Symplectic, BVD, and Palindromic Approaches to Discrete-Time Control Problems

Byers, Ralph and Mackey, D. Steven and Mehrmann, Volker and Xu, Hongguo

2008

MIMS EPrint: **2008.35**

ISSN 1749-9097

Symplectic, BVD, and Palindromic Approaches to Discrete-Time Control Problems¶

Ralph Byers * D. Steven Mackey † Volker Mehrmann‡ Hongguo Xu§

March 12, 2008

Dedicated to Mihail Konstantinov on the occasion of his 60th birthday

Abstract

We give several different formulations for the discrete-time linear-quadratic control problem in terms of structured eigenvalue problems, and discuss the relationships among the associated structured objects: symplectic matrices and pencils, BVD-pencils and polynomials, and the recently introduced classes of palindromic pencils and matrix polynomials. We show how these structured objects can be transformed into each other, and also how their eigenvalues, eigenvectors and invariant/deflating subspaces are related.

Key words. discrete-time linear-quadratic control, symplectic matrix, symplectic pencil, BVD-pencil, BVD-polynomial, palindromic pencil, palindromic matrix polynomial
AMS subject classification. 65F15, 15A21, 93B40.

1 Introduction

In this paper we consider three approaches to solving the discrete-time linear-quadratic control problem via structured eigenvalue problems, and investigate the relationships among the resulting classes of structured matrices, matrix pencils, and matrix polynomials. These include symplectic matrices and symplectic pencils [5, 26] as well as the recently introduced class of palindromic/anti-palindromic matrix polynomials [9, 10, 19]. Although these various structured objects have very similar spectral properties, they differ substantially in their representation and numerical properties.

¶Partially supported by *Deutsche Forschungsgemeinschaft*, through the DFG Research Center Matheon *Mathematics for Key Technologies* in Berlin.

*Department of Mathematics, University of Kansas, Lawrence, KS 66045, USA. The author is deceased; he was partially supported by *National Science Foundation* grants 0098150, 0112375, 9977352.

†Department of Mathematics, Western Michigan University, Kalamazoo, MI 49008, USA. `steve.mackey@wmich.edu`. Partially supported by *National Science Foundation* grant DMS-0713799.

‡Institut für Mathematik, TU Berlin, Str. des 17. Juni 136, D-10623 Berlin, FRG. `mehrmann@math.tu-berlin.de`.

§Department of Mathematics, University of Kansas, Lawrence, KS 66045, USA. `xu@math.ukans.edu`. Partially supported by *National Science Foundation* grant 0314427, and the *University of Kansas General Research Fund allocation* # 2301717.

Our motivation comes mainly from the standard discrete-time linear-quadratic optimal control problem:

$$\text{Minimize} \qquad \sum_{j=0}^{\infty} \left(x_j^T Q x_j + x_j^T Y u_j + u_j^T Y^T x_j + u_j^T R u_j \right) \tag{1}$$

subject to the kth-order discrete-time control system

$$\sum_{i=0}^{k} M_i x_{j+i+1-k} = B u_j, \quad j = 0, 1, \ldots, \tag{2}$$

with given starting values $x_0, x_{-1}, \ldots, x_{1-k} \in \mathbb{R}^n$ and coefficient matrices $Q = Q^T \in \mathbb{R}^{n,n}$, $Y \in \mathbb{R}^{n,m}$, $R = R^T \in \mathbb{R}^{m,m}$, and $M_i \in \mathbb{R}^{n,n}$ for $i = 0, \ldots, k$, $B \in \mathbb{R}^{n,m}$. Second order control problems of this form arise, for example, in the control of sampled or time-discretized multibody systems [4].

In the classical applications of this problem [12, 13, 26, 28], the matrix $\hat{R} = \begin{bmatrix} Q & Y \\ Y^T & R \end{bmatrix}$ is typically positive semi-definite, with positive definite block R. However, in other applications from discrete-time H_∞ control [37], both $\hat{R}$ and R may be indefinite and singular. When the discrete-time system (2) arises from the discretization of an ordinary differential equation then $M_k = I$, but if the underlying dynamics is constrained, as in the case of descriptor systems, see e.g. [4, 15, 26], then M_k is singular.

After introducing some notation in Section 2, we show in Section 3 how the control problem (1), (2) may first be formulated in terms of a class of structured pencils that we refer to as BVD-pencils; this acronym stands for ***B***oundary ***V***alue problem for the optimal control of ***D***iscrete systems. We then see how under certain circumstances these BVD-pencils may be reduced to symplectic pencils, and sometimes all the way to symplectic matrices. Finally we show how to formulate the solution of (1), (2) as a palindromic eigenproblem.

The next three sections explore the connections between these various types of structured matrices and matrix polynomials, including the relationships between their eigenvalues, eigenvectors and invariant/deflating subspaces. In particular, we discuss the question of whether each can be converted into the other by some simple transformation or embedding. This question is also motivated by the concerns of numerical computation. The importance of symplectic/BVD eigenvalue problems in control applications has made the development of reliable numerical methods for the solution of these eigenvalue and boundary value problems an important topic of research for the last 30 years, see e.g. [5, 25, 26, 28, 34] and the references therein. It is well known that structure-preserving methods have a much better perturbation and error analysis than unstructured methods [13, 21, 20], in particular when eigenvalues are near or on the unit circle. Thus the main focus of this research has been the development of structure-preserving methods.

However, it is not easy to preserve symplectic structure or BVD-structure in finite precision arithmetic [5, 6, 24]. BVD-structure (12) is (in part) defined by a zero block structure that is not easy to preserve by any obvious set of transformations, while the structure of a symplectic matrix or pencil is defined via the *nonlinear* relations (15), (16). By contrast, palindromic structure is defined by very simple linear relations (5), (6), and is therefore much easier to preserve in finite precision arithmetic [30, 31, 33]. Thus it would be preferable to work with palindromic rather than symplectic or BVD-structure. For this reason having suitable transformations between the structures would be especially helpful.

In Section 4 we discuss the relationship between symplectic matrices/pencils and BVD-pencils. In Section 5 we review some results of [30] that classify when a symplectic matrix $\mathcal{S}$ can be represented by a palindromic pencil, and conversely, when a palindromic pencil can be easily transformed to a symplectic matrix/pencil. In Section 6 we then study BVD-polynomials and how they can be transformed into palindromic pencils and polynomials.

2 Notation

We follow [19] in style and notation. Let $\mathbb{F}$ denote the field $\mathbb{R}$ or $\mathbb{C}$ and consider a k-th degree matrix polynomial $P(\lambda) = \sum_{i=0}^{k} \lambda^i B_i$, where $B_0, B_1, \dots B_k \in \mathbb{F}^{m,n}$ with $B_k \neq 0$. Writing $P(\lambda)$ in homogeneous form $P(\alpha,\beta) = \sum_{i=0}^{k} \alpha^i \beta^{k-i} B_i$, then the spectrum of the homogeneous matrix polynomial is defined to be the set of pairs $(\alpha,\beta) \in \mathbb{C}^2 \setminus \{(0,0)\}$ for which $P(\alpha,\beta)$ has a drop in rank. We identify pairs (α,β) with $\lambda = \frac{\alpha}{\beta}$ if $\beta \neq 0$ and with $\lambda = \infty$ if $\beta = 0$. For simplicity we use only the notation $P(\lambda)$ in the rest of the paper.

A matrix polynomial $P(\lambda) = \sum_{i=0}^{k} \lambda^i B_i$ is said to be *regular* if $m = n$ and $\det P(\lambda)$ does not vanish identically. In the following we consider only regular polynomials, although some of the results that we present can be easily extended to the singular case.

In addition, to avoid unnecessary distinctions between the real and complex case, and between the use of transpose (T) or conjugate transpose ($*$), we follow [19] and introduce the symbol $\star$ to stand for either T or $*$.

Definition 1 *Let $P(\lambda) = \sum_{i=0}^{k} \lambda^i B_i$ be a matrix polynomial, where $B_0, \dots, B_k \in \mathbb{F}^{n,n}$ with $B_k \neq 0$. Then we define the* adjoint *of $P(\lambda)$ by*

$$P^\star(\lambda) := \sum_{i=0}^{k} \lambda^i B_i^\star \,, \tag{3}$$

and the reversal *of $P(\lambda)$ by*

$$\mathrm{rev}P(\lambda) := \lambda^k P(1/\lambda) = \sum_{i=0}^{k} \lambda^i B_{k-i} \,. \tag{4}$$

A matrix polynomial $P(\lambda)$ is said to be palindromic *if*

$$\mathrm{rev}P^\star(\lambda) = P(\lambda) \,, \quad i.e. \quad B_{k-i}^\star = B_i \quad for \ i = 0, \dots, k \,, \tag{5}$$

and anti-palindromic *if*

$$\mathrm{rev}P^\star(\lambda) = -P(\lambda) \,, \quad i.e. \quad B_{k-i}^\star = -B_i \quad for \ i = 0, \dots, k \,. \tag{6}$$

Note that a *palindromic pencil* has the form $\lambda\mathcal{Z} + \mathcal{Z}^\star$, while an *anti-palindromic pencil* is of the form $\lambda\mathcal{Z} - \mathcal{Z}^\star$. It is clear that by changing λ to $-\lambda$ a palindromic pencil becomes anti-palindromic and vice-versa. To avoid having to switch signs, we use whichever structure is more convenient.

As in [19], we may sometimes add a prefix T or $*$ to the term palindromic to clarify which adjoint $\star$ we are using.

3 Linear-quadratic optimal control via structured eigenvalue problems

In this section we describe several ways to approach the discrete-time linear-quadratic optimal control problem (1), (2) via a structured eigenvalue problem.

Let us begin by assuming that M_k is invertible. In this case the classical way to solve this control problem is to turn the kth-order difference equation (2) into a first order system by introducing, for example,

$$\mathcal{E} = \begin{bmatrix} M_k & & & \\ & I & & \\ & & \ddots & \\ & & & I \end{bmatrix}, \ \mathcal{A} = \begin{bmatrix} -M_{k-1} & -M_{k-2} & \dots & -M_0 \\ I & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ & 0 & I & 0 \end{bmatrix}, \ \mathcal{R} = \mathcal{R}^\star = R,$$

$$\mathcal{Q} = \mathcal{Q}^\star = \begin{bmatrix} Q & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix}, \ \mathcal{B} = \begin{bmatrix} B \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ \mathcal{Y} = \begin{bmatrix} Y \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ z_j = \begin{bmatrix} x_j \\ x_{j-1} \\ \vdots \\ x_{j+1-k} \end{bmatrix}, \tag{7}$$

and then to apply techniques for first order systems, e.g. in [5, 12, 26, 28], to minimize

$$\sum_{j=0}^{\infty} \left(z_j^\star \mathcal{Q} z_j + z_j^\star \mathcal{Y} u_j + u_j^\star \mathcal{Y}^\star z_j + u_j^\star \mathcal{R} u_j \right) \tag{8}$$

subject to

$$\mathcal{E} z_{j+1} = \mathcal{A} z_j + \mathcal{B} u_j, \ z_0 = \begin{bmatrix} x_0^T & \dots & x_{1-k}^T \end{bmatrix}^T. \tag{9}$$

Introducing a vector of Lagrange multipliers $m_j = \begin{bmatrix} -\nu_j^T & -\tilde{\nu}_j^T \end{bmatrix}^T$ with $\nu_j \in \mathbb{F}^n$ and $\tilde{\nu}_j \in \mathbb{F}^{(k-1)n}$ and applying the Pontryagin maximum principle [16, 26] leads to the two-point boundary value problem

$$\begin{bmatrix} 0 & \mathcal{E} & 0 \\ \mathcal{A}^\star & 0 & 0 \\ \mathcal{B}^\star & 0 & 0 \end{bmatrix} \begin{bmatrix} m_{j+1} \\ z_{j+1} \\ u_{j+1} \end{bmatrix} = \begin{bmatrix} 0 & \mathcal{A} & \mathcal{B} \\ \mathcal{E}^\star & \mathcal{Q} & \mathcal{Y} \\ 0 & \mathcal{Y}^\star & \mathcal{R} \end{bmatrix} \begin{bmatrix} m_j \\ z_j \\ u_j \end{bmatrix}, \tag{10}$$

with initial condition as in (9) and terminal condition

$$\lim_{j\to\infty} \mathcal{E}^\star m_j = 0. \tag{11}$$

Remark 2 To unify notation we have again used $\star$ to denote transpose or conjugate transpose. Note, however, that in discrete-time optimal control when the resulting boundary value problem is complex then $\star$ invariably denotes conjugate transpose. We do not know of any control application where the complex transpose case arises.

