A Parametrization of Structure-Preserving Transformations for Matrix Polynomials

Garvey, Seamus D. and Tisseur, Francoise and Wang, Shujuan

2022

MIMS EPrint: **2022.12**

ISSN 1749-9097

A Parametrization of Structure-Preserving Transformations for Matrix Polynomials★

Seamus D. Garvey[a,1], Françoise Tisseur[b,3,*], Shujuan Wang[c,1]

[a]*Department of Mechanical, Materials and Manufacturing Engineering, University of Nottingham, Nottingham, NG7 2RD, United Kingdom*
[b]*Department of Mathematics, The University of Manchester, Manchester, M13 9PL, United Kingdom*
[c]*College of Mathematical Sciences, Harbin Engineering University, Harbin, 150001, Heilongjiang, China*

Abstract

Given a matrix polynomial $A(\lambda)$ of degree d and the associated vector space of pencils $\mathbb{DL}(A)$ described in Mackey, Mackey, Mehl, and Mehrmann [SIAM J. Matrix Anal. Appl., 28 (2006), pp. 971-1004], we construct a parametrization for the set of left and right transformations that preserve the block structure of such pencils, and hence produce a new matrix polynomial $\widetilde{A}(\lambda)$ that is still of degree d and is unimodularly equivalent to $A(\lambda)$. We refer to such left and right transformations as structure-preserving transformations (SPTs). Unlike previous work on SPTs, we do not require the leading matrix coefficient of $A(\lambda)$ to be nonsingular. We show that additional constraints on the parametrization lead to SPTs that also preserve extra structures in $A(\lambda)$ such as symmetric, alternating, and T-palindromic structures. Our parametrization allows easy construction of SPTs that are low-rank modifications of the identity matrix. The latter transform $A(\lambda)$ into an equivalent matrix polynomial $\widetilde{A}(\lambda)$ whose jth matrix coefficient $\widetilde{A}_j$ is a low-rank modification of A_j. We expect such SPTs to be one of the key tools for developing algorithms that reduce a matrix polynomial to Hessenberg form or tridiagonal form in a finite number of steps and without the use of a linearization.

Keywords: matrix polynomial, matrix pencil, structure-preserving transformation, symmetric matrix polynomial, Hermitian matrix polynomial, palindromic matrix polynomial, even matrix polynomial, odd matrix polynomial,

---

★Version of September 7, 2022.

*Corresponding author

Email addresses: `Seamus.Garvey@nottingham.ac.uk` (Seamus D. Garvey), `francoise.tisseur@manchester.ac.uk` (Françoise Tisseur), `wangshujuan@hrbeu.edu.cn` (Shujuan Wang)

[1]This author was supported by Engineering and Physical Sciences Research Council grants EP/E046290/1, GR/S31679/01, GR/M93062/01 and GR/M93079/01.

[2]This author was supported by Engineering and Physical Sciences Research Council grant EP/W018101, and Knowledge Transfer Partnership no. KTP12063 between the University of Manchester and Arup.

[3]This author was supported by State Scholar Fund(No.201706685065) of the China Scholarship Council(CSC).

1. Introduction

Let $A(\lambda) = \sum_{j=0}^{d} A_j \lambda^j$ with $A_j \in \mathbb{F}^{n\times n}$ be a matrix polynomial of degree d, where $\mathbb{F}$ denotes either $\mathbb{C}$ or $\mathbb{R}$. We assume throughout that $A(\lambda)$ is regular, i.e., $\det(A(\lambda)) \neq 0$ for some $\lambda \in \mathbb{C}$. The matrix polynomial $A(\lambda)$ cannot in general be reduced to simpler forms such as, for example, triangular, Hessenberg, tridiagonal, and diagonal forms with strict equivalences, that is, transformations of the form $PA(\lambda)Q$ for some constant and nonsingular matrices P and Q. Unimodular transformations $P(\lambda)A(\lambda)Q(\lambda)$, where $P(\lambda)$ and $Q(\lambda)$ have nonzero constant determinants can be used to achieve simpler forms while preserving the degree d [12], [13], [17]. Unfortunately, the λ-dependence of the unimodular transformations makes them impractical for computation. Structure-preserving transformations (SPTs) offer a way around this. They allow computation of the coefficient matrices of the matrix polynomial $\widetilde{A}(\lambda) = P(\lambda)A(\lambda)Q(\lambda)$ of degree d without explicitly forming $P(\lambda)$ and $Q(\lambda)$.

When the leading coefficient A_d of $A(\lambda)$ is nonsingular, an easy way to compute a monic matrix polynomial of degree d that is equivalent to $A(\lambda)$ is through standard pairs (X, C_A), where

$$C_A := \begin{bmatrix} -A_d^{-1}A_{d-1} & \cdots & -A_d^{-1}A_1 & -A_d^{-1}A_0 \\ I & & & \\ & \ddots & & \\ & & I & \end{bmatrix} \tag{1.1}$$

is the companion matrix associated with the monic matrix polynomial $A_d^{-1}A(\lambda)$ and $X \in \mathbb{F}^{n\times dn}$ is any matrix such that

$$\begin{bmatrix} XC_A^{d-1} \\ \vdots \\ XC_A \\ X \end{bmatrix} := T \in \mathbb{F}^{dn\times dn} \tag{1.2}$$

is nonsingular. The matrix T in (1.2) defines a structure-preserving similarity transformation for C_A in the sense that TC_AT^{-1} is the companion form of the monic matrix polynomial $\widetilde{A}(\lambda) = \lambda^d I + \lambda^{d-1}\widetilde{A}_{d-1} + \cdots + \widetilde{A}_0$ whose coefficient matrices can be read from the first block row of TC_AT^{-1} [9, Prop. 5]. This transformation is parametrized by the $n \times dn$ matrix X with the constraint that T in (1.2) is nonsingular. This class of SPTs is used in [12] and [17] to reduce matrix polynomials with nonsingular leading matrix coefficient to simpler forms such as Hessenberg, (quasi-)triangular, and (block-)diagonal forms. The computation of these simpler forms using the approach in [12] and [17] remains expensive since the construction of the parameter matrix X defining T in (1.2) requires the $dn \times dn$ transformation matrix reducing the companion form C_A to simpler form.

The SPT defined by T in (1.2) does not preserve additional properties of $A(\lambda)$, such as symmetry. To address this issue and still under the assumption that the leading coefficient matrix A_d is nonsingular, Lancaster and Prells [9] use standard triples (X, C_A, Y) with X and C_A as in (1.1)–(1.2), and $Y \in \mathbb{F}^{dn\times n}$ such that

$$\det(XC_A^{d-1}Y) \neq 0, \qquad XC_A^kY = 0, \quad k = 0\ldots, d-2,$$

to construct SPTs that preserve the block structure of the pencil

$$\lambda M_d(A) - M_{d-1}(A), \tag{1.3}$$

with

$$M_d(A) := \begin{bmatrix} & & & & A_d \\ & & & \iddots & A_{d-1} \\ & & & \iddots & \vdots \\ & \iddots & \iddots & & A_2 \\ A_d & A_{d-1} & \cdots & A_2 & A_1 \end{bmatrix}, \quad M_{d-1}(A) := \begin{bmatrix} & & & A_d & \\ & & \iddots & A_{d-1} & \\ & \iddots & \iddots & \vdots & \\ A_d & A_{d-1} & \cdots & A_2 & \\ & & & & -A_0 \end{bmatrix}.$$

The pencil in (1.3) is a linearization of $A(\lambda)$ in the sense that $\lambda M_d(A) - M_{d-1}(A)$ is equivalent to the block diagonal matrix polynomial $A(\lambda) \oplus I_{n(d-1)}$. Lancaster and Prells show that the block structure of this linearization is preserved by a pair of left and right transformations (T_L, T_R) parametrized by X and Y and taking the form

$$T_L = \begin{bmatrix} C_A^{d-1} Y & \dots & C_A Y & Y \end{bmatrix}^{-1} M_d(A)^{-1}, \qquad T_R = T^{-1} \tag{1.4}$$

with T as in (1.2) (see [9, Thm. 7]). Moreover, if the matrix coefficients of $A(\lambda)$ are Hermitian, then the matrix Y in a standard triple (X, C_A, Y) for $A(\lambda)$ has the form $Y = M_d(A)^{-1} X^*$ [4], [5] leading to $T_L = T_R^*$ in (1.4). As a result, the SPT (T_R^*, T_R) preserves the block structure of the linearization (1.3) and the Hermitian property of the blocks.

