Generalized Bezoutian and Matrix Equations

L. Lerer*
Department of Mathematics
Technion — Israel Institute of Technology
Haifa 32000, Israel

and

M. Tismenetsky
IBM — Israel Scientific Center
Haifa 32000, Israel

Submitted by Peter Lancaster

ABSTRACT

A natural generalization of the classical Bezout matrix of two polynomials is introduced for a family of several matrix polynomials. The main aim of the paper is to show that this generalized Bezoutian serves as an adequate connecting link between the class of equations in matrix polynomials $M(\lambda)Y(\lambda)+Z(\lambda)L(\lambda)=R(\lambda)$ and the class of linear matrix equations $AX-XB=C$. Each equation in one of these classes is coupled with a certain equation in the other class so that for each couple the generalized Bezoutian corresponding to a solution $(Y(\lambda),Z(\lambda))$ of the equation in matrix polynomials is a solution of the matrix equation, and conversely, any solution X of the matrix equation is a generalized Bezoutian corresponding to a certain solution of the equation in matrix polynomials. In particular, either both equations are solvable or both have no solutions. Explicit formulas connecting the solutions of the two equations are given. Also, various representation formulas for the generalized Bezoutian are derived, and its relation to the resultant matrix and the greatest common divisor of several matrix polynomials is discussed.

0. INTRODUCTION

In this paper the concept of a generalized Bezoutian matrix for several matrix polynomials is introduced, and its fundamental role in the theory of matrix equations and equations in matrix polynomials is established.

*The work of this author was supported by the Fund for Promotion of Research at the Technion.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/88/$3.50

Let

$$L_r(\lambda) = \sum_{j=0}^{\nu_r} \lambda^j l_{rj}, \quad M_r(\lambda) = \sum_{j=0}^{\mu_r} \lambda^j m_{rj} \qquad (r = 1, \dots, s) \tag{0.1}$$

be $n \times n$ matrix polynomials (i.e., the coefficients m_{rj} and l_{rj} are $n \times n$ complex matrices) which satisfy the equation

$$\sum_{r=1}^{s} M_r(\lambda) L_r(\lambda) = 0. \tag{0.2}$$

Then the matrix valued function

$$\Gamma(x, y) := (x - y)^{-1} \sum_{r=1}^{s} M_r(x) L_r(y) \tag{0.3}$$

is a polynomial in the scalar variables x and y:

$$\Gamma(x, y) = \sum_{i,j=0}^{\mu-1, \nu-1} x^i y^i \Gamma_{ij}$$

$$\left(\Gamma_{ij} \in \mathbb{C}^{n \times n}, \quad i = 0, \dots, \mu - 1, \quad j = 0, \dots, \nu - 1\right), \tag{0.4}$$

where $\mu = \max_{1 \leqslant r \leqslant s} \mu_r$, $\nu = \max_{1 \leqslant r \leqslant s} \nu_r$. The $\mu n \times \nu n$ block matrix

$$\mathbf{B} = \mathbf{B}(M_1, \dots, M_s; L_1, \dots, L_s) := \left[\Gamma_{ij}\right]_{i,j=0}^{\mu-1, \nu-1} \tag{0.5}$$

will be referred to as the *generalized Bezoutian* associated with the equation (0.2).

The matrix (0.5) coincides with the classical Bezout matrix $B(a, b)$ introduced by Sylvester [34] for two scalar $(n = 1)$ polynomials $a(\lambda)$ and $b(\lambda)$ if one puts $s = 2$ and $M_1(\lambda) = L_2(\lambda) = a(\lambda)$, $L_1(\lambda) = -M_2(\lambda) = b(\lambda)$. Setting $s = 2$ and replacing $M_2(\lambda)$ by $-M_2(\lambda)$ in (0.2)–(0.5), one obtains the Bezout matrix $B_{M_1, M_2}(L_2, L_1)$ of a quadruple of matrix polynomials introduced by Anderson and Jury [1] and subsequently studied in Bitmead, Kung, Anderson and Kailath [5] and Lerer and Tismenetsky [28]. The "tensor" Bezoutian $\mathbf{B}^{\otimes}(A, B)$ of two matrix polynomials $A(\lambda)$ and

$B(\lambda)$—introduced and studied in several papers, including Heinig [17, 18], and Barnett and Lancaster [4] (see also Heinig and Rost [19])—is obtained by setting $s=2$ and $L_1(\lambda)=M_2(\lambda)=A(\lambda)\otimes I$, $M_1(\lambda)=L_2(\lambda)=I\otimes B(\lambda)$, where $\otimes$ stands for the tensor (or Kronecker) matrix multiplication.

The Bezout matrix $B(a,b)$ enjoys wide use in various problems concerning root location of (scalar) polynomials. The classical investigations in this direction were based on the theory of symmetric and Hermitian forms and on the key property that the nullity of $B(a,b)$ is equal to the degree of the greatest common divisor of $a(\lambda)$ and $b(\lambda)$. (For details and history see the comprehensive survey of Krein and Neimark [22].) More recently Kalman [20] and Carlson and Datta [9, 6] have found certain Liapunov type matrix equations of which the Bezoutian $B(a,b)$ is a solution. This enabled the above authors to employ the well-known inertia theorems in the study of the root separation problems for scalar polynomials. We refer to the expository paper of Datta [10] for a unified presentation of proofs and further references.

In the case of matrix polynomials various eigenvalue separation problems have been solved by the present authors in [28] (see also Lerer, Rodman, and Tismenetsky [27]), where a suitable Liapunov type equation for the Bezoutian $\mathbf{B}_{M_1,M_2}(L_2,L_1)$ has been found. Furthermore, interpreting Bezoutians as solutions of certain Liapunov type equations proved to be fruitful in the inversion problem for (block) Toeplitz and (block) Hankel matrices (see Lerer and Tismenetsky [31] for details and references).

One of the main results of the present paper gives a family of Liapunov type matrix equations of which the generalized Bezoutian defined in (0.2)–(0.5) is a solution. To be specific, we adopt the notation

$$C_L=\begin{bmatrix} 0 & I & & \\ & & \ddots & \\ & & & I \\ -l_0 & -l_1 & \cdots & -l_{\nu-1} \end{bmatrix},$$

$$\hat{C}_M=\begin{bmatrix} 0 & & & -m_0 \\ I & & & -m_1 \\ & \ddots & & \vdots \\ & & I & -m_{\mu-1} \end{bmatrix} \tag{0.6}$$

for the first and the second companion matrices of the monic matrix polynomials

$$L(\lambda)=\lambda^{\nu}I+\sum_{j=0}^{\nu-1}\lambda^{j}l_{j},\qquad M(\lambda)=\lambda^{\mu}I+\sum_{j=0}^{\mu-1}\lambda^{j}m_{j},\tag{0.7}$$

respectively. Also, for a $n\times n$ matrix polynomial $A(\lambda)=\sum_{j=0}^{\alpha}\lambda^{j}a_{j}$ we set

$$A\big(X^{(0)},C_{L}\big):=\sum_{j=0}^{\alpha}a_{j}X^{(0)}C_{L}^{j},\qquad A\big(\hat{C}_{M},Z^{(0)}\big):=\sum_{j=0}^{\alpha}\hat{C}_{M}^{j}Z^{(0)}a_{j},$$

where

$$X^{(0)}=\Big[\underbrace{I\quad 0\quad\cdots\quad 0}_{\nu}\Big],\qquad Z^{(0)}=\Big[\underbrace{I\quad 0\quad\cdots\quad 0}_{\mu}\Big]^{T}.$$

Then a simplified version of Theorem 3.1 of the present paper amounts to the following matrix equation for the generalized Bezoutian $\mathbf{B}$ defined by (0.2)–(0.5):

$$\hat{C}_{M}\mathbf{B}-\mathbf{B}C_{L}=\sum_{r=1}^{s}M_{r}\big(\hat{C}_{M},Z^{(0)}\big)L_{r}\big(X^{(0)},C_{L}\big).\tag{0.8}$$

Note that by setting $s=2$ and by a suitable choice of L, M one obtains from (0.8) the matrix equations for the Bezoutians established in [20], [9], [6], [18], [28], [11].

Moreover, our results show that the generalized Bezoutian serves as an adequate connecting link between matrix equations and equations in matrix polynomials. To explain this, consider the equation in matrix polynomials

$$M(\lambda)Y(\lambda)+Z(\lambda)L(\lambda)=R(\lambda),\tag{E$_1$}$$

where the coefficient polynomials $M(\lambda)$, $L(\lambda)$ are as in (0.7), and $R(\lambda)$ is a given $n\times n$ matrix polynomial with $\deg R\leqslant\nu+\mu$. The set of all solution pairs $(Y(\lambda),Z(\lambda))$ of (E_1) with $\deg Y\leqslant\gamma$, $\deg Z\leqslant\delta$ will be denoted by $\mathbf{Y}_{\gamma,\delta}$ (E_1). Note that the right hand side $R(\lambda)=\sum_{j=0}^{\mu+\nu}\lambda^{j}r_{j}$ can always be represented in the form

$$R(\lambda)=-\sum_{r=3}^{s}M_{r}(\lambda)L_{r}(\lambda),\tag{0.9}$$

where

$$\deg M_r \leqslant \mu, \quad \deg L_r \leqslant \nu \qquad (r = 3,\dots, s). \tag{0.10}$$

[For instance, take $s = 4$ and $M_3(\lambda) = -I$, $L_3(\lambda) = \sum_{j=0}^{\nu-1}\lambda^j r_j$, $M_4(\lambda) = \sum_{j=0}^{\mu}\lambda^j r_{\nu+j}$, $L_4(\lambda) = -\lambda^\nu I$.] The equation ($E_1$) with $R(\lambda)$ written in the form (0.9)–(0.10) becomes an equality of type (0.2). Thus, assuming that a representation of the form (0.9)–(0.10) is fixed, with each solution $(Y(\lambda), Z(\lambda)) \in \mathbf{Y}_{\nu,\mu}$ (E_1) we can associate the generalized Bezoutian $\mathbf{B}_{Y,Z} :=$ $\mathbf{B}(M, Z, M_3,\dots, M_s; Y, L, L_3,\dots, L_s)$. Since $L(X^{(0)}, C_L) = 0$ and $M(\hat{C}_M, Z^{(0)}) = 0$, equation (0.8) implies that $X = \mathbf{B}_{Y,Z}$ is a solution of the equation

$$\hat{C}_M X - XC_L = R^{(0)}, \tag{E_2}$$

where

$$R^{(0)} := \sum_{r=3}^{s} M_r\big(\hat{C}_M, Z^{(0)}\big)L_r\big(X^{(0)}, C_L\big). \tag{0.11}$$

It turns out that the converse is also true: each solution $X = S = [s_{ij}]_{i,j=0}^{\mu-1,\nu-1}$ of the equation (E_2) generates a solution $(Y_S(\lambda), Z_S(\lambda))$ of (E_1) with the property $S = \mathbf{B}_{Y_S,Z_S}$. The general form of all solutions with this property is given by the formulas

$$Y_S(\lambda) = \sum_{k=0}^{\nu-1}\lambda^k s_{\mu-1,k} - \sum_{r=3}^{s} m_{r\mu}L_r(\lambda) + AL(\lambda),$$

$$Z_S(\lambda) = -\sum_{k=0}^{\mu-1}\lambda^k s_{k,\nu-1} - \sum_{r=3}^{s} M_r(\lambda)l_{r\nu} + M(\lambda)B, \tag{0.12}$$

where A and B are arbitrary $n\times n$ matrices such that $A + B = r_{\nu+\mu}$. It follows that any solution of (E_1) can be written in the form (0.12) for a suitable S in the set $\mathbf{S}$ (E_2) of all solutions of the equation (E_2). Moreover, the map $\mathscr{B}: \mathbf{Y}_{\nu,\mu}(E_1) \to \mathbf{S}(E_2)$ defined by $\mathscr{B}(Y(\lambda), Z(\lambda)) = \mathbf{B}_{Y,Z}$ maps the subsets $\mathbf{Y}_{\nu-1,\mu}(E_1)$ and $\mathbf{Y}_{\nu,\mu-1}(E_1)$ onto $\mathbf{S}(E_2)$ in a $1 \leftrightarrow 1$ way. The formulas for the corresponding inverse mappings are obtained from (0.12) by setting $B = 0$ and $A = 0$, respectively. In Sections 4–5 of the present paper these results are stated and proved in a somewhat more general setting.

It is worth noting that in the above considerations the equation in matrix polynomials (E_1) and a fixed representation (0.9)–(0.10) of $R(\lambda)$ are assumed to be given, and the results describe the interrelations between the solutions of the equations (E_1) and (E_2), provided the right hand side $R^{(0)}$ in (E_2) is defined by (0.11). An important point is that, conversely, any matrix equation of the form (E_2) with an arbitrary right hand side $R^{(0)} = [R_{ij}]_{i,j=0}^{\mu-1,\nu-1}$ can be put into this framework. Namely, taking $R(\lambda) = \sum_{i,j=0}^{\mu-1,\nu-1}\lambda^{i+j}R_{ij}$ in (E_1) and fixing the representation (0.9) with $s = \mu + 2$ and $L_{j+3}(\lambda) = \sum_{k=0}^{\nu-1}\lambda^k R_{jk}$, $M_{j+3}(\lambda) = -\lambda^j I$ ($j = 0,\dots,\mu-1$), one checks that the right hand side in (0.11) coincides with the given matrix $[R_{ij}]_{i,j=0}^{\mu-1,\nu-1}$, and hence any solution X of (E_2) coincides with $\mathbf{B}_{Y,Z}$ for a suitable solution $(Y(\lambda), Z(\lambda))$ of (E_1). It follows, in particular, that with any $\mu n \times \nu n$ block matrix X we can associate a family of $n \times n$ matrix polynomials $L_r(\lambda)$, $M_r(\lambda)$ ($r = 1,\dots,s$) satisfying (0.2) such that $X = \mathbf{B}(M_1,\dots,M_s;\ L_1,\dots,L_s)$. This observation serves as a starting point in our analysis of the inversion problem for block matrices carried out in [31].

