Linearizations of Matrix Polynomials in Bernstein Basis

Mackey, D. Steven and Perovic, Vasilije

2014

MIMS EPrint: **2014.29**

ISSN 1749-9097

Linearizations of Matrix Polynomials in Bernstein Basis

D. Steven Mackey* Vasilije Perović*

June 6, 2014

Abstract

We discuss matrix polynomials expressed in a Bernstein basis, and the associated polynomial eigenvalue problems. Using Möbius transformations of matrix polynomials, large new families of strong linearizations are generated. We also investigate matrix polynomials that are structured with respect to a Bernstein basis, together with their associated spectral symmetries. The results in this paper apply equally well to scalar polynomials, and include the development of new companion pencils for polynomials expressed in a Bernstein basis.

Key words. matrix polynomial, Bernstein polynomials, Möbius transformation, eigenvalue, partial multiplicity sequence, spectral symmetry, companion pencil, strong linearization, structured linearization.

AMS subject classification.

1 Introduction

The now-classical scalar Bernstein polynomials were first used in [4] to provide a constructive proof of the Weierstrass approximation theorem, but since then have found numerous applications in computer-aided geometric design [5, 16, 18], interpolation and least squares problems [14, 33, 34], and statistical computing [35]. For additional applications of Bernstein polynomials, as well as for more on the historical development and current research trends related to Bernstein polynomials, see [17, 20] and the references therein.

This paper focuses on *matrix polynomials* expressed in Bernstein basis, and the associated polynomial eigenvalue problems. The classical approach for solving such eigenproblems is via a linearization, hence that notion takes a central role in this paper. We provide a rich source of *new* (strong) linearizations for matrix polynomials expressed in Bernstein basis, and outline a simple procedure to easily generate them. Further, we study the impact that various matrix polynomial structures have on its spectrum, the existence of structured linearizations, and how existing structure-preserving algorithms can be applied. Along the way, we also describe a method for generating a family of new *companion pencils* for scalar polynomials expressed in Bernstein basis. It is important to emphasize that even though all the results in this paper are stated using the language of matrix polynomials, they certainly also hold for the special case of scalar polynomials expressed in Bernstein basis. This could have numerical significance when computing with a Bernstein basis at the scalar level [5, 38].

The remainder of this paper is organized as follows. Some background on matrix polynomials and linearizations, as well as a brief review of the classical (scalar) Bernstein polynomials is given in Section 2. Section 3 then introduces matrix polynomials expressed in

*Department of Mathematics, Western Michigan University, Kalamazoo, MI 49008, USA, Email: steve.mackey@wmich.edu, vasilije.perovic@wmich.edu. Supported by National Science Foundation grant DMS-1016224.

Bernstein basis, and describes a simple method for generating strong linearizations of them. Finally, Section 4 describes some special spaces of linearizations for matrix polynomials in Bernstein basis, while in Section 5 we discuss matrix polynomials that are structured with respect to a Bernstein basis.

2 Preliminaries

Throughout this paper $\mathbb{N}$ denotes the set of non-negative integers, $\mathbb{F}$ is an arbitrary field, $\overline{\mathbb{F}}$ denotes the algebraic closure of $\mathbb{F}$, and $\mathbb{F}_\infty := \mathbb{F} \cup \{\infty\}$. The ring of all univariate polynomials with coefficients from $\mathbb{F}$ is denoted by $\mathbb{F}[\lambda]$, and the field of rational functions over $\mathbb{F}$ by $\mathbb{F}(\lambda)$.

The vector space of univariate scalar polynomials of degree at most n is denoted by $\mathcal{P}_n$, and the set $\mathcal{M} = \left\{1, x, x^2, \ldots, x^n\right\}$ is referred to as the *monomial basis* or the *standard basis* for $\mathcal{P}_n$. The space of all $n \times n$ invertible matrices with entries in $\mathbb{F}$ is denoted by $GL(n, \mathbb{F})$.

Definition 2.1. (Companion pencil of a scalar polynomial)
For a scalar polynomial $p(x)$ in $\mathcal{P}_n$, any matrix pencil $xC + D \in \mathbb{F}^{n\times n}[x]$ such that $\det(xC + D) = \alpha p(x)$ for some nonzero $\alpha \in \mathbb{F}$ is called a companion pencil for $p(x)$.

Remark 2.2. The notion of a companion pencil can be extended to any *regular* matrix polynomial. In particular, a square matrix $xC + D$ is a *companion pencil* for a regular matrix polynomial $P(x)$ if $\det(xC + D) = \alpha \det\bigl(P(x)\bigr)$, where α is a nonzero scalar. While such matrix pencils have applications, see for example [6], the focus in this paper will be on the more restricted class of companion pencils that are also linearizations (see Definition 2.8 and the commentary following it).

The matrices

$$R = R_n := \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix}_{n\times n} \quad \text{and} \quad B_{a,b} := \begin{bmatrix} 1 & -a \\ -1 & b \end{bmatrix} \tag{2.1}$$

are used throughout the paper, where $a, b \in \mathbb{F}$. When there is no risk of confusion, $B_{a,b}$ will be abbreviated to just B.

2.1 Linearizations and eigenvalues

An arbitrary $m \times n$ matrix polynomial of *grade* k can be expressed in the form

$$P(\lambda) = \sum_{i=0}^{k} \lambda^i A_i \,,$$

where $A_0, A_1, \ldots, A_k \in \mathbb{F}^{m\times n}$. Here we allow any of the coefficient matrices, including A_k, to be the zero matrix. In contrast to the *degree* of a nonzero matrix polynomial, which retains its usual meaning as the largest integer j such that coefficient of λ^j in $P(\lambda)$ is nonzero, having *grade* k indicates that the polynomial $P(\lambda)$ is to be interpreted as an element of the $\mathbb{F}$-vector space of all matrix polynomials of degree less than or equal to k. Note that by this convention the grade of a matrix polynomial is an integer at least as large as its degree. Since a polynomial of grade k can also be viewed as a polynomial of any grade higher than k, the grade under consideration must be chosen; the grade of a matrix polynomial $P(\lambda)$ thus constitutes a feature of $P(\lambda)$ in addition to its degree. Throughout this paper, then, a matrix polynomial P is always accompanied by a *choice* of grade, denoted grade(P). In the context of this paper, where most polynomials are expressed in a Bernstein basis, the choice of grade will be obvious, whereas the degree will not always be so clear. See Example 3.1 for an illustration of this point. Note that throughout the paper we reserve the word *pencil* to refer only to matrix polynomials of grade 1.

