbb{B}A_2)\Delta(\mu) = -\mathbb{B}_1(\lambda)\mathbb{B}_2(\mu)$, and hence, using the observability of (Φ, U) and the controllability of (V, Ψ), we see that $\mathbb{B}$ is a solution of (13).

To complete the proof of the theorem it remains to show that the above correspondence between the set of the solutions of (13) and the set of the factorizations (16) of $G_{C,K}(\lambda)$ is bijective. Suppose that there are two factorizations of $G_{C,K}(\lambda)$,

$$L_1(\lambda)L_2(\lambda) = R_1(\lambda)R_2(\lambda), \quad L_1(\lambda)L_2(\lambda) = Q_1(\lambda)Q_2(\lambda), \quad (24)$$

where $R_1(\lambda)$ and $R_2(\lambda)$ have realizations (16), and

$$Q_1(\lambda)=E-\lambda E\Phi(I-\lambda U)^{-1}C_Q, \quad Q_2(\lambda)=D-\lambda K_Q(I-\lambda V)^{-1}\Psi D. \quad (25)$$

Assume that one and the same solution X of (13) corresponds to each of these factorizations. We claim that in this case $R_1(\lambda) = Q_1(\lambda)$ and $R_2(\lambda) = Q_2(\lambda)$. Indeed, X is the Bezoutian of the quadruple $(L_1, R_1; L_2, R_2)$ associated with (24) and realizations (14), (16). Also, X is the Bezoutian of the quadruple $(L_1, Q_1; L_2, Q_2)$ associated with (24) and realizations (14), (25). Hence

$$\begin{aligned}L_1(\lambda)L_2(\mu) - R_1(\lambda)R_2(\mu) &= L_1(\lambda)L_2(\mu) - Q_1(\lambda)Q_2(\mu)\\ &= (\lambda-\mu)E\Phi(I-\lambda U)^{-1}X(I-\mu V)^{-1}\Psi D.\end{aligned}$$

Setting $\lambda = 0$ in this equality we have $R_2(\mu) = Q_2(\mu)$, and, similarly, setting $\mu = 0$ we have $R_1(\lambda) = Q_1(\lambda)$. The theorem is thus proved. ■

Now we present some notions and results concerning column and row reduced polynomials and their realizations. Let $L(\lambda) = \sum_{j=1}^{l} l_j\lambda^j$ be an $n\times n$ matrix polynomial. Consider this polynomial as a $n\times n$ matrix whose entries are scalar polynomials. The *column index* α_j of the j-th column is defined as the maximal degree of the scalar polynomials in this column. Let $\alpha_1, \ldots, \alpha_n$ be column indices of $L(\lambda)$, and let $\hat{L}_0$ be an $n \times n$ matrix such that its j-th column consists of the coefficients of λ^{α_j} in the j-th column of $L(\lambda)$. A polynomial $L(\lambda)$ is called *column reduced* if $\hat{L}_0$ is a non-singular matrix. Any matrix polynomial can be brought to column reduced form by elementary operations on its rows and columns (see, for example, (Kailath, 1980)).

Let $\alpha_1, \ldots, \alpha_n$ be the column indices of $L(\lambda)$. We shall use the multi-index $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ when it is convenient. Without loss of generality we may (and will) suppose that $\alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_n$. For any $n \times n$ matrix polynomial $L(\lambda)$ with column indices $\alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_n$ we construct a special realization. Namely, introduce $J_\alpha = \operatorname{diag}(J_{\alpha_1}, J_{\alpha_2}, \ldots, J_{\alpha_n})$, where J_{α_i} is an $\alpha_i \times \alpha_i$ Jordan matrix ($i = 1, \ldots, n$):

$$J_{\alpha_i} = \begin{bmatrix} 0 & 1 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & 1 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & 0 & 1 \\ 0 & 0 & \cdot & \cdot & \cdot & 0 \end{bmatrix}.$$

Set $p = \sum_{i=1}^{n} \alpha_i$. Then J_α is of size $p \times p$. Define the $p \times n$ matrix

$$Q_\alpha = \operatorname{col}(Q_{\alpha_1}, Q_{\alpha_2}, \ldots, Q_{\alpha_n})$$

whose blocks Q_{α_i} are $\alpha_i \times n$ matrices of the form

$$Q_{\alpha_i} = [0 \ \ldots \ e_i \ \ldots \ 0], \quad i = 1, \ldots, n,$$

where 0 stands for the $\alpha_i \times 1$ zero-column and the i-th column of Q_{α_i} is of the form $e_i = [0 \ \ldots \ 0 \ 1]^T$. Write the polynomial $L(\lambda)$ in the form $L(\lambda) = L(0) + [L_{i,k}(\lambda)]_{i,k=1}^{n}$, where $L_{i,k}(\lambda) = \sum_{j=1}^{\alpha_k} l_{i,\alpha_1+\cdots+\alpha_{k-1}+j} \lambda^{\alpha_k - j+1}$, and define the $n \times p$ matrix $K_L = [K_1 \ K_2 \ \ldots \ K_n]$, where K_s $(s = 1, \ldots, n)$ is an $n \times \alpha_j$ matrix of the form $K_s = [l_{i,\alpha_1+\cdots+\alpha_{s-1}+j}]_{i,j=1}^{n,\alpha_s}$.

Proposition 3. *Let $L(\lambda)$ be an $n \times n$ polynomial with column indices which do not exceed $\alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_n$. Then $L(\lambda)$ has the realization*

$$L(\lambda) = L(0) + \lambda K_L (I - \lambda J_\alpha)^{-1} Q_\alpha, \tag{26}$$

where J_α, Q_α, K_L are defined by the formulae above. Moreover, the realization (26) is minimal if and only if $L(\lambda)$ is a column reduced polynomial with column indices $\alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_n$.

Proof. Computations show that

$$(I - \lambda J_\alpha)^{-1} = \operatorname{diag}\left((I - \lambda J_{\alpha_i})^{-1}\right)_{i=1}^{n} = \operatorname{diag}\left(\begin{bmatrix} 1 & \lambda & \ldots & \lambda^{\alpha_i - 1} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \ldots & 1 & \lambda \\ 0 & 0 & \ldots & 1 \end{bmatrix}\right)_{i=1}^{n}.$$

Then $(I-\lambda J_\alpha)^{-1}Q_\alpha = \mathrm{col}\,((I-\lambda J_{\alpha_i})^{-1}Q_{\alpha_i})_{i=1}^n$, where $(I-\lambda J_{\alpha_i})^{-1}Q_{\alpha_i}$ is an $\alpha_i \times n$ matrix such that the i-th column of this matrix has the form $\Gamma_i = \mathrm{col}\,(\lambda^{\alpha_i - j})_{j=1}^{\alpha_i}$ and all the other columns are equal to zero. Multiplying $(I-\lambda J_\alpha)^{-1}Q_\alpha$ by λK_L from the left and adding $L(0)$ we obtain (26). To show the minimality of the realization (26) for a column reduced polynomial, we remind that the minimal size of the state space in a realization of a matrix polynomial $L(\lambda)$ equals $\deg\det L(\lambda)$. Because of the column reduceness of $L(\lambda)$, $\deg\det L(\lambda) = \alpha_1 + \cdots + \alpha_n = p$. Conversely, suppose that (26) is minimal. Then $\deg\det L(\lambda) = p$. This implies that the column indices of $L(\lambda)$ are exactly equal to $\alpha_1, \ldots, \alpha_n$ and $L(\lambda)$ is column reduced. The proposition is proved. ■