Some further remarks are in order. Firstly, we note that the importance of matrix equations of form (E_2) with block-companion coefficients has been demonstrated by Lancaster, Lerer, and Tismenetsky [23], who show how a general matrix equation $AX - XB = C$ can be transformed into an equation of the form (E_2).

Secondly, assuming $\deg R \leqslant \mu + \nu - 2$ in (E_1) and choosing the representation (0.9) with $l_{r\nu} = m_{r\mu} = 0$ ($r = 3,\dots,s$), one sees that (0.12) (with $A = B = 0$) coincides with the formulas established earlier by S. Barnett [2], who was probably the first to observe the connection between the equation (E_1) (with $\deg R \leqslant \mu + \nu - 2$) and some matrix equation of type (E_2). Note that our general method of constructing the matrix equations (E_2) corresponding to (E_1) and the basic role of the generalized Bezoutian which is revealed in the present paper are new also in the case $\deg R \leqslant \mu + \nu - 2$.

Thirdly, we mention the work of Gohberg, Kaashoek, and Lay [13], where explicit formulas for the solution of (E_1) are obtained in terms of some contour integrals involving $L(\lambda)$, $M(\lambda)$, and $R(\lambda)$, provided the spectra of $L(\lambda)$ and $M(\lambda)$ are disjoint. These formulas are derived as an application of the general theory of linearizations of analytic operator functions developed in [13]. In the framework of this theory the equation (E_1) is reduced to its "linearization," which is again an equation in matrix polynomials of the form, say,

$$(\hat{C}_M - \lambda I)Z_2(\lambda) + Z_1(\lambda)(C_L - \lambda I) = Q(\lambda), \tag{0.13}$$

where $Q(\lambda)$ is a $\mu n \times \nu n$ matrix polynomial determined by $R(\lambda)$, $M(\lambda)$, and $L(\lambda)$. Note that in the approach of the present paper the equation (E_1) is coupled with an equation of type (0.13), where the right hand side $Q(\lambda)$ and $Z_1(\lambda)$, $Z_2(\lambda)$ do not depend on λ.

Finally, we remark that the main results of the present paper can be extended to the case of polynomials whose coefficients are bounded operators on a Banach space. In this regard we mention the works of Lerer, Rodman, and Tismenetsky [27], Clancey and Kon [8], and Lancaster and Maroulas [24], where the Bezout operator for a quadruple of operator polynomials has been introduced and studied.

The paper is organized as follows. In the first preliminary section we recall some basic facts about matrix polynomials. Section 2 contains various representation formulas for the generalized Bezoutian and a matrix relation connecting the generalized Bezoutian to the Sylvester resultant matrix. In Section 3 we present a class of Liapunov type matrix equations of which the generalized Bezoutian is a solution. In particular, equation (0.8) and its corollaries are established here. The main results concerning the linkage between the equation in matrix polynomials (E_1) and the matrix equation (E_2) are stated in Section 4 and proved in Section 5.

Preliminary versions of the results of the present paper are partially contained in the report [29] and in the Ph.D. thesis [35] of the second author guided by the first one.

1. PRELIMINARIES

We need the following notational devices.

By $\mathrm{row}(A_i)_{i=1}^k$ ($A_i \in \mathbb{C}^{n\times m}$, $i = 1,2,\dots,k$) we denote the $n \times mk$ block matrix $[A_1\ A_2\ \cdots\ A_k]$ consisting of one block row. Similarly, $\mathrm{col}(A_i)_{i=1}^k$ stands for the $nk\times m$ matrix containing one block column. The notation $A_1 \dotplus A_2 \dotplus \cdots \dotplus A_k$ as well as $\mathrm{diag}(A_i)_{i=1}^k$, is used along with $[\delta_{ij}A_j]_{i,j=1}^k$, where δ_{ij} is the Kronecker delta.

A pair of matrices (X,T) is referred to as a *right admissible pair* of order p if $X \in \mathbb{C}^{n\times p}$ and $T \in \mathbb{C}^{p\times p}$, while the pair (V,Z) with $Z \in \mathbb{C}^{q\times n}$ and $V \in \mathbb{C}^{q\times q}$ is a *left admissible pair* of order q. Note that here and elsewhere n is fixed and that, if not specified otherwise, the pairs are assumed to be right admissible. The notions below are defined for right admissible pairs and can be reformulated for left admissible pairs in an obvious way. Two pairs (X_1,T_1) and (X_2,T_2) of order p are called *similar* if there is a $p\times p$ invertible matrix Q such that $X_1 = X_2Q$ and $T_1 = Q^{-1}T_2Q$. Let the admissible pairs (X_1,T_1), (X_2,T_2) be of orders p_1 and p_2, respectively ($p_1 \geqslant p_2$). The pair (X_1,T_1) is said to be an *extension* of (X_2,T_2), or, what is equivalent, (X_2,T_2) is a *restriction* of (X_1,T_1), if there exists a $p_1\times p_2$ matrix S of full rank such that $X_1S = X_2$ and $T_1S = ST_2$. A pair (X,T) is called a *common* *restriction* of a family of admissible pairs (X_j, T_j) $(j = 1, \dots, s)$ if each of the pairs (X_j, T_j) is an extension of (X, T). A common restriction (X_0, T_0) of the pairs (X_j, T_j) $(j = 1, \dots, s)$ which is an extension of any other common restriction of these pairs is referred to as the *greatest common restriction* of the family (X_j, T_j) $(j = 1, \dots, s)$.

Passing to matrix polynomials we need the following. If ν is a nonnegative integer, l_j $(j = 0, 1, \dots, \nu)$ are $n \times n$ complex matrices, and λ is a complex variable, then the expression

$$L(\lambda) = \sum_{j=0}^{\nu} \lambda^j l_j \tag{1.1}$$

is called a *matrix polynomial of degree* ν, and in this case the matrix l_ν is referred to as the leading coefficient of $L(\lambda)$ (we don't exclude the case $l_\nu = 0$). Note that for any integer k, $k > \nu$, the function (1.1) can also be viewed as a matrix polynomial of degree k (with zero leading coefficient). By $\sigma(L)$ we denote the *spectrum* of the matrix polynomial (1.1): $\sigma(L) = \{\lambda \in \mathbb{C} : \det L(\lambda) = 0\}$. If $\sigma(L) \neq \mathbb{C}$, the polynomial $L(\lambda)$ is called *regular*. If the matrix l_ν in (1.1) is nonsingular, we say that $L(\lambda)$ is a *matrix polynomial of degree* ν *with invertible leading coefficient*. A polynomial of this kind with $l_\nu = I$, the $n \times n$ identity matrix, will be referred to as a *monic matrix polynomial of degree* ν. If $L(\lambda)$ is a matrix polynomial of the form (1.1) and (X, T) and (T, Z) are right and left admissible pairs, respectively, the following notation will be useful:

$$L(X,T) := \sum_{j=0}^{\nu} l_j X T^j, \qquad L(T,Z) := \sum_{j=0}^{\nu} T^j Z l_j.$$

We now recall some basic facts from the spectral theory of matrix polynomials (see the monograph [16] for a detailed exposition). If $L(\lambda)$ is a $n \times n$ monic matrix polynomial of degree ν, its *right standard pair* (X, T) is defined as an admissible pair of order νn such that $\operatorname{col}(XT^{j-1})_{j=1}^{\nu}$ is a nonsingular matrix and $L(X, T) = 0$. Similarly, a left admissible pair (T, Z) of order νn with $\det \operatorname{row}(T^{j-1}Z)_{j=1}^{\nu} \neq 0$ is called a *left standard pair* of $L(\lambda)$ if $L(T, Z) = 0$. Important examples of right and left standard pairs are provided by the *companion standard pairs* $(X^{(0)}, C_L)$ and $(\hat{C}_L, Z^{(0)})$, respectively. Any right (left) admissible pair which is similar to a right (left) standard pair of $L(\lambda)$ is again a right (left) standard pair of $L(\lambda)$, and conversely, any two right (left) standard pairs of the monic matrix polynomial $L(\lambda)$ are similar. If J is the Jordan form for C_L and $C_L = QJQ^{-1}$, the right standard pair (Φ, J), where $\Phi = X^{(0)}Q$, will be referred to as a *right standard Jordan pair of* $L(\lambda)$.

The spectral meaning of the notion of a standard pair lies in the fact that $\sigma(J) = \sigma(L)$ and the columns of $n \times \nu n$ matrix Φ form a canonical system of (right) Jordan chains of the matrix polynomial $L(\lambda)$ (see [16] for appropriate definitions). A *left standard Jordan pair* of a monic matrix polynomial is defined in a similar way. If $L(\lambda)$ is a matrix polynomial of degree ν with invertible leading coefficient, its right and left standard pairs are defined as right and left standard pairs of the monic polynomials $l_\nu^{-1}L(\lambda)$ and $L(\lambda)l_\nu^{-1}$, respectively.

Now, let $L(\lambda)$ be a regular matrix polynomial of degree ν with $\det L(0) \neq 0$. Let (Φ, J) be the right standard Jordan pair of the matrix polynomial $L_\infty(\lambda) := \lambda^\nu L(\lambda^{-1})$ and make the partitioning

$$J = \begin{bmatrix} J_1 & 0 \\ 0 & J_\infty \end{bmatrix},$$

where J_1 is nonsingular and J_∞ is a nilpotent matrix. Let $[\Phi_1 \;\; \Phi_\infty]$ be the corresponding partitioning of Φ. Any right admissible pair (X_F, T_F) which is similar to (Φ_1, J_1^{-1}) is referred to as a *right finite Jordan pair* of $L(\lambda)$, while any right admissible pair (X_∞, T_∞) similar to (Φ_∞, J_∞) is called a *right infinite Jordan pair* of $L(\lambda)$. Similarly one defines left finite and infinite Jordan pairs.

For an arbitrary regular matrix polynomial $L(\lambda)$ choose $\alpha \notin \sigma(L)$ and introduce the polynomial $L_\alpha(\lambda) = L(\lambda + \alpha)$, which obviously has the property $\det L_\alpha(0) \neq 0$. Any pair (X_F, T_F) with $X_F = X_{F\alpha}$, $T_F = T_{F\alpha} + \alpha I$, where $(X_{F\alpha}, T_{F\alpha})$ is a right finite Jordan pair of $L_\alpha(\lambda)$, will be called a *right finite Jordan pair of* $L(\lambda)$. A *right infinite Jordan pair* (X_∞, T_∞) of $L(\lambda)$ is defined by $X_\infty = X_{\infty\alpha}$, $T_\infty = T_{\infty\alpha}\,(I + T_{\infty\alpha})^{-1}$, where $(X_{\infty\alpha}, T_{\infty\alpha})$ is a right infinite Jordan pair of $L_\alpha(\lambda)$. Left Jordan pairs of $L(\lambda)$ are defined analogously via left Jordan pairs of $L_\alpha(\lambda)$. Although the above definitions depend on the choice of $\alpha \notin \sigma(L)$, it turns out that we obtain one and the same collection of finite (infinite) Jordan pairs for any $\alpha \notin \sigma(L)$. Also, a direct definition for any of these notions in terms of Jordan chains is possible (see [16]).

Concerning the divisibility theory for matrix polynomials we need the following notions and results (see [14–15] for details and proofs). Given matrix polynomials $L(\lambda)$ and $M(\lambda)$, we say that $M(\lambda)$ is a (*right*) *divisor* of $L(\lambda)$ if there is a matrix polynomial $Q(\lambda)$ such that $L(\lambda) = Q(\lambda)M(\lambda)$. For a regular matrix polynomial $M(\lambda)$ with right finite Jordan pair $(X_F^{(M)}, T_F^{(M)})$ to be a (right) divisor of $L(\lambda)$ it is necessary and sufficient that $L(X_F^{(M)}, T_F^{(M)}) = 0$. If the polynomial $L(\lambda)$ is also regular and $(X_F^{(L)}, T_F^{(L)})$ stands for its right finite Jordan pair, then $M(\lambda)$ is a (right) divisor of $L(\lambda)$ if and only if the pair $(X_F^{(M)}, T_F^{(M)})$ is a restriction of the pair $(X_F^{(L)}, T_F^{(L)})$.

A matrix polynomial $M(\lambda)$ is referred to as a (*right*) *common divisor* of the family of matrix polynomials $L_1(\lambda), L_2(\lambda),\dots, L_s(\lambda)$ if $M(\lambda)$ is a (right) divisor of each $L_j(\lambda)$ $(j=1,\dots,s)$. A (right) common divisor $M_0(\lambda)$ of the family $L_1(\lambda),\dots,L_s(\lambda)$ is called a (right) *greatest common divisor* of this family if any other common divisor of $L_j(\lambda)$ $(j=1,\dots,s)$ is a right divisor of $M_0(\lambda)$ as well. If $M_0(\lambda)=I$ is a (right) greatest common divisor of the family of polynomials $L_1(\lambda),\dots,L_s(\lambda)$, we say that these polynomials are (right) *coprime*. A regular matrix polynomial $M_0(\lambda)$ with right finite Jordan pair $(X_F^{(M_0)}, T_F^{(M_0)})$ is a (right) greatest common divisor of the family of regular matrix polynomials $L_1(\lambda),\dots,L_s(\lambda)$ if and only if the pair $(X_F^{(M_0)}, T_F^{(M_0)})$ is a greatest common restriction of the family of right finite Jordan pairs $(X_F^{(L_1)}, T_F^{(L_1)}),\dots,(X_F^{(L_s)}, T_F^{(L_s)})$ of the polynomials $L_j(\lambda)$ $(j=1,\dots,s)$. All the above notions and results have obvious left analogues.