This boundary value problem is usually solved, see [5, 13, 26, 28], via the computation of the deflating subspace associated with the eigenvalues inside the unit disk of the associated matrix pencil, i.e. by solving the generalized eigenvalue problem

$$\left(\lambda \begin{bmatrix} 0 & \mathcal{E} & 0 \\ \mathcal{A}^\star & 0 & 0 \\ \mathcal{B}^\star & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & \mathcal{A} & \mathcal{B} \\ \mathcal{E}^\star & \mathcal{Q} & \mathcal{Y} \\ 0 & \mathcal{Y}^\star & \mathcal{R} \end{bmatrix} \right) \begin{bmatrix} \phi \\ \psi \\ \gamma \end{bmatrix} = 0. \tag{12}$$

If $\mathcal{R}(=R)$ is invertible, then one can eliminate the variable γ corresponding to the eigenvalue infinity in (12) and also the corresponding variable u_j in the boundary value problem (10), thus reducing the problem to the computation of deflating subspaces of a smaller pencil, i.e. to the solution of the generalized eigenvalue problem

$$\begin{aligned}
\left(\lambda\hat{\mathcal{T}} - \hat{\mathcal{S}}\right)\begin{bmatrix} \phi \\ \psi \end{bmatrix} &:= \left(\lambda\begin{bmatrix} \hat{\mathcal{H}} & \mathcal{E} \\ \hat{\mathcal{F}}^\star & 0 \end{bmatrix} - \begin{bmatrix} 0 & \hat{\mathcal{F}} \\ \mathcal{E}^\star & \mathcal{G} \end{bmatrix}\right)\begin{bmatrix} \phi \\ \psi \end{bmatrix} \qquad (13)\\
&:= \left(\lambda\begin{bmatrix} -\mathcal{B}R^{-1}\mathcal{B}^\star & \mathcal{E} \\ (\mathcal{A}-\mathcal{B}R^{-1}\mathcal{Y}^\star)^\star & 0 \end{bmatrix} - \begin{bmatrix} 0 & \mathcal{A}-\mathcal{B}R^{-1}\mathcal{Y}^\star \\ \mathcal{E}^\star & \mathcal{Q}-\mathcal{Y}R^{-1}\mathcal{Y}^\star \end{bmatrix}\right)\begin{bmatrix} \phi \\ \psi \end{bmatrix} = 0\,,
\end{aligned}$$

where $\hat{\mathcal{H}}^\star = \hat{\mathcal{H}}$ and $\mathcal{G}^\star = \mathcal{G}$.

If, in addition, $\mathcal{E}$ and $\hat{\mathcal{F}}$ are invertible (which means that M_k and M_0 are invertible, respectively), then one may transform even further and instead solve the boundary value problem associated with the eigenvalue problem for the matrix $\mathcal{S}$ in

$$\begin{aligned}
&(\lambda I_{2kn} - \mathcal{S})\begin{bmatrix} \hat{\phi} \\ \psi \end{bmatrix} \qquad (14)\\
&:= \left(\lambda I_{2kn} - \begin{bmatrix} \mathcal{E}^\star & 0 \\ 0 & I_{kn} \end{bmatrix}\hat{\mathcal{T}}^{-1}\hat{\mathcal{S}}\begin{bmatrix} \mathcal{E}^{-\star} & 0 \\ 0 & I_{kn} \end{bmatrix}\right)\begin{bmatrix} \mathcal{E}^\star & 0 \\ 0 & I_{kn} \end{bmatrix}\begin{bmatrix} \phi \\ \psi \end{bmatrix}\\
&= \left(\lambda I_{2kn} - \begin{bmatrix} \mathcal{E}^{-1}\hat{\mathcal{H}}\mathcal{E}^{-\star} & I_{kn} \\ (\mathcal{E}^{-1}\hat{\mathcal{F}})^\star & 0 \end{bmatrix}^{-1}\begin{bmatrix} 0 & \mathcal{E}^{-1}\hat{\mathcal{F}} \\ I_{kn} & \mathcal{G} \end{bmatrix}\right)\begin{bmatrix} \hat{\phi} \\ \psi \end{bmatrix}\\
&:= \left(\lambda I_{2kn} - \begin{bmatrix} \mathcal{H} & I_{kn} \\ \mathcal{F}^\star & 0 \end{bmatrix}^{-1}\begin{bmatrix} 0 & \mathcal{F} \\ I_{kn} & \mathcal{G} \end{bmatrix}\right)\begin{bmatrix} \hat{\phi} \\ \psi \end{bmatrix} = 0\,,
\end{aligned}$$

where $\hat{\phi} = \mathcal{E}^\star\phi$, $\mathcal{H}^\star = \mathcal{H}$, and $\mathcal{F}$ is nonsingular. With $\mathcal{J}_{kn} := \begin{bmatrix} 0 & I_{kn} \\ -I_{kn} & 0 \end{bmatrix}$, the matrix $\mathcal{S}$ in (14) satisfies

$$\mathcal{S}\mathcal{J}_{kn}\mathcal{S}^\star = \mathcal{J}_{kn}\,. \qquad (15)$$

Matrices with this property are called *symplectic* matrices [5, 26]. If $\mathcal{E} = I_{kn}$ in (13) (which corresponds to $M_k = I$ in (2)), then regardless of whether $\hat{\mathcal{F}}$ is singular or nonsingular we have

$$\hat{\mathcal{S}}\mathcal{J}_{kn}\hat{\mathcal{S}}^\star = \hat{\mathcal{T}}\mathcal{J}_{kn}\hat{\mathcal{T}}^\star. \qquad (16)$$

Because of this property, in [5, 26, 27] pencils $\lambda\hat{\mathcal{T}} - \hat{\mathcal{S}}$ satisfying (16) are called *symplectic pencils*. Note that if $\mathcal{E} \neq I_{kn}$ in (13), then (16) will in general no longer hold.

In order to avoid any conflict with the terminology of symplectic pencils, we will denote the pencil in (12) by a name which instead emphasizes its origin in the two-point ***B***oundary ***V***alue problem for the optimal control of ***D***iscrete-time descriptor systems. Thus we call any pencil with the structure of (12) a *BVD-pencil*, and a pencil with the structure of (13) a *reduced BVD-pencil*. Note that a reduced BVD-pencil with $\mathcal{E} = I_{kn}$ is a symplectic pencil. Just as for palindromic pencils/polynomials, we sometimes add a prefix T or $*$ to indicate which adjoint we are using.

As an alternative to the reduction to a first-order problem as in (8), (9), (10), we could instead apply the Pontryagin maximum principle directly to the kth-order optimal control problem (1), (2) and thus avoid the assumption of invertibility of M_k. Using the partitioning

$\phi = \begin{bmatrix} \phi_1^T & \dots & \phi_k^T \end{bmatrix}^T$, $\psi = \begin{bmatrix} \psi_1^T & \dots & \psi_k^T \end{bmatrix}^T$ analogous to (7), leads (for $k > 2$) to a discrete-time boundary value problem

$$\begin{aligned}
&\begin{bmatrix} 0 & M_k & 0 \\ M_0^\star & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} -\nu_{j+k} \\ x_{j+1} \\ u_{j+k} \end{bmatrix} + \begin{bmatrix} 0 & M_{k-1} & 0 \\ M_1^\star & Q & 0 \\ 0 & Y^\star & 0 \end{bmatrix} \begin{bmatrix} -\nu_{j+k-1} \\ x_j \\ u_{j+k-1} \end{bmatrix} \\
&+ \dots + \begin{bmatrix} 0 & M_2 & 0 \\ M_{k-2}^\star & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} -\nu_{j+2} \\ x_{j-k+3} \\ u_{j+2} \end{bmatrix} \\
&+ \begin{bmatrix} 0 & M_1 & 0 \\ M_{k-1}^\star & 0 & 0 \\ -B^\star & 0 & 0 \end{bmatrix} \begin{bmatrix} -\nu_{j+1} \\ x_{j-k+2} \\ u_{j+1} \end{bmatrix} + \begin{bmatrix} 0 & M_0 & -B \\ M_k^\star & 0 & Y \\ 0 & 0 & R \end{bmatrix} \begin{bmatrix} -\nu_j \\ x_{j-k+1} \\ u_j \end{bmatrix} = 0
\end{aligned} \tag{17}$$

associated with the polynomial eigenvalue problem

$$\begin{aligned}
0 = P(\lambda) \begin{bmatrix} \phi_1 \\ \psi_k \\ \gamma \end{bmatrix} &:= \left(\sum_{j=0}^{k} \lambda^j \mathcal{M}_j \right) \begin{bmatrix} \phi_1 \\ \psi_k \\ \gamma \end{bmatrix} \\
&:= \left(\lambda^k \begin{bmatrix} 0 & M_k & 0 \\ M_0^\star & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \lambda^{k-1} \begin{bmatrix} 0 & M_{k-1} & 0 \\ M_1^\star & Q & 0 \\ 0 & Y^\star & 0 \end{bmatrix} + \dots \right. \\
&\qquad \left. + \lambda \begin{bmatrix} 0 & M_1 & 0 \\ M_{k-1}^\star & 0 & 0 \\ -B^\star & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & M_0 & -B \\ M_k^\star & 0 & Y \\ 0 & 0 & R \end{bmatrix} \right) \begin{bmatrix} \phi_1 \\ \psi_k \\ \gamma \end{bmatrix}.
\end{aligned} \tag{18}$$

When $k = 1$ this approach leads to (12), which after multiplication of the last two block rows by $-I$ takes the following form, consistent with (18),

$$\left(\lambda \begin{bmatrix} 0 & M_1 & 0 \\ M_0^\star & 0 & 0 \\ -B^\star & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & M_0 & -B \\ M_1^\star & Q & Y \\ 0 & Y^\star & R \end{bmatrix} \right) \begin{bmatrix} \phi_1 \\ \psi_1 \\ \gamma \end{bmatrix} = 0. \tag{19}$$

In the case $k = 2$ we obtain

$$\left(\lambda^2 \begin{bmatrix} 0 & M_2 & 0 \\ M_0^\star & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} 0 & M_1 & 0 \\ M_1^\star & Q & 0 \\ -B^\star & Y^\star & 0 \end{bmatrix} + \begin{bmatrix} 0 & M_0 & -B \\ M_2^\star & 0 & Y \\ 0 & 0 & R \end{bmatrix} \right) \begin{bmatrix} \phi_1 \\ \psi_2 \\ \gamma \end{bmatrix} = 0. \tag{20}$$

In each case (18), (19), (20), we call the corresponding matrix polynomial a *BVD-matrix polynomial*. Whenever R is invertible we can eliminate γ as in the reduction of (12) to (13), and thereby obtain a *reduced BVD-matrix polynomial*.