Assuming that $L(0)$ is invertible, define the matrix

$$B_L = J_\alpha - Q_\alpha L^{-1}(0) K_L. \tag{27}$$

We call B_L the *first Brunovsky companion matrix* for a polynomial $L(\lambda)$ with column indices $\alpha_1 \le \cdots \le \alpha_n$. Write $\tilde{L}(\lambda) = L(0)^{-1}L(\lambda) = I + [\tilde{L}_{ik}(\lambda)]_{i,k=1}^n$, where $\tilde{L}_{i,k}$ $(i,k = 1,\ldots,n)$ is the (i,k)-th entry of the matrix $\tilde{L}(\lambda)$, and represent $\tilde{L}_{i,k}(\lambda) = \sum_{j=1}^{\alpha_k} \tilde{l}_{i,\alpha_1+\cdots+\alpha_k-j+1}\lambda^j$. Substituting expressions for J_α, Q_α, K_L into (27), we have

$$B_L = \begin{bmatrix} C_{\tilde{L}_{11}} & R_{12} & \ldots & R_{1n} \\ R_{21} & C_{\tilde{L}_{22}} & \ldots & R_{2n} \\ \vdots & \vdots & \ldots & \vdots \\ R_{n1} & R_{n2} & \ldots & C_{\tilde{L}_{nn}} \end{bmatrix},$$

where

$$C_{\tilde{L}_{ii}} = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \ldots & 1 \\ -\tilde{l}_{i,\alpha_1+\cdots+\alpha_{i-1}+1} & -\tilde{l}_{i,\alpha_1+\cdots+\alpha_{i-1}+2} & \ldots & \ldots & -\tilde{l}_{i,\alpha_1+\cdots+\alpha_i} \end{bmatrix}$$

for $i = 1, \ldots, n$ and

$$R_{i,j} = \begin{bmatrix} 0 & \ldots & 0 \\ \vdots & \ldots & \vdots \\ -\tilde{l}_{i,\alpha_1+\cdots+\alpha_{j-1}+1} & \ldots & -\tilde{l}_{i,\alpha_1+\cdots+\alpha_j} \end{bmatrix}.$$

The minimality of the realization (26) implies the following result:

Proposition 4. *Let $L(\lambda)$ be a column reduced polynomial. Then the pair of matrices $(B_L^{-1}, Q_\alpha L^{-1}(0))$ is a left null pair of $L(\lambda)$.*

We also need the following notions: For an observable pair (Φ, A_1) (Φ is an $n \times p$ matrix, A_1 is a $p \times p$ matrix) define the *observability indices* $\beta_1 \leq \beta_2 \leq \cdots \leq \beta_n$ in the following way (Kailath, 1980):

$$\beta_i = \{j\colon\ n + q_{l+j-1} - q_{l+j} \leq i - 1, \quad j = 1, \ldots, s + 1 - l\},$$

where $j = 1, \ldots, n$, $q_k = \operatorname{rank} \operatorname{col} (\Phi A_1^{i-1})_{i=1}^{k}$ and $l = \max\{m \geq 1 \mid q_m = mn\}$. The *controlability indices* of a controllable pair (A_2, Ψ) can be defined, e.g., as observability indices of the observable pair (Ψ^T, A_2^T). Now we recall the solution to the following problem: Given a controllable pair (A, Ψ), construct a column reduced polynomial $L(\lambda)$ such that: (a) the pair (A^{-1}, Ψ) is a left null pair of $L(\lambda)$, (b) the column indices of $L(\lambda)$ concide with the controllability indices of (A, Ψ).

There are several solutions to this problem (Gohberg *et al.*, 1980). We present here the construction from (Ball *et al.*, 1994)).

Proposition 5. *Let (A, Ψ) be a controllable pair of matrices, where A is a $p \times p$ invertible matrix and Ψ is a $p \times n$ full rank matrix, and let $\alpha_1 \leq \cdots \leq \alpha_n$ be the controllability indices of the pair (A, Ψ). Choose a basis $\{g_{j,k}\}_{j,k=1}^{\alpha_j, n}$ of $\mathbb{C}^p$ such that it satisfies the following properties:*

1. *$\{g_{j,1}\}_{j=1}^{n}$ is a basis of $\operatorname{Im}\Psi$;*
2. *$Ag_{j,k} - g_{j,k+1} \in \operatorname{Im}\Psi$, where $g_{j,\alpha_j+1} = 0$ ($j = 1, \ldots, n$; $k = 1, \ldots, \alpha_j$).*

With the basis $\{g_{j,k}\}$ we associate the operator $V\colon \mathbb{C}^p \to \mathbb{C}^p$ defined by

$$Vg_{j,k} = \begin{cases} g_{j,k+1} & \text{if} \quad k = 1, \ldots, \alpha_j - 1, \\ 0 & \text{if} \quad k = \alpha_j, \end{cases}$$

Then there exists an operator $K:\ \mathbb{C}^p \to \mathbb{C}^n$ such that $A - \Psi K = V$. We can get linearly independent vectors $y_1, \ldots, y_n$ such that $\Psi y_j = -g_{j,1}$ for $j = 1, \ldots, n$. Finally, we define the operator $D:\ \mathbb{C}^n \to \mathbb{C}^n$ by the formula $De_j = y_j$, where $\{e_j\}_{j=1}^{n}$ is a standard basis in $\mathbb{C}^n$. Then the polynomial

$$L(\lambda) = D - \lambda K(I - \lambda V)^{-1} \Psi D \tag{28}$$

is column reduced such that (28) is its minimal realization, (A^{-1}, Ψ) is its left null pair, and $\alpha_1 \leq \cdots \leq \alpha_n$ are its column indices.

We also need the dual notion of *row reduced polynomials.* Namely, for a polynomial $M(\lambda) = \sum_{i=0}^{m} m_j \lambda^j$ we define the *row index* β_i as the maximal power of entries in the i-th row. Without loss of generality we can (and will) suppose that $\beta_1 \leq \beta_2 \leq \cdots \leq \beta_n$. Let $\hat{M}_0$ be an $n \times n$ matrix such that its i-th row consists of the coefficients of λ^{β_i} in the i-th row of $M(\lambda)$. A polynomial $M(\lambda)$ is called *row reduced* if $\hat{M}_0$ is a non-singular matrix.

Note that if $M(\lambda)$ is row reduced, then $M^T(\lambda)$ is column reduced, and hence Proposition 3 implies the following result:

Proposition 6. *Let $M(\lambda)$ be an $n \times n$ matrix polynomial with row indices which do not exceed $\beta_1 \leq \cdots \leq \beta n$. Then*

$$M(\lambda) = M(0) + \lambda \hat{Q_\beta}(I - \lambda \hat{J_\beta})^{-1} \hat{K_M}, \tag{29}$$

where

$$\hat{J_\beta} = J_\beta^T, \hat{Q_\beta} = Q_\beta^T, \hat{K_M} = K_{M^T}^T.$$

The realization (29) is minimal if and only if $M(\lambda)$ is row reduced with row indices which are equal to $\beta_1 \leq \cdots \leq \beta_n$.

Similarly, Proposition 4 implies the following result:

Proposition 7. *Let $M(\lambda)$ be a row reduced polynomial and (29) be its minimal realization. If $M(0)$ is nonsingular, then $(M(0)^{-1}\hat{Q_\beta}, (\hat{J_\beta} - \hat{K_M} M(0)^{-1} \hat{Q_\beta})^{-1})$ is a right null pair of $M(\lambda)$.*

The matrix $\hat{B_M} = \hat{J_\beta} - \hat{K_M} M(0)^{-1} \hat{Q_\beta}$ will be called the *second Brunovsky companion matrix*.

Similarly to the column reduced case, one can construct a row reduced polynomial $M(\lambda)$ from a given observable pair (Φ, A_1) such that (a) (Φ, A_1^{-1}) is a right null pair of $M(\lambda)$, (b) the row indices of $M(\lambda)$ concide with the observability indices of (Φ, A_1). To do this, we pass to the controllable pair (A_1^T, Φ^T) and construct the column reduced polynomial $M_1(\lambda)$ as described in Proposition 5. Then $M(\lambda) = M_1^T(\lambda)$ is the polynomial with desired properties.