2. EXPLICIT REPRESENTATION FORMULAS FOR THE GENERALIZED BEZOUTIAN

The following notation will be used throughout the paper. Given an $n\times n$ matrix polynomial $F(\lambda)=\sum_{i=0}^{k}\lambda^i f_i$ of degree k, we introduce the $\alpha n\times(\alpha+k)n$ block Toeplitz matrix

$$R_F^{(\alpha)}=\left.\begin{bmatrix} f_0 & f_1 & \cdots & f_k & & & \\ & f_0 & f_1 & & f_k & & \\ & & \ddots & \ddots & & \ddots & \\ & & & f_0 & f_1 & \cdots & f_k \end{bmatrix}\right\}\alpha \tag{2.1}$$

The submatrix of $R_F^{(\alpha)}$ formed by its first (respectively, last) β block columns is denoted by $\mathscr{U}_F^{(\alpha,\beta)}$ (respectively, $\mathscr{L}_F^{(\alpha,\beta)}$). Then for any two integers β_1 and β_2, $\beta_1+\beta_2=\alpha+k$, we can write

$$R_F^{(\alpha)}=\begin{bmatrix}\mathscr{U}_F^{(\alpha,\beta_1)} & \mathscr{L}_F^{(\alpha,\beta_2)}\end{bmatrix}.$$

Given matrix polynomials $M_r(\lambda)$, $L_r(\lambda)$ as in (0.1), we assume without loss of generality that $\max_{1\leqslant r\leqslant s}\nu_r:=\nu\leqslant\mu:=\max_{1\leqslant r\leqslant s}\mu_r$, and we consider $L_r(\lambda)$ and $M_r(\lambda)$ $(r=1,\dots,s)$ as matrix polynomials of degree ν and μ, respectively. Next, note that the equation (0.2) implies the following:

$$\sum_{r=1}^{s}\mathscr{L}_{M_r}^{(\alpha,\alpha)}\mathscr{L}_{L_r}^{(\alpha,\alpha)}=0,\quad \sum_{r=1}^{s}\mathscr{U}_{M_r}^{(\alpha,\alpha)}\mathscr{U}_{L_r}^{(\alpha,\alpha)}=0\qquad(\alpha=1,2,\dots). \tag{2.2}$$

In the sequel the notation P_μ is used to denote the reverse block unit matrix $[\delta_{i,\mu-i+1}I]_{i=1}^{\mu}$, where $I=I_n$ is the $n\times n$ identity matrix.

PROPOSITION 2.1. *Let* **B** *denote the Bezoutian associated with* (0.2). *Then for any right admissible pair* (X,T)

$$\mathbf{B}\,\mathrm{col}(XT^{j-1})_{j=1}^{\nu}=\sum_{r=1}^{s}P_\mu\mathscr{L}_{M_r}^{(\mu,\mu)}\,\mathrm{col}\big(L_r(X,T)T^{j-1}\big)_{j=1}^{\mu}. \tag{2.3}$$

Proof. Write the generating function $\Gamma(x,y)$ of **B** as $\Gamma(x,y)=\Sigma_{i=0}^{\mu-1}\Gamma_i(y)x^i$. From the definition of Γ we know that $(x-y)\Sigma_{j=0}^{\mu-1}\Gamma_i(y)x^i=\Sigma_{r=1}^{s}M_r(x)L_r(y)$. Comparing the coefficients of x^i on the two sides of this equality we obtain

$$\Gamma_k(y)=\sum_{r=1}^{s}\big(m_{r,k+1}+ym_{r,k+2}+\cdots+y^{\mu-k-1}m_{r,\mu}\big)L_r(y)$$

$$(k=0,\dots,\mu-1),$$

and hence

$$\Gamma_k(X,T)=\sum_{r=1}^{s}\sum_{j=k+1}^{\mu}m_{rj}L_r(X,T)T^{j-k-1}\qquad(k=0,\dots,\mu-1). \tag{2.4}$$

Representing $\Gamma_k(y)=\Sigma_{j=0}^{\nu-1}\Gamma_{kj}y^j$ and using the notation $\tilde{\Gamma}_k:=[\Gamma_{k0}\ \Gamma_{k1}\ \cdots\ \Gamma_{k,\nu-1}]$, we can write $\Gamma_k(X,T)=\tilde{\Gamma}_k\mathrm{col}(XT^{i-1})_{i=1}^{\nu}$. Since $\tilde{\Gamma}_k$ is the kth block row in **B**, it follows that

$$\mathbf{B}\,\mathrm{col}(XT^{j-1})_{j=1}^{\nu}=\mathrm{col}\big(\Gamma_k(X,T)\big)_{k=0}^{\mu-1}.$$

Substituting the expression for $\Gamma_k(X,T)$ from (2.4), one obtains (2.3). ■

Note that in the case $s=2$ the above result has been established by the authors in their earlier work [28, Proposition 1.1]. If in addition $M_1(\lambda)=L_2(\lambda)=a(\lambda)$ and $L_1(\lambda)=-M_2(\lambda)=b(\lambda)$ are scalar polynomials, one derives from (2.3) the well-known Barnett [3] factorization formula for $B(a,b)$ by setting $X=X^{(0)}$, $T=C_a$ (cf. [28, p. 400] and [25, p. 459]). For the tensor Bezoutian a block version of Barnett's formula is obtained in [4].

Now we deduce some further explicit expressions for the Bezoutian **B** in terms of the underlying matrix polynomials.

COROLLARY 2.2. *Let* **B** *denote the Bezoutian associated with* (0.2). *Then the following representations hold true*:

$$\mathbf{B}=\sum_{r=1}^{s} P_\mu \mathscr{L}_{M_r}^{(\mu,\mu)} \mathscr{U}_{L_r}^{(\mu,\nu)}, \tag{2.5}$$

$$-\mathbf{B}=\sum_{r=1}^{s} P_\mu \mathscr{U}_{M_r}^{(\mu,\mu)} \mathscr{L}_{L_r}^{(\mu,\nu)}. \tag{2.5'}$$

Proof. It is easily seen that

$$\operatorname{col}\left(L_r(X,T)T^{j-1}\right)_{j=1}^{\mu}=\mathscr{U}_{L_r}^{(\mu,\nu)}\operatorname{col}\left(XT^{j-1}\right)_{j=1}^{\nu}+\mathscr{L}_{L_r}^{(\mu,\mu)}\operatorname{col}\left(XT^{j-1}\right)_{j=1}^{\mu}T^{\nu}$$

for any right admissible pair (X,T) and for each $r=1,2,\ldots,s$.

In view of the first equation in (2.2) with $\alpha=\mu$, the formula in (2.3) implies

$$\mathbf{B}\operatorname{col}\left(XT^{j-1}\right)_{j=1}^{\nu}=\sum_{r=1}^{s} P_\mu \mathscr{L}_{M_r}^{(\mu,\mu)} \mathscr{U}_{L_r}^{(\mu,\nu)}\operatorname{col}\left(XT^{j-1}\right)_{j=1}^{\nu}.$$

Choosing (X,T) so that $\operatorname{col}(XT^{j-1})_{j=1}^{\nu}$ is invertible, one obtains (2.5).

Exploiting left admissible pairs, one derives a dual version of formula (2.3) for left admissible pairs (V,Z):

$$-\operatorname{row}\left(V^{j-1}Z\right)_{j=1}^{\mu}\mathbf{B}=\sum_{r=1}^{s}\operatorname{row}\left(V^{j-1}M_r(V,Z)\right)_{j=1}^{\nu}P_\nu \mathscr{L}_{L_r}^{(\nu,\nu)}. \tag{2.3'}$$

This implies (2.5′). ■

Note that in the case $\mu=\nu$

$$\left[\mathscr{U}_{M_r}^{(\mu,\mu)}\right]^{(T)}=P_\mu \mathscr{U}_{M_r}^{(\mu,\mu)}P_\mu,\qquad \left[\mathscr{L}_{L_r}^{(\mu,\mu)}\right]^{(T)}=P_\mu \mathscr{L}_{L_r}^{(\mu,\mu)}P_\mu, \tag{2.6}$$

where $A^{(T)}$ denotes the block transpose of A. Therefore, in case $\mu=\nu$, (2.5′) and (2.5) imply that

$$\left[\mathbf{B}(M_1,\ldots,M_s;L_1,\ldots,L_s)\right]^*=-\mathbf{B}(L_1^*,\ldots,L_s^*;M_1^*,\ldots,M_s^*),$$

where $L^*(\lambda)$ denotes the matrix polynomial whose coefficients are conjugate transposes of the coefficients of $L(\lambda)$.

Note that in the case $s = 2$ explicit expressions for the entries of $\mathbf{B}$ via the coefficients of the underlying polynomials were given in [1]. These expressions were presented in the matrix form (2.5)–(2.5′) in our earlier paper [28, pp. 397–398] (see also [12]). In the case of two scalar polynomials $a(\lambda)$ and $b(\lambda)$, explicit formulas for the entries of $B(a, b)$ appeared already in [7]. The matrix representations of type (2.5) in this case were established in [26] (see also [32] and [25]).

In certain cases we need the Bezoutians $\mathbf{B}_\infty$ and $\mathbf{B}_\alpha$ associated with the equations

$$\sum_{r=1}^{s} M_{r\infty}(\lambda) L_{r\infty}(\lambda) = 0, \tag{2.7}$$

whee $M_{r\infty} = \lambda^{\mu} M_r(\lambda^{-1})$, $L_{r\infty} = \lambda^{\nu} L_r(\lambda^{-1})$, $r = 1, 2, \ldots, s$, and

$$\sum_{r=1}^{s} M_{r\alpha}(\lambda) L_{r\alpha}(\lambda) = 0, \tag{2.8}$$

where $M_{r\alpha}(\lambda) = M_r(\lambda + \alpha)$, $L_{r\alpha}(\lambda) = L_r(\lambda + \alpha)$, $r = 1, 2, \ldots, s$, respectively. Writing (2.5) for $\mathbf{B}_\infty$ and using (2.5′), one obtains

$$\mathbf{B}_\infty = -P_\mu \mathbf{B} P_\nu.$$

A straightforward comparison of the Bezoutians $\mathbf{B}_\alpha$ and $\mathbf{B}$ based on the definition (0.2)–(0.5) gives

$$\mathbf{B}_\alpha = U_\alpha^{(\mu)} \mathbf{B} \left[U_\alpha^{(\nu)} \right]^{(T)},$$

where

$$U_\alpha^{(k)} = \left[\binom{j}{i} \alpha^{j-i} I \right]_{i,j=0}^{k-1},$$

and

$$\binom{j}{i} = 0 \qquad \text{for} \quad j < i.$$

In the case $s = 2$ the formulas for $\mathbf{B}_\infty$ and $\mathbf{B}_\alpha$ are found in [28].

We now turn to extending a relation between the classical Bezout and resultant matrices for two scalar polynomials obtained by N. Kravitsky in [21] to the case of several matrix polynomials. Note that in the case $s = 2$ the appropriate generalization has been established by the authors in [28].

Given $n \times n$ matrix polynomials $L_i(\lambda)$ $(i = 1, \dots, s)$ of degree ν, the matrix

$$\mathbf{R}_q(L) \equiv \mathbf{R}_q(L_1, L_2, \dots, L_s) := \operatorname{col}\left(R_{L_r}^{(q)}\right)_{r=1}^{s}, \tag{2.9}$$

with $R_F^{(q)}$ defined by (2.1), is referred to as the *resultant matrix* (or just *resultant*) of the matrix polynomials $L_r(\lambda)$ $(r = 1, 2, \dots, s)$. As is known (see [22], [2], [15], [16]), the resultant plays an important role in determining the common spectral data of several polynomials and constructing their greatest common divisor.

PROPOSITION 2.3. *If* $\mathbf{B}$ *denotes the Bezout matrix generated by* (0.2), *then*

$$\mathbf{R}_\mu^{(T)}(M)\operatorname{diag}\left[P_\mu, P_\mu, \dots, P_\mu\right]\mathbf{R}_\mu(L) = \begin{array}{c} \overbrace{}^{\mu}\qquad\qquad \\ \left[\begin{array}{c|c} 0 & -\mathbf{B} \\ \hline \mathbf{B} & 0 \end{array}\right]\begin{array}{l}\}\,\mu \\ \}\,\mu\end{array} \\ \qquad\qquad\underbrace{}_{\mu} \end{array} \tag{2.10}$$

Proof. Appealing to (2.2) (with $\alpha = \mu$) and to (2.5), we obtain

$$\begin{bmatrix}\mathbf{B} & 0_\mu\end{bmatrix} = \operatorname{row}\left(P_\mu \mathscr{L}_{M_r}^{(\mu,\mu)}\right)_{r=1}^{s}\left[\operatorname{col}\left(\mathscr{U}_{L_r}^{(\mu,\nu)}\right)_{r=1}^{s}\operatorname{col}\left(\mathscr{L}_{L_r}^{(\mu,\mu)}\right)_{r=1}^{s}\right], \tag{2.11}$$

where 0_μ denotes the $\mu n \times \mu n$ zero matrix. Similarly, the relations (2.2) and (2.5′) yield

$$\begin{bmatrix}0_\mu & -\mathbf{B}\end{bmatrix} = \operatorname{row}\left(P_\mu \mathscr{U}_{M_r}^{(\mu,\mu)}\right)_{r=1}^{s}\left[\operatorname{col}\left(\mathscr{U}_{L_r}^{(\mu,\mu)}\right)_{r=1}^{s}\operatorname{col}\left(\mathscr{L}_{L_r}^{(\mu,\nu)}\right)_{r=1}^{s}\right]. \tag{2.12}$$

Since

$$R_{L_r}^{(\mu)} = \left[\mathscr{U}_{L_r}^{(\mu,\mu)}\mathscr{L}_{L_r}^{(\mu,\nu)}\right] = \left[\mathscr{U}_{L_r}^{(\mu,\nu)}\mathscr{L}_{L_r}^{(\mu,\mu)}\right],$$

we see, combining (2.11) and (2.12), that

$$\begin{bmatrix}0_\mu & -\mathbf{B} \\ \mathbf{B} & 0_\mu\end{bmatrix} = \begin{bmatrix}\operatorname{row}\left(P_\mu \mathscr{U}_{M_r}^{(\mu,\mu)}\right)_{r=1}^{s} \\ \operatorname{row}\left(P_\mu \mathscr{L}_{M_r}^{(\mu,\mu)}\right)_{r=1}^{s}\end{bmatrix}\mathbf{R}_\mu(L).$$

In view of (2.6) we obtain the desired relation (2.10). ∎

PROPOSITION 2.4. *Let (X_F, T_F) and (X_∞, T_∞) denote the greatest common restrictions of the finite and infinite Jordan pairs of the regular polynomials $L_r(\lambda)$ $(r = 1,2,\dots, s)$, respectively. If $M_r(\lambda)$ $(r = 1,2,\dots,s)$ are arbitrary matrix polynomials satisfying* (0.2) *and* $\mathbf{B}$ *is the corresponding generalized Bezoutian, then*

$$\operatorname{Ker}\mathbf{B} \supset \operatorname{Im}\operatorname{col}\left(X_F T_F^{i-1}\right)_{i=1}^{\nu} \dot{+} \operatorname{Im}\operatorname{col}\left(X_\infty T_\infty^{\nu-1}\right)_{i=1}^{\nu}. \tag{2.13}$$

Proof. It is found in [15] that

$$\operatorname{Ker}\mathbf{R}_\mu(L) = \operatorname{Im}\operatorname{col}\left(X_F T_F^{i-1}\right)_{i=1}^{\nu+\mu} \dot{+} \operatorname{Im}\operatorname{col}\left(X_\infty T_\infty^{\nu+\mu-i}\right)_{i=1}^{\nu+\mu},$$

and hence (2.13) is an immediate consequence of (2.10). ∎

As shown in [28], the inclusion in (2.13) becomes equality for $s = 2$. However, if $s \geqslant 3$, the Bezoutian loses this property as the following example shows.

