Vector spaces of linearizations for matrix polynomials: a bivariate polynomial approach

Townsend, Alex and Noferini, Vanni and Nakatsukasa, Yuji

2012

Manchester Institute for Mathematical Sciences

School of Mathematics

The University of Manchester

VECTOR SPACES OF LINEARIZATIONS FOR MATRIX POLYNOMIALS: A BIVARIATE POLYNOMIAL APPROACH

ALEX TOWNSEND[*], VANNI NOFERINI[†], AND YUJI NAKATSUKASA[‡]

Abstract. We revisit the important paper [D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 971–1004] and, by viewing matrices as coefficients for bivariate polynomials, we provide concise proofs for key properties of linearizations for matrix polynomials. We also show that every pencil in the double ansatz space is intrinsically connected to a Bézout matrix, which we use to prove the eigenvalue exclusion theorem. In addition our exposition allows for any degree-graded basis, the monomials being a special case. MATLAB code is given to construct the pencils in the double ansatz space for matrix polynomials expressed in any orthogonal basis.

Key words. matrix polynomials, bivariate polynomials, Bézout matrix, degree-graded basis, structure-preserving linearizations, polynomial eigenvalue problem, matrix pencil

AMS subject classifications. 65F15, 15A18, 15A22

1. Introduction. The landmark paper by Mackey, Mackey, Mehl, and Mehrmann [11] introduced three important vector spaces of pencils for matrix polynomials: $\mathbb{L}_1(P)$, $\mathbb{L}_2(P)$, and $\mathbb{DL}(P)$. In [11] the spaces $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ generalize the companion forms of the first and second kind, respectively, and the *double ansatz space* is the intersection, $\mathbb{DL}(P) = \mathbb{L}_1(P) \cap \mathbb{L}_2(P)$.

In this article we introduce new viewpoints for these vector spaces, which are important for polynomial eigenvalue problems. The classic approach is *linearization*, i.e., computing the eigenvalues of a matrix polynomial $P(\lambda)$ by solving a generalized linear eigenvalue problem. The vector spaces we study provide a family of candidate generalized eigenvalue problems for computing the eigenvalues of a matrix polynomial. We regard a block matrix as coefficients for a bivariate matrix polynomial (see section 3), and point out that every pencil in $\mathbb{DL}(P)$ is a (generalized) Bézout matrix [10] (see section 4). These novel viewpoints allow us to obtain remarkably elegant proofs for many properties of $\mathbb{DL}(P)$ and the eigenvalue exclusion theorem, which previously required rather tedious derivations. Furthermore, our exposition includes matrix polynomials expressed in any degree-graded basis, such as the Chebyshev polynomial basis, which are beginning to become important [6].

Let us recall some basic definitions in the theory of matrix polynomials. Let $P(\lambda) = \sum_{i=0}^{k} A_i \phi_i(\lambda)$, where $A_k \neq 0$ and $A_i \in \mathbb{C}^{n \times n}$, be a matrix polynomial expressed in a degree-graded basis, i.e., $\{\phi_0, \ldots, \phi_k\}$ is a polynomial basis with ϕ_j of degree exactly j, $0 \leq j \leq k$. We assume throughout that $P(\lambda)$ is regular, i.e., $\det P(\lambda) \not\equiv 0$, which ensures the eigenvalues of $P(\lambda)$ are the roots of the scalar polynomial $\det(P(\lambda))$. We note that, although we use the field $\mathbb{C}$ for simplicity, the elements of A_i can be in any field $\mathbb{F}$, provided we work with the closure of $\mathbb{F}$.

A matrix pencil $L(\lambda) = \lambda X + Y$ where $X, Y \in \mathbb{C}^{nk \times nk}$ is a *linearization* for $P(\lambda)$ if there exist unimodular matrix polynomials $U(\lambda), V(\lambda)$ such that $L(\lambda) = U(\lambda) \operatorname{diag}(P(\lambda), I_{n(k-1)}) V(\lambda)$ and hence, $L(\lambda)$ shares its finite eigenvalues and their

[*]Mathematical Institute, University of Oxford, Oxford, England, OX1 3LB (townsend@maths.ox.ac.uk).

[†]School of Mathematics, University of Manchester, Manchester, England, M13 9PL (vanni.noferini@manchester.ac.uk).

[‡]School of Mathematics, University of Manchester, Manchester, England, M13 9PL (yuji.nakatsukasa@manchester.ac.uk).

partial multiplicities with $P(\lambda)$. If $P(\lambda)$ has a singular leading coefficient then it has an eigenvalue at infinity and to preserve any infinite eigenvalues the matrix pencil $L(\lambda)$ needs to be a *strong linearization*, i.e., $L(\lambda)$ is a linearization for $P(\lambda)$ and $\lambda Y + X$ a linearization for $\lambda^k P(1/\lambda)$.