Our interest is in SPTs that preserve the block structure of any pencil in the vector space of pencils $\mathbb{DL}(A)$ defined in [10] by

$$\begin{aligned} \mathbb{DL}(A) := \{L_v(\lambda) \in \mathbb{F}[\lambda]^{dn \times dn} :\ & L_v(\lambda) \text{ is linear in } \lambda, \text{ and} \\ & L_v(\lambda)(\Lambda \otimes I_n) = v \otimes A(\lambda), & \text{(1.5a)} \\ & (\Lambda^T \otimes I_n) L_v(\lambda) = v^T \otimes A(\lambda), & \text{(1.5b)} \\ & v \in \mathbb{F}^d \setminus \{0\}\}, \end{aligned}$$

where $\Lambda = [\lambda^{d-1}, \lambda^{d-2}, \dots, 1]^T \in \mathbb{C}^d$. Note that the pencil in (1.3) belongs to $\mathbb{DL}(A)$ with vector $v = e_d$. Here and throughout the paper, e_j denotes the jth column of the $d \times d$ identity matrix. It is shown in [7] and [10] that $\mathbb{DL}(A)$ has dimension d, that pencils in $\mathbb{DL}(A)$ have a block symmetric structure, and that for a given vector $v \in \mathbb{F}^d$ there exists a unique $L_v(\lambda) \in \mathbb{DL}(A)$. For example, for $d = 2$, any pencil in $\mathbb{DL}(A)$ is a linear combination of the two pencils

$$L_{e_1}(\lambda) := \lambda \begin{bmatrix} A_2 & 0 \\ 0 & -A_0 \end{bmatrix} + \begin{bmatrix} A_1 & A_0 \\ A_0 & 0 \end{bmatrix} =: \lambda M_1(A) - M_0(A), \tag{1.6}$$

$$L_{e_2}(\lambda) := \lambda \begin{bmatrix} 0 & A_2 \\ A_2 & A_1 \end{bmatrix} + \begin{bmatrix} -A_2 & 0 \\ 0 & A_0 \end{bmatrix} =: \lambda M_2(A) - M_1(A), \tag{1.7}$$

which we refer to as the standard basis pencils for $\mathbb{DL}(A)$, since they correspond to $v = e_1$ and $v = e_2$ in (1.5). The special block-symmetric pencils in (1.6) and (1.7) are frequently used in applications, in particular when the coefficient matrices A_i of $A(\lambda)$ are symmetric or Hermitian.

One of our main contributions is a parametrization of the set of nonsingular matrices $T_L, T_R \in \mathbb{F}^{dn\times dn}$ preserving the block structure of any pencils in $\mathbb{DL}(A)$ and thereby constructing a new matrix polynomial $\widetilde{A}(\lambda)$ that is still of degree d. Because almost all pencils in $\mathbb{DL}(A)$ preserve the finite and infinite elementary divisors of $A(\lambda)$ [10], the matrix polynomials $A(\lambda)$ and $\widetilde{A}(\lambda)$ are isospectral (i.e., they have the same finite and infinite eigenvalues, including partial multiplicities) and hence they are equivalent.

Unlike Lancaster and Prells' SPTs in (1.4), our parametrization of SPTs preserving the block structure of pencils in $\mathbb{DL}(A)$ does not rely on standard triples nor on linearizations of $A(\lambda)$, and most importantly, our parametrization does not require the leading coefficient A_d of $A(\lambda)$ to be nonsingular. For the special case $d = 2$ (the quadratic case), our parametrization of (T_L, T_R) takes the form

$$T_L = \begin{bmatrix} F_L + \frac{1}{2}A_1G_L & -A_2G_L \\ A_0G_L & F_L - \frac{1}{2}A_1G_L \end{bmatrix}^{-1}, \quad T_R = \begin{bmatrix} F_R + \frac{1}{2}G_RA_1 & G_RA_0 \\ -G_RA_2 & F_R - \frac{1}{2}G_RA_1 \end{bmatrix}^{-1} \tag{1.8}$$

for some matrices $F_L, F_R, G_L, G_R \in \mathbb{F}^{n\times n}$ such that T_L, T_R are nonsingular and $G_RF_L + F_RG_L = 0$. When A_2 is nonsingular, the parametrizations in (1.4) and (1.8) are related by

$$X = \begin{bmatrix} G_RA_2 & F_R + \frac{1}{2}G_RA_1 \end{bmatrix}, \qquad Y = \begin{bmatrix} A_2^{-1}(F_L - \frac{1}{2}A_1G_L) \\ G_L \end{bmatrix}.$$

Our parametrization of the SPTs preserving the structure of $\mathbb{DL}(A)$ allows for easy construction of SPTs that also preserve extra structures in $A(\lambda)$ such as (skew-)Hermitian, (skew-)symmetric, $*$-even, $*$-odd, and $*$-(anti)palindromic structures. For example for real symmetric quadratic matrix polynomials $A(\lambda)$, choosing $F_L = F_R^T \in \mathbb{R}^{n\times n}$ and $G_L = G_R^T \in \mathbb{R}^{n\times n}$ in (1.8) together with the parameter constraint $G_RF_L + F_RG_L = 0$ yield SPTs that transform $A(\lambda)$ into an equivalent real symmetric quadratic $\widetilde{A}(\lambda)$.

Finally, our parametrization allows for easy constructions of SPTs that are at most rank-d modifications of the identity matrix and that lead to an equivalent matrix polynomial $\widetilde{A}(\lambda)$ whose jth coefficient matrix $\widetilde{A}_j$ is simply a low rank modification of A_j, for $j = 0, \dots, d$. For example, when $d = 2$, choosing $F_L = F_R = I_n$, $G_R = ab^* = -G_L$ in (1.8) for any nonzero vectors $a, b \in \mathbb{C}^n$ implies that the parameter constraint $G_RF_L + F_RG_L = 0$ holds. This leads to

$$T_L^{-1} = I_{2n} + \begin{bmatrix} -\frac{1}{2}A_1ab^* & A_2ab^* \\ -A_0ab^* & \frac{1}{2}A_1ab^* \end{bmatrix}, \qquad T_R^{-1} = I_{2n} + \begin{bmatrix} \frac{1}{2}ab^*A_1 & ab^*A_0 \\ -ab^*A_2 & -\frac{1}{2}ab^*A_1 \end{bmatrix} \tag{1.9}$$

whose nonsingularity is ensured by choosing a, b such that

$$\det(T_L^{-1}) = \det(T_R^{-1}) = 1 - \frac{1}{4}(b^*A_1a)^2 + (b^*A_0a)(b^*A_2a) \neq 0. \tag{1.10}$$

It is not difficult to see that T_L^{-1} and T_R^{-1} are at most rank-2 modifications of I_{2n}, that their inverses can be written explicitly using the Sherman-Morrison-Woodbury formula, and when applied to (1.6) or (1.7), explicit expressions for $\widetilde{A}_j$, $j = 0, 1, 2$ can be obtained in terms of A_j, $j = 0, 1, 2$ and a, b. We expect SPTs that are low-rank modification of the identity matrix to be one of the key tools for developing algorithms that reduce a matrix polynomial to Hessenberg form or tridiagonal form in a finite number of steps without the need of a linear pencil of larger dimension.

The paper is organized as follows. Section 2 contains preliminary material that is necessary for the parametrization of the SPTs presented in section 3. Special attention is given to the quadratic case. Section 4 describes how to select the parameters so as to preserve symmetries in the matrix polynomial. Section 5 contains concluding remarks. Examples are used throughout the paper to illustrate the results.

2. Preliminaries

We summarize the properties of the vector space $\mathbb{DL}(A)$ in (1.5) that we need for the construction of the parametrization of the transformations that preserve the block structure of the pencils in $\mathbb{DL}(A)$.

The one-sided factorizations (1.5a)–(1.5b) in the definition of $\mathbb{DL}(A)$ lead to some interesting relations between the solutions of linear systems involving $A(\lambda)$ and $L_v(\lambda) \in \mathbb{DL}(A)$ with vector v.

Theorem 1. *Let $A(\lambda) \in \mathbb{F}[\lambda]^{n\times n}$ be a matrix polynomial of degree d and let $L_v(\lambda) \in \mathbb{DL}(A)$ with vector $v \in \mathbb{F}^d$. Let $\omega \in \mathbb{F}$ be such that the matrix $A(\omega)$ is nonsingular and define $\Omega = [\omega^{d-1}, \omega^{d-2}, \dots, 1]^T \in \mathbb{F}^d$.*

(a) If $x \in \mathbb{F}^n$ is the solution of the linear system $A(\omega)x = b$ for some given $b \in \mathbb{F}^n$ then $z = \Omega \otimes x \in \mathbb{F}^{dn}$ is a solution of $L_v(\omega)z = v \otimes b$.

(b) If $y \in \mathbb{F}^n$ is the solution of the linear system $y^T A(\omega) = c^T$ for some given $c \in \mathbb{F}^n$ then $w = \Omega \otimes y \in \mathbb{F}^{dn}$ is a solution of $w^T L_v(\omega) = (v \otimes b)^T$.

(c) If $z \in \mathbb{F}^{dn}$ solves the linear system $L_v(\omega)z = c$ for some given $c \in \mathbb{F}^{dn}$ then $x = (v^T \otimes I_n)z \in \mathbb{F}^n$ solves $A(\omega)x = (\Omega^T \otimes I_n)c$.