EXAMPLE. Consider the following matrix polynomials of degree one:

$$M_1(\lambda) = \begin{bmatrix} 1-\lambda & 0 \\ -1 & 2-\lambda \end{bmatrix}, \qquad M_2(\lambda) = \begin{bmatrix} \lambda-2 & 1 \\ 1 & \lambda-2 \end{bmatrix},$$

$$M_3(\lambda) = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix};$$

$$L_1(\lambda) = \begin{bmatrix} 3-\lambda & 0 \\ 0 & 2-\lambda \end{bmatrix}, \qquad L_2(\lambda) = \begin{bmatrix} 2-\lambda & 1 \\ 0 & 2-\lambda \end{bmatrix}, \qquad L_3(\lambda) = I_2.$$

It is not difficult to check that $\Sigma_{r=1}^{3} M_r(\lambda) L_r(\lambda) = 0$, and that the corresponding Bezoutian $\mathbf{B}$ is of the form

$$\mathbf{B} = \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}.$$

So $\mathbf{B}$ is singular although the polynomials $L_r(\lambda)$ $(r = 1,2,3)$ (of degree one) have neither finite nor infinite points of spectrum in common. Note that when viewing $M_2(\lambda)$, $L_2(\lambda)$ as polynomials of higher degree the relation (2.13) still remains a proper inclusion. ∎

Recall that the matrix polynomials $L_1(\lambda),\dots, L_s(\lambda)$ are right coprime if and only if they have no common eigenvector corresponding to a common (finite) eigenvalue. This means (see [14], [15]) that the right finite Jordan pairs of $L_r(\lambda)$, $r=1,\dots,s$, have no common restriction. Hence Proposition 2.4 implies the following sufficient condition of coprimeness.

COROLLARY 2.5. *If the Bezout matrix* $\mathbf{B}$ *associated with* (0.2) *is of full rank, then the matrix polynomials* $L_r(\lambda)$, $r=1,2,\dots,s$, *are right coprime, as well as the matrix polynomials* $L_{r\infty}(\lambda)$ $(r=1,\dots,s)$.

3. MODIFIED BEZOUTIANS AS SOLUTIONS OF MATRIX EQUATIONS

If (X,T) and (V,Z) are right and left admissible pairs (of arbitrary orders), respectively, and if $\mathbf{B}$ denotes the generalized Bezoutian associated with (0.2), we define the *modified Bezoutian* $\tilde{\mathbf{B}}$ by setting

$$\tilde{\mathbf{B}}=\operatorname{row}(V^{i-1}Z)_{i=1}^{\mu}\mathbf{B}\operatorname{col}(XT^{i-1})_{i=1}^{\nu}. \tag{3.1}$$

THEOREM 3.1. *Let* $\mathbf{B}$ *be the generalized Bezoutian associated with* (0.2) *and let* (X,T) *and* (V,Z) *be arbitrary right and left admissible pairs, respectively. Then the modified Bezoutian defined by* (3.1) *satisfies the following equation*:

$$V\tilde{\mathbf{B}}-\tilde{\mathbf{B}}T=\sum_{r=1}^{s}M_r(V,Z)L_r(X,T). \tag{3.2}$$

Proof. Denote

$$\tilde{X}=\operatorname{col}(XT^{i-1})_{i=1}^{\nu},\qquad \underset{\sim}{Z}=\operatorname{row}(V^{i-1}Z)_{i=1}^{\mu},\qquad \tilde{L}_r=\operatorname{col}\left(L_r(X,T)T^{i-1}\right)_{i=1}^{\mu},$$

$$W_{M_r}=\begin{bmatrix} -m_{r,0} & & & \\ & m_{r,2} & \cdots & m_{r,\mu} \\ & \vdots & \iddots & \\ & m_{r,\mu} & & \end{bmatrix}$$

In this notation $\tilde{\mathbf{B}}=\underset{\sim}{Z}\mathbf{B}\tilde{X}$, and (2.3) implies

$$V\tilde{\mathbf{B}}=V\underset{\sim}{Z}\sum_{r=1}^{s}P_{\mu}\mathscr{L}_{M_r}^{(\mu,\mu)}\tilde{L}_r.$$

Since

$$V\underset{\sim}{Z}P_\mu \mathscr{L}_{M_r}^{(\mu,\mu)} = \underset{\sim}{Z}W_{M_r} + \operatorname{row}\left(\delta_{1j}M_r(V,Z)\right)_{j=1}^{\mu},$$

it follows that

$$V\tilde{\mathbf{B}} = \sum_{r=1}^{s} \underset{\sim}{Z}W_{M_r}\tilde{L}_r + \sum_{r=1}^{s} M_r(V,Z)L_r(X,T). \tag{3.3}$$

Now observe that

$$W_{M_r}\tilde{L}_r = P_\mu \mathscr{L}_{M_r}^{(\mu,\mu)}\tilde{L}_r T - \operatorname{col}\left(\delta_{1j}M_r(L_r(X,T),T)\right)_{j=1}^{\mu},$$

and that (0.2) implies

$$\sum_{r=1}^{s} M_r(L_r(X,T),T) = \sum_{r=1}^{s} (M_rL_r)(X,T) = 0.$$

Hence (3.3) can be rewritten as follows:

$$V\tilde{\mathbf{B}} = \underset{\sim}{Z}\left(\sum_{r=1}^{s} P_\mu \mathscr{L}_{M_r}^{(\mu,\mu)}\tilde{L}_r\right)T + \sum_{r=1}^{s} M_r(V,Z)L_r(X,T).$$

In view of (2.3) this yields the desired relation (3.2). ∎

Note that if $(X,T)=(X^{(0)},C_L)$ and $(V,Z)=(\hat{C}_M,Z^{(0)})$, where $L(\lambda)$ and $M(\lambda)$ are some monic matrix polynomials of degrees ν and μ, respectively, then $\operatorname{col}(XT^{i-1})_{i=1}^{\nu}$ and $\operatorname{row}(V^{i-1}Z)_{i=1}^{\mu}$ are identity matrices, and hence in this case (3.2) becomes the equation (0.8) from the Introduction.

In the rest of this section we deduce some interesting corollaries for the case $s=2$.

PROPOSITION 3.2. *Let $M_1(\lambda)$ and $L_2(\lambda)$ be regular matrix polynomials such that*

$$M_1(\lambda)L_1(\lambda) = M_2(\lambda)L_2(\lambda) \tag{3.4}$$

for some matrix polynomials $M_2(\lambda)$ and $L_1(\lambda)$. Let (X,T) be a restriction of the right finite Jordan pair of $L_2(\lambda)$ and let (V,Z) be a restriction of the left *finite Jordan pair of* $M_1(\lambda)$. *Denote by* $\mathbf{B}$ *the generalized Bezoutian associated with* (3.4), *and let* $\tilde{\mathbf{B}}$ *be the modified Bezoutian defined by* (3.1). *Then* $\tilde{\mathbf{B}}$ *satisfies the homogeneous equation*

$$V\tilde{\mathbf{B}} = \tilde{\mathbf{B}}T. \tag{3.5}$$

In particular, if $M_1(\lambda)$ *and* $L_2(\lambda)$ *are monic matrix polynomials,*

$$\hat{C}_{M_1}\mathbf{B} = \mathbf{B}C_{L_2}. \tag{3.6}$$

Proof. Theorem 3.1 implies that

$$V\tilde{\mathbf{B}} - \tilde{\mathbf{B}}T = M_1(V,Z)L_1(X,T) - M_2(V,Z)L_2(X,T). \tag{3.7}$$

Since (X,T) is a restriction of the right finite Jordan pair of $L_2(\lambda)$, we have $L_2(X,T) = 0$. Similarly, $M_1(V,Z) = 0$, because (V,Z) is a restriction of the left finite Jordan pair of $M_1(\lambda)$. Thus, (3.5) follows from (3.7).

Furthermore, if $M_1(\lambda)$ and $L_2(\lambda)$ are monic matrix polynomials, we obtain (3.6) by setting $(X,T) = (X^{(0)}, C_{L_2})$ and $(V,Z) = (\hat{C}_{M_1}, Z^{(0)})$. ∎

Note that the formula (3.6) has frequently been found to be useful and can be found in [35], [28], [11], [29], [31], for example.

Recall that the notation $L^*(\lambda)$ is used for the matrix polynomial whose coefficients are conjugate transposes of the corresponding coefficients of the polynomial $L(\lambda)$.

COROLLARY 3.3. *Let* $L(\lambda)$ *be a regular matrix polynomial of degree* ν, *and let* (X,T) *be a restriction of its right finite Jordan pair. Assume that*

$$L^*(\lambda)M(\lambda) = M^*(\lambda)L(\lambda) \tag{3.8}$$

for some matrix polynomial $M(\lambda)$ *of degree* ν, *and let* $\mathbf{B}$ *denote the generalized Bezoutian associated with* (3.8). *Then*

$$T^*\tilde{\mathbf{B}} = \tilde{\mathbf{B}}T, \tag{3.9}$$

where $\tilde{\mathbf{B}} = [\operatorname{col}(XT^{i-1})_{i=1}^{\nu}]^*\mathbf{B}\operatorname{col}(XT^{i-1})_{i=1}^{\nu}$.

In particular, if $L(\lambda)$ *is monic, then*

$$C_L^*\mathbf{B} = \mathbf{B}C_L. \tag{3.10}$$

Proof. If (X,T) is a restriction of the right finite Jordan pair of $L(\lambda)$, then (T^*, X^*) is a restriction of the left finite Jordan pair of $L^*(\lambda)$. Thus, setting $L^* = M_1$, $M = L_1$ and $M^* = M_2$, $L = L_2$, we obtain (3.9) from (3.5) in view of Proposition 3.2. Equation (3.10) follows from (3.9) by setting $(X,T) = (X^{(0)}, C_L)$. ∎

Note that for the case of scalar polynomials with real coefficients equation (3.9) can be found in [20], [9].

COROLLARY 3.4. *Let $L(\lambda)$ and $L_1(\lambda)$ be regular matrix polynomials of degree ν, and let (X,T) be a restriction of the right finite Jordan pair of $L(\lambda)$. Assume that*

$$L_1^*(\lambda)L_1(\lambda) = L^*(\lambda)L(\lambda), \tag{3.11}$$

and let $\mathbf{B}$ *denote the generalized Bezoutian associated with* (3.11). *Then*

$$T^*\tilde{\mathbf{B}} - \tilde{\mathbf{B}}T = \left[L_1(X,T)\right]^* L_1(X,T), \tag{3.12}$$

where $\tilde{\mathbf{B}} = [\mathrm{col}(XT^{i-1})_{i=1}^{\nu}]^* \mathbf{B}\, \mathrm{col}(XT^{i-1})_{i=1}^{\nu}$.

In particular, if $L(\lambda)$ is monic, then

$$C_L^*\mathbf{B} - \mathbf{B}C_L = \left[L_1\left(X^{(0)}, C_L\right)\right]^*\left[L_1\left(X^{(0)}, C_L\right)\right]. \tag{3.13}$$

Proof. Considering (3.11) as (3.4) with $M_1 = L_1^*$, $M_2 = L^*$, $L_2 = L$, and applying (3.7) with $(V, Z) = (T^*, X^*)$, we obtain

$$T^*\tilde{\mathbf{B}} - \tilde{\mathbf{B}}T = L_1^*(T^*, X^*)L_1(X,T) - L^*(T^*, X^*)L(X,T),$$

where

$$\begin{aligned}\tilde{\mathbf{B}} &= \mathrm{row}\left[(T^*)^{i-1}X^*\right]_{i=1}^{\nu}\mathbf{B}\,\mathrm{col}\left(XT^{i-1}\right)_{i=1}^{\nu}\\ &= \left[\mathrm{col}\left(XT^{i-1}\right)_{i=1}^{\nu}\right]^*\mathbf{B}\,\mathrm{col}\left(XT^{i-1}\right)_{i=1}^{\nu}.\end{aligned}$$

Since $L(X,T) = 0$ and $L_1^*(T^*, X^*) = [L_1(X,T)]^*$, (3.12) follows from the last equation. If $L(\lambda)$ is monic, $(X^{(0)}, C_L)$ is a Jordan pair of $L(\lambda)$ and (3.13) follows from (3.12). ∎

The above result was established in [28], where it served as an important step in solving the eigenvalue separation problem for matrix polynomials. In the case of scalar polynomials equations similar to (3.13) were established in [9], [6] (see also [10]).