In the next section we extend the definitions of $\mathbb{L}_1(P)$, $\mathbb{L}_2(P)$ and $\mathbb{DL}(P)$ to allow for matrix polynomials expressed in a degree-graded basis. In section 3 we consider the same space from a new viewpoint, based on bivariate matrix polynomials, and provide concise proofs for properties of $\mathbb{DL}(P)$. Section 4 shows that every pencil in $\mathbb{DL}(P)$ is a (generalized) Bézout matrix and gives an alternative proof for the eigenvalue exclusion theorem. In section 5 we provide MATLAB code to construct the block symmetric pencils in $\mathbb{DL}(P)$ when the matrix polynomial is expressed in any orthogonal basis. In section 6 we discuss a few related observations.

2. Vector spaces and degree-graded bases. Given a matrix polynomial $P(\lambda)$ we can define a vector space $\mathbb{L}_1(P)$ [11, Def. 3.1]

$$\mathbb{L}_1(P) = \left\{ L(\lambda) = \lambda X + Y : X, Y \in \mathbb{C}^{nk \times nk}, \ L(\lambda) \cdot (\Lambda(\lambda) \otimes I_n) = v \otimes P(\lambda), v \in \mathbb{C}^k \right\},$$

where $\Lambda(\lambda) = [\phi_{k-1}(\lambda), \phi_{k-2}(\lambda), \dots, \phi_0(\lambda)]^T$ and $\otimes$ is the Kronecker product. An *ansatz vector* $v \in \mathbb{C}^k$ can be selected to generate a family of pencils in $\mathbb{L}_1(P)$, which are generically linearizations for $P(\lambda)$ [11, Thm. 4.7]. If $\{\phi_0, \dots, \phi_k\}$ is an orthogonal basis then the comrade form [13] belongs to $\mathbb{L}_1(P)$ with $v = [1, 0, \dots, 0]^T$.

The action of $L(\lambda) = \lambda X + Y \in \mathbb{L}_1(P)$ on $(\Lambda(\lambda) \otimes I_n)$ can be characterized by the *column shift sum* operator, denoted by $\boxplus\!\!\!\rightarrow$ [11, Lemma 3.4],

$$L(\lambda) \cdot (\Lambda(\lambda) \otimes I_n) = v \otimes P(\lambda) \Longleftrightarrow X \boxplus\!\!\!\rightarrow Y = v \otimes [A_k, A_{k-1}, \dots, A_0].$$

In the monomial basis $X \boxplus\!\!\!\rightarrow Y$ can be paraphrased as "insert a zero column on the right of X and a zero column on the left of Y then add them together", i.e.,

$$X \boxplus\!\!\!\rightarrow Y = \begin{bmatrix} X & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} & Y \end{bmatrix},$$

where $\mathbf{0} \in \mathbb{C}^{nk \times n}$. More generally, for degree-graded bases we define the column shift sum operator as

$$X \boxplus\!\!\!\rightarrow Y = XM + \begin{bmatrix} \mathbf{0} & Y \end{bmatrix}, \tag{2.1}$$

where $M \in \mathbb{C}^{nk \times n(k+1)}$ and $\mathbf{0} \in \mathbb{C}^{nk \times n}$. Suppose the degree-graded basis $\{\phi_0, \dots, \phi_k\}$ satisfies the recurrence relations

$$x\phi_{i-1} = \sum_{j=0}^{i} m_{k+1-i, k+1-j} \phi_j, \quad 1 \le i \le k.$$

Then the matrix M in (2.1) is given by

$$M = \begin{bmatrix} M_{11} & M_{12} & \dots & M_{1k} & M_{1(k+1)} \\ \mathbf{0} & M_{22} & \ddots & \ddots & M_{2(k+1)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{0} & M_{kk} & M_{k(k+1)} \end{bmatrix},$$

where $M_{pq} = m_{pq} I_n$, $1 \le p \le q \le k+1$, $p \ne k+1$ and I_n is the $n \times n$ identity matrix. An orthogonal basis satisfies a three term recurrence and in this case the

matrix M has only three nonzero block diagonals. For example, if $P(\lambda)$ is expressed in the Chebyshev basis $\{T_0(x), \ldots, T_k(x)\}$, $T_j(x) = \cos\left(j\cos^{-1}(x)\right)$ for $x \in [-1, 1]$, we have

$$
M = \begin{bmatrix} \frac{1}{2}I_n & 0 & \frac{1}{2}I_n \\ & \ddots & \ddots & \ddots \\ & & \frac{1}{2}I_n & 0 & \frac{1}{2}I_n \\ & & & I_n & 0 \end{bmatrix} \in \mathbb{R}^{nk \times n(k+1)}.
$$