A polynomial $P(\lambda)$ is said to be *regular* if it is invertible when viewed as a matrix over $\mathbb{F}(\lambda)$, equivalently, if $\det P(\lambda) \not\equiv 0$; otherwise it is said to be *singular*. The rank of $P(\lambda)$, sometimes called the *normal rank*, is the rank of $P(\lambda)$ when viewed as a matrix with entries in the field $\mathbb{F}(\lambda)$, or equivalently, the size of the largest nonzero minor of $P(\lambda)$.

Definition 2.3. (Reversal)
Let P be a nonzero matrix polynomial of grade $k \geq 0$. The reversal *of P is the matrix polynomial* $\operatorname{rev} P$ *given by*

$$(\operatorname{rev} P)(\lambda) := \lambda^k P(1/\lambda)\,. \tag{2.2}$$

The fact that $\operatorname{grade}(P) \geq \deg(P)$ plays a key role in Definition 2.3, since it guarantees that $\operatorname{rev} P$ is also a matrix polynomial. Further, if P is a matrix polynomial of grade k expressed in the *standard basis*, i.e., $P(\lambda) = \sum_{i=0}^{k} A_i \lambda^i$, then $\operatorname{rev} P(\lambda) = \sum_{i=0}^{k} A_{k-i} \lambda^i$. Thus taking the reversal of a matrix polynomial expressed in the standard basis has the effect of simply *reversing the order of its coefficients*. That is not usually the case if a matrix polynomial is expressed in a non-standard basis.

Theorem 2.4. (Smith form (Frobenius, 1878) [19])
Let $P(\lambda)$ be an $m \times n$ matrix polynomial over an arbitrary field $\mathbb{F}$. Then there exists $r \in \mathbb{N}$, and unimodular (i.e., with nonzero constant determinant) matrix polynomials $E(\lambda)$ and $F(\lambda)$ of size $m \times m$ and $n \times n$, respectively, such that

$$E(\lambda)P(\lambda)F(\lambda) = \operatorname{diag}(d_1(\lambda), \ldots, d_{\min\{m,n\}}(\lambda)) =: D(\lambda), \tag{2.3}$$

where $d_1(\lambda), \ldots, d_r(\lambda)$ are monic, $d_{r+1}(\lambda), \ldots, d_{\min\{m,n\}}(\lambda)$ are identically zero, and $d_1(\lambda), \ldots, d_r(\lambda)$ satisfy the divisibility chain property*, that is, $d_j(\lambda)$ is a divisor of $d_{j+1}(\lambda)$ for $j = 1, \ldots, r-1$. Moreover, $D(\lambda)$ is unique.*

The nonzero diagonal elements $d_j(\lambda)$, $j = 1, \ldots, r$ in the Smith form of $P(\lambda)$ are called the *invariant factors* or *invariant polynomials* of $P(\lambda)$.

Observe that the uniqueness of the Smith form over a field $\mathbb{F}$ implies that the Smith form is insensitive to field extensions. In particular, the Smith form of P over $\mathbb{F}$ is the same as that over $\overline{\mathbb{F}}$, the algebraic closure of $\mathbb{F}$. It is sometimes more convenient to work over $\overline{\mathbb{F}}$, since the invariant polynomials can then be completely decomposed into a product of linear factors.

Definition 2.5. (Partial multiplicity sequence [32])
Let $P(\lambda)$ be an $m \times n$ matrix polynomial over a field $\mathbb{F}$ with rank r and grade k. For any $\lambda_0 \in \mathbb{F}$, the invariant polynomials $d_i(\lambda)$ of P for $1 \leq i \leq r$ can each be uniquely factored as

$$d_i(\lambda) \;=\; (\lambda - \lambda_0)^{\alpha_i}\, p_i(\lambda) \quad \text{with} \quad \alpha_i \in \mathbb{N}\,,\; p_i(\lambda_0) \neq 0\,.$$

The sequence of exponents $(\alpha_1, \alpha_2, \ldots, \alpha_r)$ satisfies the condition

$$0 \leq \alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_r$$

by the divisibility chain property of the Smith form, and is called the partial multiplicity sequence of P at λ_0*, denoted $\mathcal{J}(P, \lambda_0)$. For $\lambda_0 = \infty$, the partial multiplicity sequence $\mathcal{J}(P, \infty)$ is defined to be identical with $\mathcal{J}(\operatorname{rev} P, 0)$. Since* $\operatorname{rank}(\operatorname{rev} P) = \operatorname{rank}(P)$*, we see that the sequence $\mathcal{J}(P, \infty)$ is also of length r.*

Definition 2.6. (Eigenvalue)
An eigenvalue *of P is an element $\lambda_0 \in \overline{\mathbb{F}}_\infty$ such that $\mathcal{J}(P, \lambda_0)$ does* not *consist of all zeroes. If an eigenvalue λ_0 has $\mathcal{J}(P, \lambda_0) = (\alpha_1, \alpha_2, \ldots, \alpha_r)$, then the* algebraic multiplicity *of λ_0 is just the sum $\alpha_1 + \alpha_2 + \cdots + \alpha_r$, and the* geometric multiplicity *of λ_0 is the number of* positive *α_j's in $\mathcal{J}(P, \lambda_0)$.*