The following result will be important in the sequel:

Proposition 8. *Let $L(\lambda)$ be a column reduced polynomial with column indices $\alpha_1 \leq \cdots \leq \alpha_n$ and a minimal realization (28). Then any polynomial $R(\lambda)$ which is equal to D at $\lambda = 0$ with column indices which do not exceed $\alpha_1 \leq \cdots \leq \alpha_n$ has the realization of the form*

$$R(\lambda) = D - \lambda K_1 (I - \lambda V)^{-1} \Psi D \tag{30}$$

for some K_1. Conversely, any polynomial $R(\lambda)$ with realization (30) has the above properties (i.e. its column indices do not exceed $\alpha_1 \leq \cdots \leq \alpha_n$ and $R(0) = D$). The realization (30) is minimal if and only if $R(\lambda)$ is column reduced.

Similarly, let $M(\lambda)$ be a row reduced polynomial with row indices $\beta_1 \leq \cdots \leq \beta_n$ and a minimal realization

$$M(\lambda) = E - \lambda E \Phi (I - \lambda U)^{-1} C. \tag{31}$$

Then any polynomial $N(\lambda)$ which is equal to E at $\lambda = 0$ and has row indices which do not exceed $\beta_1 \leq \cdots \leq \beta_n$ has the realization

$$N(\lambda) = E - \lambda E \Phi (I - \lambda U)^{-1} C_1$$

for some matrix C_1. Conversely, any polynomial $N(\lambda)$ with realization (31) has row indices which do not exceed $\beta_1 \le \cdots \le \beta_n$ and $N(0) = E$. The realization (31) is minimal if and only if $N(\lambda)$ is row reduced.

Proof. Since the realizations (26) and (28) have the same size of the state space, there is an invertible matrix S such that

$$J_\alpha = S^{-1}VS, \quad SQ_\alpha = \Psi D, \quad K_L S^{-1} = K,$$

and $L(0) = D$. On the other hand, $R(\lambda)$ admits a realization of the type (26) as well:

$$R(\lambda) = R(0) + \lambda K_R (I - \lambda J_\alpha)^{-1} Q_\alpha.$$

Substituting the above similarity relations into the last expressions for $R(\lambda)$, we obtain

$$\begin{aligned} R(\lambda) &= R(0) + \lambda K_R S^{-1}(I - \lambda V)^{-1} S Q_\alpha \\ &= R(0) + \lambda K_R S^{-1}(I - \lambda V)^{-1} \Psi D. \end{aligned}$$

If $R(\lambda)$ is column reduced, then this realization as well as the realization of the type (26) are minimal. Since $R(0) = D$, by setting $K_R S^{-1} = K_1$ we obtain the first part of the proposition. The second part can be obtained in a similar way. ■

Based on the construction of column and row reduced polynomials for given controllable and observable pairs, we now refine Theorem 3. Namely, for (13) consider two pairs of matrices: the controllable pair (A_2, Ψ) and the observable pair (Φ, A_1). Applying the procedure of Proposition 5 to the pair (A_2, Ψ) with controllability indices $\alpha_1 \le \cdots \le \alpha_n$ yields a column reduced polynomial $L_2(\lambda)$ with column indices $\alpha_1 \le \cdots \le \alpha_n$ such that the pair (A_2^{-1}, Ψ) is a left null pair of $L_2(\lambda)$. For this polynomial $L_2(\lambda)$ write the minimal realization

$$L_2(\lambda) = D - \lambda K (I - \lambda V)^{-1} \Psi D, \tag{32}$$

where $A_2 = V + \Psi K$.

In a similar way, for the observable pair (Φ, A_1) with observability indices $\beta_1 \le \cdots \le \beta_n$ we can also construct a row reduced polynomial $L_1(\lambda)$ whose right null pair is (Φ, A_1^{-1}) and the row indices are $\beta_1 \le \cdots \le \beta_n$. The polynomial $L_1(\lambda)$ has a minimal realization

$$L_1(\lambda) = E - \lambda E \Phi (I - \lambda U)^{-1} C, \tag{33}$$

where $A_1 = U + C\Phi$.

Now consider the matrix polynomial

$$G(\lambda) = L_1(\lambda) L_2(\lambda). \tag{34}$$

Let P_c^α be the set of all column reduced polynomials with column indices $\alpha_1 \le \cdots \le \alpha_n$ which are equal to D at 0, and let P_r^β be the set of all row reduced polynomials with row indices $\beta_1 \leq \cdots \leq \beta_n$ which are equal to E at 0. We are interested in factorizations of the polynomial $G(\lambda)$ into a product of two polynomials

$$G(\lambda) = L_1(\lambda)L_2(\lambda) = R_1(\lambda)R_2(\lambda) \tag{35}$$

such that $R_2(\lambda)$ belongs to P_c^α and $R_1(\lambda)$ belongs to P_r^β. Let $D_{\alpha\beta}(G)$ be the set of all such factorizations. The next theorem establishes the connection between the set of solutions of (13) and the class $D_{\alpha\beta}$ of the factorizations of $G(\lambda)$ into a product of row and column reduced polynomials with row and column indices defined by the given pairs (Φ, A_1) and (A_2, Ψ).

Theorem 4. *For (13) consider column and row reduced polynomials*

$$L_1(\lambda) = E - \lambda E\Phi(I - \lambda U)^{-1}C, \quad L_2(\lambda) = D - \lambda K(I - \lambda V)^{-1}\Psi D \tag{36}$$

as described in the construction of (33) and (32), respectively. Let X be a solution of (13). Then X generates a factorization of $G(\lambda) = L_1(\lambda)L_2(\lambda)$ into a product of two polynomials

$$L_1(\lambda)L_2(\lambda) = R_1(\lambda)R_2(\lambda), \tag{37}$$

where

$$\begin{aligned} R_1(\lambda) &= E - \lambda E\Phi(I - \lambda U)^{-1}(C + X\Psi), \\ R_2(\lambda) &= D - \lambda(K - \Phi X)(I - \lambda V)^{-1}\Psi D. \end{aligned} \tag{38}$$

This factorization belongs to the class $D_{\alpha\beta}(G)$. Here X is the Bezoutian of the quadruple $(L_1, R_1; L_2, R_2)$ associated with the equality (37) and the realizations (36), (38).

Conversely, consider any factorization of $G(\lambda)$ belonging to the class $D_{\alpha\beta}(G)$: $G(\lambda) = R_1(\lambda)R_2(\lambda)$, and hence

$$R_1(\lambda) = E - \lambda E\Phi(I - \lambda U)^{-1}C_1, \quad R_2(\lambda) = D - \lambda K_1(I - \lambda V)^{-1}\Psi D. \tag{39}$$

Then the Bezoutian of $(L_1, R_1; L_2, R_2)$ associated with (37) and realizations (36), (39) is a solution of (13). The correspondence between the solutions of (13) and the factorizations $G(\lambda)$ in the class $D_{\alpha\beta}(G)$ is bijective.

Proof. All the assertions of the theorem follow from Theorem 3 except the fact that the polynomials $R_1(\lambda)$ and $R_2(\lambda)$ which are defined by the realization (38) are row and column reduced, respectively. Indeed, from Proposition 8 it is clear that the row indices of $R_1(\lambda)$ do not exceed $\beta_1 \leq \cdots \leq \beta_n$ and the column indices of $R_2(\lambda)$ do not exceed $\alpha_1 \leq \cdots \leq \alpha_n$. We claim that $R_1(\lambda)$ is row reduced with row indices which are exactly equal to $\beta_1, \ldots, \beta_n$ and $R_2(\lambda)$ is column reduced with column indices $\alpha_1, \ldots, \alpha_n$. Indeed, since $L_1(\lambda)$ is row reduced with row indices as above, the degree of the polynomial $\det L_1(\lambda)$ is exactly equal to $t = \beta_1 + \cdots + \beta_n$. Analogously, because of the column reduceness of $L_2(\lambda)$, the degree of $\det L_2(\lambda)$ is exactly equal to $p = \alpha_1 + \cdots + \alpha_n$. Clearly,

$$\deg \det G(\lambda) = t + p = \deg \det R_1(\lambda) + \deg \det R_2(\lambda). \tag{40}$$

Since the degree of $\det R_1(\lambda)$ does not exceed t and the degree of $\det R_2(\lambda)$ does not exceed p, (40) implies $\deg\det R_1(\lambda) = t$, $\deg\det R_2(\lambda) = p$, and therefore $R_1(\lambda)$ belongs to $P_r^{(\beta)}$ and $R_2(\lambda)$ belongs to $P_c^{(\alpha)}$. ■

Note that Theorem 4 can be proved without using Theorem 3 by a direct proof for the case when the coefficients are in Brunovsky form and by applying a similarity in order to pass to the general case.