The properties of the vector space $\mathbb{L}_2(P)$ are analogous to $\mathbb{L}_1(P)$ [8]. If $L(\lambda) = \lambda X + Y$ is in $\mathbb{L}_2(P)$ then $L(\lambda) = \lambda X^{\mathcal{B}} + Y^{\mathcal{B}}$ belongs to $\mathbb{L}_1(P)$, where the superscript $^{\mathcal{B}}$ represents blockwise transpose[1]. This connection means the action of $L(\lambda) \in \mathbb{L}_2(P)$ is characterized by a *row shift sum* operator, denoted by $\boxplus\!\!\downarrow$,

$$
X \boxplus\!\!\downarrow Y = \left(X^{\mathcal{B}} \boxplus\!\!\rightarrow Y^{\mathcal{B}}\right)^{\mathcal{B}} = M^{\mathcal{B}} X + Y.
$$

2.1. Extending results to general bases. Many of the derivations in [11] are specifically for when $P(\lambda)$ is expressed in a monomial basis, though the lemmas and theorems can be generalized to any degree-graded bases. One approach to generalize [11] is to use the change of basis matrix S such that $\Lambda(\lambda) = S[\lambda^{k-1}, \ldots, \lambda, 1]^T$ and to define the mapping (see also [5])

$$
\mathcal{C}\left(\hat{L}(\lambda)\right) = \hat{L}(\lambda)(S^{-1} \otimes I_n) = L(\lambda), \tag{2.2}
$$

where $\hat{L}(\lambda)$ is a pencil in $\mathbb{L}_1$ for the matrix polynomial $P(\lambda)$ expressed in the monomial basis. In particular, the strong linearization theorem holds for any degree-graded basis.

THEOREM 2.1 (Strong Linearization Theorem). *Let $P(\lambda)$ be a regular matrix polynomial (expressed in any degree-graded basis), and let $L(\lambda) \in \mathbb{L}_1(P)$. Then the following statements are equivalent:*

1. $L(\lambda)$ is a linearization for $P(\lambda)$.
2. $L(\lambda)$ is a regular pencil.
3. $L(\lambda)$ is a strong linearization for $P(\lambda)$.

Proof. It is a corollary of [11, Theorem 4.3]. In fact, the mapping $\mathcal{C}$ in (2.2) is a strict equivalence between $\mathbb{L}_1(P)$ expressed in the monomial basis and $\mathbb{L}_1(P)$ expressed in a degree-graded basis. Therefore, $L(\lambda)$ has one of the three properties above if and only if $\hat{L}(\lambda)$ has the corresponding property. The properties are equivalent for $\hat{L}(\lambda)$ because they are equivalent for $L(\lambda)$. $\square$

This strict equivalence can be used to generalize many other properties of $\mathbb{L}_1(P)$, $\mathbb{L}_2(P)$ and $\mathbb{DL}(P)$, though we use an approach based on bivariate polynomials resulting in concise derivations.

3. Recasting to bivariate matrix polynomials. A block matrix $X \in \mathbb{C}^{nk \times nk}$ with $n \times n$ blocks can provide the coefficients for a bivariate matrix polynomial of degree k. For example, the bivariate matrix polynomial corresponding to the coefficients X is $F(x, y)$, where

$$
X = \begin{bmatrix} X_{11} & \ldots & X_{1k} \\ \vdots & \ddots & \vdots \\ X_{k1} & \ldots & X_{kk} \end{bmatrix}, \quad X_{ij} \in \mathbb{C}^{n \times n}, \qquad F(x, y) = \sum_{i=0}^{k-1}\sum_{j=0}^{k-1} X_{k-i,k-j}\phi_i(y)\phi_j(x).
$$

---

[1]If $X = (X_{ij})_{1 \le i,j \le k}$, $X_{ij} \in \mathbb{C}^{n \times n}$ then $X^{\mathcal{B}} = (X_{ji})_{1 \le i,j \le k}$.