It is well known, see e.g. [26], that the spectrum of real symplectic matrices is symmetric with respect to the unit circle. In other words, whenever λ is an eigenvalue then $1/\lambda$, $\bar{\lambda}$, and $1/\bar{\lambda}$ are also eigenvalues. Analogous reciprocal pairings (either $(\lambda, 1/\lambda)$ or $(\lambda, 1/\bar{\lambda})$) are also present in the spectra of complex symplectic matrices and pencils [5, 26]; here 0 and ∞ are regarded as reciprocals of each other. Note that if $|\lambda| = 1$ then eigenvalue pairs may coalesce, leading to degenerate situations. Thus it can be expected that eigenvalues on the unit circle, and especially the eigenvalues 1 and -1, will play a special role in our analysis.

We call this reciprocal pairing of eigenvalues *symplectic eigen-symmetry.* We will see that this eigen-symmetry also partially carries over to the eigenvalues of BVD-pencils and BVD-matrix polynomials, as well as to their reduced versions.

It has been shown in [19] that palindromic and anti-palindromic eigenvalue problems also have symplectic eigen-symmetry, see also [7]. Later it was shown in [33] and also in a different context in [29] how to obtain palindromic eigenvalue problems directly from discrete-time optimal control problems; in light of this fact the symplectic eigen-symmetry of palindromic eigenvalue problems is perhaps not so surprising.

We present next yet another, even simpler way to obtain a palindromic formulation of the discrete-time control problem. Begin by rewriting the boundary value problem (10) as

$$\begin{bmatrix} \mathcal{E} & 0 \end{bmatrix} \begin{bmatrix} z_{j+1} \\ u_{j+1} \end{bmatrix} = \begin{bmatrix} \mathcal{A} & \mathcal{B} \end{bmatrix} \begin{bmatrix} z_j \\ u_j \end{bmatrix}, \tag{21}$$

$$\begin{bmatrix} \mathcal{A}^\star \\ \mathcal{B}^\star \end{bmatrix} m_{j+1} = \begin{bmatrix} \mathcal{E}^\star \\ 0 \end{bmatrix} m_j + \begin{bmatrix} \mathcal{Q} & \mathcal{Y} \\ \mathcal{Y}^\star & \mathcal{R} \end{bmatrix} \begin{bmatrix} z_j \\ u_j \end{bmatrix}. \tag{22}$$

Then replace j with $j+1$ in (22) to get

$$\begin{bmatrix} \mathcal{A}^\star \\ \mathcal{B}^\star \end{bmatrix} m_{j+2} = \begin{bmatrix} \mathcal{E}^\star \\ 0 \end{bmatrix} m_{j+1} + \begin{bmatrix} \mathcal{Q} & \mathcal{Y} \\ \mathcal{Y}^\star & \mathcal{R} \end{bmatrix} \begin{bmatrix} z_{j+1} \\ u_{j+1} \end{bmatrix}. \tag{23}$$

Introducing the new variable $w_j = m_j - m_{j+1}$, subtracting (23) from (22) and rearranging the terms we obtain

$$\begin{bmatrix} \mathcal{A}^\star \\ \mathcal{B}^\star \end{bmatrix} w_{j+1} + \begin{bmatrix} \mathcal{Q} & \mathcal{Y} \\ \mathcal{Y}^\star & \mathcal{R} \end{bmatrix} \begin{bmatrix} z_{j+1} \\ u_{j+1} \end{bmatrix} = \begin{bmatrix} \mathcal{E}^\star \\ 0 \end{bmatrix} w_j + \begin{bmatrix} \mathcal{Q} & \mathcal{Y} \\ \mathcal{Y}^\star & \mathcal{R} \end{bmatrix} \begin{bmatrix} z_j \\ u_j \end{bmatrix}.$$

Combining this equation with (21) we have the boundary value problem

$$\underbrace{\begin{bmatrix} 0 & \mathcal{E} & 0 \\ \mathcal{A}^\star & \mathcal{Q} & \mathcal{Y} \\ \mathcal{B}^\star & \mathcal{Y}^\star & \mathcal{R} \end{bmatrix}}_{\mathcal{Z}} \begin{bmatrix} w_{j+1} \\ z_{j+1} \\ u_{j+1} \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & \mathcal{A} & \mathcal{B} \\ \mathcal{E}^\star & \mathcal{Q} & \mathcal{Y} \\ 0 & \mathcal{Y}^\star & \mathcal{R} \end{bmatrix}}_{\mathcal{Z}^\star} \begin{bmatrix} w_j \\ z_j \\ u_j \end{bmatrix}, \quad j = 0, \ldots, \infty, \tag{24}$$

with the initial condition (9) and $\sum_j^\infty \mathcal{E}^\star w_j = \mathcal{E}^\star m_0$. This boundary value problem can then be solved by decoupling the forward and backward iteration via the solution of the associated anti-palindromic eigenvalue problem $(\lambda \mathcal{Z} - \mathcal{Z}^\star)x = 0$.

4 Symplectic matrices/pencils and BVD-pencils

In Section 3 we have shown (under some nonsingularity assumptions) how to reduce the BVD-pencil (12) first to (13), and then even further to the symplectic matrix $\mathcal{S}$ in (14). We discuss now how this process can be reversed, i.e. how a symplectic matrix can be embedded into a BVD-pencil, beginning with the following lemma.

Lemma 3 *Let $\mathcal{S} \in \mathbb{F}^{2n,2n}$ be symplectic. Then there exists a symplectic matrix*

$$X = \begin{bmatrix} I_n & D \\ 0 & I_n \end{bmatrix}$$

where D is diagonal with diagonal elements either 0 *or* 1 *such that*

$$\mathcal{K} := X\mathcal{S}X^{-1} = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix}, \quad K_{ij} \in \mathbb{F}^{n,n}$$

is symplectic with K_{11} nonsingular.

Proof. Partition $\mathcal{S}$ as

$$\mathcal{S} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}, \quad S_{ij} \in \mathbb{F}^{n,n},$$

and assume $\operatorname{rank} S_{11} = r$. By [3] there exists a subset $\alpha \subseteq \{1, \ldots, n\}$ with $|\alpha| = r$ such that the rows of S_{11} indexed by α are linearly independent, and together with the rows of S_{21} indexed by α', the complement of α, form a basis of $\mathbb{F}^n$. Let D be the $n \times n$ diagonal matrix such that $d_{ii} = 0$ if $i \in \alpha$, and $d_{ii} = 1$ if $i \in \alpha'$. Then for the matrix X with this D we have

$$K_{11} = S_{11} + DS_{21}\,,$$

and it can be easily verified that K_{11} is nonsingular. Note that $\mathcal{K}$ is similar to $\mathcal{S}$ and, since the symplectic matrices form a group, $\mathcal{K}$ is symplectic. ∎

We start our construction of a BVD-pencil by introducing $G := K_{21}K_{11}^{-1}$, satisfying $G = G^\star$ because $\mathcal{K}$ is symplectic. Let $m = \operatorname{rank} G$, and let $G = BR^{-1}B^\star$ be a full rank decomposition, where $R = R^\star \in \mathbb{F}^{m,m}$ is nonsingular, see e.g. [8]. Now embed $\lambda I - \mathcal{K}$ into a matrix pencil of the form

$$\lambda\tilde{\mathcal{T}}_1 - \tilde{\mathcal{S}}_1 = \lambda \begin{bmatrix} I_n & 0 & 0 \\ 0 & I_n & 0 \\ R^{-1}B^\star & 0 & 0 \end{bmatrix} - \begin{bmatrix} K_{11} & K_{12} & 0 \\ K_{21} & K_{22} & 0 \\ 0 & 0 & I_m \end{bmatrix}.$$

By this embedding we have added the eigenvalue ∞ to the pencil $\lambda I - \mathcal{K}$ with algebraic and geometric multiplicity m. It is clear that eigenvectors and invariant subspaces for the eigenvalues of $\mathcal{K}$ (and $\mathcal{S}$) can be directly obtained from those of $\lambda\tilde{\mathcal{T}}_1 - \tilde{\mathcal{S}}_1$.

A sequence of equivalence transformations applied from the left now converts $\lambda\tilde{\mathcal{T}}_1 - \tilde{\mathcal{S}}_1$ into a BVD-pencil. Multiplying from the left with

$$W_1 = \begin{bmatrix} I_n & 0 & 0 \\ -K_{21}K_{11}^{-1} & I_n & 0 \\ 0 & 0 & I_m \end{bmatrix} = \begin{bmatrix} I_n & 0 & 0 \\ -G & I_n & 0 \\ 0 & 0 & I_m \end{bmatrix}$$

gives the pencil

$$\lambda\tilde{\mathcal{T}}_2 - \tilde{\mathcal{S}}_2 := W_1(\lambda\tilde{\mathcal{T}}_1 - \tilde{\mathcal{S}}_1) = \lambda \begin{bmatrix} I_n & 0 & 0 \\ -G & I_n & 0 \\ R^{-1}B^\star & 0 & 0 \end{bmatrix} - \begin{bmatrix} K_{11} & K_{12} & 0 \\ 0 & \tilde{K}_{22} & 0 \\ 0 & 0 & I_m \end{bmatrix},$$

with $\tilde{K}_{22} = K_{22} - K_{21}K_{11}^{-1}K_{12}$. Since $\mathcal{K}$ is symplectic, we have $\tilde{K}_{22} = K_{11}^{-\star}$. Multiplying $\lambda\tilde{\mathcal{T}}_2 - \tilde{\mathcal{S}}_2$ from the left with

$$W_2 = \begin{bmatrix} K_{11}^{-1} & 0 & 0 \\ 0 & I_n & 0 \\ 0 & 0 & R \end{bmatrix}$$

gives the pencil

$$\lambda\tilde{\mathcal{T}}_3 - \tilde{\mathcal{S}}_3 := W_2(\lambda\tilde{\mathcal{T}}_2 - \tilde{\mathcal{S}}_2) = \lambda \begin{bmatrix} K_{11}^{-1} & 0 & 0 \\ -G & I_n & 0 \\ B^\star & 0 & 0 \end{bmatrix} - \begin{bmatrix} I & Q & 0 \\ 0 & K_{11}^{-\star} & 0 \\ 0 & 0 & R \end{bmatrix}$$

with $Q = Q^\star := K_{11}^{-1}K_{12}$. Multiplying $\lambda\tilde{\mathcal{T}}_3 - \tilde{\mathcal{S}}_3$ from the left with

$$W_3 = \begin{bmatrix} I_n & 0 & 0 \\ 0 & I_n & BR^{-1} \\ 0 & 0 & I_m \end{bmatrix}$$

gives the pencil

$$\lambda\tilde{\mathcal{T}}_4 - \tilde{\mathcal{S}}_4 := W_3(\lambda\tilde{\mathcal{T}}_3 - \tilde{\mathcal{S}}_3) = \lambda \begin{bmatrix} A^\star & 0 & 0 \\ 0 & I_n & 0 \\ B^\star & 0 & 0 \end{bmatrix} - \begin{bmatrix} I_n & Q & 0 \\ 0 & A & B \\ 0 & 0 & R \end{bmatrix},$$

with $A := K_{11}^{-\star} = \tilde{K}_{22}$. Finally, multiplying with the matrix $W_4 := \begin{bmatrix} 0 & I_n & 0 \\ I_n & 0 & 0 \\ 0 & 0 & I_m \end{bmatrix}$ from the left we get the BVD-pencil

$$P(\lambda) := \lambda\tilde{\mathcal{T}} - \tilde{\mathcal{S}} := W_4(\lambda\tilde{\mathcal{T}}_4 - \tilde{\mathcal{S}}_4) := \lambda \begin{bmatrix} 0 & I_n & 0 \\ A^\star & 0 & 0 \\ B^\star & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & A & B \\ I_n & Q & 0 \\ 0 & 0 & R \end{bmatrix}. \tag{25}$$