Remark 2.7. It is worth noting that the sequence $\mathcal{J}(P, \lambda_0)$ is nontrivial only for a finite subset of λ_0's in $\overline{\mathbb{F}}_\infty$. From the Smith form the following properties of partial multiplicity sequences are easily deduced, and will be used freely throughout the rest of the paper:

$$\begin{aligned} \mathcal{J}(P, \lambda_0) &= \mathcal{J}(P^T, \lambda_0) \\ \text{and } \; \mathcal{J}(cP, \lambda_0) &= \mathcal{J}(P, \lambda_0)\,, \quad \text{for all } \lambda_0 \in \overline{\mathbb{F}}_\infty \text{ and any } c \neq 0. \end{aligned} \tag{2.4}$$

Now given a matrix polynomial P, we are interested in solving the associated polynomial eigenvalue problem $P(\lambda)x = 0$, that is, to find eigenpairs (λ, x) such that $P(\lambda)x = 0$ holds. The classical approach to this is to first reduce the given polynomial eigenproblem to an equivalent eigenproblem for a matrix pencil, i.e., to find a linearization for the polynomial P, as in the following definition.

Definition 2.8. (Linearization)
Let $P(\lambda)$ be an $m \times n$ nonzero matrix polynomial of grade $k \geq 1$. A matrix pencil $L(\lambda) = \lambda X + Y$ is a linearization *of $P(\lambda)$ if there exist unimodular matrix polynomials $E(\lambda)$, $F(\lambda)$ and $s \in \mathbb{N}$ such that*

$$E(\lambda)L(\lambda)F(\lambda) = \left[\begin{array}{c|c} P(\lambda) & 0 \\ \hline 0 & I_{s\times s} \end{array}\right]. \tag{2.5}$$

A linearization $L(\lambda)$ is called a strong linearization *if* $\operatorname{rev} L(\lambda)$ *is also a linearization of* $\operatorname{rev} P(\lambda)$.

From Definition 2.8 it follows immediately that any linearization of a regular matrix polynomial P is also a companion pencil for P. But the converse is not true. For example, consider the scalar polynomial $p(x) = (\lambda - 1)^2$ and the 2×2 pencils

$$L_1(\lambda) := \begin{bmatrix} \lambda - 1 & 0 \\ 0 & \lambda - 1 \end{bmatrix} \quad \text{and} \quad L_2(\lambda) := \begin{bmatrix} \lambda - 1 & 1 \\ 0 & \lambda - 1 \end{bmatrix}.$$

Clearly both L_1 and L_2 are companion pencils for $p(\lambda)$, but it can be shown that only L_2 is a linearization for $p(\lambda)$. This example illustrates the fact that linearizations constitute a restricted class of companion pencils. Just as for companion pencils, a linearization for a regular polynomial P always possesses the same finite eigenvalues with the same algebraic multiplicities as P. But a linearization also has the same geometric and partial multiplicities as P, which a companion pencil may not. In addition, a strong linearization for P has the same infinite eigenvalues with the same algebraic, geometric, and partial multiplicities as P. It is also worth noting that the concept of (strong) linearization extends to singular matrix polynomials, whereas that of a companion pencil does not.

2.2 Scalar Bernstein polynomials

For references on scalar Bernstein polynomials, see for example [3, 17, 36]. Here we only establish the notation we will be using, and state some relevant properties of Bernstein polynomials.

Definition 2.9. *For two distinct elements a, b of an arbitrary field $\mathbb{F}$, we define the* Bernstein polynomials of grade $n \geq 1$ *to be of the form*

$$\beta_{i,n}(\lambda; a, b) := \frac{1}{(b-a)^n}\binom{n}{i}(\lambda - a)^i (b-\lambda)^{n-i}, \quad \textit{for } i = 0, 1, \ldots, n. \tag{2.6}$$

The binomial coefficient $\binom{n}{i}$ in (2.6) is to be interpreted [15] as the sum of the identity element in $\mathbb{F}$ taken $\frac{n!}{i!(n-i)!}$ times, while $\frac{1}{(b-a)^n}$ denotes the multiplicative inverse of $(b-a)^n$ in $\mathbb{F}$.

The most interesting and familiar case of Bernstein polynomials is for $\mathbb{F} = \mathbb{R}$ and finite a and b, which is the case originally considered by Bernstein himself [4]. Note that there are $n+1$ Bernstein polynomials of degree n. We will often use the shorter notation $\beta_{i,n}(\lambda)$ instead of $\beta_{i,n}(\lambda; a, b)$, unless the truth of a statement depends on a particular choice of a and b.

It is often convenient to consider the *scaled Bernstein polynomials* of degree n, denoted $\phi_{i,n}(\lambda; a, b)$, and defined as

$$\phi_{i,n}(\lambda; a, b) := (\lambda - a)^i (b - \lambda)^{n-i}. \tag{2.7}$$

Again, we will often use the abbreviated notation $\phi_{i,n}(\lambda)$ for the scaled Bernstein polynomials, whenever there is no risk of confusion. In particular, scaled Bernstein polynomials of *grade one* will be used extensively in this paper, and so are highlighted here for ease of reference:

$$\phi_{0,1}(\lambda; a, b) := b - \lambda \qquad \text{and} \qquad \phi_{1,1}(\lambda; a, b) := \lambda - a. \tag{2.8}$$

Remark 2.10. It is important to emphasize that Bernstein and scaled Bernstein polynomials have some nontrivial differences. For example, Bernstein polynomials form a partition of unity over $\mathbb{F}$, which is an important fact in applications where $\mathbb{F} = \mathbb{R}$. On the other hand, when $\mathbb{F}$ is a field of finite characteristic, then Bernstein polynomials of grade n may no longer form a basis for $\mathcal{P}_n$, since some of the binomial coefficients $\binom{n}{i}$ in (2.6) may be zero. By contrast, scaled Bernstein polynomials of grade n always form a basis for $\mathcal{P}_n$, over any field $\mathbb{F}$. Thus we establish the following convention for the rest of the paper — any result stated for a Bernstein basis is to be understood to be for an arbitrary field of characteristic zero, while results for a scaled Bernstein basis hold for every field.