The matrix M in (2.1) is such that the bivariate matrix polynomial corresponding to the coefficients XM is $F(x,y)x$, i.e., M applied on the right of X represents multiplication of $F(x,y)$ by x. This gives an equivalent definition for the column shift sum operator: if the block matrices X and Y are the coefficients for $F(x,y)$ and $G(x,y)$ then the coefficients of $H(x,y)$ are Z, where

$$Z = X \boxright Y, \qquad H(x,y) = F(x,y)x + G(x,y).$$

Therefore, in terms of bivariate matrix polynomials, we can define $\mathbb{L}_1(P)$ as

$$\mathbb{L}_1(P) = \left\{ L(\lambda) = \lambda X + Y : F(x,y)x + G(x,y) = v(y)P(x), v \in \Pi_{k-1}(\mathbb{C}) \right\},$$

where $\Pi_{k-1}(\mathbb{C})$ is the space of polynomials in $\mathbb{C}[x]$ of degree $k-1$, or less.

Regarding the space $\mathbb{L}_2(P)$, the coefficient matrix $M^{\mathcal{B}}X$ corresponds to the bivariate matrix polynomial $yF(x,y)$, i.e., $M^{\mathcal{B}}$ applied on the left of X represents multiplication of $F(x,y)$ by y. Hence, we can define $\mathbb{L}_2(P)$ as

$$\mathbb{L}_2(P) = \left\{ L(\lambda) = \lambda X + Y : yF(x,y) + G(x,y) = P(y)w(x), w \in \Pi_{k-1}(\mathbb{C}) \right\}.$$

The space $\mathbb{DL}(P)$ is the intersection of $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$, and is of importance because it contains block symmetric linearizations. A pencil $L(\lambda) = \lambda X + Y$ belongs to $\mathbb{DL}(P)$ with ansätze $v(y)$ and $w(x)$ if the following $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ conditions are satisfied:

$$F(x,y)x + G(x,y) = v(y)P(x), \qquad yF(x,y) + G(x,y) = P(y)w(x). \qquad (3.1)$$

It appears that $v(y)$ and $w(x)$ could be chosen independently; however, if we substitute $y = x$ into (3.1) we obtain the *compatibility condition*

$$v(x)P(x) = F(x,x)x + G(x,x) = xF(x,x) + G(x,x) = P(x)w(x)$$

and hence, $v = w$ as elements of $\Pi_{k-1}(\mathbb{C})$ since $P(x)(v(x) - w(x))$ is the zero matrix. This shows the double ansatz space is actually a single ansatz space; a fact that required two quite technical proofs in [11, Prop. 5.2, Thm. 5.3].

The bivariate matrix polynomials $F(x,y)$ and $G(x,y)$ are uniquely defined by the ansatz $v(x)$ since they satisfy the explicit formulas

$$yF(x,y) - F(x,y)x = P(y)v(x) - v(y)P(x), \qquad (3.2)$$

$$yG(x,y) - G(x,y)x = yv(y)P(x) - P(y)v(x)x. \qquad (3.3)$$

In other words, there is an isomorphism between $\Pi_k(\mathbb{C})$ and $\mathbb{DL}(P)$. It also follows from (3.2) and (3.3) that $F(x,y) = F(y,x)$ and $G(x,y) = G(y,x)$. This shows that all the pencils in $\mathbb{DL}(P)$ are block symmetric. Furthermore, if $F(x,y)$ and $G(x,y)$ are symmetric and satisfy $F(x,y)x + G(x,y) = P(x)v(y)$ then we also have $F(y,x)x + G(y,x) = P(x)v(y)$, and by swapping x and y we obtain the $\mathbb{L}_2(P)$ condition, $yF(x,y) + G(x,y) = P(y)v(x)$. This shows all block symmetric pencils in $\mathbb{L}_1(P)$ belong to $\mathbb{L}_2(P)$ and hence, also belong to $\mathbb{DL}(P)$. Thus, $\mathbb{DL}(P)$ can be defined as the space of block symmetric pencils in $\mathbb{L}_1(P)$ [8, Thm. 3.4].

4. Eigenvalue exclusion theorem. The eigenvalue exclusion theorem [11, Thm. 6.9] shows that if $L(\lambda) \in \mathbb{DL}(P)$ with ansatz $v \in \Pi_{k-1}(\mathbb{C})$ then $L(\lambda)$ is a linearization for the matrix polynomial $P(\lambda)$ if and only if $v(\lambda)I_n$ and $P(\lambda)$ do not share an eigenvalue. This theorem is important because, generically, $v(\lambda)I_n$ and $P(\lambda)$ do not share eigenvalues and almost all choices for $v \in \Pi_{k-1}(\mathbb{C})$ correspond to linearizations in $\mathbb{DL}(P)$ for $P(\lambda)$. This theorem also includes infinite eigenvalues.