Combining Theorem 4 with Proposition 2, we obtain the following result:

Corollary 1. *Let X be a solution of (13), and let this solution X generate the factorization $L_1(\lambda)L_2(\lambda) = R_1(\lambda)R_2(\lambda)$ from the class $D_{\alpha\beta}(G)$, where L_1, L_2, R_1, R_2 are defined by (36), (39). Then* $\operatorname{rank} X = p - \deg\det D_r(\lambda)$, *where $D_r(\lambda)$ denotes the greatest right common divisor of L_2 and R_2. In particular, X is invertible if and only if $L_2(\lambda)$ and $R_2(\lambda)$ are right coprime.*

Theorem 4 allows us to obtain conditions for the existence of an invertible solution of the Sylvester equation

$$ZA_1 - A_2Z + W = 0, \tag{41}$$

where A_1, A_2 and W are $p \times p$ matrices.

Corollary 2. *Given the Sylvester equation (41), construct a row reduced polynomial $L_1(\lambda)$ and a column reduced polynomial $L_2(\lambda)$ defined by (36). If there is a factorization $L_1(\lambda)L_2(\lambda) = R_1(\lambda)R_2(\lambda)$ from the class $D_{\alpha\beta}(G)$ such that $L_2(\lambda)$ and $R_2(\lambda)$ are right coprime, then (41) has an invertible solution Z and*

$$R_1(\lambda) = E - \lambda E\Phi(I - \lambda U)^{-1}(C + Z^{-1}\Psi), \tag{42}$$

$$R_2(\lambda) = D - \lambda(K - \Phi Z^{-1})(I - \lambda V)^{-1}\Psi D. \tag{43}$$

Conversely, if (41) has an invertible solution Z, then $L_1(\lambda)L_2(\lambda)$ can be factored as $L_1(\lambda)L_2(\lambda) = R_1(\lambda)R_2(\lambda)$, where $L_2(\lambda)$ and $R_2(\lambda)$ are right coprime, $R_1(\lambda)$ and $R_2(\lambda)$ can be written in the form (42) and (43), and $X = Z^{-1}$ is the Bezoutian of the quadruple (L_1, R_1, L_2, R_2) associated with the realizations involved and the equation $L_1(\lambda)L_2(\lambda) = R_1(\lambda)R_2(\lambda)$.

Furthermore, any invertible solution Z of (41) can be found by the formula $Z = X^{-1}$, where X is the Bezoutian of the quadruple $(L_1, R_1; L_2, R_2)$ associated with the equality $L_1(\lambda)L_2(\lambda) = R_1(\lambda)R_2(\lambda)$ and the realizations (36), (38), provided that $L_2(\lambda)$ and $R_2(\lambda)$ are right coprime.

For other results concerning the existence of invertible solutions of (41), see, e.g., (Gohberg *et al.*, 1981; Hearon, 1977; Lerer, 1989; Lerer and Rodman, 1999) and references therein.

4. An Inertia Theorem for Matrix Polynomials

In this section we deal with inertia (i.e. localization of zeros with respect to $\mathbb{R}$) of matrix polynomials. This problem has been addressed previously in (Lerer and Tismenetsky, 1982), where the localization of zeroes of a matrix polynomial is described in terms of the 'coefficient' Bezoutian in the extended complex plane $\mathbb{C}\bigcup\infty$ (i.e. taking into account possible zeroes at infinity). In this section we use the notion of the Bezoutian based on the representation of polynomials in realized form (as defined in Section 1) to obtain results on zero localization in the finite complex plane. Some results of this sort were also obtained in (Dym, 1991; Dym and Young, 1990), in other terms and by different methods.

Here a *symmetric factorization* of a matrix polynomial $G(\lambda)$ means a factorization of the type $G(\lambda) = P^*(\lambda)P(\lambda)$, where $P(\lambda)$ is a matrix polynomial and $P^*(\lambda) = (P(\bar{\lambda}))^*$. We start with some general auxiliary facts. Assume that there are two symmetric factorizations

$$L^*(\lambda)L(\lambda) = R^*(\lambda)R(\lambda), \tag{44}$$

where $L(\lambda)$ and $R(\lambda)$ are polynomials. Let $L(\lambda)$ be written in realized form,

$$L(\lambda) = D + \lambda K(I - \lambda V)^{-1}\Psi D. \tag{45}$$

Clearly, (45) is controllable if and only if the realization

$$L^*(\lambda) = D^* + \lambda D^*\Psi^*(I - \lambda V^*)^{-1}K^*$$

is observable.

Proposition 9. *Let $L(\lambda)$ and $R(\lambda)$ be $n \times n$ matrix polynomials such that (44) holds; we have a controllable realization*

$$\begin{bmatrix} L(\lambda) \\ R(\lambda) \end{bmatrix} = \begin{bmatrix} D \\ D \end{bmatrix} + \lambda \begin{bmatrix} K_L \\ K_R \end{bmatrix} (I - \lambda V)^{-1}\Psi D \tag{46}$$

and hence a (necessarily observable) realization

$$\big[L^*(\lambda)\ \ R^*(\lambda)\big] = [D^*\ \ D^*] + \lambda D^*\Psi^*(I - \lambda V^*)^{-1}[K_L^*\ \ K_R^*]. \tag{47}$$

Then the Bezoutian $\mathbb{B}$ associated with (44), (46) and (47) is a skew-hermitian matrix.

Proof. Write $\Gamma(\lambda,\mu) = L^*(\lambda)L(\mu) - R^*(\lambda)R(\mu)$. Then the Bezoutian $\mathbb{B}$ associated with (44) and realizations (46), (47) is defined by the equation

$$(\lambda - \mu)\Gamma(\lambda,\mu) = D^*\Psi^*(I - \lambda V^*)^{-1}\mathbb{B}(I - \mu V)^{-1}\Psi D. \tag{48}$$

One easily sees that

$$\big((\Gamma(\lambda,\mu)\big)^* = \Gamma(\bar{\mu},\bar{\lambda}). \tag{49}$$

From (48) we have

$$\frac{\Gamma(\bar{\mu},\bar{\lambda})}{\bar{\mu}-\bar{\lambda}} = D^*\Psi^*(I-\bar{\mu}V^*)^{-1}\mathbb{B}(I-\bar{\lambda}V)^{-1}\Psi D.$$

Taking adjoints of both the sides of this equality and using (49), we obtain

$$\frac{\Gamma(\lambda,\mu)}{\mu-\lambda} = D^*\Psi^*(I-\lambda V^*)^{-1}\mathbb{B}^*(I-\mu V)^{-1}\Psi D. \tag{50}$$

Comparing (50) with (48) and using the uniqueness of the Bezoutian, we infer that $\mathbb{B}^* = -\mathbb{B}$. ■

Let $L(\lambda) \in P_c^{(\alpha)}$, i.e. $L(\lambda)$ is an $n\times n$ column reduced polynomial with column indices $\alpha_1 \le \cdots \le \alpha_n$, which is equal to D at $\lambda = 0$. It is obvious that the $n\times n$ matrix polynomial $L^*(\lambda)$ is row reduced with row indices $\alpha_1 \le \cdots \le \alpha_n$ (i.e. it belongs to P_r^{α} with $E = D^*$). Set

$$G(\lambda) = L^*(\lambda)L(\lambda) \tag{51}$$

and consider symmetric factorizations of $G(\lambda)$,

$$G(\lambda) = R^*(\lambda)R(\lambda), \tag{52}$$

such that $R(0) = D$. Denote the set of such factorizations of $G(\lambda)$ by $D_\alpha^{\mathrm{sym}}(G)$.