The following proposition gathers together some well known and easily proven facts about (scaled) Bernstein polynomials that are important for this paper.

Proposition 2.11. (Properties of Bernstein polynomials [3, 17])

(a) *For any field $\mathbb{F}$ and any choice of distinct $a, b \in \mathbb{F}$, the scaled Bernstein polynomials $\phi_{i,n}(\lambda; a, b)$ of grade n form a* basis *for the vector space $\mathcal{P}_n$. When $\mathbb{F}$ has characteristic zero, then the Bernstein polynomials $\beta_{i,n}(\lambda; a, b)$ also form a basis for $\mathcal{P}_n$.*

(b) *The polynomials $\beta_{i,n}(\lambda; a, b)$ and $\beta_{n-i,n}(\lambda; a, b)$ are* mirror images *of each other about the midpoint $\lambda = \frac{b+a}{2}$, i.e., $\beta_{n-i,n}(b + a - \lambda; a, b) \equiv \beta_{i,n}(\lambda; a, b)$. The polynomials $\phi_{i,n}(\lambda; a, b)$ and $\phi_{n-i,n}(\lambda; a, b)$ also have this mirror image property.*

3 Matrix Polynomials and Linearizations in Bernstein Basis

An $m \times n$ matrix polynomial $P(\lambda)$ of *grade* k expressed in a Bernstein basis is of the form

$$P(\lambda) = \sum_{i=0}^{k} A_i \beta_{i,k}(\lambda) \tag{3.1}$$

where $A_i \in \mathbb{F}^{m\times n}$ for $i = 0, 1, \dots, k$. Equivalently, any matrix polynomial expressed as in (3.1) can be rewritten in terms of scaled Bernstein polynomials; i.e.,

$$P(\lambda) \;=\; \sum_{i=0}^{k} A_i \beta_{i,k}(\lambda) \;=\; \sum_{i=0}^{k} \widehat{A}_i \phi_{i,k}(\lambda)\,, \quad \text{where} \quad \widehat{A}_i := \frac{1}{(b-a)^k}\binom{k}{i} A_i\,. \tag{3.2}$$

When a matrix polynomial is expressed in Bernstein basis it is important to keep in mind the distinction between its *grade* and its *degree*. Note that the degree of a matrix polynomial is determined by finding the first leading *nonzero* matrix coefficient when the polynomial is expressed in the standard basis. However, for a matrix polynomial in Bernstein basis, none of the individual matrix coefficients determines the degree of the matrix polynomial.

Example 3.1. Consider the 2×1 matrix polynomial G of grade 5 in Bernstein basis with $a = 0$ and $b = 1$, given by

$$\begin{aligned} G(\lambda) \;=\; & \begin{bmatrix} 1 \\ -14 \end{bmatrix} \beta_{0,5}(\lambda) \,+\, \begin{bmatrix} 1 \\ -7 \end{bmatrix} \beta_{1,5}(\lambda) \,+\, \begin{bmatrix} 1 \\ -2 \end{bmatrix} \beta_{2,5}(\lambda) \\ & \qquad + \begin{bmatrix} 1 \\ 1 \end{bmatrix} \beta_{3,5}(\lambda) \,+\, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \beta_{4,5}(\lambda) \,+\, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \beta_{5,5}(\lambda)\,. \end{aligned}$$

Observe that all matrix coefficients of G in Bernstein basis are nonzero, but when rewritten in the standard basis we have

$$G(\lambda) \;=\; \begin{bmatrix} 0 \\ -20 \end{bmatrix} \lambda^2 \,+\, \begin{bmatrix} 0 \\ 35 \end{bmatrix} \lambda \,+\, \begin{bmatrix} 1 \\ -14 \end{bmatrix},$$

so that $\deg(G) = 2$ even though $\operatorname{grade}(G) = 5$. This example clearly illustrates that the degree of a matrix polynomial expressed in Bernstein basis is not at all obvious from simply looking at the coefficient matrices. Hence the use of the grade is *fundamental* when working in Bernstein basis.

3.1 Linearizations in Bernstein basis

The classical approach to solving a polynomial eigenvalue problem $P(\lambda)x = 0$ is to first find a (strong) linearization for $P(\lambda)$, and then do computations with the linearization. It is well known that if L is a strong linearization for P, then L and P have the same finite and infinite elementary divisors including partial multiplicities [12, 21, 22]. In many recently investigated examples of strong linearizations [9, 10, 11, 27], eigenvectors and minimal bases of P are easily recoverable from L, and there is a simple relation between the minimal indices of P and L. These concepts have been extensively studied when P is expressed in standard basis.

By contrast, the situation is not nearly so well understood for polynomials expressed in Bernstein basis. The eigenvalue problem for companion pencils associated to *scalar* polynomials in Bernstein basis has been studied in [24, 38]; these results have been partially extended to matrix polynomials in [1]. In this paper we not only show how to generate large families of companion pencils for scalar polynomials in Bernstein basis, but indeed how to *characterize all strong linearizations* for matrix polynomials in Bernstein basis. As a consequence we easily recover the results from [1, 24, 38] as special cases of the more general construction provided here.

The following very simple procedure for producing strong linearizations for *any* matrix polynomial expressed in Bernstein basis is the heart of this paper:

Linearization Procedure (for polynomials in Bernstein basis):

1. Given a matrix polynomial $P(\lambda) = \sum_{i=0}^{k} A_i\beta_{i,k}(\lambda)$ of grade k in Bernstein basis, first rewrite it in terms of the corresponding scaled Bernstein basis as in (3.2), i.e., $P(\lambda) = \sum_{i=0}^{k} \widehat{A}_i\phi_{i,k}(\lambda)$.

2. Define a new matrix polynomial $\widehat{P}(\lambda)$ of grade k in the standard basis, *using the same coefficients* $\widehat{A}_i$ as in step 1, i.e., $\widehat{P}(\lambda) := \sum_{i=0}^{k} \widehat{A}_i\lambda^i$.