The matrix polynomial $P(\lambda)$ expressed in a degree-graded basis has an infinite eigenvalue if its leading matrix coefficient is singular. In order to correctly take care of infinite eigenvalues we write $P(\lambda) = \sum_{i=0}^{g} A_i \phi_i(\lambda)$, where the integer $g \geq k$ is called the *grade* [12]. If the grade of $P(\lambda)$ is larger than the degree then $P(\lambda)$ has at least one infinite eigenvalue. Usually, and unless stated otherwise, the grade is equal to the degree.

We prove the eigenvalue exclusion theorem by noting that a $\mathbb{DL}(P)$ pencil is a (generalized) Bézout matrix. We first recall the definition of a Bézout matrix and Bézoutian function [3, p. 277], [4, sec. 2.9].

DEFINITION 4.1 (Bézout matrix and Bézoutian function). *Let $p_1(x), p_2(x)$ be scalar polynomials*

$$p_1(x) = \sum_{i=0}^{k} a_i \phi_i(x), \qquad p_2(x) = \sum_{i=0}^{k-1} c_i \phi_i(x)$$

with $a_k \neq 0$ (c_{k-1} can be zero, i.e., we regard $p_2(x)$ as a polynomial of grade $k-1$), then the Bézoutian function is the bivariate function

$$\mathcal{B}(p_1, p_2) = \frac{p_1(y)p_2(x) - p_2(y)p_1(x)}{x - y} = \sum_{i,j=1}^{k} b_{ij} \phi_{k-i}(y) \phi_{k-j}(x).$$

The $k \times k$ Bézout matrix associated to $p_1(x)$ and $p_2(x)$ is defined via the coefficients of the Bézoutian function

$$B(p_1, p_2) = (b_{ij})_{1 \leq i,j \leq k}.$$

Similarly, for $n \times n$ regular matrix polynomials $P_1(x), P_2(x)$ of grades k and $k-1$ respectively, the associated Bézoutian function $\mathcal{B}_{M_2,M_1}$ is defined by [2, 10]

$$\mathcal{B}_{M_2,M_1}(P_1, P_2) = \frac{M_2(y)P_2(x) - M_1(y)P_1(x)}{x - y} = \sum_{i,j=1}^{k} B_{ij} \phi_{k-i}(y) \phi_{k-j}(x), \qquad (4.1)$$

where $M_1(x)$ and $M_2(x)$ are regular matrix polynomials such that $M_1(x)P_1(x) = M_2(x)P_2(x)$, [7, Ch. 9]. The $nk \times nk$ Bézout block matrix is defined by $B_{M_2,M_1}(P_1, P_2) = (B_{ij})_{1 \leq i,j \leq k}$.

Note that for matrix polynomials the Bézoutian function and the Bézout block matrix are not unique as there are many choices of M_1 and M_2. However, when $P_1(x)$ and $P_2(x)$ commute, i.e., $P_2(x)P_1(x) = P_1(x)P_2(x)$, the natural choice is $M_1 = P_2$ and $M_2 = P_1$ and we write $\mathcal{B}(P_1, P_2) = \mathcal{B}_{P_1,P_2}(P_1, P_2)$.

Here are some standard properties of a Bézoutian function and Bézout matrix:
1. The Bézoutian function is skew-symmetric with respect to its polynomial arguments: $\mathcal{B}(P_1, P_2) = -\mathcal{B}(P_2, P_1)$.
2. $\mathcal{B}(P_1, P_2)$ is bilinear with respect to its polynomial arguments.

3. In the scalar case, $B(p_1, p_2)$ is nonsingular if and only if p_1 and p_2 have no common roots.

4. $B(P_1, P_2)$ is a (block) symmetric matrix.

THEOREM 4.2 (Eigenvalue Exclusion Theorem). *Suppose that $P(\lambda)$ is a regular matrix polynomial of degree k and $L(\lambda)$ is in $\mathbb{DL}(P)$ with a nonzero ansatz polynomial $v(\lambda)$. Then $L(\lambda)$ is a linearization for $P(\lambda)$ if and only if $v(\lambda)I_n$ (with grade $k-1$) and $P(\lambda)$ do not share an eigenvalue.*

Proof. To highlight the connection with the classic Bézout matrix we first consider scalar polynomials, where $n = 1$. Let $P(\lambda)$ be a scalar polynomial of degree (and grade) k and $v(\lambda)$ a scalar polynomial of grade $k - 1$. We first solve the relations in (3.2) and (3.3) to obtain

$$F(x, y) = \frac{P(y)v(x) - v(y)P(x)}{x - y}, \qquad G(x, y) = \frac{yv(y)P(x) - P(y)v(x)x}{x - y}$$

and thus, by Definition 4.1, $F(x, y) = \mathcal{B}(v, P)$ and $G(x, y) = \mathcal{B}(P, vx)$. Moreover, $\mathcal{B}$ is skew-symmetric and bilinear with respect to its polynomial arguments so we have

$$L(\lambda) = \lambda X + Y = \lambda B(v, P) + B(P, xv) = -\lambda B(P, v) + B(P, xv) = B(P, (x - \lambda)v).$$

Therefore, since B is a Bézout matrix, $\det(L(\lambda)) = \det(B(P, (x - \lambda)v)) = 0$ for all λ if and only if P and v share a root. Finally, by Theorem 2.1, $L(\lambda)$ is a linearization for $P(\lambda)$ if and only if P and v do not share a root.