Proposition 10. *Let $G(\lambda)$ be a matrix polynomial defined by (51), where $L(\lambda)$ belongs to the set P_c^{α}. Then $D_\alpha^{\mathrm{sym}}(G)$ consists of all the symmetric factorizations (52), where $R(\lambda)$ belongs to $P_c^{(\alpha)}$.*

Proof. Let $L(\lambda) \in P_c^{(\alpha)}$. First, note that the degrees of the entries $(G(\lambda))_{j,j}$ of the matrix polynomial $G(\lambda)$ are $2\alpha_j$. This follows from the equality

$$\big(G(\lambda)\big)_{j,j} = \big(L^*(\lambda)\big)_j\big(L(\lambda)\big)_j,$$

where $(L(\lambda))_j$ is the j-th column of $L(\lambda)$ and $(L^*(\lambda))_j$ is the j-th row of $L^*(\lambda)$. Suppose that for some j the column index of the j-th column of $R(\lambda)$ is not α_j. Then the equality $(G(\lambda))_{jj} = (R^*(\lambda))_j(R(\lambda))_j$ implies that the degree of the polynomial $(G(\lambda))_{j,j} \ne 2\alpha_j$. Hence the j-th column index of $R(\lambda)$ is exactly equal to α_j. The fact that $R(\lambda)$ is column reduced is known from Theorem 4. ■

Now recall some notions. For a matrix A its *inertia* is defined as the triple of integers $\operatorname{In} A = (\pi(A), \nu(A), \delta(A))$, where $\pi(A), \nu(A), \delta(A)$ denote the number of eigenvalues of A, counting multiplicities, with positive, negative and zero real part, respectively. Similarly, introduce $\widetilde{\operatorname{In}}\,(A) = (\gamma_+(A), \gamma_-(A), \gamma_0(A))$ with respect to the real axis. It is clear that

$$\pi(iA) = \gamma_-(A), \quad \nu(iA) = \gamma_+(A), \quad \delta(iA) = \gamma_0(A).$$

For a matrix polynomial $L(\lambda)$ introduce in a similar manner $\widetilde{\operatorname{In}}\,(L) = (\gamma_+(L), \gamma_-(L), \gamma_0(L))$, where $\gamma_+(L)$ is the number of eigenvalues (counting multiplicities) in the open upper half-plane, $\gamma_-(L)$ is the number of eigenvalues (counting multiplicities) in the open lower half-plane and $\gamma_0(L)$ is the number of those lying on the real axis. It is clear that if $L(\lambda)$ is a matrix polynomial with minimal realization

$$L(\lambda) = D + \lambda C(I - \lambda A)^{-1}B,$$

where D is an invertible matrix, then for $A^\times = A - BD^{-1}C$ we have $\gamma_-(L) = \gamma_+(A^\times)$, $\gamma_+(L) = \gamma_-(A^\times)$, $\gamma_0(L) = \gamma_0(A^\times)$.

For a given column reduced polynomial $L(\lambda)$ with column indices $\alpha_1 \le \cdots \le \alpha_n$ and minimal realization

$$L(\lambda) = D - \lambda K(I - \lambda V)^{-1}\Psi D \tag{53}$$

define the non-negative matrix polynomial $G(\lambda) = L^*(\lambda)L(\lambda)$. Consider symmetric factorizations of this polynomial

$$L^*(\lambda)L(\lambda) = R^*(\lambda)R(\lambda), \tag{54}$$

where $R(0) = D$, and where the column indices of $R(\lambda)$ do not exceed $\alpha_1 \le \cdots \le \alpha_n$. According to Propositions 8 and 10, $R(\lambda)$ has the minimal realization

$$R(\lambda) = D - \lambda K_R(I - \lambda V)^{-1}\Psi D \tag{55}$$

for some matrix K_R of the appropriate size. In this and in the next section, we denote by $\mathbb{B}$ the Bezoutian of $(L^*, R^*; L, R)$ associated with (54) and realizations (53) and (55). Let $L_0(\lambda)$ be the greatest right common divisor of $L(\lambda)$ and $R(\lambda)$. The following theorem is the main result of this section. Let $L(\lambda)$ be a column reduced polynomial.

Theorem 5. *Let $L(\lambda)$ be a column reduced polynomial. Preserving the above notations, we have*

$$\gamma_+(L) = \pi(i\mathbb{B}) + \gamma_+(L_0), \quad \gamma_-(L) = \nu(i\mathbb{B}) + \gamma_-(L_0),$$
$$\gamma_0(L) = \delta(i\mathbb{B}) - \gamma_+(L_0) - \gamma_-(L_0).$$

To prove this theorem, remind the following results.

Theorem 6. (Carlson and Shneider, 1963). *Let A be an $s \times s$ matrix and $\delta(A) = 0$. Suppose that there exists a non-singular hermitian matrix H such that*

$$AH + HA^* \ge 0. \tag{56}$$

Then $\mathrm{In}\,(A) = \mathrm{In}\,(H)$.

Lemma 1. (Lerer and Tismenetsky, 1982) *Let $L(\lambda)$ and $M(\lambda)$ be matrix polynomials such that*

$$L^*(\lambda)L(\lambda) = M^*(\lambda)M(\lambda),$$

and let λ_0 be a real eigenvalue of $L(\lambda)$. Then λ_0 is an eigenvalue of $M(\lambda)$ and the Jordan chains of $L(\lambda)$ corresponding to λ_0 are also the Jordan chains of $M(\lambda)$ corresponding to λ_0, i.e. they are the common Jordan chains of $L(\lambda)$ and $M(\lambda)$ corresponding to λ_0.

Proof of Theorem 5. Decompose the space $\mathbb{C}^p$ into a direct sum

$$\mathbb{C}^p = N \oplus \operatorname{Ker}\mathbb{B}, \tag{57}$$

where N is the orthogonal complement of $\operatorname{Ker}\mathbb{B}$ in $\mathbb{C}^p$. (Recall that p is the minimal state space dimension in the realization (45).) Since $\mathbb{B}$ is skew-hermitian, it is clear that N is $\mathbb{B}$-invariant. Therefore, with respect to the decomposition (57) of $\mathbb{C}^p$, the operator $\mathbb{B}$ has the representation

$$\mathbb{B} = \begin{bmatrix} \mathbb{B}_1 & 0 \\ 0 & 0 \end{bmatrix}. \tag{58}$$

Here $\mathbb{B}_1$ acts from N into N and $\operatorname{Ker}\mathbb{B}_1 = (0)$, i.e. $\mathbb{B}_1$ is invertible. In a similar way, we have the representation of the matrices A, K and K_R with respect to the direct sum (57). But $A = V^\times = V - \Psi K$. According to Proposition 2, $\operatorname{Ker}\mathbb{B}$ is the maximal A-invariant subspace contained in $\operatorname{Ker}(K - K_R)$. Hence we have the following representation of $A\colon\ N \oplus \operatorname{Ker}\mathbb{B} \to N \oplus \operatorname{Ker}\mathbb{B}$:

$$A = \begin{bmatrix} A_1 & 0 \\ \times & A_0 \end{bmatrix}. \tag{59}$$

Similarly, K and K_R acting from $N \oplus \operatorname{Ker}\mathbb{B}$ into $\mathbb{C}^n$ have the representations

$$K = [K_1\ \ K_0], \tag{60}$$

$$K_R = [K_{1R}\ \ K_0] \tag{61}$$

with the same K_0 since K and K_R coincide on $\operatorname{Ker}\mathbb{B}$. According to Proposition 2, the matrix $\mathbb{B}$ satisfies the equation

$$A^*\mathbb{B} - \mathbb{B}A = -(K^* - K_R^*)(K - K_R). \tag{62}$$

Rewriting (62) for the matrix representations (58)–(61), we obtain

$$A_1^*\mathbb{B}_1 - \mathbb{B}_1 A_1 = -(K_1^* - K_{1R}^*)(K_1 - K_{1R}). \tag{63}$$

We claim that A_1 has no real eigenvalues. Indeed, according to Lemma 1, if λ_0 is a real eigenvalue of multiplicity k_0 for $L(\lambda)$, then λ_0 is an eigenvalue of the same multiplicity k_0 for $R(\lambda)$, and the Jordan chains corresponding to the real eigenvalue λ_0 of $L(\lambda)$ coincide with the Jordan chains corresponding to the same eigenvalue of $R(\lambda)$. Then, according to Theorem 2, λ_0^{-1} is an eigenvalue of multiplicity k_0 of $A\,|_{\operatorname{Ker}\mathbb{B}} = A_0$. Thus all the real eigenvalues of A are eigenvalues of A_0 and hence A_1 has no real eigenvalues.