3. Find a strong linearization $\widehat{L}(\lambda) = \lambda X + Y$ for $\widehat{P}(\lambda)$ by *any technique whatsoever*, see for example [2, 7, 8, 9, 10, 11, 22, 27, 28, 37].

4. Use exactly the *same coefficients* X and Y from step 3 to build the new pencil

$$L(\lambda) := \phi_{1,1}(\lambda)X \, + \, \phi_{0,1}(\lambda)Y \,,$$

   expressed in scaled Bernstein basis of grade 1. By Theorem 3.3 the pencil $L(\lambda)$ will be a strong linearization for $P(\lambda)$.

5. (Optional) Multiply the pencil $L(\lambda)$ from step 4 by the scalar $\left(\frac{1}{b-a}\right)$ to obtain the pencil $\beta_{1,1}(\lambda)X + \beta_{0,1}(\lambda)Y$ expressed in Bernstein basis of grade 1, which is also a strong linearization for $P(\lambda)$.

Remark 3.2. Several features of this procedure deserve special attention:

(a) Not only is the procedure simple and easy to apply, but more importantly, it leverages the large body of existing knowledge about linearizations for matrix polynomials expressed in the standard basis.

(b) The strong linearizations obtained in step 4 are expressed in scaled Bernstein basis for $\mathcal{P}_1$, in contrast to the usual practice of expressing linearizations only in the standard basis. The linearizations from step 4 can also be easily re-expressed in Bernstein basis as in step 5 of the Procedure. Since either of these linearizations are equally adequate, in this paper we emphasize the use of *strong linearizations expressed in scaled Bernstein basis*, mainly for simplicity and ease of exposition.

(c) From the point of view of doing computations on pencils, the linearizations from Step 4 (or Step 5) of the procedure are just as good as pencils expressed in the standard basis. Indeed, any algorithm that works by treating a pencil simply as a *pair of matrices* can be easily adapted to find the eigenvalues of any regular pencil of the form $L(\lambda) = \phi_{1,1}(\lambda)X \, + \, \phi_{0,1}(\lambda)Y$. For example, using the QZ algorithm to simultaneously reduce X and Y to upper triangular form T_X and T_Y, then the pencil $T(\lambda) = \phi_{1,1}(\lambda)T_X \, + \, \phi_{0,1}(\lambda)T_Y$ has the same spectrum as $L(\lambda)$. But the eigenvalues of $T(\lambda)$ are now readily computed from the diagonal entries to be

$$\lambda_i \, = \, \frac{a(T_X)_{ii} - b(T_Y)_{ii}}{(T_X)_{ii} - (T_Y)_{ii}}\,, \qquad i = 1, \, \dots \, , \, \ell\,.$$

If $(T_X)_{ii} - (T_Y)_{ii} = 0$ for some i, then of course $\lambda_i = \infty$.

To complete this section we state the theorem that justifies the Procedure, and then illustrate the Procedure with several concrete examples. The proof of Theorem 3.3 (the main result of the paper) is postponed until Section 3.2, so that all the necessary background can first be established.

Theorem 3.3. (Strong linearizations in Bernstein basis)
Let $P(\lambda)$ be an arbitrary $m \times n$ matrix polynomial of grade k expressed in Bernstein basis, i.e.,

$$P(\lambda) = \sum_{i=0}^{k} A_i \beta_{i,k}(\lambda) = \sum_{i=0}^{k} \widehat{A}_i \phi_{i,k}(\lambda)\,, \qquad \text{where} \qquad \widehat{A}_i := \frac{1}{(b-a)^k} \binom{k}{i} A_i \tag{3.3}$$

for $i = 0, \dots, k$. Then a pencil $L(\lambda) = \phi_{1,1}(\lambda)X + \phi_{0,1}(\lambda)Y$ is a strong linearization for $P(\lambda)$ if and only if *the pencil $\widehat{L}(\lambda) := \lambda X + Y$ is a strong linearization for the polynomial $\widehat{P}(\lambda) := \sum_{i=0}^{k} \widehat{A}_i \lambda^i$.*

It is worth stressing that Theorem 3.3 establishes a *bijection* between the set of *all* strong linearizations of P and the set of *all* strong linearizations of $\widehat{P}$. This bijection is the basis for the Linearization Procedure, which we now illustrate with several examples.

Example 3.4. Let $P(\lambda) = \sum_{i=0}^{k} \widehat{A}_i \phi_{i,k}(\lambda)$ be an $m \times n$ matrix polynomial of grade k, and consider the new matrix polynomial $\widehat{P}(\lambda) := \sum_{i=0}^{k} \widehat{A}_i \lambda^i$ of the same grade. If

$$X_1 = \operatorname{diag}(\widehat{A}_k, I_{(k-1)n}), \qquad\qquad X_2 = \operatorname{diag}(\widehat{A}_k, I_{(k-1)m}),$$

$$Y_1 = \begin{bmatrix} \widehat{A}_{k-1} & \widehat{A}_{k-2} & \cdots & \widehat{A}_0 \\ -I_n & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & -I_n & 0 \end{bmatrix}, \quad \text{and} \quad Y_2 = \begin{bmatrix} \widehat{A}_{k-1} & -I_m & \cdots & 0 \\ \widehat{A}_{k-2} & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & -I_m \\ \widehat{A}_0 & \cdots & 0 & 0 \end{bmatrix}, \tag{3.4}$$

then $\widehat{C}_1(\lambda) := \lambda X_1 + Y_1$ and $\widehat{C}_2(\lambda) := \lambda X_2 + Y_2$ are the first and the second Frobenius companion forms for $\widehat{P}(\lambda)$, respectively. Since $\widehat{C}_1(\lambda)$ and $\widehat{C}_2(\lambda)$ are always strong linearizations, for regular or singular $\widehat{P}(\lambda)$ [11, 21], then by Theorem 3.3 the pencils

$$\phi_{1,1}(\lambda)X_i \,+\, \phi_{0,1}(\lambda)Y_i \quad \text{for} \quad i = 1, 2 \tag{3.5}$$

are strong linearizations of $P(\lambda)$.