Proof. Follows from [6, Corollary 4.2]. □

As already mentioned in the introduction, $\mathbb{DL}(A)$ is a vector space of dimension d. It is shown in [7, Thm. 3.5] that

$$\mathbb{DL}(A) = \Big\{ \sum_{k=1}^{d} v_k L_{e_k}(\lambda) : v \in \mathbb{F}^d,\ L_{e_k}(\lambda) = \lambda M_k(A) - M_{k-1}(A),\ 1 \le k \le d \Big\}, \tag{2.1}$$

where

$$M_k(A) = \begin{bmatrix} \mathcal{L}_k(A) & 0 \\ 0 & -\mathcal{U}_{d-k}(A) \end{bmatrix} \in \mathbb{F}^{dn\times dn}, \qquad 0 \le k \le d \tag{2.2}$$

with block Hankel matrices $\mathcal{L}_j(A), \mathcal{U}_j(A) \in \mathbb{F}^{jn\times jn}$ given by

$$\mathcal{L}_j(A) := \begin{bmatrix} & & & A_d \\ & & \iddots & A_{d-1} \\ & \iddots & \iddots & \vdots \\ A_d & A_{d-1} & \dots & A_{d-j+1} \end{bmatrix}, \qquad \mathcal{U}_j(A) := \begin{bmatrix} A_{j-1} & \dots & A_1 & A_0 \\ \vdots & \iddots & \iddots & \\ A_1 & \iddots & & \\ A_0 & & & \end{bmatrix}. \tag{2.3}$$

The pencils $L_{e_k}(\lambda)$, $k = 1, \ldots, d$ in (2.1) form the standard basis pencils for $\mathbb{DL}(A)$ and correspond to choosing $v = e_k$, $k = 1, \ldots, d$ in (1.5). All the pencils in $\mathbb{DL}(A)$ have a block symmetric structure—see for example, the two standard basis pencils for $\mathbb{DL}(A)$ provided in (1.6)–(1.7) when $d = 2$. The pencils L_{e_k} first appear in [8] for scalar polynomials. They are used in [3] to define the class of SPTs of interest in this paper.

3. SPTs preserving the block structure of $\mathbb{DL}(A)$

Our main objective is to parametrize the set of nonsingular matrices $T_L, T_R \in \mathbb{F}^{dn\times dn}$ such that the statement

$$L_v(\lambda) \in \mathbb{DL}(A) \textit{ with vector } v \textit{ if and only if } T_L L_v(\lambda) T_R \in \mathbb{DL}(\widetilde{A}) \textit{ with vector } v \quad (3.1)$$

holds for any $v \in \mathbb{F}^d$, where $\widetilde{A}(\lambda) \in \mathbb{F}[\lambda]^{n\times n}$ is a matrix polynomial of degree d. It follows from (2.1) that asserting that (3.1) holds for all $v \in \mathbb{F}^d$ is equivalent to asserting that (3.1) holds for $v = e_k$, $k = 1, \ldots, d$. In other words, we are looking for nonsingular matrices T_L, T_R that preserve the block structure of the standard basis pencils of $\mathbb{DL}(A)$.

Let $\widetilde{B}_L \in \mathbb{F}^{n\times n}$ be any nonsingular matrix and consider the linear system

$$\widetilde{A}(\omega)\widetilde{X} = \widetilde{B}_L,$$

with ω such that the matrix $\widetilde{A}(\omega)$ is nonsingular. Let us define

$$\widetilde{L}_{e_k}(\lambda) := T_L L_{e_k}(\lambda) T_R, \quad k = 1, \ldots, d.$$

Then by (3.1), $\widetilde{L}_{e_k}(\lambda) \in \mathbb{DL}(\widetilde{A})$ with vector e_k. With the notation

$$\Omega = [\omega^{d-1}, \omega^{d-2}, \ldots, 1]^T \in \mathbb{F}^d$$

it follows from Theorem 1(a) that $\Omega \otimes \widetilde{X} \in \mathbb{F}^{dn\times n}$ is a solution of the d linear systems

$$\widetilde{L}_{e_k}(\omega)(\Omega \otimes \widetilde{X}) = e_k \otimes \widetilde{B}_L, \quad k = 1, \ldots d,$$

or, equivalently, that

$$Z := T_R(\Omega \otimes \widetilde{X}) \in \mathbb{F}^{dn\times n} \quad (3.2)$$

solves

$$L_{e_k}(\omega)Z = T_L^{-1}(e_k \otimes \widetilde{B}_L), \quad k = 1, \ldots d.$$

Hence since $\widetilde{B}_L$ is nonsingular,

$$T_L^{-1} = \begin{bmatrix} L_{e_1}(\omega)Z & L_{e_2}(\omega)Z & \cdots & L_{e_d}(\omega)Z \end{bmatrix} (I_d \otimes \widetilde{B}_L^{-1}). \quad (3.3)$$

In a similar way, if we consider the linear system

$$\widetilde{Y}^T \widetilde{A}(\omega) = \widetilde{B}_R$$

with $\widetilde{B}_R \in \mathbb{F}^{n\times n}$ nonsingular then it follows from Theorem 1(b) that

$$(\Omega \otimes \widetilde{Y})^T \widetilde{L}_{e_k}(\omega) = e_k^T \otimes \widetilde{B}_R, \quad k = 1, \ldots, d,$$

or equivalently, that

$$W^T = (\Omega \otimes \widetilde{Y})^T T_L \in \mathbb{F}^{n\times dn} \tag{3.4}$$

solves the d linear systems

$$W^T L_{e_k}(\omega) = (e_k^T \otimes \widetilde{B}_R) T_R^{-1}, \quad k = 1, \ldots, d.$$

Since $\widetilde{B}_R$ is nonsingular,

$$T_R^{-1} = (I_d \otimes \widetilde{B}_R^{-1}) \begin{bmatrix} W^T L_{e_1}(\omega) \\ W^T L_{e_2}(\omega) \\ \vdots \\ W^T L_{e_d}(\omega) \end{bmatrix}. \tag{3.5}$$

We show in what follows that the blocks $L_{e_k}(\omega)Z$ and $W^T L_{e_k}(\omega)$ in (3.3) and (3.5) have a special structure.

Lemma 2. *Let $A(\lambda) = \sum_{j=0}^{d} \lambda^j A_j \in \mathbb{F}[\lambda]^{n\times n}$, $L_{e_k}(\lambda)$ be the kth standard basis pencil of $\mathbb{DL}(A)$, and $Z \in \mathbb{F}^{dn\times n}$ be partitioned into $n\times n$ blocks Z_j according to $Z_j := (e_j^T \otimes I_n)Z$. Then*

$$L_{e_k}(\lambda)Z = e_k \otimes f_A(Z,\lambda) + \sum_{\ell=1}^{d-1} P_\ell(A)(e_k \otimes (Z_\ell - \lambda Z_{\ell+1})), \tag{3.6}$$

where

$$f_A(Z,\lambda) = A_0 Z_d + \lambda A_d Z_1 + \frac{1}{2}\sum_{\ell=1}^{d-1} A_{d-\ell}(Z_\ell + \lambda Z_{\ell+1}), \tag{3.7}$$

and

$$(P_\ell(A))_{ij} = \begin{cases} \frac{1}{2}A_{d-\ell} & \text{if } i = j \le \ell, \\ -\frac{1}{2}A_{d-\ell} & \text{if } i = j > \ell, \\ A_{d-\ell-i+j} & \text{if } i > j,\ j \le \ell,\ \ell + i - j \le d, \\ -A_{d-\ell-i+j} & \text{if } i < j,\ j > \ell,\ \ell + i - j \ge 0, \\ 0 & \text{otherwise.} \end{cases} \tag{3.8}$$

Proof. The proof is not difficult but readers may find helpful to go through it with $d = 2$ using $P_1(A)$ in (3.20) or with $d = 3$ using $P_1(A)$ and $P_2(A)$ in (3.29).

Recall from (2.1) that $L_{e_k}(\lambda) = \lambda M_k(A) - M_{k-1}(A)$, where $M_k(A)$ is defined in (2.2)–(2.3), and let us partitioned $L_{e_k}(\lambda)Z$ into d $n\times n$ blocks $(L_{e_k}(\lambda)Z)_i$. Then

$$(L_{e_k}(\lambda)Z)_i = \begin{cases} \displaystyle\sum_{\ell=k-i}^{k-1} -A_{d+k-i-\ell}(Z_\ell - \lambda Z_{\ell+1}) & \text{if } i < k, \\ \displaystyle f_A(Z,\lambda) + \sum_{\ell=1}^{d-1} \frac{1}{2} A_{d-\ell}(Z_\ell - \lambda Z_{\ell+1}) & \text{if } i = k, \\ \displaystyle\sum_{\ell=k}^{d+k-i} A_{d+k-i-\ell}(Z_\ell - \lambda Z_{\ell+1}) & \text{if } i > k. \end{cases}$$

The expression for $L_{e_k}(\lambda)Z$ in (3.6) follows. $\square$

Since the pencils in $\mathbb{DL}(A)$ are block symmetric, $L_{e_k}^T(\lambda)$ is the kth standard basis pencil of $\mathbb{DL}(A^T)$ and it follows from Lemma 2 that for $W \in \mathbb{F}^{dn\times n}$,

$$W^T L_{e_k}(\lambda) = e_k^T \otimes (f_{A^T}(W,\lambda))^T + \sum_{\ell=1}^{d-1} (e_k^T \otimes (W_\ell - \lambda W_{\ell+1})^T)(P_\ell(A^T))^T. \tag{3.9}$$

We are now ready to state the main result of this section.