Note that the symplectic matrix $\mathcal{S}$ has no eigenvalues at 0 or ∞; thus the constructed BVD-pencil $P(\lambda) = \lambda\tilde{\mathcal{T}} - \tilde{\mathcal{S}}$ has the same finite eigenvalues as $\mathcal{S}$ plus m extra eigenvalues at ∞. For a general BVD-pencil, however, there may be more than m infinite eigenvalues. Note also that from the transformation matrices we obtain a direct relationship between the eigenvectors and deflating/invariant subspaces of $\mathcal{S}$ and $\lambda\tilde{\mathcal{T}} - \tilde{\mathcal{S}}$.

One can construct many other ways of transforming symplectic matrices into BVD-pencils. The transformation introduced above, however, is precisely the inverse of the reduction in Section 3; that is, if the reduction in Section 3 is applied to (25) we will recover $\mathcal{K}$, therefore also $\mathcal{S}$. Thus we see that BVD-pencils constitute a genuine extension of symplectic matrix structure.

We turn next to the more complicated problem of embedding general symplectic pencils into BVD-pencils. For this task we first need the following two lemmas, proved previously in [22, 23] and using the notation $\Sigma_{p,q} := \mathrm{diag}(I_p, -I_q)$.

Lemma 4 *Let $B \in \mathbb{F}^{p,2n}$ satisfy* $\operatorname{rank} B = p$ *and* $\operatorname{rank} B\mathcal{J}_n B^T = 2n_0 \le p$, *i.e., the dimension of the null space of BJ_nB^T is $\delta_0 = p - 2n_0 \ge 0$. Then there exists an invertible matrix $X \in \mathbb{F}^{2n}$ such that*

$$BX = \begin{array}{c} \begin{array}{ccc} 2n_1 & \delta_0 & p \end{array} \\ \begin{bmatrix} 0 & 0 & B_0 \end{bmatrix} \end{array}, \quad X^T\mathcal{J}_nX = \mathcal{J}_{n_1} \oplus \begin{bmatrix} 0 & 0 & I_{\delta_0} \\ 0 & \mathcal{J}_{n_0} & 0 \\ -I_{\delta_0} & 0 & 0 \end{bmatrix},$$

where $B_0 \in \mathbb{F}^{p,p}$ is nonsingular and $n_1 = n - n_0 - \delta_0$.

Lemma 5 *Let $B \in \mathbb{C}^{p,m}$ with $m \geq p$, and let $\pi, \nu \geq 0$ be integers such that $\pi + \nu = m$. Suppose that* $\operatorname{rank} B = p$ *and that the inertia index (i.e. the number of positive, negative and zero eigenvalues) of the Hermitian matrix $B\Sigma_{\pi,\nu}B^*$ is (π_0, ν_0, δ_0). Then $\pi_0 + \nu_0 + \delta_0 = p$ and there exists an invertible matrix $X \in \mathbb{C}^{m,m}$ such that*

$$BX = \begin{array}{c} \begin{array}{ccc} \pi_1+\nu_1 & \delta_0 & p \end{array} \\ \left[\begin{array}{ccc} 0 & 0 & B_0 \end{array}\right] \end{array}, \quad X^*\Sigma_{\pi,\nu}X = \Sigma_{\pi_1,\nu_1} \oplus \begin{bmatrix} & & I_{\delta_0} \\ & \Sigma_{\pi_0,\nu_0} & \\ I_{\delta_0} & & \end{bmatrix},$$

where $B_0 \in \mathbb{C}^{p,p}$ is nonsingular, $\pi_1 = \pi - \pi_0 - \delta_0 \geq 0$, and $\nu_1 = \nu - \nu_0 - \delta_0 \geq 0$.

Based on these two lemmas we now show how to transform a regular symplectic pencil $\lambda\mathcal{T} - \mathcal{S} \in \mathbb{F}^{2n,2n}$ to a (reduced) BVD-pencil.

We first consider the case when $\star = T$. Let $\mathcal{Q}_0$ be nonsingular such that $\mathcal{Q}_0^{-1}\mathcal{S} = \begin{bmatrix} W \\ 0 \end{bmatrix}$, where $W \in \mathbb{F}^{p,2n}$ has full row rank, i.e. $\operatorname{rank} W = p$. By Lemma 4, there exists a nonsingular matrix $\mathcal{X}_0$ such that

$$W\mathcal{X}_0 = \begin{array}{c} \begin{array}{ccc} 2n_1 & \delta_0 & p \end{array} \\ \left[\begin{array}{ccc} 0 & 0 & W_0 \end{array}\right] \end{array}, \quad \mathcal{X}_0^T\mathcal{J}_n\mathcal{X}_0 = \mathcal{J}_{n_1} \oplus \begin{bmatrix} 0 & 0 & I_{\delta_0} \\ 0 & \mathcal{J}_{n_0} & 0 \\ -I_{\delta_0} & 0 & 0 \end{bmatrix} =: \Delta,$$

where W_0 is nonsingular and $n = n_1 + n_0 + \delta_0$, $p = 2n_0 + \delta_0$. Let $\mathcal{Q}_1 = \mathcal{Q}_0 \operatorname{diag}(W_0, I_{2n-p})$, Then

$$\mathcal{S}_1 := \mathcal{Q}_1^{-1}\mathcal{S}\mathcal{X}_0 = \begin{array}{c} \\ p \\ 2n-p \end{array} \begin{array}{c} \begin{array}{ccc} 2n_1 & \delta_0 & p \end{array} \\ \begin{bmatrix} 0 & 0 & I_p \\ 0 & 0 & 0 \end{bmatrix} \end{array}$$

and $\mathcal{T}_1 := \mathcal{Q}_1^{-1}\mathcal{T}$ satisfies

$$\mathcal{T}_1\mathcal{J}_n\mathcal{T}_1^T = \mathcal{Q}_1^{-1}\mathcal{S}\mathcal{J}_n\mathcal{S}^T\mathcal{Q}_1^{-T} = \mathcal{S}_1(\mathcal{X}_0^{-1}\mathcal{J}_n\mathcal{X}_0^{-T})\mathcal{S}_1^T = \mathcal{S}_1\Delta\mathcal{S}_1^T = \mathcal{J}_{n_0} \oplus 0_{2(n-n_0)}.$$

If we partition $\mathcal{T}_1 = \begin{bmatrix} T_1 \\ T_2 \end{bmatrix}$ with $T_1 \in \mathbb{F}^{2n_0,2n}$, then $T_1\mathcal{J}_nT_1^T = \mathcal{J}_{n_0}$, which implies that $\operatorname{rank} T_1 = n_0$. By applying Lemma 4 again, then there exists a nonsingular matrix $\mathcal{Y}_0$ such that

$$T_1\mathcal{Y}_0 = \begin{array}{c} \begin{array}{cc} 2(n-n_0) & 2n_0 \end{array} \\ \left[\begin{array}{cc} 0 & C_0 \end{array}\right] \end{array}, \quad \mathcal{Y}_0^T\mathcal{J}_n\mathcal{Y}_0 = \mathcal{J}_{n-n_0} \oplus \mathcal{J}_{n_0}, \quad C_0\mathcal{J}_{n_0}C_0^T = \mathcal{J}_{n_0}.$$

With $\mathcal{Y}_1 = \mathcal{Y}_0 \operatorname{diag}(I_{2(n-n_0)}, C_0^{-1})$, then

$$T_1\mathcal{Y}_1 = \begin{array}{c} \begin{array}{cc} 2(n-n_0) & 2n_0 \end{array} \\ \left[\begin{array}{cc} 0 & I_{2n_0} \end{array}\right] \end{array}, \quad \mathcal{Y}_1^T\mathcal{J}_n\mathcal{Y}_1 = \mathcal{J}_{n-n_0} \oplus \mathcal{J}_{n_0}$$

and $\mathcal{T}_1\mathcal{Y}_1$ can be partitioned as

$$\mathcal{T}_1\mathcal{Y}_1 = \begin{bmatrix} 0 & I_{2n_0} \\ T_{21} & T_{22} \end{bmatrix}.$$

Then, from

$$\mathcal{J}_{n_0} \oplus 0_{2(n-n_0)} = \mathcal{T}_1\mathcal{J}_n\mathcal{T}_1^T = (\mathcal{T}_1\mathcal{Y}_1)(\mathcal{Y}_1^{-1}\mathcal{J}_n\mathcal{Y}_1^{-T})(\mathcal{T}_1\mathcal{Y}_1^{-1})^T = \begin{bmatrix} \mathcal{J}_{n_0} & \mathcal{J}_{n_0}T_{22}^T \\ T_{22}\mathcal{J}_{n_0} & T_{21}\mathcal{J}_{n-n_0}T_{21}^T \end{bmatrix}$$

we obtain

$$T_{22} = 0, \quad T_{21}\mathcal{J}_{n-n_0}T_{21}^T = 0.$$

If we partition T_{21} further as

$$T_{21} = \begin{array}{c} \delta_0 \\ 2n-p \end{array} \begin{bmatrix} \hat{T}_1 \\ \hat{T}_2 \end{bmatrix},$$

then applying the same procedure that was used to compute $\mathcal{S}_1$ from $\mathcal{S}$ now to $\hat{T}_2$ and using that $\hat{T}_2\mathcal{J}_{n-n_0}\hat{T}_2^T = 0$, we obtain

$$V_0^{-1}\hat{T}_2Y_0 = \begin{array}{c} \\ \delta_1 \\ 2n-p-\delta_1 \end{array} \begin{array}{c} \begin{array}{ccc} 2(n-n_0-\delta_1) & \delta_1 & \delta_1 \end{array} \\ \begin{bmatrix} 0 & 0 & I_{\delta_1} \\ 0 & 0 & 0 \end{bmatrix} \end{array}, \quad Y_0^T\mathcal{J}_{n-n_0}Y_0 = \mathcal{J}_{n-n_0-\delta_1} \oplus \mathcal{J}_{\delta_1},$$

where $V_0 \in \mathbb{F}^{2n-p,2n-p}$ is nonsingular. Setting $V_1 = I_{\delta_0} \oplus V_0$, using the partitioning

$$V_1^{-1}T_{21}Y_0 = \begin{bmatrix} \hat{T}_{11} & \hat{T}_{12} & \hat{T}_{13} \\ 0 & 0 & I_{\delta_1} \\ 0 & 0 & 0 \end{bmatrix},$$

and setting

$$V_2 = V_1 \operatorname{diag}\left(\begin{bmatrix} I_{\delta_0} & \hat{T}_{13} \\ 0 & I_{\delta_1} \end{bmatrix}, I_{2n-p-\delta_1} \right),$$

we obtain

$$V_2^{-1}T_{21}Y_0 = \begin{bmatrix} \hat{T}_{11} & \hat{T}_{12} & 0 \\ 0 & 0 & I_{\delta_1} \\ 0 & 0 & 0 \end{bmatrix}.$$