Note that when $m = n$ and $k = 5$, the pencil in (3.5) with $i = 2$ almost recovers the example appearing in [1]. More precisely, the linearization $L(\lambda)$ from [1] is simply related to the pencil in (3.5) by

$$L(\lambda) = (R_5 \otimes I_n) \cdot D \cdot \left[\phi_{1,1}(\lambda)X_2 + \phi_{0,1}(\lambda)Y_2\right] \cdot (R_5 \otimes I_n)\,, \tag{3.6}$$

where R_5 is the 5×5 reverse identity and $D = \operatorname{diag}(1, \frac{1}{5}, \frac{1}{10}, \frac{1}{10}, \frac{1}{5}) \otimes I_n$. The fact that $L(\lambda)$ from [1] is a strong linearization for $P(\lambda) = \sum_{i=0}^{5} A_i \beta_{i,5}(\lambda)$ is now an easy consequence of Theorem 3.3; by contrast, the argument in [1] requires some nontrivial special knowledge of the LU factors of $L(\lambda)$. Furthermore, Theorem 3.3 guarantees that $L(\lambda)$ is a strong linearization for $P(\lambda)$ over an arbitrary field, whereas in [1] the underlying field is required to be algebraically closed, due to the use of the local Smith form as an essential tool in the argument.

Finally we would like to note the importance in this example of expressing $L(\lambda)$ in a scaled Bernstein basis rather than the standard basis (as was done in [1]). Doing this reveals the connection between L and the second companion form for $\widehat{P}$ more transparently; together with similar examples this helped us to clarify our view that studying strong linearizations for matrix polynomials in Bernstein basis is most effective if the linearizations themselves are also expressed in (scaled) Bernstein basis.

Example 3.5. Consider a scalar polynomial of grade n on $[0,1]$, i.e.,

$$p(\lambda) = \sum_{i=0}^{k} a_i \beta_{i,k}(\lambda;0,1) = \sum_{i=0}^{k} \widehat{a}_i \phi_{i,k}(\lambda;0,1)\,, \qquad \text{where} \qquad \widehat{a}_i = \binom{n}{i} a_i \in \mathbb{F}\,.$$

Jónsson and Vavasis in [24] and Winkler in [38] both provided the same companion pencil for $p(\lambda)$ and studied its properties. Using the notation in this paper the companion pencil in [24, 38] can be expressed as

$$L(\lambda) \;=\; R_n \cdot D_1 \cdot \left[\, \phi_{1,1}(\lambda) X_2 \;+\; \phi_{0,1}(\lambda) Y_2 \,\right] \cdot D_2 \cdot R_n\,,$$

where X_2 and Y_2 are given by (3.4), R_n is the $n \times n$ reverse identity, and

$$D_1 = \operatorname{diag}\left(\tbinom{n}{n-1}, \tbinom{n}{n-2}, \ldots, \tbinom{n}{0}\right)^{-1} \qquad \text{and} \qquad D_2 = \operatorname{diag}\left(\tbinom{n}{n}, \tbinom{n}{n-1}, \ldots, \tbinom{n}{1}\right).$$

Note that $\det(D_1 \cdot D_2) = 1$.

In contrast to [38], where a significant amount of work was needed to show that $\det L(\lambda) = p(\lambda)$, here that fact is almost immediate from Theorem 3.3. On the other hand, the approach in [24] is similar to the argument provided here, but it is only applied to an isolated case. Most importantly, though, using the procedure described in this paper one can easily generate many more companion pencils for a scalar polynomial expressed in Bernstein basis, and study their numerical properties.

Example 3.6. Let $W(\lambda) = \sum_{i=0}^{5} A_i \beta_{i,5}(\lambda) = \sum_{i=0}^{5} \widehat{A}_i \phi_{i,5}(\lambda)$ be an $n \times n$ matrix polynomial of grade five. Then applying the Linearization Procedure to pencils expressed in standard basis from [2, 30], we see that the following pencil is always a strong linearization of $W(\lambda)$:

$$L(\lambda) \;=\; \phi_{1,1}(\lambda) \begin{bmatrix} \widehat{A}_1 & I_n & & & \\ I_n & 0 & 0 & & \\ & 0 & \widehat{A}_3 & I_n & \\ & & I_n & 0 & 0 \\ & & & 0 & \widehat{A}_5 \end{bmatrix} + \phi_{0,1}(\lambda) \begin{bmatrix} \widehat{A}_0 & 0 & & & \\ 0 & 0 & I_n & & \\ & I_n & \widehat{A}_2 & 0 & \\ & & 0 & 0 & I_n \\ & & & I_n & \widehat{A}_4 \end{bmatrix}. \tag{3.7}$$

In general, for an arbitrary matrix polynomial of *odd* grade expressed in Bernstein basis, we can always find a strong linearization with the block-tridiagonal structure analogous to (3.7). If the grade of a matrix polynomial V is *even*, analogous block-tridiagonal linearizations exist provided that the leading coefficient is invertible [2]. For example, if $V(\lambda) = \sum_{i=0}^{4} A_i \beta_{i,4}(\lambda) = \sum_{i=0}^{4} \widehat{A}_i \phi_{i,4}(\lambda)$ and A_4 is nonsingular, then a variation of a pencil in [2] together with Theorem 3.3 gives a strong linearization for $V(\lambda)$:

$$\phi_{1,1}(\lambda) \begin{bmatrix} 0 & I & & \\ I & \widehat{A}_3 & & \\ & & 0 & I \\ & & I & \widehat{A}_1 \end{bmatrix} + \phi_{0,1}(\lambda) \begin{bmatrix} -\widehat{A}_4^{-1} & & & \\ & \widehat{A}_2 & I & \\ & I & 0 & \\ & & & \widehat{A}_0 \end{bmatrix}. \tag{3.8}$$

On the other hand, even if the leading coefficient of an even grade matrix polynomial is singular, then a strong linearization with block-pentadigonal structure can still be found [10].