Theorem 3. *Let $A(\lambda) \in \mathbb{F}[\lambda]^{n\times n}$ be a matrix polynomial of degree d. The nonsingular matrices $T_L, T_R \in \mathbb{F}^{dn\times dn}$ that preserve the block structure of any pencil in $\mathbb{DL}(A)$ have the form*

$$T_L = \Big(I_d \otimes F_L + \sum_{\ell=1}^{d-1} P_\ell(A)(I_d \otimes G_{L\ell})\Big)^{-1}, \tag{3.10}$$

$$T_R = \Big(I_d \otimes F_R + \sum_{\ell=1}^{d-1} (I_d \otimes G_{R\ell})P_\ell(A^T)^T\Big)^{-1}, \tag{3.11}$$

with $P_\ell(A)$ as in (3.8) *and parameter matrices F_L, F_R, $G_{L\ell}$, $G_{R\ell} \in \mathbb{F}^{n\times n}$, $\ell = 1,\dots,d-1$ such that*

(i) *T_L and T_R are nonsingular, and*
(ii) *$F_R G_{Lk} + G_{Rk} F_L = G_R H_k(A) G_L$, $k = 1,\dots,d-1$, where*

$$G_R := \begin{bmatrix} G_{R1} & \cdots & G_{R(d-1)} \end{bmatrix}, \quad G_L := \begin{bmatrix} G_{L1} \\ \vdots \\ G_{L(d-1)} \end{bmatrix},$$

and $H_k(A)$ is a block-symmetric $(d-1)\times(d-1)$ matrix with $n\times n$ blocks defined by

$$H_k(A)_{ij} = \begin{cases} -A_{d-i-j+k} & \text{if } k < j,\ k < i \text{ and } i+j-k \le d, \\ A_{d-i-j+k} & \text{if } k > j,\ k > i \text{ and } i+j-k \ge 0, \\ -\frac{1}{2}A_{d-j} & \text{if } k = i \text{ and } k < j, \\ -\frac{1}{2}A_{d-i} & \text{if } k = j \text{ and } k < i, \\ \frac{1}{2}A_{d-j} & \text{if } k = i \text{ and } k > j, \\ \frac{1}{2}A_{d-i} & \text{if } k = j \text{ and } k > i, \\ 0 & \text{otherwise}. \end{cases}$$

Proof. Readers may find helpful to check the proof with $d = 3$ using $P_1(A)$ and $P_2(A)$ in (3.29), and $H_1(A) = \begin{bmatrix} 0 & -\frac{1}{2}A_1 \\ -\frac{1}{2}A_1 & -A_0 \end{bmatrix}$, $H_2(A) = \begin{bmatrix} A_3 & \frac{1}{2}A_2 \\ \frac{1}{2}A_2 & 0 \end{bmatrix}$.

If we let

$$F_L = f_A(Z,\omega)\widetilde{B}_L^{-1}, \qquad G_{L,\ell} = (Z_\ell - \omega Z_{\ell+1})\widetilde{B}_L^{-1}, \quad \ell = 1,\dots,d-1, \tag{3.12}$$

$$F_R = \widetilde{B}_R^{-1} f_{A^T}(W,\omega)^T, \qquad G_{R,\ell} = \widetilde{B}_R^{-1}(W_\ell - \omega W_{\ell+1})^T, \quad \ell = 1,\dots,d-1 \tag{3.13}$$

then the expression for T_L^{-1} in the theorem follows from (3.3) and (3.6), and that for T_R^{-1} follows from (3.5) and (3.9).

The matrices F_L, F_R, $G_{L,\ell}$, and $G_{R,\ell}$, $\ell = 1, \ldots, d-1$ depend on Z in (3.2) and W in (3.4), i.e., $Z = T_R(\Omega \otimes \widetilde{X})$ and $W^T = (\Omega \otimes \widetilde{Y})^T T_L$. It is easy to see that these two matrix equations are equivalent to

$$((e_k - \omega e_{k+1})^T \otimes I_n) T_R^{-1} Z = 0, \quad k = 1, \ldots, d-1, \qquad (e_d^T \otimes I_n) T_R^{-1} Z = \widetilde{X}, \tag{3.14}$$

$$W^T T_L^{-1}((e_k - \omega e_{k+1}) \otimes I_n) = 0, \quad k = 1, \ldots, d-1, \qquad W T_L^{-1}(e_d \otimes I_n) = \widetilde{Y}. \tag{3.15}$$

The matrix equations on the right hand-side of (3.14) and of (3.15) do not impose any constraint on the parameter matrices F_L, F_R, $G_{L,\ell}$, and $G_{R,\ell}$ since $\widetilde{X}$ and $\widetilde{Y}$ are free to choose. On the other hand, the first $d-1$ matrix equations in (3.14) and the first $d-1$ matrix equations in (3.15) do impose constraints as we now show.

We start with the $d-1$ first equations in (3.14). It follows from (3.12) that

$$Z_\ell = \sum_{j=1}^{d-\ell} \omega^{j-1} G_{L(\ell+j-1)} \widetilde{B}_L + \omega^{d-\ell} Z_d, \quad 1 \le \ell \le d.$$

Then on using the parametrization of T_R in (3.11), we find that $((e_k - \omega e_{k+1})^T \otimes I_n) T_R^{-1} Z = 0$ is equivalent to

$$F_R G_{Lk} \widetilde{B}_L + G_{Rk} A(\omega) Z_d + G_{Rk}\, h(\omega, A, G_L) \widetilde{B}_L + \sum_{i=1}^{d-1} \sum_{j=1}^{d-1} G_{Ri} P_i(A)_{jk} G_{Lj} \widetilde{B}_L = 0, \tag{3.16}$$

where

$$h(\omega, A, G_L) = \sum_{i=1}^{d-1} \sum_{j=1}^{d-j} \omega^j A_{d-i+1} G_{L(i+j-1)}.$$

Since Z solves $L_d(\omega) Z = C$ with $C = T_L^{-1}(e_d \otimes \widetilde{B}_L)$, it follows from Theorem 1(c) that $X = (e_d^T \otimes I) Z = Z_d$ solves $A(\omega) X = (\Omega^T \otimes I) C$, i.e.,

$$\begin{aligned} A(\omega) Z_d &= (\Omega^T \otimes I) T_L^{-1}(e_d \otimes \widetilde{B}_L) \\ &= -h(\omega, A, G_L) \widetilde{B}_L + \sum_{j=1}^{d-1} P_j(A)_{dd} G_{Lj} \widetilde{B}_L + F_L \widetilde{B}_L. \end{aligned} \tag{3.17}$$

Replacing $A(\omega) Z_d$ in(3.16) by (3.17) and using the fact that $\widetilde{B}_L$ is nonsingular yields

$$\begin{aligned} F_R G_{Lk} + G_{Rk} F_L &= -\sum_{i=1}^{d-1} \sum_{j=1}^{d-1} G_{Ri} P_i(A)_{jk} G_{Lj} - \sum_{j=1}^{d-1} G_{Rk} P_j(A)_{dd} G_{Lj} \\ &= \sum_{i=1}^{d-1} \sum_{j=1}^{d-1} G_{Ri} H_k(A)_{ij} G_{Lj} \\ &= G_R H_k(A) G_L \end{aligned}$$

where

$$H_k(A)_{ij} = \begin{cases} -P_i(A)_{jk} & \text{if } i \neq k, \\ -P_k(A)_{jk} - P_j(A)_{dd} & \text{if } i = k. \end{cases}$$

Then, the expression for $H_k(A)_{ij}$ in the theorem follows from (3.8) and it is not difficult to check that $H_k(A)$ is block-symmetric, i.e., that $(H_k(A))_{ij} = (H_k(A))_{ji}$. □

Setting $G_{L\ell} = G_{L\ell} = 0$ in (3.10)–(3.11) leads to SPTs (T_L, T_R) that transform $A(\lambda)$ into a strictly equivalent matrix polynomial $\widetilde{A}(\lambda) = F_L A(\lambda) F_R$. We refer to such SPTs as trivial SPTs.

3.1. Quadratic case

When $d = 2$, the constraint equation in Theorem 3(ii) does not depend on $A(\lambda)$.