Because $T_{21}\mathcal{J}_{n-n_0}T_{21}^T = 0$, from

$$0 = (V_2^{-1}T_{21}Y_0)(Y_0^{-1}\mathcal{J}_{n-n_0}Y_0^{-T})(V_2^{-1}T_{21}Y_0)^T,$$

we have

$$\hat{T}_{12} = 0, \quad \hat{T}_{11}\mathcal{J}_{n-n_0-\delta_1}\hat{T}_{11}^T = 0.$$

Let $\mathcal{Q}_2 = \mathcal{Q}_1 \operatorname{diag}(I_{2n_0}, V_2)$ and $\mathcal{Y}_2 = \mathcal{Y}_1 \operatorname{diag}(Y_0, I_{2n_0})$. Then

$$\begin{aligned}
\mathcal{Q}_2^{-1}\mathcal{S}\mathcal{X}_0 &= \begin{array}{c} \\ 2n_0 \\ \delta_0 \\ \delta_1 \\ 2n-p-\delta_1 \end{array} \begin{array}{c} \begin{array}{cccc} 2n_1 & \delta_0 & 2n_0 & \delta_0 \end{array} \\ \begin{bmatrix} 0 & 0 & I_{2n_0} & 0 \\ 0 & 0 & 0 & I_{\delta_0} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{array} \\
\mathcal{X}_0^T\mathcal{J}_n\mathcal{X}_0 &= \mathcal{J}_{n_1} \oplus \begin{bmatrix} 0 & 0 & I_{\delta_0} \\ 0 & \mathcal{J}_{n_0} & 0 \\ -I_{\delta_0} & 0 & 0 \end{bmatrix} \\
\mathcal{Q}_2^{-1}\mathcal{T}\mathcal{Y}_2 &= \begin{array}{c} \\ 2n_0 \\ \delta_0 \\ \delta_1 \\ 2n-p-\delta_1 \end{array} \begin{array}{c} \begin{array}{cccc} 2(n-n_0-\delta_1) & \delta_1 & \delta_1 & 2n_0 \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & I_{2n_0} \\ \hat{T}_{11} & 0 & 0 & 0 \\ 0 & 0 & I_{\delta_1} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{array} \\
\mathcal{Y}_2^T\mathcal{J}_n\mathcal{Y}_2 &= \mathcal{J}_{n-n_0-\delta_1} \oplus \mathcal{J}_{\delta_1} \oplus \mathcal{J}_{n_0}.
\end{aligned}$$

Note that based on the block structure of V_2, $\mathcal{Q}_2^{-1}\mathcal{S}\mathcal{X}_0$ has the same form as $\mathcal{Q}_1^{-1}\mathcal{S}\mathcal{X}_0$.

Because $\lambda\mathcal{T}-\mathcal{S}$ is regular, we have $2n-p-\delta_1=0$, since otherwise the intersection of the null spaces of $\mathcal{T}^T$ and $\mathcal{S}^T$ is non-empty, indicating that $\lambda\mathcal{T}-\mathcal{S}$ is singular. Since $p=2n_0+\delta_0$, we have $2n=2n_0+\delta_0+\delta_1$ and by the number of columns in the partitioning of $\mathcal{Q}_2^{-1}\mathcal{S}X_0$, we obtain $\delta_1=2n_1+\delta_0\geq\delta_0$. Similarly, by considering the number of columns in the partitioning of $\mathcal{Q}_2^{-1}\mathcal{T}Y_2$, we have $\delta_0\geq\delta_1$. Therefore $\delta_1=\delta_0$, and the factorizations are reduced to

$$
\begin{aligned}
\mathcal{Q}_2^{-1}\mathcal{S}\mathcal{X}_0 &= \begin{array}{c} \\ 2n_0 \\ \delta_0 \\ \delta_0 \end{array}\begin{array}{c} \begin{array}{ccc} \delta_0 & 2n_0 & \delta_0 \end{array} \\ \left[\begin{array}{ccc} 0 & I_{2n_0} & 0 \\ 0 & 0 & I_{\delta_0} \\ 0 & 0 & 0 \end{array}\right] \end{array}, \quad \mathcal{X}_0^T\mathcal{J}_n\mathcal{X}_0 = \left[\begin{array}{ccc} 0 & 0 & I_{\delta_0} \\ 0 & \mathcal{J}_{n_0} & 0 \\ -I_{\delta_0} & 0 & 0 \end{array}\right], \\
\mathcal{Q}_2^{-1}\mathcal{T}\mathcal{Y}_2 &= \begin{array}{c} \\ 2n_0 \\ \delta_0 \\ \delta_0 \end{array}\begin{array}{c} \begin{array}{ccc} \delta_0 & \delta_0 & 2n_0 \end{array} \\ \left[\begin{array}{ccc} 0 & 0 & I_{2n_0} \\ 0 & 0 & 0 \\ 0 & I_{\delta_0} & 0 \end{array}\right] \end{array}, \quad \mathcal{Y}_2^T\mathcal{J}_n\mathcal{Y}_2 = \mathcal{J}_{\delta_0}\oplus\mathcal{J}_{n_0}.
\end{aligned} \tag{26}
$$

With

$$
\begin{aligned}
\mathcal{Q} &= \mathcal{Q}_2\left[\begin{array}{cccc} I_{n_0} & 0 & 0 & 0 \\ 0 & 0 & I_{n_0} & 0 \\ 0 & 0 & 0 & I_{\delta_0} \\ 0 & I_{\delta_0} & 0 & 0 \end{array}\right], \\
\mathcal{X}_1 &= \mathcal{X}_0\,\mathrm{diag}\left(\left[\begin{array}{cc} 0 & I_{\delta_0} \\ I_{n_0} & 0 \end{array}\right], I_{n_0+\delta_0}\right), \quad \mathcal{Y}_3 = \mathcal{Y}_2\left[\begin{array}{cccc} 0 & 0 & 0 & -I_{\delta_0} \\ 0 & I_{\delta_0} & 0 & 0 \\ I_{n_0} & 0 & 0 & 0 \\ 0 & 0 & I_{n_0} & 0 \end{array}\right],
\end{aligned}
$$

then $\mathcal{X}_1^T\mathcal{J}_n\mathcal{X}_1=\mathcal{Y}_3^T\mathcal{J}_n\mathcal{Y}_3=\mathcal{J}_n$, i.e., $\mathcal{X}_1$ and $\mathcal{Y}_3$ are symplectic, and

$$
\mathcal{Q}^{-1}\mathcal{T}\mathcal{Y}_3 = \left[\begin{array}{cc} I_n & 0 \\ 0 & F \end{array}\right], \quad \mathcal{Q}^{-1}\mathcal{S}\mathcal{X}_1 = \left[\begin{array}{cc} F & 0 \\ 0 & I_n \end{array}\right], \quad F = \left[\begin{array}{cc} I_{n_0} & 0 \\ 0 & 0 \end{array}\right].
$$

Define $\mathcal{X}=\mathcal{X}_1^{-1}\mathcal{Y}_3$, which is again symplectic. Then

$$
\lambda\mathcal{T}_1-\mathcal{S}_1 := \mathcal{Q}^{-1}(\lambda\mathcal{T}-\mathcal{S})\mathcal{Y}_3 = \lambda\left[\begin{array}{cc} I_n & 0 \\ 0 & F \end{array}\right] - \left[\begin{array}{cc} F & 0 \\ 0 & I_n \end{array}\right]\mathcal{X}. \tag{27}
$$

By using Lemma 3, $\mathcal{X}$ has a factorization

$$
\mathcal{X} = \left[\begin{array}{cc} I_n & D \\ 0 & I_n \end{array}\right]\left[\begin{array}{cc} I_n & 0 \\ G & I_n \end{array}\right]\left[\begin{array}{cc} K_{11} & 0 \\ 0 & K_{11}^{-T} \end{array}\right]\left[\begin{array}{cc} I_n & Q \\ 0 & I_n \end{array}\right]\left[\begin{array}{cc} I_n & -D \\ 0 & I_n \end{array}\right],
$$

where D is a diagonal matrix with diagonal elements either 1 or 0, and $G=G^T$, $Q=Q^T$.

By pre- and post-multiplying $\lambda\mathcal{T}_1-\mathcal{S}_1$ by

$$
\left[\begin{array}{cc} I_n & -FD \\ 0 & I_n \end{array}\right] \quad \text{and} \quad \left[\begin{array}{cc} I_n & D \\ 0 & I_n \end{array}\right]\left[\begin{array}{cc} I_n & -Q \\ 0 & I_n \end{array}\right]\left[\begin{array}{cc} K_{11}^{-1} & 0 \\ 0 & I_n \end{array}\right],
$$

respectively, and using the fact that $FD = DF$ is diagonal and $F^2 = F$, it follows that by permuting the block-columns we finally have obtained the reduced BVD-pencil

$$\lambda\hat{\mathcal{T}} - \hat{\mathcal{S}} = \begin{bmatrix} (I-F)D - Q & K_{11}^{-1} \\ F & 0 \end{bmatrix} - \begin{bmatrix} 0 & F \\ K_{11}^{-T} & G \end{bmatrix}.$$

For the case that $\star = *$ and $\mathbb{F} = \mathbb{C}$ we again assume that $\lambda\mathcal{T} - \mathcal{S}$ is regular and satisfies

$$\mathcal{S}\mathcal{J}_n\mathcal{S}^* = \mathcal{T}\mathcal{J}_n\mathcal{T}^*.$$

With

$$\Pi_n = \frac{\sqrt{2}}{2}\begin{bmatrix} I_n & I_n \\ iI_n & -iI_n \end{bmatrix}$$

we then have

$$\Pi_n^*\Pi_n = I_{2n}, \quad \Pi_n^*\mathcal{J}_n\Pi_n = i\Sigma_{n,n},$$

and for $\tilde{\mathcal{S}} = \mathcal{S}\Pi_n$, $\tilde{\mathcal{T}} = \mathcal{T}\Pi_n$ it follows that

$$\tilde{\mathcal{S}}\Sigma_{n,n}\tilde{\mathcal{S}}^* = \tilde{\mathcal{T}}\Sigma_{n,n}\tilde{\mathcal{T}}^*.$$