Note that for restrictions of infinite Jordan pairs of the underlying polynomials one can state results similar to Proposition 3.2 and Corollaries 3.3–3.4. To this end one has to replace the equation of type (0.2) by a suitable equation of type (2.7), and the Bezoutian $\mathbf{B}$ by $\mathbf{B}_\infty$. For instance, equation (3.11) has to be replaced by the equation

$$L_{1\infty}^*(\lambda)L_{1\infty}(\lambda) = L_\infty^*(\lambda)L_\infty(\lambda), \tag{3.14}$$

where $L_\infty(\lambda) := \lambda^\nu L(\lambda^{-1})$, $L_{1\infty}(\lambda) := \lambda^\nu L_1(\lambda^{-1})$.

Finally, we establish a result used in [30] (Lemma 2.4).

COROLLARY 3.5. *Let $\tilde{L}(\lambda) := L(\lambda) + FL_1(\lambda)$, where $F \in \mathbb{C}^{n\times n}$ and $L(\lambda), L_1(\lambda)$ are $n\times n$ regular matrix polynomials of degree ν satisfying (3.11). Assume that the matrix $L(0)+FL_1(0)$ is nonsingular, and let (X,T) be a restriction of the right Jordan pair of $\tilde{L}_\infty(\lambda) = \lambda^\nu \tilde{L}(\lambda^{-1})$. Then*

$$T^*\tilde{\mathbf{B}}_\infty - \tilde{\mathbf{B}}_\infty T = [L_{1\infty}(X,T)](1 - F^*F)L_{1\infty}(X,T),$$

where

$$\tilde{\mathbf{B}}_\infty = [\operatorname{col}(XT^{i-1})_{i=1}^{\nu}]^* \mathbf{B}_\infty \operatorname{col}(XT^{i-1})_{i=1}^{\nu},$$

and $\mathbf{B}_\infty$ stands for the Bezoutian associated with (3.14).

Proof. Applying Theorem 3.1 with $(V,Z) = (T^*, X^*)$ to the equation (3.14), we obtain

$$T^*\tilde{\mathbf{B}}_\infty - \tilde{\mathbf{B}}T = L_{1\infty}^*(T^*,X^*)L_{1\infty}(X,T) - L_\infty^*(T^*,X^*)L_\infty(X,T). \tag{3.15}$$

Since (X,T) is a restriction of the Jordan pair of $\tilde{L}_\infty$, we have

$$0 = \tilde{L}_\infty(X,T) = L_\infty(X,T) + FL_{1\infty}(X,T),$$

and hence $L_\infty(X,T) = -FL_{1\infty}(X,T)$. Substituting this in (3.15), we obtain the desired result. ∎

4. EQUATIONS IN MATRIX POLYNOMIALS VERSUS LINEAR MATRIX EQUATIONS

In this section we present the main results of the paper concerning solvability of the equation

$$M(\lambda)Y(\lambda)+Z(\lambda)L(\lambda)=R(\lambda), \tag{E_1}$$

of which the coefficients $M(\lambda)$ and $L(\lambda)$ are regular $n\times n$ matrix polynomials and the right hand term $R(\lambda)$ is also an $n\times n$ matrix polynomial (not necessarily regular). Recall that the set of all solution pairs $(Y(\lambda), Z(\lambda))$ of (E_1) such that $\deg Y\leqslant\gamma$, $\deg Z\leqslant\delta$ is denoted by $\mathbf{Y}_{\gamma,\delta}\,(\mathrm{E}_1)$. If for some integers γ,δ the set $\mathbf{Y}_{\gamma,\delta}\,(\mathrm{E}_1)$ is nonempty, we say that the equation (E_1) is (γ,δ)-solvable, and a pair of matrix polynomials $(Y(\lambda), Z(\lambda))\in\mathbf{Y}_{\gamma,\delta}\,(\mathrm{E}_1)$ will be referred to as a (γ,δ)-solution of (E_1).

We typically assume that an *a priori* estimation of the degree of $R(\lambda)$ is fixed—say, $\deg R\leqslant\rho$—and that $R(\lambda)$ is represented in the form

$$R(\lambda)=-\sum_{r=3}^{s}M_r(\lambda)L_r(\lambda), \tag{4.1}$$

where $M_r(\lambda)$, $L_r(\lambda)$ $(r=3,\dots,s)$ are $n\times n$ matrix polynomials such that

$$\deg M_r\leqslant\rho_1,\ \deg L_r\leqslant\rho_2,\quad \rho_1+\rho_2=\rho\qquad (r=3,\dots,s). \tag{4.2}$$

Modifying slightly the argument in the Introduction, one sees that any matrix polynomial $R(\lambda)$ with $\deg R\leqslant\rho$ can be represented in the form (4.1) with the estimation (4.2). As we shall see below, the above representations of $R(\lambda)$ provide the possibility of viewing (E_1) as an equation of type (0.2) and enables us to associate with (E_1) a Bezoutian of the smallest possible size.

Throughout this section we assume that $L(\lambda)$ and $M(\lambda)$ are matrix polynomials with invertible leading coefficients of degree ν and μ $(\nu\leqslant\mu)$, respectively, i.e.

$$M(\lambda)=\sum_{j=0}^{\mu}\lambda^j m_j,\quad L(\lambda)=\sum_{j=0}^{\nu}\lambda^j l_j,\qquad \det m_\mu\neq 0\neq\det l_j, \tag{4.3}$$

and we fix a right [respectively, left] standard pair (Φ, U) [respectively,

$(V, \Psi)]$ of the polynomial $L(\lambda)$ [respectively, $M(\lambda)$]. Next, we assume that $\deg R \leqslant \nu + \mu$ and that $R(\lambda)$ is represented in the form (4.1) with

$$\deg M_r(\lambda) \leqslant \mu, \quad \deg L_r(\lambda) \leqslant \nu \qquad (r = 3, \dots, s). \tag{4.4}$$

Then any (ν, μ)-solution $(Y(\lambda), Z(\lambda))$ of (E_1) (if it exists) generates the Bezoutian

$$\mathbf{B}_{Y,Z} := \mathbf{B}(M, Z, M_3, \dots, M_s; Y, L, L_3, \dots, L_s),$$

and we can define the mapping $\mathscr{B}: \mathbf{Y}_{\nu,\mu}(E_1) \to \mathbb{C}^{\mu n \times \nu n}$ by setting

$$\mathscr{B}(Y(\lambda), Z(\lambda)) = \operatorname{row}(V^{i-1}\Psi)_{i=1}^{\mu} \mathbf{B}_{Y,Z} \operatorname{col}(\Phi U^{i-1})_{i=1}^{\nu}. \tag{4.5}$$

Now we associate with the equation (E_1) the linear matrix equation

$$VS - SU = R, \tag{E$_3$}$$

where

$$R = \sum_{r=3}^{s} M_r(V, \Psi) L_r(\Phi, U). \tag{4.6}$$

The set of all solutions of the equation (E_3) is denoted by $S(E_3)$. Note that the coefficients of the equation (E_3) depend on the choice of the representation (4.1) for $R(\lambda)$. In what follows the standard pairs $(\Phi, U), (V, \Psi)$ and the polynomials $L_r(\lambda), M_r(\lambda)$ $(r = 3, \dots, s)$, satisfying (4.1) and (4.4), are assumed to be fixed.

THEOREM 4.1. *Preserving the above notation and assumptions, the equation* (E_1) *is solvable if and only if the equation* (E_3) *is solvable.*

Using the well-known Roth theorem (see [33]) one can rephrase the above result as follows.

THEOREM 4.1′. *The equation* (E_1) *is solvable if and only if the matrices*

$$\begin{bmatrix} V & 0 \\ 0 & U \end{bmatrix} \quad \textit{and} \quad \begin{bmatrix} V & R \\ 0 & U \end{bmatrix}$$

are similar.

A more detailed and deep connection between the solutions of Equations (E_1) and (E_3) can be described in terms of the mapping $\mathscr{B}$ defined by (4.5).

THEOREM 4.2. *The mapping* $\mathscr{B}$ *maps* $Y_{\nu,\mu}$ (E_1) *onto* $S(E_3)$. *If* $\hat{\mathscr{B}}: Y_{\nu,\mu}$ $(E_1) \to S(E_3)$ *denotes the corresponding restriction of* $\mathscr{B}$, *then a right inverse* $\hat{\mathscr{B}}^{[-1]}: S(E_3) \to Y_{\nu,\mu}$ (E_1) *of* $\hat{\mathscr{B}}$ *can be defined as follows: for any* $S \in S$ (E_3) *set* $\mathscr{B}^{[-1]}(S) = (Y_S(\lambda), Z_S(\lambda))$ *with*

$$Y_S(\lambda) = \sum_{k=0}^{\nu-1} \lambda^k m_\mu^{-1} Q_{\mu-1,k} - \sum_{r=3}^{s} m_\mu^{-1} m_{r\mu} L_r(\lambda) + m_\mu^{-1} A L(\lambda) l_\nu^{-1},$$

$$Z_S(\lambda) = -\sum_{k=0}^{\mu-1} \lambda^k Q_{k,\nu-1} l_\nu^{-1} - \sum_{r=3}^{s} M_r(\lambda) l_{r\nu} l_\nu^{-1} + m_\mu^{-1} M(\lambda) B l_\nu^{-1}, \quad (4.7)$$

where

$$[Q_{jk}]_{j,k=0}^{\mu-1,\nu-1} = [\mathrm{row}(V^{i-1}\Psi)_{i=1}^{\mu}]^{-1} S [\mathrm{col}(\Phi U^{i-1})_{i=1}^{\nu}]^{-1}, \quad (4.8)$$

and A, B *are arbitrary (fixed)* $n \times n$ *matrices such that*

$$A + B = \sum_{r=3}^{s} m_{r\mu} l_{r\nu}. \quad (4.9)$$

Furthermore, any (ν, μ)*-solution of* (E_1) *can be written in the form* (4.7) *for some choice of* A *and* B *satisfying* (4.9).

THEOREM 4.3. *The mapping* $\mathscr{B}$ *maps* $Y_{\nu-1,\mu}$ (E_1) *onto* $S(E_3)$ *in a* $1 \leftrightarrow 1$ *way. If* $\tilde{\mathscr{B}}: Y_{\nu-1,\mu}$ $(E_1) \to S(E_3)$ *denotes the corresponding restriction of* $\mathscr{B}$, *then the inverse* $\tilde{\mathscr{B}}^{-1}: S(E_3) \to Y_{\nu-1,\mu}$ *of* $\tilde{\mathscr{B}}$ *is defined by* $\tilde{\mathscr{B}}^{-1}$ $(S) = (Y_S(\lambda), Z_S(\lambda))$, *where* $Y_S(\lambda)$ *and* $Z_S(\lambda)$ *are given by* (4.7) *with* $B = 0$, $A = \sum_{r=3}^{s} m_{r\mu} l_{rn}$. *Similarly, the mapping* $\mathscr{B}$ *considered as acting* $Y_{\nu,\mu-1}$ $(E_1) \to S$ (E_3) *is bijective, and its inverse maps any* $S \in S$ (E_3) *into the pair* $(Y_S(\lambda), Z_S(\lambda))$ *defined by* (4.7) *with* $B = \sum_{r=3}^{s} m_{r\mu} l_{r\nu}$, $A = 0$.

Theorem 4.3 in conjunction with the well-known results on uniqueness of solutions of Liapunov type equations implies immediately the following result.

COROLLARY 4.4. *Preserving the notation and assumptions of this section, the following assertions are equivalent*:

(i) $\sigma(L)\cap\sigma(M)=\varnothing$;
(ii) *Equation* (E_1) *has a unique* $(\nu-1,\mu)$*-solution*;
(iii) *Equation* (E_1) *has a unique* $(\nu,\mu-1)$*-solution.*

Observe that if $\deg R\leqslant\nu+\mu-1$, then $\mathbf{Y}_{\nu-1,\mu}\ (\mathrm{E}_1)=\mathbf{Y}_{\nu,\mu-1}\ (\mathrm{E}_1)=\mathbf{Y}_{\nu-1,\mu-1}(\mathrm{E}_1)$. Thus, in this case condition (i) is equivalent to the existence of a unique $(\nu-1,\mu-1)$-solution of (E_1). The last assertion was established in [2] under the assumption $\deg R\leqslant\nu+\mu-2$.

In the general case of regular matrix polynomials $L(\lambda)$ and $M(\lambda)$, the relation between the equation (E_1) and the matrix equations is more complicated and will appear in a forthcoming publication. Some preliminary results in this direction are contained in [29].

Some remarks about the matrix equation (E_3) are in order. In the statements of Theorems 4.1–4.3 it is assumed that an equation of type (E_1) in matrix polynomials is given, and by fixing a representation (4.1), (4.4) and by choosing the standard pairs (Φ,U) and (V,Ψ) of $L(\lambda)$ and $M(\lambda)$, respectively, we associate with (E_1) the Liapunov type matrix equation (E_3), whose right hand side is defined by (4.6). The next simple result shows that any $\mu n\times\nu n$ block matrix can be represented in the form (4.6). We shall state this result in the framework of monic polynomials $L(\lambda)$ and $M(\lambda)$ and corresponding companion standard pairs. The passage from this case to the general one is not difficult (see the beginning of Section 5).