Corollary 4. *The pair of matrices $T_L, T_R \in \mathbb{F}^{2n\times 2n}$ that preserve the block structure of any pencil in $\mathbb{DL}(A)$ when $A(\lambda) = \lambda^2 A_2 + \lambda A_1 + A_0 \in \mathbb{F}[\lambda]^{n\times n}$ have the form*

$$T_L^{-1} = I_2 \otimes F_L + \begin{bmatrix} \frac{1}{2}A_1 G_L & -A_2 G_L \\ A_0 G_L & -\frac{1}{2}A_1 G_L \end{bmatrix}, \quad T_R^{-1} = I_2 \otimes F_R + \begin{bmatrix} \frac{1}{2}G_R A_1 & G_R A_0 \\ -G_R A_2 & -\frac{1}{2}G_R A_1 \end{bmatrix} \tag{3.18}$$

where F_L, F_R, G_L, G_R are such that T_L^{-1} and T_R^{-1} are nonsingular and

$$F_R G_L + G_R F_L = 0. \tag{3.19}$$

Proof. We have from (3.8) that

$$P_1(A) = \begin{bmatrix} \frac{1}{2}A_1 & -A_2 \\ A_0 & -\frac{1}{2}A_1 \end{bmatrix}, \qquad P_1(A^T)^T = \begin{bmatrix} \frac{1}{2}A_1 & A_0 \\ -A_2 & -\frac{1}{2}A_1 \end{bmatrix}. \tag{3.20}$$

The expressions for T_L^{-1} and T_R^{-1} in (3.18) are then a simple application of Theorem 3 with $d = 2$. We also have that $H_1(A) = -(P_1(A^T)^T)_{11} + \frac{1}{2}A_1 = 0$ so that (3.19) follows from Theorem 3(ii) by setting $G_{L1} \equiv G_L$ and $G_{R1} \equiv G_R$. □

Example 5. *Consider the 2×2 quadratic*

$$A(\lambda) = \lambda^2 \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \lambda \begin{bmatrix} 1 & 2 \\ 2 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ -1 & 3 \end{bmatrix}$$

and construct the SPT (T_L, T_R) parametrized by

$$F_L = \begin{bmatrix} -1 & -1 \\ 0 & 1 \end{bmatrix}, \quad G_R = \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix}, \quad F_R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

and

$$G_L = -F_R^{-1} G_R F_L = \begin{bmatrix} 0 & -1 \\ 1 & 2 \end{bmatrix}$$

so that the constraint (3.19) *holds. Then for* $L_{e_2}(\lambda)$ *in* (1.7)*, we find that*

$$T_L L_{e_2}(\lambda) T_R = \lambda \begin{bmatrix} 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & 0 \\ 5 & 4 & 16 & 4 \\ 0 & 0 & -8 & -4 \end{bmatrix} - \begin{bmatrix} 5 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -11 & -4 \\ 0 & 0 & 12 & 4 \end{bmatrix}$$
$$= \lambda \begin{bmatrix} 0 & \widetilde{A}_2 \\ \widetilde{A}_2 & \widetilde{A}_1 \end{bmatrix} - \begin{bmatrix} \widetilde{A}_2 & 0 \\ 0 & -\widetilde{A}_0 \end{bmatrix} = \widetilde{L}_{e_2}(\lambda).$$

So the SPT (T_L, T_R) *transforms* $A(\lambda)$ *into the equivalent quadratic matrix polynomial*

$$\widetilde{A}(\lambda) = \lambda^2 \widetilde{A}_2 + \lambda \widetilde{A}_1 + \widetilde{A}_0 = \lambda^2 \begin{bmatrix} 5 & 4 \\ 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} 16 & 4 \\ -8 & -4 \end{bmatrix} + \begin{bmatrix} 11 & 4 \\ -12 & -4 \end{bmatrix}.$$

We remark here that to construct $\widetilde{A}(\lambda)$*, we used the pencil* $L_{e_2}(\lambda)$*, which is not a linearization of* $A(\lambda)$ *since* A_2 *in this example is singular.*

We expect the next result to be useful when transforming quadratics to simpler forms.

Proposition 6. *Let* (T_L, T_R) *be the SPT parameterized by* F_L, G_L, F_R, G_R *as in* (3.18) *mapping the quadratic* $A(\lambda)$ *to the quadratic* $\widetilde{A}(\lambda)$*. Then* $(\widetilde{T}_L, \widetilde{T}_R) = (T_L^{-1}, T_R^{-1})$ *is an SPT mapping* $\widetilde{A}(\lambda)$ *to* $A(\lambda)$ *that is parameterized by the four matrices* $\widetilde{F}_L, \widetilde{G}_L, \widetilde{F}_R, \widetilde{G}_R$ *with* $\widetilde{G}_L$ *and* $\widetilde{G}_R$ *such that*

$$\widetilde{G}_L = G_R, \qquad \widetilde{G}_R = G_L.$$

Proof. It suffices to show that $\widetilde{G}_L = G_R$. Let us partition any $2n \times 2n$ matrix T into $n \times n$ blocks as $T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}$. It follows from (1.8) that

$$\begin{bmatrix} (T_L^{-1})_{21} \\ (T_L^{-1})_{11} - (T_L^{-1})_{22} \\ -(T_L^{-1})_{12} \end{bmatrix} = \begin{bmatrix} A_0 \\ A_1 \\ A_2 \end{bmatrix} G_L, \tag{3.21}$$

$$\widetilde{G}_R \begin{bmatrix} \widetilde{A}_0 & \widetilde{A}_1 & \widetilde{A}_2 \end{bmatrix} = \begin{bmatrix} (\widetilde{T}_R^{-1})_{12} & (\widetilde{T}_R^{-1})_{11} - (\widetilde{T}_R^{-1})_{22} & -(\widetilde{T}_R^{-1})_{21} \end{bmatrix}. \tag{3.22}$$

Since (T_L, T_R) preserves the block structure of the standard basis of $\mathbb{DL}(A)$, we have that $M_k(A)\widetilde{T}_R^{-1} = T_L^{-1} M_k(\widetilde{A})$, $k = 0, 1, 2$. We partition these $2n \times 2n$ matrix equalities into $n \times n$ blocks as above and denote by $(k)_{ij}$ the $n \times n$ matrix equality $(M_k(A)\widetilde{T}_R^{-1})_{ij} = (T_L^{-1} M_k(\widetilde{A}))_{ij}$. Then the following linear combination of these matrix equations

$$\begin{pmatrix} -(0)_{22} & -(0)_{21} + (1)_{22} & (1)_{21} \\ -(0)_{12} + (1)_{22} & -(0)_{11} + (1)_{21} - (2)_{22} + (1)_{12} & -(2)_{21} + (1)_{11} \\ (1)_{12} & -(2)_{12} + (1)_{11} & -(2)_{11} \end{pmatrix}$$

leads to

$$\begin{bmatrix} A_0 \\ A_1 \\ A_2 \end{bmatrix} \begin{bmatrix} (\widetilde{T}_R^{-1})_{12} & (\widetilde{T}_R^{-1})_{11} - (\widetilde{T}_R^{-1})_{22} & -(\widetilde{T}_R^{-1})_{21} \end{bmatrix} = \begin{bmatrix} (T_L^{-1})_{21} \\ (T_L^{-1})_{11} - (T_L^{-1})_{22} \\ -(T_L^{-1})_{12} \end{bmatrix} \begin{bmatrix} \widetilde{A}_0 & \widetilde{A}_1 & \widetilde{A}_2 \end{bmatrix}. \tag{3.23}$$

Since $A(\lambda)$ is a regular quadratic matrix polynomial, $\det(\lambda^2 A_2 + \lambda A_1 + A_0) \neq 0$ for some $\lambda \in \mathbb{F}$. As a result,

$$\begin{aligned} n = \operatorname{rank}(\lambda^2 A_2 + \lambda A_1 + A_0) &= \operatorname{rank}\left(\begin{bmatrix}\lambda^2 I & \lambda I & I\end{bmatrix}\begin{bmatrix}A_2\\A_1\\A_0\end{bmatrix}\right) \\ &\leq \min\left(\operatorname{rank}\left(\begin{bmatrix}\lambda^2 I & \lambda I & I\end{bmatrix}\right), \operatorname{rank}\left(\begin{bmatrix}A_2\\A_1\\A_0\end{bmatrix}\right)\right), \end{aligned}$$

which implies that $\begin{bmatrix}A_2\\A_1\\A_0\end{bmatrix}$ is of full column rank. Similarly, $\begin{bmatrix}\widetilde{A}_0 & \widetilde{A}_1 & \widetilde{A}_2\end{bmatrix}$ is of full row rank and we have that

$$\begin{bmatrix}A_0\\A_1\\A_2\end{bmatrix}^+ \begin{bmatrix}A_0\\A_1\\A_2\end{bmatrix} = I_n, \qquad \begin{bmatrix}\widetilde{A}_0 & \widetilde{A}_1 & \widetilde{A}_2\end{bmatrix}\begin{bmatrix}\widetilde{A}_0 & \widetilde{A}_1 & \widetilde{A}_2\end{bmatrix}^+ = I_n,$$

where B^+ denotes the pseudoinverse of B. Then it follows from (3.22), (3.23), and (3.21) that

$$\begin{aligned} \widetilde{G}_R &= \begin{bmatrix}(\widetilde{T}_R^{-1})_{21} & (\widetilde{T}_R^{-1})_{11} - (\widetilde{T}_R^{-1})_{22} & -(\widetilde{T}_R^{-1})_{12}\end{bmatrix}\begin{bmatrix}\widetilde{A}_0 & \widetilde{A}_1 & \widetilde{A}_2\end{bmatrix}^+ \\ &= \begin{bmatrix}A_0\\A_1\\A_2\end{bmatrix}^+ \begin{bmatrix}(T_L^{-1})_{21}\\(T_L^{-1})_{11} - (T_L^{-1})_{22}\\-(T_L^{-1})_{12}\end{bmatrix} \\ &= G_L \qquad \square \end{aligned}$$

3.2. At most rank-d modification of the identity matrix SPTs

Our parametrization in Theorem 3 is for T_L^{-1} and T_R^{-1}. Fortunately, the parametrization allows easy constructions of nontrivial SPTs that are low rank modification of I_{dn} so that they can be inverted efficiently or even explicitly when the update is of very low rank. We illustrated this when $d = 2$ in the introduction with T_L, T_R in (1.9). We refer to [2] for early work on this topic for quadratics with nonsingular leading coefficients and a more complicated set of constraints on the parameters than (3.19). Such SPT are used in [15] to deflate eigenpairs from quadratic matrix polynomials.