Note that (3.7) and (3.8) are particularly suitable for *scalar* polynomials in Bernstein basis, since the invertibility of the leading coefficient is trivial to check. These low bandwidth linearizations not only are aesthetically pleasing, but they also are closely connected to the construction of various types of structure-preserving linearizations [30, 31]. In addition, they may have a significant computational payoff [13].

In the preceding examples we have seen how easy it is to apply Theorem 3.3 and the resulting Linearization Procedure to find strong linearizations for matrix polynomials expressed in Bernstein basis. Now that we have a better feel for how to generate these strong linearizations, we turn next to the proof of Theorem 3.3. The next section begins by establishing the necessary background.

3.2 Möbius transformations and the proof of Theorem 3.3

Let $P(\lambda) = \sum_{i=0}^{k} A_i \beta_{i,k}(\lambda) = \sum_{i=0}^{k} \widehat{A}_i \phi_{i,k}(\lambda)$ be an $m \times n$ matrix polynomial of grade k, and consider the new matrix polynomial $\widehat{P}(\lambda) := \sum_{i=0}^{k} \widehat{A}_i \lambda^i$ of the same grade, built using the same coefficients $\widehat{A}_i$ as appear in $P(\lambda)$. The key ingredient in the Linearization Procedure for generating strong linearizations for P turns out to be the relationship between P and $\widehat{P}$. The most effective way to capture this relationship is via a *Möbius transformation of matrix polynomials*, which we now define.

Definition 3.7. (Möbius Function and Möbius Transformation [32])
Let V be the vector space of all $m \times n$ matrix polynomials of grade k over the field $\mathbb{F}$, and let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in GL(2, \mathbb{F})$. Then

(a) *the* Möbius function on $\mathbb{F}_\infty := \mathbb{F} \cup \{\infty\}$ induced by A *is the map* $\mathsf{m}_A : \mathbb{F}_\infty \to \mathbb{F}_\infty$ *defined by*

$$\mathsf{m}_A(\lambda) := \frac{a\lambda + b}{c\lambda + d}, \tag{3.9}$$

(b) *the* Möbius transformation on V induced by A *is the map* $\mathbf{M}_A : V \to V$ *defined by*

$$\mathbf{M}_A\Big(\sum_{i=0}^{k} B_i \lambda^i\Big)(\mu) := \sum_{i=0}^{k} B_i (a\mu + b)^i (c\mu + d)^{k-i}. \tag{3.10}$$

Remark 3.8. Using a standard convention for working with the Möbius function m_A, we set $\frac{1}{0} = \infty$ and $\mathsf{m}_A(\infty) = \frac{a}{c}$. Also notice that the relation between the Möbius function and the Möbius transformation induced by A is given by

$$\mathbf{M}_A(P)(\mu) = (c\mu + d)^k P\big(\mathsf{m}_A(\mu)\big). \tag{3.11}$$

It is worth emphasizing that a Möbius transformation *acts on graded polynomials*, returning polynomials of the same grade (although the degree may increase, decrease, or stay the same, depending on the input polynomial).

Example 3.9. As an example of a Möbius transformation consider the matrix $R = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$, and let P be any matrix polynomial of grade k expressed in the standard basis. Using (3.11) it is easy to see that $\mathbf{M}_R(P) = \operatorname{rev} P$, i.e., reversal is the Möbius transformation $\mathbf{M}_R$ acting on the vector space of all matrix polynomials of grade k.

The following proposition summarizes some of the fundamental properties of Möbius transformations; proofs can be found in [32].

Proposition 3.10. *Let V be the vector space of all $m \times n$ polynomials over $\mathbb{F}$ of grade k, and let $A \in GL(2, \mathbb{F})$. Then*

(a) $\mathbf{M}_A$ *is an $\mathbb{F}$-linear operator on V, i.e.,*

$$\mathbf{M}_A(P + Q) = \mathbf{M}_A(P) + \mathbf{M}_A(Q), \quad \textit{and} \quad \mathbf{M}_A(cP) = c\mathbf{M}_A(P),$$

for any $c \in \mathbb{F}$ and for all $P, Q, \in V$.

(b) $(\mathbf{M}_A)^{-1} = \mathbf{M}_{A^{-1}}$, *so* $\mathbf{M}_A$ *is a bijection on* V*, indeed a linear isomorphism from* V *back into itself.*

(c) *Let* P *be any* $n \times n$ *matrix polynomial of grade* k*; by convention, we assume that* $\operatorname{grade}\big(\det(P)\big) = kn$*. Then* $\det\big(\mathbf{M}_A(P)\big) = \mathbf{M}_A\big(\det(P)\big)$*.*

(d) *For any* $P \in V$*, we have* $\operatorname{rank} \mathbf{M}_A(P) = \operatorname{rank}(P)$*.*

(e) *Let* P *and* Q *be matrix polynomials of grades* k *and* ℓ*, respectively. Provided that the grades of* PQ *and* $P \otimes Q$ *are chosen to be* $k + \ell$*, and* PQ *is defined, then*

$$\mathbf{M}_A(PQ) \;=\; \mathbf{M}_A(P)\mathbf{M}_A(Q)\,, \tag{3.12}$$

$$\mathbf{M}_A(P \otimes Q) \;=\; \mathbf{M}_A(P) \otimes \mathbf{M}_A(Q)\,. \tag{3.13}$$

The essential result needed to prove Theorem 3.3 is the following:

Theorem 3.11. ([32], Corollary 8.6)
Let P *be any* $m \times n$ *matrix polynomial over an arbitrary field* $\mathbb{F}$*, and* $A \in GL(2, \mathbb{F})$ *with the associated Möbius transform* $\mathbf{M}_A$*. Then* L *is a strong linearization for* P *if and only if* $\mathbf{M}_A(L)$ *is a strong linearization for* $\mathbf{M}_A(P)$*.*

This now gives all the tools needed to prove Theorem 3.3.