Write $A_2 = -iA_1$ and $\mathbb{B}_2 = -i\mathbb{B}_1$. Then (63) can be rewritten as

$$A_2^*(-\mathbb{B}_2) + (-\mathbb{B}_2 A_2) = (K_1^* - K_{1R}^*)(K_1 - K_{1R}). \tag{64}$$

Clearly, the matrix $-\mathbb{B}_2$ is hermitian and invertible. We know that A_1 has no real eigenvalues. Hence $\operatorname{In} A_2 = (\pi(A_2), \nu(A_2), 0)$. Then, in view of Theorem 6, $\pi(\mathbb{B}_2) = \pi(A_2)$, $\nu(\mathbb{B}_2) = \nu(A_2)$ or

$$\pi(i\mathbb{B}_1) = \pi(iA_1) = \gamma_-(A_1), \quad \nu(i\mathbb{B}_1) = \nu(iA_1) = \gamma_+(A_1). \tag{65}$$

We return to the initial matrices $\mathbb{B}$ and A. Since $i\mathbb{B}$ is hermitian, it has only real eigenvalues. Then it is clear from the representations (58) and (59) that $\delta(i\mathbb{B}) = \dim\operatorname{Ker}\mathbb{B} = p_0$, where p_0 is the size of A_0. Therefore $\delta(i\mathbb{B}) = \gamma_-(A_0) + \gamma_+(A_0) + \gamma_0(A_0)$. Recall that λ_i is an eigenvalue of A_0 of multiplicity k_i if and only if λ_i^{-1} is a common eigenvalue of $L(\lambda)$ and $R(\lambda)$ of common multiplicity k_i. Therefore we have $\gamma_-(L_0) = \gamma_+(A_0)$, $\gamma_+(L_0) = \gamma_-(A_0)$ and $\gamma_0(L_0) = \gamma_0(A_0)$, where $L_0(\lambda)$ is the greatest right common divisor of $L(\lambda)$ and $R(\lambda)$. Hence

$$\delta(i\mathbb{B}) = \gamma_+(L_0) + \gamma_-(L_0) + \gamma_0(L_0).$$

It remains to rewrite (65) as follows:

$$\gamma_-(A) = \pi(i\mathbb{B}) + \gamma_-(A_0), \quad \gamma_+(A) = \nu(i\mathbb{B}) + \gamma_+(A_0),$$

$$\gamma_0(A) = \delta(i\mathbb{B}) - \gamma_+(A_0) - \gamma_-(A_0)$$

or

$$\gamma_+(L) = \pi(i\mathbb{B}) + \gamma_+(L_0), \quad \gamma_-(L) = \nu(i\mathbb{B}) + \gamma_-(L_0),$$

$$\gamma_0(L) = \delta(i\mathbb{B}) - \gamma_+(L_0) - \gamma_-(L_0),$$

and the theorem is proved. ■

5. Spectral Factorizations and Extremal Solutions of the Riccati Equation

In this section we consider the algebraic Riccati equation of the form

$$A^*X + XA = XWX, \tag{66}$$

where A and W are $p\times p$ matrices, and A is invertible, W is a non-negative definite hermitian matix, and (A, W) is a controllable pair. Here we establish an explicit connection between the problem of determining hermitian solutions of (66) and the problem of symmetric factorizations for non-negative definite matrix polynomials. Special attention is paid to the description of extremal solutions of (66) in terms of spectral factorizations.

First, construct the matrix polynomial $L(\lambda)$ as in Proposition 5. Write a rank decomposition of W: $W = \Psi\Psi^*$, where Ψ is a $p\times n$ matrix. Let $\alpha_1 \le \cdots \le \alpha_n$ be the controllability indices of the controllable pair (iA, Ψ). For the given pair (iA, Ψ), define a column reduced polynomial $L(\lambda)$ with column indices $\alpha_1 \le \cdots \le \alpha_n$ as described in Proposition 5. Let

$$L(\lambda) = D - \lambda iK(I - \lambda iV)^{-1}\Psi D \tag{67}$$

be a minimal realization of such a polynomial $L(\lambda)$ belonging to $P_c^{(\alpha)}$. Write the $n\times n$ row reduced polynomial

$$L^*(\lambda) = D^* + \lambda iD^*\Psi^*(I + i\lambda V^*)^{-1}K^* \tag{68}$$

with row indices $\alpha_1 \leq \cdots \leq \alpha_n$. Set

$$G(\lambda) = L^*(\lambda)L(\lambda) \tag{69}$$

and consider symmetric factorizations of $G(\lambda)$,

$$G(\lambda) = R^*(\lambda)R(\lambda), \tag{70}$$

such that $R(0) = D$.

Theorem 7. *Preserving the above notation, assume that X is a hermitian solution of (66). Then X generates a symmetric factorization of the polynomial $G(\lambda) = L^*(\lambda)L(\lambda)$ as follows:*

$$L^*(\lambda)L(\lambda) = R^*(\lambda)R(\lambda), \tag{71}$$

where

$$R(\lambda) = D - \lambda i(K - \Psi^* X)(I - \lambda iV)^{-1}\Psi D \tag{72}$$

and, obviously,

$$R^*(\lambda) = D^* + i\lambda D^*\Psi^*(I + i\lambda V^*)^{-1}(K^* - X\Psi). \tag{73}$$

In this case the Bezoutian of the quadruple $(L^*, R^*; L, R)$ associated with (71) and realizations (68), (73), (67), (72) equals iX.

Conversely, let any symmetric factorization of $G(\lambda)$ be given, i.e. let a matrix polynomial $R(\lambda)$ satisfying (70) be given. Then $R(\lambda)$ admits a minimal realization of the form

$$R(\lambda) = D - \lambda iK_R(I - \lambda iV)^{-1}\Psi D \tag{74}$$

for some K_R and, consequently,

$$R^*(\lambda) = D^* + \lambda iD^*\Psi^*(I + \lambda iV^*)^{-1}K_R^*. \tag{75}$$

If $\mathbb{B}$ denotes the Bezoutian of $(L^*, R^*; L, R)$ associated with (71) and (68), (75), (67), (74), then the matrix $-i\mathbb{B}$ is a hermitian solution of (66).

The above correspondence between the set of hermitian solutions of (66) and the set of all the symmetric factorizations of the polynomial $G(\lambda)$ is one-to-one.

The proof of this theorem is immediately obtained from Proposition 10 and Theorem 4 if we set $A_1 = iA$, $X_1 = iX$.

For two matrix polynomials $L(\lambda)$ and $R(\lambda)$ satisfying (71) we define the numbers $\gamma_+(L, R)$, $\gamma_-(L, R)$, $\gamma_0(L, R)$ as the total common multiplicity of the common eigenvalues of $L(\lambda)$ and $R(\lambda)$ in the upper half-plane, in the lower half-plane and on the real axis, respectively, i.e. $\gamma_-(L_0) = \gamma_-(L, R)$, $\gamma_+(L_0) = \gamma_+(L, R)$ and $\gamma_0(L_0) = \gamma_0(L, R)$. Then, combining Theorem 5 with Theorem 7, we obtain the following result.