We show in this section that for arbitrary degree d we can construct SPTs that are at most rank-d modification of the $dn \times dn$ identity matrix. A description of all the SPTs that are low rank modification of I_{dn} is outside the scope of this work.

For given nonzero vectors $a, b \in \mathbb{C}^n$, consider T_L^{-1} in (3.10) and T_R^{-1} in (3.11) with

$$F_L = F_R = I_n, \qquad G_{R\ell} = ab^*, \qquad G_{L\ell} = g_\ell ab^*, \quad \ell = 1, \ldots, d-1, \tag{3.24}$$

where the scalars g_ℓ are to be determined so that the constraints in Theorem 3(ii) hold. These constraints simplify to

$$\sum_{j=1}^{d-1} g_j b^* \left(\sum_{i=1}^{d-1} H_k(A)_{ij} \right) a - g_k = 1, \qquad k = 1, \ldots, d-1,$$

which can be rewritten as the linear system $Cg = e$ with g the vector of scalars g_ℓ, e the vector of all ones, and

$$C_{kj} = \begin{cases} \sum_{i=1}^{d-1} b^* H_k(A)_{ij} a & \text{if } k \neq j, \\ \sum_{i=1}^{d-1} b^* H_k(A)_{ij} a - 1 & \text{if } k = j. \end{cases} \tag{3.25}$$

With the choice (3.24), the inverse of T_L and T_R can be rewritten as

$$T_L^{-1} = I_{dn} + \left(\sum_{\ell=1}^{d-1} g_\ell P_\ell(A)(I_d \otimes a) \right)(I_d \otimes b^*) =: I_{dn} + U_L V_L^*, \tag{3.26}$$

$$T_R^{-1} = I_{dn} + (I_d \otimes a)\left(\sum_{\ell=1}^{d-1} (I_d \otimes b^*) P_\ell(A^*)^* \right) =: I_{dn} + U_R V_R^*, \tag{3.27}$$

where $V_L = I_d \otimes b$ and $U_R := I_d \otimes a$ are $dn \times d$ matrices of rank d, and $U_L := \sum_{\ell=1}^{d-1} g_\ell P_\ell(A)(I_d \otimes a)$ and $V_R := \sum_{\ell=1}^{d-1} P_\ell(A^*)(I_d \otimes b)$ are $dn \times d$ matrices of rank at most d. It follows from Theorem 3 that as long as the nonzero vectors a and b are chosen such that $\det(I_d + V_L^* U_L) \neq 0$ and $\det(I_d + V_R^* U_R) \neq 0$, and the linear system $Cg = e$ with C as in (3.25) has a solution, then

$$T_L = I_{dn} - U_L(I_d + V_L^* U_L)^{-1} V_L^*, \qquad T_R = I_{dn} - U_R(I_d + V_R^* U_R)^{-1} V_R^*$$

defines an SPT for pencils in $\mathbb{DL}(A)$ that is made up of two at most rank-d modifications of the identity matrix.

Example 7. *As an illustration, consider the cubic matrix polynomial $A(\lambda) = \lambda^3 A_3 + \lambda^2 A_2 + \lambda A_1 + A_0$. For any nonzero vectors a, b such that*

(a) the linear system

$$\begin{bmatrix} -\frac{1}{2} b^* A_1 a - 1 & -\frac{1}{2} b^* A_1 a - b^* A_0 a \\ \frac{1}{2} b^* A_2 a + b^* A_3 a & \frac{1}{2} b^* A_2 a - 1 \end{bmatrix} \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{3.28}$$

has a solution $\left[\begin{smallmatrix} g_1 \\ g_2 \end{smallmatrix}\right]$, and

(b) $\det(I_3 + V_L^* U_L) \neq 0$ *and* $\det(I_3 + V_R^* U_R) \neq 0$ *with*

$$V_L = I_3 \otimes b, \qquad U_L = (g_1 P_1(A) + g_2 P_2(A))(I_3 \otimes a),$$
$$U_R = I_d \otimes a, \qquad V_R = (P_1(A^*) + P_2(A^*))(I_3 \otimes b),$$

where the matrices $P_\ell(A)$ are given by (see (3.8)*)*

$$P_1(A) = \begin{bmatrix} \frac{1}{2} A_2 & -A_3 & 0 \\ A_1 & -\frac{1}{2} A_2 & -A_3 \\ A_0 & 0 & -\frac{1}{2} A_2 \end{bmatrix}, \qquad P_2(A) = \begin{bmatrix} \frac{1}{2} A_1 & 0 & -A_3 \\ A_0 & \frac{1}{2} A_1 & -A_2 \\ 0 & A_0 & -\frac{1}{2} A_1 \end{bmatrix} \tag{3.29}$$

then $T_L = I_{3n} - U_L(I_3 + V_L^* U_L)^{-1} V_L^*$, $T_R = I_{3n} - U_R(I_3 + V_R^* U_R)^{-1} V_R^*$ *define an SPT for any pencil in* $\mathbb{DL}(A)$.

Now for

$$A_3 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} -6 & 0 \\ 0 & 1 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 11 & 0 \\ 0 & 3 \end{bmatrix}, \quad A_0 = \begin{bmatrix} -6 & 0 \\ 0 & 2 \end{bmatrix}$$

and $a = \left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]$, $b = \left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$, *we find that the linear system* (3.28) *has solution* $g = \left[\begin{smallmatrix} -5 \\ 1 \end{smallmatrix}\right]$ *leading to*

$$U_L = \begin{bmatrix} 20.5 & 5 & -1 \\ 1 & 0 & 0 \\ -61 & -9.5 & 11 \\ 13 & -4 & 1 \\ 30 & -6 & -20.5 \\ 10 & -2 & -1 \end{bmatrix}, \qquad V_R = \begin{bmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 0 \\ 5 & 1 & -1 \\ 0 & 0 & 0 \\ 2 & 2 & -2 \end{bmatrix}.$$

Then it is easy to check that the two matrices

$$V_L^* U_L = \begin{bmatrix} 1 & 0 & 0 \\ 13 & -4 & 1 \\ 10 & -2 & -1 \end{bmatrix}. \qquad V_R^* U_R = \begin{bmatrix} -2 & -5 & -2 \\ 0 & -1 & -2 \\ 0 & 1 & 2 \end{bmatrix}$$

do not have -1 *as an eigenvalue so that* $T_L = I_{3n} - U_L(I_3 + V_L^* U_L)^{-1} V_L^*$ *and* $T_R = I_{3n} - U_R(I_3 + V_R^* U_R)^{-1} V_R^*$ *form an SPT, which when applied to, for example,*

$$\lambda M_1(A) + M_0(A) = \lambda \begin{bmatrix} A_3 & 0 & 0 \\ 0 & -A_1 & -A_0 \\ 0 & -A_0 & 0 \end{bmatrix} + \begin{bmatrix} A_2 & A_1 & A_0 \\ A_1 & A_0 & 0 \\ A_0 & 0 & 0 \end{bmatrix}$$

transforms $A(\lambda)$ *into the equivalent cubic polynomial* $\widetilde{A}(\lambda)$, *where* $\widetilde{A}_j = A_j$+*rank*-1 *update. We find that*

$$\widetilde{A}_3 = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}, \quad \widetilde{A}_2 = \begin{bmatrix} -6 & 4.75 \\ 0 & -0.5 \end{bmatrix}, \quad \widetilde{A}_1 = \begin{bmatrix} 11 & -11.75 \\ 0 & -1.5 \end{bmatrix}, \quad \widetilde{A}_0 = \begin{bmatrix} -6 & -26.5 \\ 0 & -1 \end{bmatrix}.$$

Note that the SPT $(\widetilde{T}_L, \widetilde{T}_R)$ *with* $\widetilde{T}_L = I_{3n} + U_L V_L^*$ *and* $\widetilde{T}_R = I_{3n} + U_R V_R^*$ *diagonalizes the triangular cubic matrix polynomial* $\widetilde{A}(\lambda)$ *into* $A(\lambda)$.