PROPOSITION 4.5. *Let* $L(\lambda)$ *and* $M(\lambda)$ *be* $n\times n$ *monic matrix polynomials of degree* ν *and* μ, *respectively, and let* $R^{(0)}$ *be an arbitrary* $\mu n\times\nu n$ *block matrix. Let* $R^{(0)}$ *be represented in the form*

$$R^{(0)}=\sum_{j=1}^{\beta}\Omega_j\Lambda_j,\tag{4.10}$$

where

$$\Omega_j=\begin{bmatrix}\Omega_{j0}\\ \Omega_{j1}\\ \vdots\\ \Omega_{j,\mu-1}\end{bmatrix},\quad \Lambda_j=\begin{bmatrix}\Lambda_{j0} & \Lambda_{j1} & \cdots & \Lambda_{j,\nu-1}\end{bmatrix}\qquad(j=1,\dots,\beta),$$

and Ω_{jk}, Λ_{jk} *are* $n\times n$ *matrices. Then*

$$R^{(0)}=\sum_{i=3}^{\beta+2} M_i\big(\hat{C}_M, Z^{(0)}\big)L_i\big(X^{(0)}, C_L\big),$$

where

$$L_{j+2}(\lambda)=\sum_{k=0}^{\nu-1}\lambda^k\Lambda_{jk},\quad M_{j+2}(\lambda)=\sum_{k=0}^{\mu-1}\lambda^k\Omega_{jk}\qquad (j=1,\dots,\beta). \tag{4.11}$$

The *proof* follows immediately from the following easily verified equalities

$$L_{j+2}\big(X^{(0)},C_L\big)=\Lambda_j,\quad M_{j+2}\big(\hat{C}_M, Z^{(0)}\big)=\Omega_j\qquad (j=1,\dots,\beta).$$

Note that for any $\mu n\times \nu n$ block matrix $R^{(0)}=[R_{ij}]_{i,j=0}^{\mu-1,\nu-1}$ a decomposition of the form (4.10) is easily obtained by setting $\beta=\mu$ and

$$\Omega_j=\begin{bmatrix}0\\ \vdots\\ 0\\ I\\ \vdots\\ 0\end{bmatrix}\begin{matrix}\left.\vphantom{\begin{matrix}0\\ \vdots\\ 0\end{matrix}}\right\} j-1\\ \\ \\ \\ \end{matrix},\quad \Lambda_j=\begin{bmatrix}R_{j-1,0} & R_{j-1,1} & \cdots & R_{j-1,\nu-1}\end{bmatrix}$$

$$(j=1,\dots,\mu). \tag{4.12}$$

Proposition 4.5 allows one to put any matrix equation

$$\hat{C}_M X - XC_L = R^{(0)} \tag{E_2}$$

with an arbitrary right hand side $R^{(0)}$ into the framework of Theorems 4.1–4.3. To this end one has to consider an equation of the form (E_1) with $R(\lambda)=-\sum_{j=3}^{\beta+2}M_j(\lambda)L_j(\lambda)$, where $M_j(\lambda)$ and $L_j(\lambda)$ are defined by (4.11). For instance, if Ω_j and Λ_j are chosen as in (4.12), the right hand side in (E_1) becomes

$$R(\lambda)=-\sum_{j=1}^{\mu}\lambda^{j-1}\sum_{i=1}^{\nu}\lambda^{i-1}R_{j-1,i-1}=-\sum_{k=0}^{\mu+\nu-2}\lambda^k\sum_{i=0}^{k}R_{i,k-i}. \tag{4.13}$$

Thus, according to Theorem 4.1, the equation (E_2) is solvable if and only if the equation (E_1) with $R(\lambda)$ given by (4.13) is solvable. Note that since $\deg R \leqslant \mu+\nu-2$, the solvability of the equation (E_1) is the same as its $(\nu-1,\mu-1)$-solvability. Furthermore, if $(Y(\lambda), Z(\lambda))$ is a $(\nu-1,\mu-1)$-solution of (E_1), then a solution of (E_2) is given by $X=\mathbf{B}_{Y,Z}$, the Bezoutian generated by (E_1) with $R(\lambda)$ represented as in (4.13). Applying the formula (2.3) with $(X,T)=(X^{(0)},C_L)$, and using the equalities $L(X^{(0)},C_L)=0$, $Y(X^0,C_L)=[y_0y_1\cdots y_{\nu-1}]$ for $Y(\lambda)=\sum_{i=0}^{\nu-1}\lambda^i y_i$, we obtain the following result.

COROLLARY 4.6. *The equation*

$$\hat{C}_M X - XC_L = \left[R_{ij}\right]_{i,j=0}^{\mu-1,\nu-1} \tag{4.14}$$

is solvable if and only if the equation

$$M(\lambda)Y(\lambda)+Z(\lambda)L(\lambda) = -\sum_{k=0}^{\mu+\nu-2}\lambda^k\sum_{i=0}^{k}R_{i,k-i} \tag{4.15}$$

is $(\nu-1,\mu-1)$-solvable. If $(Y(\lambda), Z(\lambda))$ is a $(\nu-1,\mu-1)$-solution of (4.15) *and $Y(\lambda)=\sum_{i=0}^{\nu-1}\lambda^i y_i$, then*

$$X = \begin{bmatrix} m_1 & m_2 & \cdot & \cdot & \cdot & m_{\mu-1} & I \\ m_2 & & & & \cdot & \cdot & \\ \cdot & & & \cdot & \cdot & & \\ \cdot & & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & & & & \\ m_{\mu-1} & \cdot & & & & & \\ I & & & & & & \end{bmatrix}\begin{bmatrix} Y \\ YC_L \\ \vdots \\ YC_L^{\mu-1}\end{bmatrix} + \begin{bmatrix} R_1+R_2C_L+\cdots+R_{\mu-1}C_L^{\mu-2} \\ R_2+R_3C_L+\cdots+R_{\mu-1}C_L^{\mu-3} \\ \vdots \\ R_{\mu-1} \\ 0 \end{bmatrix} \tag{4.16}$$

is a solution of (4.15). *Here* $Y := [y_0y_1\cdots y_{\nu-1}]$ *and* $R_j := [R_{j0}R_{j1}\cdots R_{j,\nu-1}]$ $(j=0,1,\ldots,\mu-1)$.

Equation (4.16) shows that Y is the last block row in the solution matrix X. Representing the Bezoutian $\mathbf{B}_{Y,Z}$ according to the formula (2.3′), one can derive a dual formula for X in terms of its last column $Z := [z_0^T\ z_1^T\ \cdots\ z_{\mu-1}^T]^T$, where $Z(\lambda)=\sum_{i=0}^{\mu-1}\lambda^i z_i$ comes from the $(\nu-1,\mu-1)$-solution $(Y(\lambda), Z(\lambda))$ of (4.15). Additional formulas for X can be obtained by exploiting (4.5) and (4.5′). Also, representations of $[R_{ij}]_{i,j=0}^{\mu-1,\nu-1}$ of type (4.10) different from (4.12) lead to new formulas for the solutions X of (4.14).

5. PROOFS OF THE MAIN RESULTS

In this section we shall prove Theorems 4.1–4.3. Our first observation is that in these proofs we can assume without loss of generality that the polynomials $L(\lambda)$ and $M(\lambda)$ are monic. Indeed, considering the equation

$$\tilde{M}(\lambda)Y(\lambda)+Z(\lambda)\tilde{L}(\lambda)=R(\lambda) \qquad (\tilde{\mathrm{E}}_1)$$

with $\tilde{L}(\lambda) := l_\nu^{-1}L(\lambda)$ and $\tilde{M}(\lambda) := M(\lambda)m_\mu^{-1}$, one readily sees that $(Y(\lambda), Z(\lambda)) \in \mathbf{Y}_{\gamma,\delta}\ (\mathrm{E}_1)$ if and only if $(m_\mu Y(\lambda), Z(\lambda)l_\nu) \in \mathbf{Y}_{\gamma,\delta}\ (\tilde{\mathrm{E}}_1)$. So, let $M(\lambda)$ and $L(\lambda)$ be monic polynomials. Our next observation is that the standard pairs (Φ, U) and (V, Ψ) in the statements of Theorems 4.1–4.3 can be replaced by the companion pairs $(X^{(0)}, C_L)$ and $(\hat{C}_M, Z^{(0)})$, respectively. Indeed, setting $K := \operatorname{col}(\Phi U^{i-1})_{i=1}^{\nu}$, $K_i := \operatorname{row}(V^{i-1}\Psi)_{i=1}^{\mu}$, one has

$$\Phi = X^{(0)}K, \qquad U = K^{-1}C_L K,$$

$$\Psi = K_1 Z^{(0)}, \qquad V = K_1\hat{C}_M K_1^{-1}.$$

Substituting in (E_3), we see that $\mathrm{S}\in \mathbf{S}(\mathrm{E}_3)$ if and only if $K_1^{-1}\mathrm{S}K^{-1}$ is a solution of the equation

$$\hat{C}_M X - XC_L = R^{(0)}, \qquad (\mathrm{E}_2)$$

where

$$R^{(0)} := \sum_{r=3}^{s} M_r(\hat{C}_M, Z^{(0)})L_r(X^{(0)}, C_L)\ \left(=K_1^{-1}RK^{-1}\right). \qquad (5.1)$$

Thus, in Lemmas 5.1–5.2, in Proposition 5.3, and in the proofs of Theorems 4.1–4.3, we assume that the polynomials $L(\lambda)$ and $M(\lambda)$ are monic and we replace the equation (E_3) by the equation (E_2). Note that in this case the mapping $\mathscr{B}$ defined by (4.5) becomes very simple:

$$\mathscr{B}(Y(\lambda), Z(\lambda)) = \mathbf{B}_{Y,Z}. \tag{5.2}$$

We need some further preparations.

LEMMA 5.1. *Let the matrix polynomial* $R(\lambda)$, $\deg R \leqslant \mu + \nu$, *be represented as in* (4.1) *with*

$$\deg M_r \leqslant \alpha \leqslant \mu, \quad \deg L_r \leqslant \beta \leqslant \nu \qquad (r = 3, \dots, s).$$

Consider the equation

$$M(\lambda)Y(\lambda) + Z(\lambda)L(\lambda) = -\sum_{r=3}^{s} \tilde{M}_r(\lambda)L_r(\lambda) \tag{E$_4$}$$

with

$$\tilde{M}_r(\lambda) := M_r(\lambda) - M(\lambda)m_{r\alpha} \qquad (r = 3, \dots, s). \tag{5.3}$$

Then $(\tilde{Y}(\lambda), \tilde{Z}(\lambda))$ *is a solution of* (E_4) *if and only if*

$$\tilde{Z}(\lambda) \equiv Z(\lambda), \qquad \tilde{Y}(\lambda) = Y(\lambda) + \sum_{r=3}^{s} m_{r\alpha}L_r(\lambda), \tag{5.4}$$

where $(Y(\lambda), Z(\lambda))$ *is a solution of the equation* (E_1), *and in this case the Bezoutian* $\mathbf{B}_{\tilde{Y},\tilde{Z}}$ *generated by* (E_4) *coincides with the Bezoutian* $\mathbf{B}_{Y,Z}$ *generated by* (E_1) *(for corresponding* $Y(\lambda)$, $Z(\lambda)$*).*

Proof. The assertion about the relation between the solutions of equations (E_1) and (E_4) is easily verified. Now assume that the solutions $(Y(\lambda), Z(\lambda))$ and $(\tilde{Y}(\lambda), \tilde{Z}(\lambda))$ of the equations (E_1) and (E_4), respectively, are related by (5.4). Making use of (2.3) with $(X, T) = (X^{(0)}, C_L)$ and taking into consideration the equality $L(X^{(0)}, C_L) = 0$, we have

$$\mathbf{B}_{\tilde{Y},\tilde{Z}} = P_\mu \mathscr{L}_M^{(\mu,\mu)} \operatorname{col}\left(\tilde{Y}\left(X^{(0)}, C_L\right)C_L^{i-1}\right)_{i=1}^{\mu} + \tilde{\Omega}, \tag{5.5}$$

with

$$\tilde{\Omega} = \sum_{r=3}^{s} P_\mu \mathscr{L}_{\tilde{M}_r}^{(\mu,\mu)} \operatorname{col}\left(L_r\left(X^{(0)}, C_L\right)C_L^{i-1}\right)_{i=1}^{\mu}. \tag{5.6}$$

In view of (5.3),

$$\mathscr{L}_{\tilde{M}_r}^{(\mu,\mu)} = \mathscr{L}_{M_r}^{(\mu,\mu)} - \mathscr{L}_{M}^{(\mu,\mu)} \operatorname{diag}(m_{r\alpha},\dots,m_{r\alpha}) \qquad (r = 3,\dots,s),$$

and hence (5.6) becomes

$$\begin{aligned}\tilde{\Omega} = {} & \sum_{r=3}^{s} P_\mu \mathscr{L}_{M_r}^{(\mu,\mu)} \operatorname{col}\left(L_r\left(X^{(0)}, C_L\right)C_L^{i-1}\right)_{i=1}^{\mu} \\ & - \sum_{r=3}^{s} P_\mu \mathscr{L}_{M}^{(\mu,\mu)} \operatorname{diag}(m_{r\alpha},\dots,m_{r\alpha}) \operatorname{col}\left(L_r\left(X^{(0)}, C_L\right)C_L^{i-1}\right)_{i=1}^{\mu}.\end{aligned}$$

Substituting this expression for $\tilde{\Omega}$ in (5.5) and using the equality

$$\begin{aligned}&\operatorname{col}\left(\tilde{Y}\left(X^{(0)}, C_L\right)C_L^{i-1}\right)_{i=1}^{\mu} \\ &\quad = \operatorname{col}\left(Y\left(X^{(0)}, C_L\right)C_L^{i-1}\right)_{i=1}^{\mu} \\ &\qquad - \sum_{r=3}^{s} \operatorname{diag}(m_{r\alpha},\dots,m_{r\alpha}) \operatorname{col}\left(L_r\left(X^{(0)}, C_L\right)C_L^{i-1}\right)_{i=1}^{\mu},\end{aligned}$$

which follows easily from (5.4), one obtains $\mathbf{B}_{\tilde{Y},\tilde{Z}} = \mathbf{B}_{Y,Z}$. ■

LEMMA 5.2. *Let the matrix polynomial* $R(\lambda)$, $\deg R \leqslant \mu + \nu$, *be represented in the form* (4.1) *with estimations* (4.4). *Let* $S = [s_{jk}]_{j,k=0}^{\mu-1,\nu-1}$ *be a solution of the equation* (E_2), *and set*

$$Y(\lambda) := \sum_{k=0}^{\nu-1} \lambda^k s_{\mu-1,k} - \sum_{r=3}^{s} m_{r\mu} L_r(\lambda). \tag{5.7}$$

Then there is a matrix polynomial $Z(\lambda)$, $\deg Z \leqslant \mu - 1$, *such that* $(Y(\lambda), Z(\lambda))$ *is a solution of* (E_1) *and* $\mathbf{B}_{Y,Z} = S$.