For the matrix case $n > 1$, we let $P_1 = P(x)$ and $P_2 = (x - \lambda)v(x)I_n$ in (4.1). Then P_1 and P_2 commute for all λ, so we take $M_1 = P_2$ and $M_2 = P_1$ and obtain

$$\mathcal{B}(P(x), (x - \lambda)v(x)I_n) = \frac{P(y)(x - \lambda)v(x) - (y - \lambda)v(y)P(x)}{x - y}$$

$$= \sum_{i,j=1}^{k} B_{ij}\phi_{k-i}(y)\phi_{k-j}(x).$$

This gives the $nk \times nk$ Bézout block matrix $B(P, (x - \lambda)vI) = (B_{ij})_{1 \le i,j \le k}$. Completely analogously to the scalar case, we have $L(\lambda) = B(P, (x - \lambda)vI)$.

The kernel of the Bézout block matrix is

$$\ker B(P, (x - \lambda)vI) = \operatorname{Im} \begin{bmatrix} X_F\phi_{k-1}(T_F) \\ \vdots \\ X_F\phi_0(T_F) \end{bmatrix} \oplus \operatorname{Im} \begin{bmatrix} X_\infty\phi_0(T_\infty) \\ \vdots \\ X_\infty\phi_{k-1}(T_\infty) \end{bmatrix}, \qquad (4.2)$$

and does not depend on the choice of M_1 and M_2, as shown in Theorem 1.1 of [10] for the monomial case. Equation (4.2) can be obtained from [10, Theorem 1.1] via a congruence transformation involving the mapping $\mathcal{C}$ in (2.2). Here (X_F, T_F), (X_∞, T_∞) are the greatest common restrictions [7, Ch. 9] of the finite and infinite spectral pairs of $P(x)$ and $(x - \lambda)v(x)I_n$. The infinite spectral pairs are defined regarding both polynomials as grade k. We recall that two matrix polynomials have a nonempty greatest common restriction if and only if they share both an eigenvalue and the corresponding eigenvector [7, Ch. 7,9].

If vI_n and P share a finite eigenvalue λ_0 and $P(\lambda_0)w = 0$ for a nonzero w then $(\lambda_0 - \lambda)v(\lambda_0)w = 0$ for all λ. Hence, the kernel of $L(\lambda) = B(P, (x - \lambda)v)$ is nonempty for all λ and $L(\lambda)$ is singular. An analogous argument holds for a shared infinite

6

eigenvalue. Conversely, suppose $v(\lambda)I_n$ and $P(\lambda)$ have no common eigenvalues. If λ_0 is an eigenvalue of P then $(\lambda_0 - \lambda)v(\lambda_0)I$ is nonsingular unless $\lambda = \lambda_0$. It follows that if λ is not an eigenvalue for P then the common restriction is empty, which means $L(\lambda)$ is nonsingular. Hence, $L(\lambda)$ is regular and a linearization by Theorem 2.1. □