Similarly as in the previous case, by using Lemma 5 we have nonsingular matrices $\tilde{\mathcal{Q}}_2$, $\tilde{\mathcal{X}}_0$, and $\tilde{\mathcal{Y}}_2$ such that

$$\begin{aligned} \tilde{\mathcal{Q}}_2^{-1}\tilde{\mathcal{S}}\tilde{\mathcal{X}}_0 &= \begin{array}{c} \\ 2n_0 \\ \delta_0 \\ \delta_0 \end{array}\!\!\begin{array}{c} \begin{array}{ccc} \delta_0 & 2n_0 & \delta_0 \end{array} \\ \begin{bmatrix} 0 & I_{2n_0} & 0 \\ 0 & 0 & I_{\delta_0} \\ 0 & 0 & 0 \end{bmatrix} \end{array}, \quad \tilde{\mathcal{X}}_0^*\Sigma_{n,n}\tilde{\mathcal{X}}_0 = \begin{bmatrix} 0 & 0 & I_{\delta_0} \\ 0 & \Sigma_{n_0,n_0} & 0 \\ I_{\delta_0} & 0 & 0 \end{bmatrix}, \\ \tilde{\mathcal{Q}}_2^{-1}\tilde{\mathcal{T}}\tilde{\mathcal{Y}}_2 &= \begin{array}{c} \\ 2n_0 \\ \delta_0 \\ \delta_0 \end{array}\!\!\begin{array}{c} \begin{array}{ccc} \delta_0 & \delta_0 & 2n_0 \end{array} \\ \begin{bmatrix} 0 & 0 & I_{2n_0} \\ 0 & 0 & 0 \\ 0 & I_{\delta_0} & 0 \end{bmatrix} \end{array}, \quad \tilde{\mathcal{Y}}_2^*\Sigma_{n,n}\tilde{\mathcal{Y}}_2 = \Sigma_{\delta_0,\delta_0} \oplus \Sigma_{n_0,n_0}. \end{aligned}$$

With

$$\mathcal{Q}_2 = \tilde{\mathcal{Q}}_2 \operatorname{diag}(\Pi_{n_0}^*, I_{2\delta_0}), \quad \mathcal{X}_0 = \Pi_n\tilde{\mathcal{X}}_0 \operatorname{diag}(iI_{\delta_0}, \Pi_{n_0}^*, I_{\delta_0}), \quad \mathcal{Y}_2 = \Pi_n\tilde{Y}_2 \operatorname{diag}(iI_{\delta_0}, I_{\delta_0}, \Pi_{n_0}^*),$$

then $\mathcal{S}$, $\mathcal{T}$, $\mathcal{Q}_2$, $\mathcal{X}_0$, and $\mathcal{Y}_2$ satisfy analogous relations to those in (26) and hence we can proceed analogously.

In this way we have shown that every regular symplectic pencil is equivalent to a reduced BVD-pencil, and any reduced BVD-pencil of the form (13) can be easily extended to a BVD-pencil by performing a full rank factorization of $\hat{H}$ and then reversing the transformation that led from (12) to (13).

For singular symplectic pencils $\lambda\mathcal{T} - \mathcal{S}$, however, there may be no equivalent reduced BVD-pencil, as shown by the following example.

Example 6 Let

$$\lambda\mathcal{T} - \mathcal{S} = \lambda\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Clearly the two coefficient matrices will remain equal under any equivalence transformation. Thus any equivalent reduced BVD-pencil

$$\lambda \begin{bmatrix} h & e \\ \bar{f} & 0 \end{bmatrix} - \begin{bmatrix} 0 & f \\ \bar{e} & g \end{bmatrix}$$

will be forced to have $h = g = 0$ and $e = f$.

But then the reduced BVD-pencil would be either $\lambda 0_2 - 0_2$ or

$$\lambda \begin{bmatrix} 0 & e \\ \bar{e} & 0 \end{bmatrix} - \begin{bmatrix} 0 & e \\ \bar{e} & 0 \end{bmatrix}, \quad e \neq 0\,,$$

which cannot be equivalent to $\lambda\mathcal{T} - \mathcal{S}$ since $\operatorname{rank}\mathcal{T} = 1 \neq 0, 2$.

Note that a symplectic matrix $\mathcal{S}$ can be consided as a symplectic pencil $\lambda I_n - \mathcal{S}$, which is just $\lambda\mathcal{T}_1 - \mathcal{S}_1$ in (27) with $F = I_n$. So the above transformation provides another way to transform $\mathcal{S}$ to a reduced BVD-pencil.

We have seen in this section that every symplectic matrix and every regular symplectic pencil can be embedded into a BVD-pencil. Note that this relationship between symplectic matrices/pencils and BVD-pencils, in particular this embedding construction, does not extend to BVD-polynomials, since there is no known notion of symplectic structure for higher degree matrix polynomials.

5 Palindromic pencils and symplectic pencils

In this section we study the relationship between symplectic matrices/pencils and anti-palindromic pencils. In particular we discuss when a symplectic matrix $\mathcal{S}$ can be factored as $\mathcal{Z}^{-1}\mathcal{Z}^\star$, hence can be represented by an anti-palindromic pencil $\lambda\mathcal{Z} - \mathcal{Z}^\star$, and conversely, when an anti-palindromic pencil is equivalent to a symplectic matrix. The equivalence of symplectic pencils and anti-palindromic pencils is also examined. These results were proved in [30] based on palindromic Kronecker canonical forms that were derived independently in [11, 30].

Theorem 7 *[30]*

1. *A complex T-symplectic matrix $\mathcal{S}$ admits a factorization $\mathcal{S} = \mathcal{Z}^{-1}\mathcal{Z}^T$ if and only if for any even integer p, the number of Jordan blocks of size p associated with the eigenvalue 1 of $\mathcal{S}$ is even.*

   Conversely, let $\mathcal{Z} \in \mathbb{C}^{n,n}$ be nonsingular. Then $\mathcal{Z}^{-1}\mathcal{Z}^T$ is similar to a T-symplectic matrix if and only if n is even, and for any odd integer q the number of blocks of size q in the (anti-palindromic) Kronecker canonical form of the pair $\lambda\mathcal{Z} - \mathcal{Z}^T$ associated with the eigenvalue 1 is even.

2. *Every complex $*$-symplectic matrix admits a factorization $\mathcal{Z}^{-1}\mathcal{Z}^*$.*

   Conversely, let $\mathcal{Z} \in \mathbb{C}^{n,n}$ be nonsingular. Then $\mathcal{Z}^{-1}\mathcal{Z}^*$ is similar to a $*$-symplectic matrix if and only if n is even.

3. *A real symplectic matrix $\mathcal{S}$ admits a real factorization $\mathcal{Z}^{-1}\mathcal{Z}^T$ if and only if for any even integer p, the number of Jordan blocks of size p associated with the eigenvalue 1 of $\mathcal{S}$ is even. Conversely, let $\mathcal{Z} \in \mathbb{R}^{n,n}$ be nonsingular. Then $\mathcal{Z}^{-1}\mathcal{Z}^T$ is similar to a real symplectic matrix if and only if n is even, and for any odd integer q the number of blocks of size q in the (anti-palindromic) Kronecker canonical form of the pair $\lambda\mathcal{Z} - \mathcal{Z}^T$ corresponding to the eigenvalue 1 is even.*

Note that to obtain corresponding results with palindromic instead of anti-palindromic pencils, we have to replace 1 by -1 and $\mathcal{Z}^{-1}\mathcal{Z}^\star$ by $-\mathcal{Z}^{-1}\mathcal{Z}^\star$.

This result can be easily extended to symplectic pencils by making use of the structured Kronecker form for symplectic pencils, see [18].

Corollary 8 1. *A real (complex) T-symplectic pencil $\lambda\mathcal{T} - \mathcal{S}$ is equivalent to a real (complex) T-anti-palindromic pencil $\lambda\mathcal{Z} - \mathcal{Z}^T$, i.e. there exist nonsingular matrices V, W such that*

$$\lambda V\mathcal{T}W - V\mathcal{S}W = \lambda\mathcal{Z} - \mathcal{Z}^T,$$

if and only if for any even integer p, the number of Jordan blocks of size p associated with the eigenvalue 1 of $\lambda\mathcal{T} - \mathcal{S}$ is even.

Conversely, the real (complex) T-anti-palindromic pencil $\lambda\mathcal{Z} - \mathcal{Z}^T$ of size n is equivalent to a real (complex) T-symplectic pencil $\lambda\mathcal{T} - \mathcal{S}$ if and only if n is even, and for any odd integer q the number of blocks of size q in the (anti-palindromic) Kronecker canonical form of the pair $\lambda\mathcal{Z} - \mathcal{Z}^T$ associated with the eigenvalue 1 is even.

2. *Every complex $*$-symplectic pencil $\lambda\mathcal{T} - \mathcal{S}$ is equivalent to a $*$-anti-palindromic pencil $\lambda\mathcal{Z} - \mathcal{Z}^*$ of even dimension. Conversely, a complex $*$-anti-palindromic pencil $\lambda\mathcal{Z} - \mathcal{Z}^*$ is equivalent to a complex $*$-symplectic pencil $\lambda\mathcal{T} - \mathcal{S}$ if and only if the size n of the pencil is even.*

Proof. The proof follows directly from the fact that ∞ and 0 are nicely paired. □

These two results show that anti-palindromic pencils and symplectic matrices/pencils of even size are closely related, with differences mainly due to partial multiplicities associated with the eigenvalue 1. Symplectic matrices and pencils of odd size do not exist, so in this sense the notion of palindromic structure represents a significant extension of the idea of symplectic structure.

6 BVD and palindromic matrix polynomials

It was shown in section 3 of this paper, and earlier in [29, 33], how to obtain palindromic formulations of the optimal control problem. Thus BVD and palindromic matrix polynomials are clearly related to each other, albeit indirectly, via the control problem. In this section, however, we show several ways to algebraically transform BVD-polynomials directly into palindromic polynomials. One result of this development is that the "almost" symplectic eigen-symmetry of BVD-polynomials is revealed; their spectra have perfect reciprocal pairing except for the eigenvalues 0 and ∞. Consequently BVD and palindromic polynomials cannot be strictly equivalent, and to convert one into the other requires the use of non-equivalence, indeed even non-unimodular, transformations. Extensive use of simple block-diagonal transformations that are not unimodular is made throughout this section. These transformations change the spectrum of the polynomials, but only in very simple ways that can be easily tracked.

Let us begin with a $(2n+m)\times(2n+m)$ BVD-pencil of the form

$$P(\lambda) := \lambda\tilde{\mathcal{T}} - \tilde{\mathcal{S}} = \lambda \begin{bmatrix} 0 & E & 0 \\ A^\star & 0 & 0 \\ B^\star & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & A & B \\ E^\star & Q & Y \\ 0 & Y^\star & R \end{bmatrix}. \tag{28}$$

Setting $\lambda = -\mu$ and multiplying the pencil with $-\operatorname{diag}(I_n, \mu I_n, \mu I_m)$ from the left, we obtain a second order palindromic matrix polynomial

$$\hat{P}(\mu) := \mu^2 \begin{bmatrix} 0 & 0 & 0 \\ A^\star & 0 & 0 \\ B^\star & 0 & 0 \end{bmatrix} + \mu \begin{bmatrix} 0 & E & 0 \\ E^\star & Q & Y \\ 0 & Y^\star & R \end{bmatrix} + \begin{bmatrix} 0 & A & B \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{29}$$

By this procedure (which is closely connected to the so-called logarithmic reduction [17]) we have increased the degree of the matrix polynomial by one and added $n+m$ copies of the eigenvalue 0 as well as n copies of the eigenvalue ∞ to the matrix polynomial. From the relation $\det \hat{P}(\mu) = (-1)^m \mu^{m+n} \det P(-\mu)$ we see that the other eigenvalues of (29) are just the negatives of the eigenvalues of (28); the relationship between the eigenvectors and deflating subspaces is also easily obtained.