Similarly, for the matrix polynomial

$$Z_1(\lambda) := -\sum_{j=0}^{\mu-1} \lambda^k s_{j,\nu-1} - \sum_{r=3}^{s} M_r(\lambda) l_{r\nu} \tag{5.8}$$

there is a matrix polynomial $Y_1(\lambda)$, $\deg Y_1 \leqslant \nu - 1$, *such that* $\mathbf{B}_{Y_1, Z_1} = S$.

Proof. Let us write the solution S and the right hand term $R^{(0)}$ in (E_2) as block row matrices:

$$S = \operatorname{col}(S_i)_{i=0}^{\mu-1}, \quad R = \operatorname{col}(R_i)_{i=0}^{\mu-1} \qquad \left(S_i, R_i \in \mathbb{C}^{n \times \nu n}, \quad i = 0, 1, \dots, \mu - 1\right).$$

Now, compare, starting with the last, the block rows of the expressions on the left and right in (E_2) and obtain for $i = \mu - 2, \mu - 3, \dots, 0$

$$S_i = S_{\mu-1} C_L^{\mu-i-1} + \sum_{j=1}^{\mu-i-1} \left(m_{\mu-j} S_{\mu-1} + R_{\mu-j}\right) C_L^{\mu-i-j-1}. \tag{5.9}$$

Subsequently, the comparison of the first block rows gives

$$S_{\mu-1} C_L^{\mu} + \sum_{j=0}^{\mu-1} \left(m_j S_{\mu-1} + R_j\right) C_L^j = 0. \tag{5.10}$$

Recalling the definition (5.1) of $R^{(0)}$ and observing that $M_r(\hat{C}_M, Z^{(0)}) = \operatorname{col}(m_{rj} - m_j m_{r\mu})_{j=0}^{\mu-1}$, we have

$$R_j = \sum_{r=3}^{s} (m_{rj} - m_j m_{r\mu}) L_r\left(X^{(0)}, C_L\right), \tag{5.11}$$

which, being substituted in (5.10), yields

$$\sum_{j=0}^{\mu} m_j S_{\mu-1} C_L^j + \sum_{r=3}^{s} \sum_{j=0}^{\mu-1} \tilde{m}_{rj} L_r\left(X^{(0)}, C_L\right) C_L^j = 0, \tag{5.12}$$

where $\tilde{m}_{rj} := m_{rj} - m_j m_{r\mu}$ $(0 \leqslant j \leqslant \mu - 1,\ 3 \leqslant r \leqslant s)$ are the coefficients of the polynomials $\tilde{M}_r(\lambda)$, defined by (5.3) (with $\alpha = \mu$), and $m_\mu := I$.

Now, decompose the last row in S as follows:

$$S_{\mu-1}=\operatorname{row}(y_j)_{j=0}^{\nu-1} \qquad \left(y_j\in\mathbb{C}^{n\times n},\quad j=0,1,\dots,\nu-1\right),$$

and observe that for the matrix polynomial $Y(\lambda)=\sum_{j=0}^{\nu-1}\lambda^j y_j$ the equation $\tilde{Y}(X^{(0)},C_L)=S_{\mu-1}$ holds. Denote

$$K(\lambda)=M(\lambda)\tilde{Y}(\lambda)+\sum_{r=3}^{s}\tilde{M}_r(\lambda)L_r(\lambda), \tag{5.13}$$

and compute the value of $K(\lambda)$ at the pair $(X^{(0)},C_L)$.

To do this, note the following fact: If $N(\lambda)$, $P(\lambda)$, $Q(\lambda)$ are matrix polynomials such that $N(\lambda)=P(\lambda)Q(\lambda)$, then for any admissible pair (X,T)

$$N(X,T)=P(Q(X,T),T)=\sum_{i=0}^{\rho}p_iQ(X,T)T^i,$$

where $P(\lambda)=\sum_{i=0}^{\rho}\lambda^i p_i$. Using this rule and (5.12), we easily see that $K(X^{(0)},C_L)=0$, which implies that $K(\lambda)$ is divisible by $L(\lambda)$ on the right. Thus, there exists an $n\times n$ matrix polynomial $\tilde{Z}(\lambda)$ such that $K(\lambda)=-\tilde{Z}(\lambda)L(\lambda)$. It follows, in view of (5.13), that the equality

$$M(\lambda)\tilde{Y}(\lambda)+\tilde{Z}(\lambda)L(\lambda)=-\sum_{r=3}^{s}\tilde{M}_r(\lambda)L_r(\lambda) \tag{5.14}$$

holds.

Note that $\deg\tilde{M}_r\leqslant\mu-1$, and hence $\deg K(\lambda)\leqslant\mu+\nu-1$. Since $L(\lambda)$ is monic of degree ν, it follows that $\deg\tilde{Z}(\lambda)\leqslant\mu-1$. Thus, $(\tilde{Y}(\lambda),\tilde{Z}(\lambda))$ is a $(\nu-1,\mu-1)$-solution of the equation (E_4), and, as is proved in Lemma 5.1, $(Y(\lambda),Z(\lambda))$ is a $(\nu-1,\mu-1)$-solution of (E_1), where $Z(\lambda)=\tilde{Z}(\lambda)$ and $Y(\lambda)=\tilde{Y}(\lambda)-\sum_{r=3}^{s}m_{r\mu}L_r(\lambda)$.

Furthermore, (5.9) implies

$$S=\operatorname{col}(S_i)_{i=0}^{\mu-1}=P_\mu\mathscr{L}_M^{(\mu,\mu)}\operatorname{col}\left(S_{\mu-1}C_L^{i-1}\right)_{i=1}^{\mu}+\operatorname{col}\left(\tilde{R}_i\right)_{i=0}^{\mu-1}, \tag{5.15}$$

where for $i=0,1,\dots,\mu-2$

$$\tilde{R}_i=R_{\mu-1}C_L^{\mu-i-2}+R_{\mu-2}C_L^{\mu-i-3}+\cdots+R_{i+1},$$

and $\tilde{R}_{\mu-1}=0$. By (5.11)

$$\operatorname{col}\left(\tilde{R}_i\right)_{i=0}^{\mu-1}=\sum_{r=3}^{s} P_\mu \mathscr{L}_{M_r}^{(\mu,\mu)} \operatorname{col}\left(L_r\left(X^{(0)},C_L\right)C_L^{i-1}\right)_{i=1}^{\mu}.$$

Substituting this expression in (5.15) and observing that $S_{\mu-1}=\tilde{Y}(X^{(0)},C_L)$ we obtain, in view of (5.5), that $S=\mathbf{B}_{\tilde{Y},\tilde{Z}}$, and in view of Lemma 5.1 we have $S=\mathbf{B}_{Y,Z}$, which proves the first assertion of the lemma. The proof of the second assertion and the construction of the solution $(Y_1(\lambda), Z_1(\lambda))$ go along the same lines, by comparison of the block columns of the matrices S and $R^{(0)}$. ∎

Now we are able to prove Theorem 4.1.

Proof of Theorem 4.1. The "if" part of the Theorem follows immediately from Lemma 5.2. To prove the "only if" part we first observe that if the equation (E_1) is solvable, then it is (ν,μ)-solvable. Indeed, let $Y(\lambda)=\sum_{j=0}^{\gamma}\lambda^j y_j$ $(y_\gamma\neq 0)$, $Z(\lambda)=\sum_{j=0}^{\delta}\lambda^j z_j$ $(z_\delta\neq 0)$ be a solution of (E_1), and assume that at least one of the inequalities

$$\gamma>\nu,\qquad \delta>\mu \tag{5.16}$$

holds true. Since $\deg R(\lambda)\leqslant \nu+\mu$, we conclude that $\gamma+\mu=\delta+\nu$, and hence both inequalities in (5.16) hold true. Furthermore, $y_\gamma=-z_\delta$, and hence the polynomials

$$\tilde{Y}(\lambda)=Y(\lambda)-\lambda^{\gamma-\nu}y_\gamma L(\lambda),\qquad \tilde{Z}(\lambda)=Z(\lambda)+\lambda^{\delta-\mu}M(\lambda)z_\delta$$

form a $(\gamma-1,\delta-1)$-solution of (E_1).

Proceeding in this way we obtain a (ν,μ)-solution of (E_1). Now, the "only if" part of the theorem follows immediately from Theorem 3.1 and the equalities

$$M\left(\hat{C}_M, Z^{(0)}\right)\equiv M_1\left(\hat{C}_M, Z^{(0)}\right)=0,\qquad L\left(X^{(0)},C_L\right)\equiv L_1\left(X^{(0)},C_L\right)=0.$$ ∎

The following proposition is an important ingredient in the proofs of Theorems 4.2–4.3.

PROPOSITION 4.3. *Let the matrix polynomial* $R(\lambda)$, $\deg R\leqslant \nu+\mu-1$, *be represented in the form* (4.1) *with*

$$\deg L_r\leqslant \nu-1,\quad \deg M_r\leqslant \mu\qquad (r=3,\dots,s). \tag{5.17}$$

Then the mapping $\mathscr{B}$ *defined by* (5.2) *maps* $\mathbf{Y}_{\nu-1,\mu-1}\,(\mathrm{E}_1)$ *onto* $\mathrm{S}(\mathrm{E}_2)$ *in a* $1 \leftrightarrow 1$ *way. If* $\mathscr{B}_0 : \mathbf{Y}_{\nu-1,\mu-1}\,(\mathrm{E}_1) \to \mathrm{S}(\mathrm{E}_2)$ *denotes the corresponding restriction of* $\mathscr{B}$, *then the inverse mapping* $\mathscr{B}_0^{-1}$ *acts as follows: if* $\mathrm{S} = [s_{jk}]_{j,k=0}^{\mu-1,\nu-1} \in \mathrm{S}(\mathrm{E}_2)$, *then* $\mathscr{B}_0^{-1}(\mathrm{S}) = (Y_{\mathrm{S}}(\lambda), Z_{\mathrm{S}}(\lambda))$ *with*

$$Y_{\mathrm{S}}(\lambda) = \sum_{k=0}^{\nu-1} \lambda^k s_{\mu-1,k} - \sum_{r=3}^{s} m_{r\mu} L_r(\lambda), \qquad Z_{\mathrm{S}}(\lambda) = -\sum_{j=0}^{\mu-1} \lambda^j s_{j,\nu-1}. \tag{5.18}$$

Proof. In view of Lemma 5.2 we have, in particular, $\mathscr{B}(\mathbf{Y}_{\nu-1,\mu-1}\,(\mathrm{E}_1)) \supset \mathrm{S}(\mathrm{E}_2)$. From Theorem 4.1 and its proof we also know that $\mathscr{B}(\mathbf{Y}_{\nu-1,\mu-1}\,(\mathrm{E}_1)) \subset \mathscr{B}(\mathbf{Y}_{\nu,\mu}\,(\mathrm{E}_1)) \subset \mathrm{S}(\mathrm{E}_2)$. Hence, $\mathscr{B}$ maps $\mathbf{Y}_{\nu-1,\mu-1}\,(\mathrm{E}_1)$ onto $\mathrm{S}(\mathrm{E}_2)$. Now we shall prove that $\mathscr{B}_0$ is injective. To this end assume that $\mathscr{B}_0(Y_1(\lambda), Z_1(\lambda)) = \mathscr{B}_0(Y_2(\lambda), Z_2(\lambda))$ for some $(\nu-1, \mu-1)$-solutions $(Y_1(\lambda), Z_1(\lambda))$ and $(Y_2(\lambda), Z_2(\lambda))$ of the equation (E_1). Making use of (2.3) [with $(X,T) = (X^{(0)}, C_L)$], one readily sees that this is equivalent to the equality

$$P_\mu \mathscr{L}_M^{(\mu,\mu)} \operatorname{col}\left(Y_1\left(X^{(0)}, C_L\right) C_L^{i-1}\right)_{i=1}^{\mu} = P_\mu \mathscr{L}_M^{(\mu,\mu)} \operatorname{col}\left(Y_2\left(X^{(0)}, C_L\right) C_L^{i-1}\right)_{i=1}^{\mu},$$

which, in view of the invertibility of $P_\mu \mathscr{L}_M^{(\mu,\mu)}$, yields

$$Y_1\left(X^{(0)}, C_L\right) = Y_2\left(X^{(0)}, C_L\right). \tag{5.19}$$

A simple calculation gives

$$Y_i\left(X^{(0)}, C_L\right) = \left[\, y_0^{(i)} \quad \cdots \quad y_{\nu-1}^{(i)} \,\right] \qquad (i = 1,2,),$$

where $Y_i(\lambda) = \sum_{j=0}^{\nu-1} \lambda^j y_j^{(i)}$ $(i = 1,2)$. So (5.19) means that $Y_1(\lambda) \equiv Y_2(\lambda)$. But then,

$$0 = M(\lambda)\left[Y_1(\lambda) - Y_2(\lambda)\right] = \left[Z_2(\lambda) - Z_1(\lambda)\right] L(\lambda),$$

and since $L(\lambda)$ is monic, we conclude that $Z_1(\lambda) \equiv Z_2(\lambda)$, which proves the injectivity of $\mathscr{B}_0$.