5. Construction. The formulas (3.2) and (3.3) can be used to construct any pencil in $\mathbb{DL}(P)$ without basis conversion, which can be numerically important [1]. We provide a MATLAB code that constructs pencils in $\mathbb{DL}(P)$ when the matrix polynomial is expressed in any orthogonal basis. If $P(\lambda)$ is expressed in the monomials then $\mathtt{a} = [\mathtt{ones(k, 1)}]; \mathtt{b} = \mathtt{zeros(k, 1)}; \mathtt{c} = \mathtt{zeros(k, 1)};$ and if expressed in the Chebyshev basis then $\mathtt{a} = [\mathtt{ones(k - 1, 1); 2]/2}; \mathtt{b} = \mathtt{zeros(k, 1)}; \mathtt{c} = \mathtt{ones(k, 1)/2};$.

```
function [X Y] = DLP(AA,v,a,b,c)
%DLP constructs the DL pencil with ansatz vector v.
% [X,Y] = DLP(AA,v,a,b,c) returns the DL pencil lambda*X + Y
%    corresponding to the matrix polynomial with coefficients AA in an
%    orthogonal basis defined by the recurrence relations a, b, c.

[n m] = size(AA); k=m/n-1; s=n*k;               % matrix size & degree
M = spdiags([a b c],[0 1 2],k,k+1);
M = kron(M,eye(n));                             % multiplication matrix

S = kron(v,AA);
for j=0:k-1, jj=n*j+1:n*j+n; AA(:,jj)=AA(:,jj)';end % block transpose
T = kron(v.',AA'); R=M'*S-T*M;                  % construct RHS

% The Bartel-Stewart algorithm on M'Y+YM=R
X = zeros(s); Y=X; ii=n+1:s+n; nn=1:n;          % useful indices
Y(nn,:)=R(nn,ii)/M(1); X(nn,:)=T(nn,:)/M(1);    % 1st column of X and Y
Y(nn+n,:)=(R(nn+n,ii)-M(1,n+1)*Y(nn,:)+Y(nn,:)*M(:,n+1:s+n))/M(n+1,n+1);
X(nn+n,:)=(T(nn+n,:)-Y(nn,:)-M(1,n+1)*X(nn,:))/M(n+1,n+1); % 2nd cols

for i = 3:k                                     % backwards subs
    ni=n*i; jj=ni-n+1:ni; j0=jj-2*n; j1=jj-n;   % useful indices
    M0=M(ni-2*n,ni); M1=M(ni-n,ni); m=M(ni,ni); % consts of 3-term
    Y0=Y(j0,:); Y1=Y(j1,:); X0=X(j0,:); X1=X(j1,:); % vars in 3-term
    Y(jj,:)=(R(jj,ii)-M1*Y1-M0*Y0+Y1*M(:,n+1:s+n))/m;
    X(jj,:)=(T(jj,:)-Y1-M1*X1-M0*X0)/m;         % use Y to solve for X
end
```

If $P(\lambda)$ is expressed in the monomial basis we have (see [4, Eqn. 2.9.3] for scalar polynomials)

$$L(\lambda) = \begin{bmatrix} A_{k-1} & \dots & A_0 \\ \vdots & \cdot^{\cdot^{\cdot}} & \\ A_0 & & \end{bmatrix} \begin{bmatrix} \hat{v}_k I_n & \dots & \hat{v}_1 I_n \\ & \ddots & \vdots \\ & & \hat{v}_k I_n \end{bmatrix} - \begin{bmatrix} \hat{v}_{k-1} I_n & \dots & \hat{v}_0 I_n \\ \vdots & \cdot^{\cdot^{\cdot}} & \\ \hat{v}_0 I_n & & \end{bmatrix} \begin{bmatrix} A_k & \dots & A_1 \\ & \ddots & \vdots \\ & & A_k \end{bmatrix},$$

where $\hat{v}_i = (v_{i-1} - \lambda v_i)$. This relation can be used to obtain expressions for the block matrices X and Y. For other orthogonal bases the relation is more complicated.

Matrix polynomials expressed in the Legendre or Chebyshev basis are of practical importance, for example, for a nonlinear eigenvalue solvers based on Chebyshev interpolation [6]. Table 5.1 depicts three $\mathbb{DL}(P)$ pencils for the cubic matrix polynomial

$P(\lambda) = A_3 T_3(\lambda) + A_2 T_2(\lambda) + A_1 T_1(\lambda) + A_0 T_0(\lambda)$, where $T_j(\lambda)$ is the jth Chebyshev polynomial.

v	$L(\lambda) \in \mathbb{DL}(P)$ for given v	Linearization condition
$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$	$\lambda \begin{bmatrix} 2A_3 & 0 & 0 \\ 0 & 2A_3 - 2A_1 & -2A_0 \\ 0 & -2A_0 & A_3 - A_1 \end{bmatrix} + \begin{bmatrix} A_2 & A_1 - A_3 & A_0 \\ A_1 - A_3 & 2A_0 & A_1 - A_3 \\ A_0 & A_1 - A_3 & A_0 \end{bmatrix}$	$\det(A_0 + \frac{-A_3 + A_1}{\sqrt{2}}) \neq 0$ $\det(A_0 - \frac{-A_3 + A_1}{\sqrt{2}}) \neq 0$
$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$	$\lambda \begin{bmatrix} 0 & 2A_3 & 0 \\ 2A_3 & 2A_2 & 2A_3 \\ 0 & 2A_3 & A_2 - A_0 \end{bmatrix} + \begin{bmatrix} -A_3 & 0 & -A_3 \\ 0 & A_1 - 3A_3 & A_0 - A_2 \\ -A_3 & A_0 - A_2 & -A_3 \end{bmatrix}$	$\det(-A_2 + A_0) \neq 0$ $\det(A_3) \neq 0$
$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$	$\lambda \begin{bmatrix} 0 & 0 & 2A_3 \\ 0 & 4A_3 & 2A_2 \\ 2A_3 & 2A_2 & A_1 + A_3 \end{bmatrix} + \begin{bmatrix} 0 & -2A_3 & 0 \\ -2A_3 & -2A_2 & -2A_3 \\ 0 & -2A_3 & A_0 - A_2 \end{bmatrix}$	$\det(A_3) \neq 0$