For the eigenvalue 0 the eigenvectors of $\hat{P}(\mu)$ are contained in $\mathcal{N} := \ker \begin{bmatrix} 0 & A & B \end{bmatrix}$, while those of $P(\lambda)$ are contained in

$$\mathcal{N}_0 := \ker \begin{bmatrix} 0 & A & B \\ E^\star & Q & Y \\ 0 & Y^\star & R \end{bmatrix} = \ker \begin{bmatrix} 0 & A & B \end{bmatrix} \cap \ker \begin{bmatrix} E^\star & Q & Y \\ 0 & Y^\star & R \end{bmatrix} \subseteq \mathcal{N}.$$

Hence we can obtain those of $P(\lambda)$ from those of $\hat{P}(\mu)$ by restriction to $\mathcal{N}_0$.

The sum of this deflating subspace and that associated with all nonzero eigenvalues of $P(\lambda)$ inside the unit circle forms the important *stable deflating subspace* which is used for the decoupling of the forward and backward integration in the solution of the two-point boundary value problem (10). Similar computations can be done for the left eigenvectors and deflating subspaces. Note that a different palindromic quadratic can be obtained in a similar way from $P(\lambda)$ by setting $\lambda = -\mu$ and multiplying by $-\operatorname{diag}(I_n, \mu I_n, \mu I_m)$ from the right.

For BVD-polynomials of degree $k \geq 2$ there is an analogous but slightly different construction. If we multiply the polynomial $P(\lambda)$ in (18) by $\operatorname{diag}(\lambda^{k-2} I_n, I_n, \lambda^{k-2} I_m)$ from the left and by $\operatorname{diag}(I_n, I_n, \lambda I_m)$ from the right, we obtain the palindromic matrix polynomial

$$\begin{aligned}
\hat{P}(\lambda) &:= \sum_{j=0}^{2k-2} \lambda^j \hat{\mathcal{M}}_j \\
&:= \lambda^{2k-2} \begin{bmatrix} 0 & M_k & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \lambda^{2k-3} \begin{bmatrix} 0 & M_{k-1} & 0 \\ 0 & 0 & 0 \\ 0 & Y^\star & 0 \end{bmatrix} + \lambda^{2k-4} \begin{bmatrix} 0 & M_{k-2} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \cdots \\
&+ \lambda^k \begin{bmatrix} 0 & M_2 & 0 \\ M_0^\star & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \lambda^{k-1} \begin{bmatrix} 0 & M_1 & -B \\ M_1^\star & Q & 0 \\ -B^\star & 0 & R \end{bmatrix} + \lambda^{k-2} \begin{bmatrix} 0 & M_0 & 0 \\ M_2^\star & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\
&+ \ldots + \lambda^2 \begin{bmatrix} 0 & 0 & 0 \\ M_{k-2}^\star & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} 0 & 0 & 0 \\ M_{k-1}^\star & 0 & Y \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ M_k^\star & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.
\end{aligned}$$

Here we have added $n(k-2)+m(k-1)$ copies of the eigenvalue 0, and added or deleted $|n(k-2)-m|$ copies of the eigenvalue ∞ depending on whether $n(k-2)-m$ is positive or negative. Using the block-diagonal transformation matrices we can directly relate the right eigenvectors and deflating subspaces associated with the nonzero eigenvalues of $P(\lambda)$ and $\hat{P}(\lambda)$ to each other.

For $k \geq 2$, the eigenvectors of $\hat{P}(\lambda)$ associated with the eigenvalue 0 are contained in $\mathcal{N} = \ker \begin{bmatrix} M_k^\star & 0 & 0 \end{bmatrix}$, while those of $P(\lambda)$ are contained in

$$\mathcal{N}_0 = \ker \begin{bmatrix} 0 & M_0 & -B \\ M_k^\star & 0 & Y \\ 0 & 0 & R \end{bmatrix}.$$

In this case, however, $\mathcal{N}_0 \not\subseteq \mathcal{N}$.

Let us evaluate this way of transforming BVD-polynomials into palindromic polynomials from the point of view of numerical methods. First of all, we have seen that the described procedures only add copies of the eigenvalues 0 and ∞, and that the eigenvalues, eigenvectors and deflating subspaces of the BVD-pencil/polynomial can be easily retrieved from the palindromic polynomials. Thus we have obtained a palindromic formulation that contains the desired information. Since the palindromic structure is defined via linear relations that are easier to preserve in finite arithmetic, see [14, 31, 33], this is a big advantage.

On the other hand, for $k=1$ and $k>2$ we have increased the degree of the matrix polynomial to 2 and $2k-2$, respectively. However, aside from a few exceptional cases, it is known that palindromic matrix polynomials of degree $k \geq 2$ can be expressed as linear palindromic matrix polynomials [19]. Furthermore, it has been shown in [1] how to obtain so-called trimmed linearizations that first deflate some of the possible critical eigenvalues (such as $0, 1, -1, \infty$) before carrying out the formulation as a linear eigenvalue problem. Thus, with some extra effort, we are able to express a BVD-polynomial as a palindromic pencil from which all the information concerning eigenvalues, eigenvectors and deflating subspaces can be determined. Implementation of these methods and a detailed error and perturbation analysis is currently under investigation.

There are alternatives to the constructions described so far in this section that do not increase the degree of the matrix polynomial. Consider again the BVD-pencil (28). By applying a Cayley transformation to this $P(\lambda)$, see [35], we get a pencil $\nu(\tilde{\mathcal{S}} - \hat{\mathcal{T}}) - (\tilde{\mathcal{S}} + \tilde{\mathcal{T}})$, where $\nu = (\lambda+1)/(\lambda-1)$ is the scalar Cayley transformation. By deleting the submatrix

$$\begin{bmatrix} Q & Y \\ Y^\star & R \end{bmatrix}$$

from the matrix $\tilde{\mathcal{S}} - \tilde{\mathcal{T}}$ we obtain a so-called *even pencil* ([19]),

$$P_1(\nu) = \nu\tilde{\mathcal{T}}_1 - \tilde{\mathcal{S}}_1 = \nu \begin{bmatrix} 0 & A-E & B \\ -(A-E)^\star & 0 & 0 \\ -B^\star & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & A+E & B \\ (A+E)^\star & Q & Y \\ B^\star & Y^\star & R \end{bmatrix},$$

i.e., $\tilde{\mathcal{T}}_1 = -\tilde{\mathcal{T}}_1^\star$ and $\tilde{\mathcal{S}}_1 = \tilde{\mathcal{S}}_1^\star$. This transformation was introduced in [35, 36], where the precise eigenvalue and eigenvector/deflating subspace relations between $P(\lambda)$ and $P_1(\nu)$ were derived for the case $\star = *$. The results for the complex case with $\star = T$ can be derived in a similar way. With this transformation a finite eigenvalue λ_k of $P(\lambda)$ is transformed to the eigenvalue $\nu_k = (\lambda_k + 1)/(\lambda_k - 1)$ of $P_1(\nu)$. The infinite eigenvalues of $P(\lambda)$ split into two groups: one group is transformed to the eigenvalue 1 of $P_1(\nu)$ to match a possible eigenvalue -1, and another group is transformed to the eigenvalue ∞ of $P_1(\nu)$.

Now performing the inverse Cayley transformation to $P_1(\nu)$ followed by an equivalence transformation with $\mathrm{diag}((1/2)I_n, I_{n+m})$ on both sides, we obtain the anti-palindromic pencil

$$\begin{aligned}\hat{P}(\lambda) &= \lambda \mathcal{Z} - \mathcal{Z}^\star = \begin{bmatrix} \frac{1}{2}I_n & 0 \\ 0 & I_{n+m} \end{bmatrix} (\lambda(\tilde{\mathcal{S}}_1 - \tilde{\mathcal{T}}_1) - (\tilde{\mathcal{S}}_1 + \tilde{\mathcal{T}}_1)) \begin{bmatrix} \frac{1}{2}I_n & 0 \\ 0 & I_{n+m} \end{bmatrix} \\ &= \lambda \begin{bmatrix} 0 & E & 0 \\ A^\star & Q & Y \\ B^\star & Y^\star & R \end{bmatrix} - \begin{bmatrix} 0 & A & B \\ E^\star & Q & Y \\ 0 & Y^\star & R \end{bmatrix}.\end{aligned}$$

The relation between $P(\lambda)$ and $\hat{P}(\lambda)$ can be concisely described by the simple transformations

$$\hat{P}(\lambda) = \begin{bmatrix} (1-\lambda)^{-1}I_n & 0 \\ 0 & I_{n+m} \end{bmatrix} P(\lambda) \begin{bmatrix} I_n & 0 \\ 0 & (1-\lambda)I_{n+m} \end{bmatrix},$$

or

$$\hat{P}(\lambda) = \begin{bmatrix} I_n & 0 \\ 0 & (1-\lambda)I_{n+m} \end{bmatrix} P(\lambda) \begin{bmatrix} (1-\lambda)^{-1}I_n & 0 \\ 0 & I_{n+m} \end{bmatrix},$$

from which we immediately see that $\det \hat{P}(\lambda) = (1-\lambda)^m \det P(\lambda)$. Thus m copies of the eigenvalue ∞ for $P(\lambda)$ (note that the algebraic multiplicity of ∞ is at least m) have been moved to the eigenvalue 1, and the remaining spectrum is unchanged. The advantage of this approach to converting BVD-structure into palindromic structure is that we do not increase the degree of the polynomial. Again, the relationship between the eigenvectors and deflating subspaces of the two formulations is easily obtained, see [35, 36].

A similar approach also works for the case $k \geq 2$, but we need to distinguish the even and odd degree cases. When $k = 2\ell$, by pre- and post-multiplying $P(\lambda)$ in (18) with $\mathrm{diag}(\lambda^{\ell-1}I_n, I_n, I_m)$ and $\mathrm{diag}(I_n, \lambda^{1-\ell}I_n, \lambda^\ell I_m)$, respectively, we obtain the palindromic polynomial

$$\begin{aligned}\hat{P}(\lambda) &:= \sum_{j=0}^{k} \lambda^j \hat{\mathcal{M}}_j \\ &:= \lambda^{2\ell} \begin{bmatrix} 0 & M_{2\ell} & 0 \\ M_0^\star & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \lambda^{2\ell-1} \begin{bmatrix} 0 & M_{2\ell-1} & -B \\ M_1^\star & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \lambda^{2\ell-2} \begin{bmatrix} 0 & M_{2\ell-2} & 0 \\ M_2^\star & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ &+ \cdots + \lambda^{\ell+1} \begin{bmatrix} 0 & M_{\ell+1} & 0 \\ M_{\ell-1}^\star & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \lambda^\ell \begin{bmatrix} 0 & M_\ell & 0 \\ M_\ell^\star & Q & Y \\ 0 & Y^\star & R \end{bmatrix} + \lambda^{\ell-1} \begin{bmatrix} 0 & M_{\ell-1} & 0 \\ M_{\ell+1}^\star & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ &+ \ldots + \lambda^2 \begin{bmatrix} 0 & M_2 & 0 \\ M_{2\ell-2}^\star & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} 0 & M_1 & 0 \\ M_{2\ell-1}^\star & 0 & 0 \\ -B^\star & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & M_0 & 0 \\ M_{2\ell}^\star & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.\end{aligned}$$

Note that other palindromic polynomials of degree $k = 2\ell$ can be obtained by pre- and post-multiplying $P(\lambda)$ in (18) with $\mathrm{diag}(\lambda^{\ell-1}I_n, I_n, \lambda^{\ell-j}I_m)$ and $\mathrm{diag}(I_n, \lambda^{1-\ell}I_n, \lambda^j I_m)$, respectively, for any fixed $j = 0, \ldots, \ell+1$.