Corollary 3. *Let X be a hermitian solution of (66) generating the symmetric factorization (71), where $R(\lambda)$ is defined by (72). Then* $\pi(X) = \pi(A) - \gamma_-(L,R)$, $\nu(X) = \nu(A) - \gamma_+(L,R)$, $\delta(X) = \delta(A) + \gamma_+(L,R) + \gamma_-(L,R)$.

Indeed, for example, $\pi(X) = \pi(-i\mathbb{B}) = \nu(i\mathbb{B}) = \gamma_+(iA) - \gamma_+(A_0) = \pi(A) - \gamma_-(L,R)$. From this corollary we obtain the next proposition concerning coprime factorizations of $G(\lambda)$.

Theorem 8. *Let X be a hermitian invertible solution of (66). Then* $\mathrm{In}\,(X) = \mathrm{In}\,(A)$ *and X generates a coprime symmetric factorization $L^*(\lambda)L(\lambda) = R^*(\lambda)R(\lambda)$, where $L(\lambda)$ and $R(\lambda)$ have the minimal realizations (67) and (72), respectively.*

Conversely, let

$$L^*(\lambda)L(\lambda) = R^*(\lambda)R(\lambda) \tag{76}$$

be a symmetric factorization, where $L(\lambda)$ and $R(\lambda)$ have realizations (67) and (74) such that $L(\lambda)$ and $R(\lambda)$ are right coprime. Then (66) has an invertible solution which coincides with the Bezoutian $\mathbb{B}$ of the quadruple $(L^*, R^*; L, R)$ associated with (76) and realizations (67), (74).

Among all the hermitian solutions of (66) we pay special attention to *extremal* (i.e. *maximal* and *minimal*) solutions. (For details on extremal solutions, see, e.g., (Lancaster and Rodman, 1995; Rodman, 1980; 1983; Shayman, 1983)). A hermitian solution X_+ of (66) is called *maximal* if for any hermitian solution X of (66) the inequality $X \leq X_+$ holds true. Similarly, a hermitian solution of (66) X_- is *minimal* if $X \geq X_-$.

We show in this section that extremal solutions correspond to spectral factorizations of the matrix polynomial $G(\lambda)$ under the bijective correspondence described in Theorem 9. Recall (Gohberg *et al.*, 1982; 1983) that the factorization $G(\lambda) = R_+^*(\lambda)R_+(\lambda)$ of a non-negative matrix polynomial $G(\lambda)$ is called *right spectral* if all the eigenvalues of $R_+(\lambda)$ are in the closed upper half-plane, while the factorization $G(\lambda) = R_-^*(\lambda)R_-(\lambda)$ is called *left spectral* if the spectrum of $R_-(\lambda)$ is situated in the closed lower half-plane. The polynomials $R_+(\lambda)$ and $R_-(\lambda)$ are called the *right* and *left spectral divisors* of G, respectively.

The connection between the extremal solutions of (66) and the spectral factorizations of $G(\lambda)$ is explored in the following result:

Theorem 9. *Let*

$$L^*(\lambda)L(\lambda) = R_+^*(\lambda)R_+(\lambda) \tag{77}$$

be a right spectral factorization of $L^*(\lambda)L(\lambda)$. Then $R_+(\lambda)$ has the realization

$$R_+(\lambda) = D - i\lambda K_{R_+}(I - i\lambda V)^{-1}\Psi D, \tag{78}$$

and $X_+ = -i\mathbb{B}_+$, where $\mathbb{B}_+$ is the Bezoutian of $(L^*, L; R_+^*, R_+)$ based on (77) and associated with realizations (67) and (78), is the maximal hermitian solution of (66). In this case we have $\mathrm{In}\,(X_+) = (\pi(A), 0, \nu(A) + \delta(A))$.

In a similar way, let

$$L^*(\lambda)L(\lambda) = R_-^*(\lambda)R_-(\lambda) \tag{79}$$

be a left spectral factorization of $L^*(\lambda)L(\lambda)$. *Then*

$$R_-(\lambda) = D - \lambda i K_{R_-}(I - \lambda i V)^{-1}\Psi D, \tag{80}$$

and $X_- = -i\mathbb{B}_-$, *where* $\mathbb{B}_-$ *is the Bezoutian of* $(L^*, R^*; L, R)$ *associated with factorization (79) and realizations (67), (80), is the minimal solution of (66). In this case,*

$$\mathrm{In}\,(X_-) = \big(0, \nu(A), \pi(A) + \delta(A)\big). \tag{81}$$

Conversely, let X_+ *be the maximal solution of (66). Then* X_+ *generates the right spectral factorization (77), where*

$$R_+(\lambda) = D - \lambda(K - \Psi^* X_+)(I - \lambda V)^{-1}\Psi D$$

and

$$\mathrm{In}\,(A - WX_+) = \big(0, \pi(A) + \nu(A), \delta(A)\big).$$

Analogously, the minimal solution X_- *generates the left spectral factorization (79), where*

$$R_-(\lambda) = D - \lambda i(K - \Psi^* X_-)(I - \lambda i V)^{-1}\Psi D.$$

Moreover,

$$\mathrm{In}\,(A - WX_-) = \big(\pi(A) + \nu(A), 0, \delta(A)\big).$$

Proof. First we observe that any $R(\lambda)$ satisfying (76), with minimal realization (74), has a left null pair $(-iA_R^{-1}, \Psi)$, where $A_R = A - WX_R$. From Theorem 7 we know that $K_R = K - \Psi^* X_R$, where X_R is a solution of (66) generated by the given factorization (76) in the sense of Theorem 7. The last equality implies that $A_R = V + i\Psi(K - \Psi^* X_R) = A - WX_R$. Now let us prove the first part of the theorem. Suppose that (77) holds and $X_1 = -i\mathbb{B}_+$ is the solution of (66) corresponding to (77). For an arbitrary hermitian the solution X of (66) set $Y = X - X^+$. Then one easily checks that Y satisfies

$$A_{R_+}^* Y + Y A_{R_+} = YWY, \tag{82}$$

where $A_{R_+} = A - WX^+$. In view of Theorem 7, the solution Y of (82) generates a factorization of the matrix polynomial $R_+^*(\lambda)R_+(\lambda)$:

$$R_+^*(\lambda)R_+(\lambda) = M^*(\lambda)M(\lambda).$$

Now Theorem 5 implies $\pi(Y) = \gamma_-(R_+) - \gamma_+(R, M)$, $\nu(Y) = \gamma_+(R_+) - \gamma_+(R_+, M)$, $\delta(Y) = \gamma_+(R_+, M) + \gamma_-(R_+, M) + \gamma_0(R_+, M)$. Since the spectrum of $R_+(\lambda)$ lies in the closed upper half-plane, $\gamma_-(R_+) = 0$, and hence $\gamma_-(R_+, M) = 0$. We obtain

$\text{In}\,(Y) = (0, \gamma_+(R_+) - \gamma_+(R_+, M), \gamma_+(R_+, M) + \gamma_0(R_+, M))$, i.e. for any solution X of (66) we have $X - X_+ \le 0$, i.e. X_+ is the maximal solution of (66). As $-X_+$ is also a solution of (82), corresponding to the factorization $R_+^*(\lambda)R_+(\lambda) = L^*(\lambda)L(\lambda)$, we have

$$\text{In}\,(-X_+) = \big(0, \gamma_+(R_+) - \gamma_+(R_+, L), \gamma_+(R_+, L) + \gamma_0(R_+, L)\big). \tag{83}$$

On the other hand, since X_+ is a solution of (66),

$$\text{In}\,(X_+) = \big(\gamma_-(L) - \gamma_-(R_+, L), \gamma_+(L) - \gamma_+(R_+, L), \gamma_+(R_+, L) + \gamma_0(R_+, L)\big). \tag{84}$$

But $\gamma_-(R_+, L) = 0$, and we see from (83) and (84) that $\gamma_+(R_+, L) = \gamma_+(L)$ and hence $\text{In}\,(X_+) = (\gamma_-(L), 0, \gamma_+(L) + \gamma_0(L))$, which proves the assertion on $\text{In}\,(X_+)$.