It remains to compute the inverse mapping $\mathscr{B}_0^{-1}$. Let $\mathrm{S} = [s_{jk}]_{j,k=0}^{\mu-1,\nu-1} \in \mathrm{S}(\mathrm{E}_3)$. From Lemma 5.2 we know that if $Y(\lambda)$ is defined by (5.7), then there is a matrix polynomial $Z(\lambda)$, $\deg Z \leqslant \mu - 1$, such that $(Y(\lambda), Z(\lambda))$ is a solution of (E_1) and $\mathbf{B}_{Y,Z} = \mathrm{S}$. But $\deg R \leqslant \mu + \nu - 1$, i.e. $\sum_{r=3}^{s} m_{r\mu} l_{r\nu} = 0$, and consequently the polynomial $\sum_{r=3}^{s} m_{r\mu} L_r(\lambda)$ has degree $\leqslant \nu - 1$. Thus

$(Y(\lambda), Z(\lambda)) \in \mathbf{Y}_{\nu-1,\mu-1}\,(\mathrm{E}_1)$. Using the second assertion of Lemma 5.2, we conclude that $(Y_1(\lambda), Z_1(\lambda))$ is a solution of (E_1) such that $\mathscr{B}_{Y_1, Z_1} = \mathrm{S}$, where $Z_1(\lambda)$ is given by (5.8) and $Y_1(\lambda)$ is some polynomial of degree $\leqslant \nu - 1$. In view of (5.17), $l_{r\nu} = 0$ $(r = 3,\dots, s)$ and hence (5.8) becomes

$$Z_1(\lambda) = -\sum_{j=0}^{\mu-1} \lambda^j s_{j,\nu-1}. \tag{5.20}$$

So $(Y_1(\lambda), Z_1(\lambda)) \in \mathbf{Y}_{\nu-1,\mu-1}\,(\mathrm{E}_1)$. We know already that $\mathscr{B}_0 : \mathbf{Y}_{\nu-1,\mu-1}\,(\mathrm{E}_1) \to \mathbf{S}(\mathrm{E}_2)$ is injective, and hence the solutions $(Y(\lambda), Z(\lambda))$ and $(Y_1(\lambda), Z_1(\lambda))$ must coincide. Now from (5.7) and (5.20) we obtain the formulas (5.18) for the inverse mapping. ∎

Proof of Theorem 4.2. If $(Y(\lambda), Z(\lambda)) \in \mathbf{Y}_{\nu,\mu}\,(\mathrm{E}_1)$, then it follows from Theorem 3.1 that $\mathscr{B}(Y(\lambda), Z(\lambda)) \in \mathbf{S}(\mathrm{E}_2)$ (cf. the proof of Theorem 4.1). Furthermore, using Lemma 5.2 one can associate with each $\mathrm{S} \in \mathbf{S}(\mathrm{E}_2)$ a $(\nu, \mu - 1)$-solution $(Y(\lambda), Z(\lambda))$ of (E_1) such that $\mathrm{S} = \mathbf{B}_{Y,Z}$. It follows, in particular, that $\mathscr{B}$ maps $\mathbf{Y}_{\nu,\mu}\,(\mathrm{E}_1)$ onto $\mathbf{S}(\mathrm{E}_2)$.

Now we pass to determining the formula for the right inverse $\hat{\mathscr{B}}^{[-1]}$ of $\hat{\mathscr{B}} : \mathbf{Y}_{\nu,\mu}\,(\mathrm{E}_1) \to \mathbf{S}(\mathrm{E}_2)$. Introduce the polynomials

$$\tilde{L}_r(\lambda) = L_r(\lambda) - l_{r\nu} L(\lambda) \qquad (r = 3,\dots, s) \tag{5.21}$$

of degree $\leqslant \nu - 1$, and consider the equation

$$M(\lambda)Y(\lambda) + Z(\lambda)L(\lambda) = -\sum_{r=3}^{s} M_r(\lambda)\tilde{L}_r(\lambda). \tag{E_5}$$

Using the suitable analogue of Lemma 5.1, one can show that $(Y(\lambda), Z(\lambda)) \in \mathbf{Y}_{\nu,\mu}\,(\mathrm{E}_1)$ if and only if $(\tilde{Y}(\lambda), \tilde{Z}(\lambda)) \in \mathbf{Y}_{\nu,\mu}\,(\mathrm{E}_5)$, where $\tilde{Y}(\lambda) = Y(\lambda)$ and

$$\tilde{Z}(\lambda) = Z(\lambda) + \sum_{r=3}^{s} M_r(\lambda) l_{r\nu}, \tag{5.22}$$

and in this case

$$\mathbf{B}_{Y,Z} = \mathbf{B}_{\tilde{Y},\tilde{Z}}. \tag{5.23}$$

The right hand term in the equation (E_5) has degree $\leqslant \nu + \mu - 1$. We claim that any (ν, μ)-solution $(\hat{Y}(\lambda), \hat{Z}(\lambda))$ of such an equation can be written in the form

$$\hat{Y}(\lambda) = Y_1(\lambda) + GL(\lambda), \qquad \hat{Z}(\lambda) = Z_1(\lambda) - M(\lambda)G, \tag{5.24}$$

where $(Y_1(\lambda), Z_1(\lambda))$ is a $(\nu-1, \mu-1)$-solution of (E_5) and G is some $n \times n$ matrix. Indeed, one easily checks that $(\hat{Y}(\lambda), \hat{Z}(\lambda))$ given by (5.24) is a solution of (E_5). Conversely, if $(\hat{Y}(\lambda), \hat{Z}(\lambda)) \in \mathbf{Y}_{\nu,\mu}$ (E_5), then clearly $\hat{y}_\nu = -\hat{z}_\mu$, where $\hat{y}_\nu$ and $\hat{z}_\mu$ denote the leading coefficients of $\hat{Y}(\lambda)$ and $\hat{Z}(\lambda)$, respectively. Setting

$$Y_1(\lambda) = \hat{Y}(\lambda) - \hat{y}_\nu L(\lambda), \qquad Z_1(\lambda) = Z(\lambda) + M(\lambda)\hat{y}_\nu,$$

we obtain a $(\nu-1, \mu-1)$-solution of (E_5). Observe also that $Y(X^{(0)}, C_L) = Y_1(X^{(0)}, C_L)$ and therefore, using formulas of type (5.5) for the Bezoutians, we conclude that

$$\mathbf{B}_{\hat{Y},\hat{Z}} = \mathbf{B}_{Y_1,Z_1}. \tag{5.25}$$

Now, all $(\nu-1, \mu-1)$-solutions of (E_5) are described in Proposition 5.3. Namely, any $S = [s_{jk}]_{j,k=0}^{\mu-1,\nu-1} \in \mathrm{S}(E_2)$ generates the solution

$$Z_1(\lambda) = -\sum_{j=0}^{\mu-1} \lambda^j s_{j,\nu-1},$$

$$Y_1(\lambda) = \sum_{k=0}^{\nu-1} \lambda^k s_{\mu-1,k} - \sum_{r=3}^{s} m_{r\mu}\tilde{L}_r(\lambda)$$

$$= \sum_{k=0}^{\nu-1} \lambda^k s_{\mu-1,k} - \sum_{r=3}^{s} m_{r\mu}L_r(\lambda) + \sum_{r=3}^{s} m_{r\mu}l_{r\nu}L(\lambda) \tag{5.26}$$

such that

$$\mathbf{B}_{Y_1,Z_1} = S. \tag{5.27}$$

Combining (5.26), (5.24), and (5.22), we see that any (ν, μ)-solution of (E_1)

can be written in the form

$$Y(\lambda) = \sum_{k=0}^{\nu-1} \lambda^k s_{\mu-1,k} - \sum_{r=3}^{s} m_{r\mu} L_r(\lambda) + \left(\sum_{r=3}^{s} m_{r\mu} l_{r\nu} + G \right) L(\lambda),$$

$$Z(\lambda) = - \sum_{j=0}^{\mu-1} \lambda^j s_{j,\nu-1} - \sum_{r=3}^{s} M_r(\lambda) l_{r\nu} - M(\lambda) G,$$

and from (5.27), (5.25), and (5.23) it follows that $\mathbf{B}_{Y,Z} = S$. This proves the assertions of Theorem 4.2 about the generalized inverse and the general form of (ν, μ)-solutions. ∎

Proof of Theorem 4.3. We know already from Lemma 5.2 that $\mathscr{B}$ maps $\mathbf{Y}_{\nu,\mu-1}$ onto $\mathrm{S}(\mathrm{E}_2)$. Similarly, $\mathscr{B}$ maps $\mathbf{Y}_{\nu-1,\mu}$ onto $\mathrm{S}(\mathrm{E}_2)$. Using the same argument as in the proof of injectivity of $\mathscr{B}_0$ in Proposition 5.3, one shows that the map $\tilde{\mathscr{B}}: \mathbf{Y}_{\nu-1,\mu} \to \mathrm{S}(\mathrm{E}_2)$ is injective. Similarly, the mapping $\tilde{\tilde{\mathscr{B}}}: \mathbf{Y}_{\nu,\mu-1} \to \mathrm{S}(\mathrm{E}_2)$ is injective. The formulas for $\tilde{\mathscr{B}}^{-1}$ and $\tilde{\tilde{\mathscr{B}}}^{-1}$ are checked by inspection of the degrees of $Y_S(\lambda)$ and $Z_S(\lambda)$ in (4.7). ∎

It is our pleasure to thank Professor I. Gohberg for valuable remarks and fruitful discussions. We are also grateful to Professors M. A. Kaashoek and P. Lancaster for constructive comments.

REFERENCES

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

2 S. Barnett, Regular polynomial matrices having relatively prime determinants, *Proc. Cambridge Philos. Soc.* 65:585–590 (1969).

3 S. Barnett, A note on the Bezoutian matrix, *SIAM J. Appl. Math.* 22:84–86 (1972).

4 S. Barnett and P. Lancaster, Some properties of the Bezoutian for polynomial matrices, *Linear and Multilinear Algebra* 9:99–111 (1980).

5 R. R. Bitmead, S. Y. Kung, B. D. O. Anderson, and T. Kailath, Greatest common divisors via generalized Sylvester and Bezout matrices, *IEEE Trans. Automat. Control* AC-23:1043–1047 (1978).

6 D. Carlson and B. N. Datta, On the effective computation of the inertia of a non-hermitian matrix, *Numer. Math.* 33:315–322 (1979).

7 A. Cayley, Note sur la méthode d'elimination de Bezout, *J. Reine Angew. Math.* 53:366–376 (1857).

8 K. F. Clancey and B. A. Kon, The Bezoutian and the algebraic Riccati equation, *Linear and Multilinear Algebra* 15:265–278 (1984).

9 B.N. Datta, On the Routh-Hurwitz-Fujiwara and the Schur-Cohn-Fujiwara theorems for the root-separation problem, *Linear Algebra Appl.* 22:235–246 (1978).

10 B.N. Datta, On the root separation of polynomials and inertia theorems for matrices, to appear.

11 P. A. Fuhrmann, On a symmetric rational transfer functions, *Linear Algebra Appl.* 50:167–250 (1983).

12 P. A. Fuhrmann, Block Hankel inversion—the polynomial approach, in *Operator Theory: Advances and Applications*, Vol. 19, Birkhäuser, Basel, 1986, pp. 207–230.

13 I. C. Gohberg, M. A. Kaashoek, and D. C. Lay, Equivalence, linearization, and decomposition of holomorphic operator functions, *J. Functional Anal.* 28: 102–144 (1978).

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

15 I. Gohberg, M. A. Kaashoek, L. Lerer, and L. Rodman, Common multiples and common divisors of matrix polynomials, II. The Vandermonde and the resultant matrix, *Linear and Multilinear Algebra* 12:159–203 (1982).

16 I. Gohberg, P. Lancaster, and L. Rodman, *Matrix Polynomials*, Academic, 1982.

17 G. Heinig, On the concepts of Bezoutian and resultant for operator bundles (transl. from Russian), *Functional Anal. Appl.* 11:241–243 (1977).

18 G. Heinig, Resultante, Bezoutiante und Spektralverteilungsprobleme fur Operatorpolynome, *Math. Nachr.* 91:23–43 (1979).

19 G. Heinig and K. Rost, *Algebraic Methods for Toeplitz-like Matrices and Operators*, Birkhäuser, Basel, 1984.

20 R. E. Kalman, On the Hermite-Fujiwara theorem in stability theory, *Quart. Appl. Math.* 23:279–282 (1965).

21 N. Kravitsky, On the discriminant function of two commuting nonselfadjoint operators, *Integral Equations Operator Theory* 3:97–125 (1980).

22 M. G. Krein and M. A. Naimark, The method of symmetric and hermitian forms in the theory of separation of the roots of algebraic equations (transl. from Russian), *Linear and Multilinear Algebra* 10:265–308 (1981).

23 P. Lancaster, L. Lerer, and M. Tismenetsky, Factored forms of solutions of the equation $AX - XB = C$ in matrices, *Linear Algebra Appl.* 62:19–49 (1984).

24 P. Lancaster and J. Maroulas, The kernel of the Bezoutian for operator polynomials, *Linear and Multilinear Algebra* 17:181–201 (1985).

25 P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, Academic, 1985.

26 F.L. Lander, The Bezoutian and inversion of Hankel and Toeplitz matrices (in Russian), *Mat. Issled.* 9(32):49–57 (1974).

27 L. Lerer, L. Rodman, and M. Tismenetsky, Bezoutian and the Schur-Cohn problem for operator polynomials, *J. Math. Anal. Appl.* 103:83–102 (1984).

28 L. Lerer and M. Tismenetsky, The eigenvalue-separation problem for matrix polynomials, *Integral Equations Operator Theory* 5:386–445 (1982).

29 L. Lerer and M. Tismenetsky, Bezoutian for Several Matrix Polynomials and Matrix Equations, TR88.145, IBM—Israel Scientific Center, Haifa, Nov. 1984.
30 L. Lerer and M. Tismenetsky, On the location of spectrum of matrix polynomials, *Contemp. Math.* 47:287–297 (1985).
31 L. Lerer and M. Tismenetsky, Generalized Bezoutian and the inversion problem for block matrices, I. General scheme, *Integral Equations Operator Theory* 9:790–819 (1986).
32 V. Pták, Explicit expressions for Bezoutians, *Linear Algebra Appl.* 59:43–54 (1984).
33 W. Roth, The equations $AX - YB = C$ and $AX - XB = C$ in matrices, *Proc. Amer. Math. Soc.* 2:392–396 (1951).
34 I. Sylvester, On a theory of the syzygetic relations of two rational integral functions, comprising an application to the theory of Sturm's functions, and that of the greatest algebraical common measure, *Philos. Trans. Roy. Soc. London* 143:407–548 (1853).
35 M. Tismenetsky, Bezoutians, Toeplitz and Hankel Matrices in the Spectral Theory of Matrix Polynomials, Ph.D. Thesis, Technion, Haifa, 1981.

Received 20 January 1987; final manuscript accepted 12 April 1987