When $k = 2\ell + 1$, pre- and post-multiplying $P(\lambda)$ in (18) with $\mathrm{diag}(\lambda^\ell I_n, (\lambda+1)I_n, (\lambda+1)I_m)$ and $\mathrm{diag}((\lambda+1)^{-1}I_n, \lambda^{-\ell}I_n, \lambda^\ell I_m)$, respectively, yields

$$
\begin{aligned}
\hat{P}(\lambda) &:= \sum_{j=0}^{k} \lambda^j \hat{\mathcal{M}}_j \\
&:= \lambda^{2\ell+1} \begin{bmatrix} 0 & M_{2\ell+1} & 0 \\ M_0^\star & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \lambda^{2\ell} \begin{bmatrix} 0 & M_{2\ell} & -B \\ M_1^\star & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \lambda^{2\ell-1} \begin{bmatrix} 0 & M_{2\ell-1} & 0 \\ M_2^\star & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\
&+ \cdots + \lambda^{\ell+1} \begin{bmatrix} 0 & M_{\ell+1} & 0 \\ M_\ell^\star & Q & Y \\ 0 & Y^\star & R \end{bmatrix} + \lambda^{\ell} \begin{bmatrix} 0 & M_\ell & 0 \\ M_{\ell+1}^\star & Q & Y \\ 0 & Y^\star & R \end{bmatrix} + \dots \\
&+ \lambda^2 \begin{bmatrix} 0 & M_2 & 0 \\ M_{2\ell-1}^\star & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \lambda \begin{bmatrix} 0 & M_1 & 0 \\ M_{2\ell}^\star & 0 & 0 \\ -B^\star & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & M_0 & 0 \\ M_{2\ell+1}^\star & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.
\end{aligned}
$$

In both cases the resulting palindromic polynomial has the *same* degree as $P(\lambda)$. For $k = 2\ell$, we have moved ℓm copies of the eigenvalue ∞ of $P(\lambda)$ to the eigenvalue 0 of $\hat{P}(\lambda)$. For $k = 2\ell+1$, we have moved ℓm and m copies of the eigenvalue ∞ of $P(\lambda)$ to the eigenvalues 0 and -1, respectively, of $\hat{P}(\lambda)$. The other eigenvalues remain unchanged.

Let us evaluate this alternative approach, again from the point of view of numerical methods. The big advantage compared to the previous construction is certainly that the degree of the matrix polynomial is not increased. The disadvantage is that for odd degree polynomials, the eigenvalue ∞ is also mapped to -1 in the palindromic pencil representation.

Furthermore, it is well-known that eigenvalues at ± 1 may cause difficulties for numerical methods. What is worse is that in optimal control one is typically interested in the eigenvectors/deflating subspaces associated with eigenvalues in the open/closed unit disk. So if there already exists an eigenvalue ± 1, then the multiplicity of this eigenvalue is changed and it is necessary to distinguish the original eigenvalues at ± 1 from the artificially created eigenvalues. The same is true for the eigenvalue 0 in the previous approach. There is certainly a fix for this problem in exact arithmetic. In finite precision arithmetic, however, extra difficulties may arise. A perturbation and error analysis of these potential problems is currently under investigation.

7 Conclusion

We have presented several structured eigenvalue problems (symplectic, BVD, and palindromic) that arise from linear-quadratric optimal control problems. We have discussed how these different representations are related and also the advantages and disadvantages of these representations from a numerical point of view.

In general, a palindromic representation is to be preferred because its structure is defined by simple linear relationships that can be easily retained in finite precision arithmetic. We have also discussed different palindromic representations and their advantages and disadvantages.

A variety of numerical methods for palindromic problems are currently being constructed and analyzed [2, 14, 20, 31, 32, 33], so that soon these techniques can replace the methods based on representations via symplectic matrices or symplectic pencils.

References

[1] R. Byers, V. Mehrmann, and H. Xu. Staircase forms and trimmed linearizations for structured matrix polynomials. Preprint 395, DFG Research Center MATHEON, TU Berlin, Berlin, Germany, 2007.

[2] K.-W. E. Chu, T.-M. Hwang, W.-W. Lin, and C.-T. Wu. Vibration of fast trains, palindromic eigenvalue problems and structure-preserving doubling algorithms. *J. Comput. Appl. Math.*, in press, 2007.

[3] F.M. Dopico and C.R. Johnson. Complementary bases in symplectic matrices and a proof that their determinant is one. *Linear Algebra Appl.*, pages 772–778, 2006.

[4] E. Eich-Soellner and C. Führer. *Numerical Methods in Multibody Systems.* B. G. Teubner Stuttgart, 1998.

[5] H. Faßbender. *Symplectic Methods for the Symplectic Eigenproblem.* Kluwer Academic, New York, 2000.

[6] U. Flaschka, V. Mehrmann, and D. Zywietz. An analysis of structure preserving methods for symplectic eigenvalue problems. *RAIRO Automatique Productique Informatique Industrielle*, 25:165–190, 1991.

[7] I. Gohberg, P. Lancaster, and L. Rodman. *Indefinite Linear Algebra and Applications.* Birkhäuser, Basel, 2005.

[8] G.H. Golub and C.F. Van Loan. *Matrix Computations.* Johns Hopkins University Press, Baltimore, third edition, 1996.

[9] A. Hilliges. Numerische Lösung von quadratischen Eigenwertproblemen mit Anwendung in der Schienendynamik. Diplomarbeit, TU Berlin, Inst. f. Mathematik, 2004.

[10] A. Hilliges, C. Mehl, and V. Mehrmann. On the solution of palindromic eigenvalue problems. In *Proceedings of the 4th European Congress on Computational Methods in Applied Sciences and Engineering (ECCOMAS).* Jyväskylä, Finland, 2004. CD-ROM.

[11] R.A. Horn and V. V. Sergeichuk. Canonical forms for complex matrix congruence and $*$ congruence. *Linear Algebra Appl.*, 416:1010–1032, 2006.

[12] T. Kailath. *Systems Theory.* Prentice-Hall, Englewood Cliffs, NJ, 1980.

[13] M.M. Konstantinov, D.W. Gu, V. Mehrmann, and P.Hr. Petkov. *Perturbation Theory for Matrix Equations.* Elsevier, North Holland, 2003.

[14] D. Kressner, C. Schröder, and D.S. Watkins. Implicit QR algorithms for palindromic and even eigenvalue problems. Preprint 432, TU Berlin, MATHEON, Germany, January 2008.

[15] P. Kunkel and V. Mehrmann. *Differential-Algebraic Equations. Analysis and Numerical Solution.* EMS Publishing House, Zürich, Switzerland, 2006.

[16] P. Kunkel and V. Mehrmann. Necessary and sufficient conditions in the optimal control for general nonlinear differential-algebraic equations. Preprint 355, DFG Research Center MATHEON, TU Berlin, Berlin, Germany, 2006.

[17] G. Latouche and V. Ramaswami. *Introduction to Matrix Analytic Methods in Stochastic Modeling*. SIAM Publications, Philadelphia, 1999.

[18] W.-W. Lin and V. Mehrmann und H. Xu. Canonical forms for Hamiltonian and symplectic matrices and pencils. *Linear Algebra Appl.*, 301-303:469–533, 1999.

[19] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Structured polynomial eigenvalue problems: Good vibrations from good linearizations. *SIAM J. Matrix Anal. Appl.*, 28:1029–1051, 2006.

[20] D.S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Numerical methods for palindromic eigenvalue problems: Computing the anti-triangular Schur form. Technical Report 409, DFG Research Center MATHEON, *Mathematics for key technologies* in Berlin, Str. des 17. Juni 136, D-10623 Berlin, Germany, 2007.

[21] C. Mehl, V. Mehrmann, A. Ran, and L. Rodman. Perturbation analysis of Lagrangian invariant subspaces of symplectic matrices. *Linear Mulitilin. Alg.*, to appear, 2008.

[22] C. Mehl, V. Mehrmann and H. Xu. Singular value decomposition for complex matrix triples PREPRINT 412, MATHEON, DFG Research Center *Mathematics for key technologies* in Berlin. `url: http://www.matheon.de/`. *J. Comput. Appl. Math.*, to appear, 2008.

[23] C. Mehl, V. Mehrmann and H. Xu. Structured decompositions for matrix triples: SVD-like concepts for structured matrices. In preparation.

[24] V. Mehrmann. A symplectic orthogonal method for single input or single output discrete time optimal linear quadratic control problems. *SIAM J. Matrix Anal. Appl.*, 9:221–248, 1988.

[25] V. Mehrmann. Existence, uniqueness and stability of solutions to singular, linear-quadratic control problems. *Linear Algebra Appl.*, 121:291–331, 1989.

[26] V. Mehrmann. *The Autonomous Linear Quadratic Control Problem, Theory and Numerical Solution*. Number 163 in Lecture Notes in Control and Information Sciences. Springer-Verlag, Heidelberg, July 1991.

[27] T. Pappas, A.J. Laub, and N.R. Sandell. On the numerical solution of the discrete-time algebraic Riccati equation. *IEEE Trans. Automat. Control*, AC-25:631–641, 1980.

[28] P.H. Petkov, N.D. Christov, and M.M. Konstantinov. *Computational Methods for Linear Control Systems*. Prentice-Hall, Hertfordshire, UK, 1991.

[29] L. Rodman. Bounded and stably bounded palindromic difference equations of first order. *Electr. Journ. Lin. Alg.*, 15:22–49, 2006.

[30] C. Schröder. A canonical form for palindromic pencils and palindromic factorizations. Preprint 316, TU Berlin, MATHEON, Germany, 2006.

[31] C. Schröder. A QR-like algorithm for the palindromic eigenvalue problem. Preprint 388, TU Berlin, Matheon, Germany, 2007.

[32] C. Schröder. URV decomposition based structured methods for palindromic and even eigenvalue problems. Preprint 375, TU Berlin, Matheon, Germany, March 2007. submitted.

[33] C. Schröder. *Palindromic and Even Eigenvalue Problems - Analysis and Numerical Methods*. Phd thesis, Technical University Berlin, Germany, 2008.

[34] P. Van Dooren. A generalized eigenvalue approach for solving Riccati equations. *SIAM J. Sci. Statist. Comput.*, 2:121–135, 1981.

[35] H. Xu. On equivalence of pencils from discrete-time and continuous-time control. *Linear Algebra Appl.*, 414:97–124, 2006.

[36] H. Xu. Transformations between discrete-time and continuous-time algebraic Riccati equations. *Linear Algebra Appl.*, 425:77–101, 2007.

[37] K. Zhou, J.C. Doyle, and K. Glover. *Robust and Optimal Control*. Prentice-Hall, Upper Saddle River, NJ, 1996.