4. Preserving symmetries in matrix polynomials

The quadratic matrix polynomials that arise in applications often have additional structures that come from the physics of the problem and that should be preserved by the SPTs. To be concise with the presentation, we use the $\star$-adjoint $A^\star(\lambda) = \sum_{j=0}^{d} \lambda^j A_j^\star$ of the matrix polynomial $A(\lambda) = \sum_{j=0}^{d} \lambda^j A_j \in \mathbb{F}[\lambda]^{n\times n}$, where the symbol $\star$ denotes transpose T in the real case $\mathbb{F} = \mathbb{R}$, and either the transpose T or conjugate transpose $*$ in the complex case $\mathbb{F} = \mathbb{C}$.

The three most important matrix polynomial structures are

(a) Hermitian/symmetric when $A^\star(\lambda) = A(\lambda)$ and skew-Hermitian/skew-symmetric when $A^\star(\lambda) = -A(\lambda)$,
(b) $\star$-alternating when $A^\star(-\lambda) = \varepsilon A(\lambda)$ with $\varepsilon \in \{1, -1\}$, also called $\star$-even when $\varepsilon = 1$ and $\star$-odd when $\varepsilon = -1$, and
(c) $\star$-palindromic when $\mathrm{rev}A^\star(\lambda) = \varepsilon A(\lambda)$ with $\varepsilon \in \{1, -1\}$.

Quadratic matrix polynomials with symmetric coefficient matrices arise frequently in the vibration analysis of structural systems [16]. Gyroscopic systems leads to T-even quadratics $A(\lambda) = \lambda^2 A_2 + \lambda A_1 + A_0$ with A_0, A_2 symmetric and A_1 skew-symmetric. We refer to the NLEVP collection of nonlinear eigenvalue problems [1] and references therein for concrete examples of matrix polynomials having one of the structures described above. To preserve these structures, the parameter matrices defining the SPT (T_L, T_R) in (3.10)–(3.11) must satisfy additional constraints as shown in the following theorem.

Theorem 8. *Let $A(\lambda) \in \mathbb{F}[\lambda]^{n\times n}$ be a matrix polynomial of degree d and let $T_L, T_R \in \mathbb{F}^{dn\times dn}$ be as in* (3.10) *and* (3.11) *with parameter matrices F_L, F_R, $G_{L,\ell}$, $G_{R,\ell} \in \mathbb{F}^{n\times n}$, $\ell = 1, \dots, d-1$ such that* (i) *and* (ii) *in Theorem* 3 *hold. Let $\widetilde{A}(\lambda)$ be the matrix polynomial of degree d that results from applying T_L, T_R to any pencil in $\mathbb{DL}(A)$. Let $\varepsilon \in \{-1, 1\}$.*

(a) *If $A^\star(\lambda) = \varepsilon A(\lambda)$ and $F_L = F_R^\star$, $G_{L,\ell} = \varepsilon G_{R,\ell}^\star$, $\ell = 1, \dots, d-1$ then $\widetilde{A}^\star(\lambda) = \varepsilon\widetilde{A}(\lambda)$.*
(b) *If $A^\star(-\lambda) = \varepsilon A(\lambda)$ and $F_L = F_R^\star$, $G_{L,\ell} = \varepsilon(-1)^{d-\ell} G_{R,\ell}^\star$, $\ell = 1, \dots, d-1$ then $\widetilde{A}^\star(-\lambda) = \varepsilon\widetilde{A}(\lambda)$.*
(c) *If* $\mathrm{rev}A^\star(\lambda) = \varepsilon A(\lambda)$ *and $F_L = -F_R^\star$, $G_{L,\ell} = \varepsilon G_{R,\ell}^\star$, $\ell = 1, \dots, d-1$ then* $\mathrm{rev}\widetilde{A}^\star(\lambda) = \varepsilon\widetilde{A}(\lambda)$.

Proof. We note that for $P_\ell(A)$ in (3.8), $P_\ell(A^T)^T = P_\ell(A^\star)^\star$, $\ell = 1, \dots, d-1$, and since the $M_k(A)$ in (2.2) are block symmetric, $M_k(A^\star) = M_k(A)^\star$, $k = 0, \dots, d$.

(a) Assume that $A^\star(\lambda) = \varepsilon A(\lambda)$, or equivalently, that $A_j^\star = \varepsilon A_j$, $j = 0, \dots, d$. Then $P_\ell(A^\star) = \varepsilon P_\ell(A)$ so that

$$(T_R^{-1})^\star = \left(I_d \otimes F_R + \sum_{\ell=1}^{d-1} (I_d \otimes G_{R,\ell}) P_\ell(A^\star)^\star\right)^\star = I_d \otimes F_L + \sum_{\ell=1}^{d-1} \varepsilon P_\ell(A)(I_d \otimes \varepsilon G_{L,\ell}) = T_L^{-1}.$$

Since $M_k(A)^\star = \varepsilon M_k(A)$,

$$M_k(\widetilde{A}^\star) = M_k(\widetilde{A})^\star = (T_R^\star M_k(A) T_R)^\star = T_R^\star M_k(A)^\star T_R = \varepsilon T_R^\star M_k(A) T_R = \varepsilon M_k(\widetilde{A}),$$

which implies that $\widetilde{A}^\star(\lambda) = \varepsilon\widetilde{A}(\lambda)$.

(b) Assume that $A^\star(-\lambda) = \varepsilon A(\lambda)$, or equivalently, that $A_j^\star = (-1)^j \varepsilon A_j$, $j = 0, \dots, d$. Then we find that $P_\ell(A^\star) = (-1)^{d-\ell}\varepsilon(D \otimes I_n) P_\ell(A)(D \otimes I_n)$, where

$$D = \mathrm{diag}(1, -1, \dots, (-1)^{d-1}) \in \mathbb{R}^{d\times d}$$

so that

$$
\begin{aligned}
(T_R^{-1})^\star &= I_d \otimes F_R^\star + \sum_{\ell=1}^{d-1} P_\ell(A^\star)(I_d \otimes G_{R,\ell}^\star) \\
&= I_d \otimes F_L + \sum_{\ell=1}^{d-1} (-1)^{d-\ell}\varepsilon (D\otimes I_n) P_\ell(A)(D\otimes I_n)(I_d \otimes \varepsilon(-1)^{d-\ell} G_{L,\ell}) \\
&= (D\otimes I_n) T_L^{-1} (D\otimes I_n).
\end{aligned}
$$

Also, $A_j^\star = (-1)^j \varepsilon A_j$, $j = 0,\ldots,d$ is equivalent to

$$
\left(M_k(A)(D\otimes I_n)\right)^\star = -\varepsilon(-1)^{d-k} M_k(A)(D\otimes I_n)
$$

(i.e., the pencils $M_k(A)(D\otimes I_n)$ alternates between being Hermitian (or symmetric) and skew-Hermitian (or skew-symmetric)). Now,

$$
\begin{aligned}
(M_k(\widetilde{A})(D\otimes I_n))^\star &= (T_L M_k(A) T_R (D\otimes I_n))^\star \\
&= (T_L M_k(A)(D\otimes I_n) T_L^\star)^\star \\
&= -\varepsilon(-1)^{d-k} T_L M_k(A)(D\otimes I_n) T_L^\star \\
&= -\varepsilon(-1)^{d-k} T_L M_k(A) T_R (D\otimes I_n) \\
&= -\varepsilon(-1)^{d-k} M_k(\widetilde{A})(D\otimes I_n),
\end{aligned}
$$

which implies that $\widetilde{A}_j^\star = (-1)^j \varepsilon \widetilde{A}_j$, $j = 0,\ldots,d$.

(c) Assume that $\mathrm{rev}A^\star(\lambda) = \varepsilon A(\lambda)$, or equivalently, that $A_j^\star = \varepsilon A_{d-j}$, $j = 0,\ldots,d$. If we denote by

$$
S = \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix}
$$

the $d\times d$ standard involutory permutation matrix then

$$
P_\ell(A^\star) = -\varepsilon(S\otimes I_n) P_{d-\ell}(A)(S\otimes I_n), \quad \ell = 1,\ldots,d-1,
$$

so that

$$
(T_R^{-1})^\star = -(S\otimes I_n) T_L^{-1} (S\otimes I_n).
$$

It follows from (2.2)–(2.3) that $A_j^\star = \varepsilon A_{d-j}$, $j = 0,\ldots,d$ is equivalent to

$$
M_0(A)(S\otimes I_n) = -\left(M_d(A)(S\otimes I_n)\right)^\star.
$$

Then,

$$
\begin{aligned}
M_0(\widetilde{A})(S \otimes I_n) &= T_L M_0(A) T_R (S \otimes I_n) \\
&= T_L M_0(A)(S \otimes I_n)(S \otimes I_n) T_R (S \otimes I_n) \\
&= -T_L M_0(A)(S \otimes I_n) T_L^\star \\
&= T_L (M_d(A)(S \otimes I_n))^\star T_L^\star \\
&= (T_L M_d(A)(S \otimes I_n) T_L^\star)^\star \\
&= -(T_L M_d(A) T_R (S \otimes I_n))^\star \\
&= -(M_d(\widetilde{A})(S \otimes I_n))^\star,
\end{aligned}
$$

which implies that $\widetilde{A}_j^\star = \varepsilon \widetilde{A}_{d-j}$, $j = 0, \ldots, d$. □

The next result is a direct consequence of Corollary 4 and Theorem 8.