In a similar way one proves that the left spectral factorization generates the minimal solution X_-, where X_- satisfies (81).

Conversely, let X_+ be the maximal solution of (66). Then for any hermitian solution X of (66), $Y = X - X_+$ is a non-positive definite solution of (82). Using again Theorem 7, we infer that Y generates the factorization of $R_+^*(\lambda)R_+(\lambda)$ into the product

$$R_+^*(\lambda)R_+(\lambda) = M^*(\lambda)M(\lambda). \tag{85}$$

In view of Theorem 5,

$$\text{In}\,(Y) = \big(0, \nu(Y), \delta(Y)\big) = \gamma_-(R_+) - \gamma_-(R_+, M),$$

$$\gamma_+(R_+) - \gamma_+(R_+, M), \gamma_0(R_+) + \gamma_+(R_+, M) + \gamma_-(R_+, M)\big). \tag{86}$$

We see from (86) that $\gamma_-(R_+) = \gamma_-(R_+, M)$ for any factorization (85). In particular, this relation holds true for $M(\lambda)$ with the spectrum located in the closed upper half-plane. It is known that such a factorization exists. Thus we have $\gamma_-(R_+) = 0$. The hermitian solution $Y = -X_+$ of (82) generating the factorization

$$R_+^*(\lambda)R_+(\lambda) = L^*(\lambda)L(\lambda)$$

satisfies the equation

$$\text{In}\,(-X_+) = \big(0, \gamma_+(R_+) - \gamma_+(R_+, L), \gamma_0(R_+) + \gamma_+(R_+, L) + \gamma_-(R_+, L)\big). \tag{87}$$

Comparing (87) and (84), we conclude that $\gamma_+(R_+, L) = \gamma_+(L)$ and therefore

$$\text{In}\,(-X_+) = \big(0, \gamma_+(R_+) - \gamma_+(L), \gamma_0(R_+)) + \gamma_+(L)\big). \tag{88}$$

Finally, since (88) holds, we have $\text{In}\,(A - WX_+) = (0, \pi(X_+) + \nu(A); \delta(X_+) - \nu(A))$, or, in view of Corollary 3,

$$\text{In}\,(A - WX_+) = \big(0, \pi(A) + \nu(A), \delta(A)\big).$$

The assertions about X_- are proved in a similar way. ■

References

Ando T. (1988): *Matrix Qudratic Equations.* — Sapporo, Japan: Hokkaido University Press.

Anderson B.D.O. and Jury E.I. (1976): *Generalized Bezoutian and Sylvester matrices in multivariable linear control.* — IEEE Trans. Automat. Contr., Vol.AC-21, pp.551–556.

Ball J.A., Groenewald G., Kaashoek M.A. and Kim J. (1994): *Column reduced rational matrix functions with given null-pole data in the complex plane.* — Lin. Alg. Appl., Vol.203/204, pp.67–110.

Bart H., Gohberg I. and Kaashoek M.A. (1979): *Minimal Factorization of Matrix and Operator Functions.* — Basel: Birkhäuser.

Carlson D. and Shneider H. (1963): *Inertia theorem for matrices: The semidefinite case.* — Math. Anal. Appl., Vol.6, pp.430–446.

Dym H. (1991): *A Hermite theorem for matrix polynomials*, In: Operator Theory: Advances and Applications (H. Bart, I. Gohberg and M.A. Kaashoek, Eds), pp.191–214.

Dym H. and Young N.Y. (1990): *A Shur-Cohn theorem for matrix polynomials.* — Proc. Edinburgh Math. Soc., Vol.33, pp.337–366.

Gohberg I., Kaashoek M.A., Lerer L. and Rodman L. (1981): *Common multiples and common divisors of matrix polynomials, I.: Spectral method.* — Indiana Univ. Math. J., Vol.30, pp.321–356.

Gohberg I., Kaashoek M.A., Lerer L. and Rodman L. (1984): *Minimal divisors of rational matrix functions with prescribed zero and pole structure*, In: Operator theory: Advances and Applications. — Basel: Birkhäuser, pp.241–275.

Gohberg I., Kaashoek M.A. and Lancaster P. (1988): *General theory of regular matrix polynomials and band Toeplitz operators.* — Int. Eqns. Oper. Theory, Vol.6, pp.776–882.

Gohberg I., Lancaster P. and Rodman L. (1982): *Matrix Polynomials.* — New York: Academic Press.

Gohberg I., Lancaster P. and Rodman L. (1983): *Matrices and Indefinite Scalar Products.* — Basel: Birkhäuser.

Gohberg I., Lerer L. and Rodman L. (1980): *On factorization indices and completely decomposable matrix polynomials.* — Tech. Rep., Tel-Aviv University, pp.47–80.

Gomez G. and Lerer L. (1994): *Generalized Bezoutian for analytic operator functions and inversion of stuctured operators*, In: System and Networks: Mathematical Theory and Applications (U. Helmke, R. Mennicken and J. Saures, Eds.), Academie Verlag, pp.691–696.

Haimovici J. and Lerer L. (1995): *Bezout operators for analytic operator functions I: A general concept of Bezout operators.* — Int. Eqns. Oper. Theory, Vol.21, pp.33–70.

Haimovici J. and Lerer L. (2001): *Bezout operators for analytic operator functions II.* — In preparation.

Hearon J.Z. (1977): *Nonsingular solutions of $TA - TB = C$.* — Lin. Alg. Appl., Vol.16, pp.57–65.

Ionescu V. and Weiss M. (1993): *Continuous and discrete-time Riccati theory: A Popov-function approach.* — Lin. Appl., Vol.193, pp.173–209.

Kailath T. (1980): *Linear systems.* — Engelwood Cliffs, N.J.: Prentice Hall.

Karelin I. and Lerer L. (2001): *Generalized Bezoutian, factorization of rational matrix functions and matrix quadratic equations.* — Oper. Theory Adv. Appl., Vol.122, pp.303–321.

Karelin I., Lerer L. and Ran A.C.M. (2001): *J-symmetric factorizations and algebraic Riccati equation.* — Oper. Theory: Adv. Appl., Vol.124, pp.319–360.

Lerer L. (1989): *The matrix quadratic equations and factorization of matrix polynomials.* — Oper. Theory: Adv. Appl., Vol.40, pp.279–324.

Lancaster P. and Rodman L. (1995): *Algebraic Riccati Equations.* — Oxford: Oxford University Press.

Lerer L. and Ran A.C.M. (1996): *J-pseudo spectral and J-inner-pseudo-outer factorizations for matrix polynomials.* — Int. Eqns. Oper. Theory, Vol.29, pp.23–51.

Lerer L. and Rodman L. (1996a): *Common zero structure of rational matrix functions.* — J. Funct. Anal., Vol.136, pp.1–38.

Lerer L. and Rodman L. (1996b): *Bezoutians of rational matrix functions.* — J. Funct. Anal., Vol.141, pp.1–36.

Lerer L. and Rodman L. (1996c): *Symmetric factorizations and locations of zeroes of rational matrix functions.* — Lin. Multilin. Alg., Vol.40, pp.259–281.

Lerer L. and Rodman L. (1999): *Bezoutian of rational matrix functions, matrix equations and factorizations.* — Lin. Alg. Appl., Vol.302–303, pp.105–133.

Lerer L. and Tismenetsky M. (1982): *The Bezoutian and the eigenvalue separation problem.* — Int. Eqns. Oper. Theory, Vol.5, pp.386–445.

Rodman L. (1980): *On extremal solutions of the algebraic Riccati equations*, In: A.M.S. Lectures on Applied Math., Vol.18, pp.311–327.

Rodman L. (1983): *Maximal invariant neutral subspaces and an application to the algebraic Riccati equation.* — Manuscript Math., Vol.43, pp.1–12.

Shayman M.A. (1983): *Geometry of the algebraic Riccati equations. I, II.* — SIAM J. Contr., Vol.21, pp.375–394 and 395–409.