TABLE 5.1

Three instances of pencils in $\mathbb{DL}(P)$ and their linearization condition for the cubic matrix polynomial $P(\lambda) = A_3 T_3(\lambda) + A_2 T_2(\lambda) + A_1 T_1(\lambda) + A_0 T_0(\lambda)$, expressed in the Chebyshev basis of the first kind. These three pencils form a basis for the vector space $\mathbb{DL}(P)$.

6. Further discussion. One might expect that the definition for $\mathbb{L}_1(P)$ could be generalized by relaxing the $\mathbb{L}_1(P)$ condition to

$$F(x, y)x + G(x, y) = h(x, y).$$

However, it turns out we must still have $h(x, y) = v(y)P(x)$, at least in the scalar case. For example, if P is a scalar polynomial with roots $\lambda_1, \ldots, \lambda_k$ then we still require

$$F(\lambda_i, y)\lambda_i + G(\lambda_i, y) = h(\lambda_i, y) = 0, \quad i = 1, \ldots, k$$

and hence, $h(x, y) = q(x, y)P(x)$. Furthermore, $h(x, y)$ must be a bivariate polynomial of degree at most k in x so $q(x, y) = v(y)$ for some polynomial v and hence, we must have $h(x, y) = v(y)P(x)$.

Matrix polynomials in general orthogonal bases are likely to become more important when considering high degree matrix polynomials as in [6].

The exposition of this article can also be applied to study definite pencils in $\mathbb{DL}(P)$ [9] and, for instance, Lemma 4.5 in [9] can be shown using L'Hôpital's rule.

REFERENCES

[1] A. AMIRASIANI, R. M. CORLESS AND P. LANCASTER, *Linearization of matrix polynomials expressed in polynomial bases*, IMA J. Numer. Anal., 29 (2009), pp. 141–157.

[2] B. ANDERSON AND E. JURY, *Generalized Bezoutian and Sylvester matrices in multivariable linear control*, IEEE Trans. Autom. Control, 21 (1976), pp. 551–556.

[3] D. S. BERNSTEIN, *Matrix Mathematics: Theory, Facts, and Formulas*, Princeton University Press, (2nd edition), 2009.

[4] D. BINI AND V. PAN, *Polynomial and Matrix Computations, Vol. 1: Fundamental Algorithms*, Birkhauser, Boston, 1994.

[5] F. DE TERÁN, F. M. DOPICO AND D. S. MACKEY, *Linearizations of singular matrix polynomials and the recovery of minimal indices*, Electron. J. Linear Algebra, 18 (2009), pp. 371–402.

[6] C. EFFENBERGER AND D. KRESSNER, *Chebyshev interpolation for nonlinear eigenvalue problems*, BIT, 52 (2012), pp. 933–951.

[7] I. GOHBERG, P. LANCASTER, AND L. RODMAN, *Matrix Polynomials*, SIAM, Philadelphia, USA, 2009, (unabridged republication of book first published by Academic Press in 1982).

[8] N. J. HIGHAM, D. S. MACKEY, N. MACKEY AND F. TISSEUR, *Symmetric linearizations for matrix polynomials*, SIAM J. Matrix Anal. Appl., 29 (2006), pp. 143–159.

[9] N. J. HIGHAM, D. S. MACKEY AND F. TISSEUR, *Definite matrix polynomials and their linearization by definite pencils*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 478–502.

[10] L. LERER AND M. TISMENETSKY, *The Bézoutian and the eigenvalue-separation problem for matrix polynomials*, Integral Equations Operator Theory, 5 (1982), pp. 386–445.

[11] D. S. MACKEY, N. MACKEY, C. MEHL AND V. MEHRMANN, *Vector spaces of linearizations for matrix polynomials*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 971–1004.

[12] D. S. MACKEY, N. MACKEY, C. MEHL AND V. MEHRMANN, *Smith forms of palindromic matrix polynomials*, Electron. J. Linear Algebra, 22 (2011), pp. 53–91.

[13] W. SPECHT, *Die Lage der Nullstellen eines Polynoms. III*, Math. Nachr., 16 (1957), pp. 369–389.