Corollary 9. *Let $A(\lambda) = \lambda^2 A_2 + \lambda A_1 + A_0 \in \mathbb{F}[\lambda]^{n\times n}$ and*

$$
T^{-1} = I_2 \otimes F + \begin{bmatrix} \frac{1}{2} G A_1 & G A_0 \\ -G A_2 & -\frac{1}{2} G A_1 \end{bmatrix} \tag{4.1}
$$

with $G, F \in \mathbb{F}^{n\times n}$. Let $\varepsilon \in \{+1, -1\}$.

(a) *If $A_j^\star = \varepsilon A_j$, $j = 0, 1, 2$ then $(T^\star, T)$ with G, F such that $FG^\star = -\varepsilon (FG^\star)^\star$ is an SPT that transforms $A(\lambda)$ into $\widetilde{A}(\lambda)$ whose coefficient matrices are such that $\widetilde{A}_j^\star = \varepsilon \widetilde{A}_j$, $j = 0, 1, 2$.*

(b) *If $A_j^\star = \varepsilon (-1)^j A_j$, $j = 0, 1, 2$ then $((D \otimes I_n) T^\star (D \otimes I_n), T)$ with $D = diag(1, -1)$ and G, F such that $FG^\star = \varepsilon (FG^\star)^\star$ is an SPT that transforms $A(\lambda)$ into $\widetilde{A}(\lambda)$ whose coefficient matrices are such that $\widetilde{A}_j^\star = \varepsilon \widetilde{A}_j$, $j = 0, 1, 2$.*

(c) *If $A_2^\star = \varepsilon A_0$ and $A_1^\star = \varepsilon A_1$ then $(-(S \otimes I_n) T^\star (S \otimes I_n), T)$ with $S = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$ and G, F such that $FG^\star = \varepsilon (FG^\star)^\star$ is an SPT that transforms $A(\lambda)$ into $\widetilde{A}(\lambda)$ whose coefficient matrices are such that $\widetilde{A}_2^\star = \varepsilon \widetilde{A}_0$ and $\widetilde{A}_1^\star = \varepsilon \widetilde{A}_1$.*

Example 10. *To preserve the symmetry of the quadratic matrix polynomial in Example 5, we apply Corollary 9 (a) with $\star = T$ and $\varepsilon = 1$, and choose*

$$
F = I_2, \qquad G = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}
$$

so that FG^T is skew-symmetric. Then the SPT (T^T, T) transforms $A(\lambda)$ into the symmetric quadratic

$$
\widetilde{A}(\lambda) = \lambda^2 \begin{bmatrix} -\frac{1}{3} & 0 \\ 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} 4/9 & -4/9 \\ -4/9 & -4/3 \end{bmatrix} + \begin{bmatrix} -7/27 & -16/27 \\ -16/27 & -52/27 \end{bmatrix}.
$$

Note that when $\mathbb{F} = \mathbb{C}$ and $\star = *$, the SPT in (1.9) preserves Hermitian structure if $b = ia \in \mathbb{C}^n$ since $G_L = -ab^* = iaa^* = ba^* = G_R^*$. On the other hand, when $\mathbb{F} = \mathbb{R}$ and $\star = T$ then the constraint $G_L \neq G_R^T$ for all nonzero $a, b \in \mathbb{R}^n$. So the SPT in (1.9) does not in general preserve symmetry when $A(\lambda)$ is symmetric. Although the parametrization in Corollary 9 was not known at the time, the SPT used in [15] to deflate eigenpairs of symmetric quadratic while preserving the symmetry correspond to choosing

$$F = I + af^T, \qquad G = aa^T,$$

in (4.1) for some nonzero vectors $a, f \in \mathbb{R}^n$ such that

$$a^T f = -1, \qquad (a^T A_2 a)(a^T A_0 a) - \frac{1}{4}(a^T A_0 a)^2 \neq 0.$$

We remark that the above constraints on the parameter a and f is much simpler than that in [15].

5. Concluding remarks

We have constructed a parametrization for the inverse of the left and right transformations that preserve the block structure of pencils in $\mathbb{DL}(A)$, and hence produce a new matrix polynomial $\widetilde{A}(\lambda)$ that is still of degree d and is unimodularly equivalent to $A(\lambda)$. We have also identified constraints on the parametrization that lead to SPTs that preserve existing structures in $A(\lambda)$ such as symmetric, alternating and palindromic structures.

We have shown that our parametrization allows constructions of SPTs that are low rank modifications of the identity. The latter are easy to invert and when applied to any pencil in $\mathbb{DL}(A)$, lead to a new matrix polynomial $\widetilde{A}(\lambda)$ whose matrix coefficients $\widetilde{A}_j$ are low rank modifications of A_j. SPTs of this type can be used to deflate d eigenpairs with distinct eigenvalues and linearly independent eigenvectors (see [14] and [15] for $d = 2$). How to identify among this class of SPTs, transformations that have specific actions such as that of introducing zeros in specific entries or columns of the matrix polynomial is the subject of ongoing work.

We concentrated here on matrix polynomials $A(\lambda)$ expressed in the monomial basis $1, \lambda, \lambda^2 \ldots, \lambda^d$. The definition of the vector space of pencils $\mathbb{DL}(A)$ can however be generalized to other bases such as for example the Legendre basis or the Chebyshev basis [11]. Then the one-sided factorizations (1.5a)–(1.5b) hold but for a different Λ. These factorizations lead to standard basis pencils that differ from those in (2.1). But as long as we have access to one-sided factorizations of the type (1.5a)–(1.5b) and the corresponding standard basis for $\mathbb{DL}(A)$, the procedure we followed to construct the SPTs that preserve the block structure of pencils in $\mathbb{DL}(A)$ still applies.

References

[1] Timo Betcke, Nicholas J. Higham, Volker Mehrmann, Christian Schröder, and Françoise Tisseur. NLEVP: A collection of nonlinear eigenvalue problems. *ACM Trans. Math. Software*, 39(2):7:1–7:28, February 2013.

[2] S. D. Garvey, M. I. Friswell, and U. Prells. Co-ordinate transformations for second order systems. part II: Elementary structure-preserving transformations. *Journal of Sound and Vibration*, 258(5):911–930, 2002.

[3] Seamus Garvey, Uwe Prells, Michael I. Friswell, and Zheng Chen. General isospectral flows for linear dynamic systems. *Linear Algebra Appl.*, 385:335–368, 2004.

[4] Israel Gohberg, Peter Lancaster, and Leiba Rodman. Spectral analysis of selfadjoint matrix polynomials. *Ann. of Math. (2)*, 112(1):33–71, 1980.

[5] Israel Gohberg, Peter Lancaster, and Leiba Rodman. *Indefinite Linear Algebra and Applications*. Birkhäuser, Basel, Switzerland, 2005.

[6] Laurence Grammont, Nicholas J. Higham, and Françoise Tisseur. A framework for analyzing nonlinear eigenproblems and parametrized linear systems. *Linear Algebra Appl.*, 435(3):623–640, 2011.

[7] Nicholas J. Higham, D. Steven Mackey, Niloufer Mackey, and Françoise Tisseur. Symmetric linearizations for matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 29(1):143–159, 2006.

[8] Peter Lancaster. Symmetric transformations of the companion matrix. *NABLA: Bulletin of the Malayan Math. Soc.*, 8:146–148, 1961.

[9] Peter Lancaster and Uwe Prells. Isospectral families of high-order systems. *Z. Angew. Math. Mech.*, 87(3):219–234, 2007.

[10] D. Steven Mackey, Niloufer Mackey, Christian Mehl, and Volker Mehrmann. Vector spaces of linearizations for matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 28(4):971–1004, 2006.

[11] Yuji Nakatsukasa, Vanni Noferini, and Alex Townsend. Vector spaces of linearizations for matrix polynomials: A bivariate polynomial approach. *SIAM J. Matrix Anal. Appl.*, 38(1):1–29, 2017.

[12] Yuji Nakatsukasa, Leo Taslaman, Françoise Tisseur, and Ion Zaballa. Reduction of matrix polynomials to simpler forms. *SIAM J. Matrix Anal. Appl.*, 39(1):148–177, 2018.

[13] Leo Taslaman, Françoise Tisseur, and Ion Zaballa. Triangularizing matrix polynomials. *Linear Algebra Appl.*, 439.

[14] Françoise Tisseur. Deflation of matrix polynomials. Manuscript in preparation, 2022.

[15] Françoise Tisseur, Seamus D. Garvey, and Christopher Munro. Deflating quadratic matrix polynomials with structure preserving transformations. *Linear Algebra Appl.*, 435(3):464–479, 2011.

[16] Françoise Tisseur and Karl Meerbergen. The quadratic eigenvalue problem. *SIAM Rev.*, 43(2):235–286, 2001.

[17] Françoise Tisseur and Ion Zaballa. Triangularizing quadratic matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 34(2):312–337, 2013.