Proof of Theorem 3.3: Let $P(\lambda)$ be an $m \times n$ matrix polynomial of grade k expressed as in (3.3), and consider the matrix polynomial $\widehat{P}(\lambda) = \sum_{i=0}^{k} \widehat{A}_i \lambda^i$. Let $B = \left[\begin{smallmatrix} 1 & -a \\ -1 & b \end{smallmatrix}\right]$, and note that B is nonsingular since $a \neq b$. Then using Definition 3.7(b) we have

$$\mathbf{M}_B(\widehat{P})(\lambda) \;=\; \sum_{i=0}^{k} \widehat{A}_i (\lambda - a)^i (b - \lambda)^{k-i} \;=\; \sum_{i=0}^{k} \widehat{A}_i \phi_{i,k}(\lambda) \;=\; P(\lambda)\,. \tag{3.14}$$

Now let $\widehat{L}(\lambda) = \lambda X + Y$ be any pencil expressed in the standard basis. By Theorem 3.11, $\widehat{L}(\lambda)$ is a strong linearization for $\widehat{P}(\lambda)$ if and only if

$$L(\lambda) := \mathbf{M}_B(\widehat{L})(\lambda) \;=\; \phi_{1,1}(\lambda) X \;+\; \phi_{0,1}(\lambda) Y$$

is a strong linearization for $P(\lambda) = \mathbf{M}_B(\widehat{P})(\lambda)$. This completes the proof of Theorem 3.3. $\square$

The proof of Theorem 3.3 makes it clear that the essential fact underlying the Linearization Procedure is the relation of the monomial basis to any scaled Bernstein basis via a Möbius transformation. But scaled Bernstein bases are not the only polynomial bases that can be related to the standard basis via some Möbius transformation. In particular, if $G = \left[\begin{smallmatrix} t & u \\ v & w \end{smallmatrix}\right]$ is any nonsingular matrix, then the Möbius transformation $\mathbf{M}_G$ applied to the space $\mathcal{P}_k$ of 1×1 polynomials of grade k implies by Proposition 3.10(b) that the image of the monomial basis, i.e.,

$$\mathcal{G} := \Big\{ \gamma_i(\lambda) := (t\lambda + u)^i (v\lambda + w)^{k-i} \Big\}_{i=0}^{k}\,, \tag{3.15}$$

is also a basis for $\mathcal{P}_k$. Thus we have the following mild generalization of Theorem 3.3.

Corollary 3.12. *Let* $\mathcal{G} = \big\{\gamma_i(\lambda) := (t\lambda + u)^i (v\lambda + w)^{k-i}\big\}_{i=0}^{k}$ *with* $tw - vu \neq 0$ *be a polynomial basis for* $\mathcal{P}_k$*, and consider an* $m \times n$ *matrix polynomial* P *of grade* k *expressed in* $\mathcal{G}$*-basis, i.e,* $P(\lambda) = \sum_{i=0}^{k} A_i \gamma_i(\lambda)$ *where* $A_i \in \mathbb{F}^{m \times n}$ *for all* i*. Then a pencil* $L(\lambda) = (t\lambda + u)X + (v\lambda + w)Y$ *is a strong linearization for* $P(\lambda)$ if and only if *the pencil* $\lambda X + Y$ *is a strong linearization for the polynomial* $\widehat{P}(\lambda) := \sum_{i=0}^{k} A_i \lambda^i$*.*

Proof. Observe that we have $\mathbf{M}_G(\widehat{P})(\lambda) = \sum_{i=0}^{k} A_i(t\lambda+u)^i(v\lambda+w)^{k-i} = P(\lambda)$ and $\mathbf{M}_G\big(\lambda X+Y\big) = L(\lambda)$, so the desired result follows from Theorem 3.11. □

Remark 3.13. It is important to note that the Linearization Procedure described in Section 3.1 *does not work* if a "weak" linearization (i.e., a linearization that is not strong) is used in Step 3. The essential difficulty is the use of Theorem 3.11, which is only valid for strong linearizations, in the proof of Theorem 3.3. Indeed, as a consequence of Theorem 8.7 in [32] and the fact that B is not upper triangular, it follows that if $\widehat{L}(\lambda)$ in Step 3 is a weak linearization, then $L(\lambda)$ in Step 4 will *never* be even a linearization for $P(\lambda)$, let alone a strong linearization. Thus the necessity of using strong linearizations is the "weak link" in the Linearization Procedure.

4 Spaces of Strong Linearizations in Bernstein Basis

We have seen in Proposition 3.10 that for any fixed $B \in GL(2,\mathbb{F})$, in particular for $B = \left[\begin{smallmatrix} 1 & -a \\ -1 & b \end{smallmatrix}\right]$, the Möbius transformation $\mathbf{M}_B$ defines a linear isomorphism from the space V of all $m \times n$ matrix polynomials over a field $\mathbb{F}$ of fixed grade k back into itself. Thus the symbol $\mathbf{M}_B$ actually denotes a whole *family* of isomorphisms, one for each choice of size, grade, and field. In the specific case of pencils, i.e., for $k = 1$, the transformation $\mathbf{M}_B$ is simply described by the formula

$$\lambda X + Y \ \stackrel{\mathbf{M}_B}{\longmapsto}\ \phi_{1,1}(\lambda)X + \phi_{0,1}(\lambda)Y = (\lambda-a)X + (b-\lambda)Y\,, \tag{4.1}$$

which defines a bijection between the space of *all* pencils and itself. For any two partner polynomials

$$\widehat{P}(\lambda) = \sum_{i=0}^{k} \widehat{A}_i\lambda^i \quad \text{and} \quad P(\lambda) = \sum_{i=0}^{k} \widehat{A}_i\phi_{i,k}(\lambda) \tag{4.2}$$

of grade k, Theorem 3.11 shows that the bijection (4.1) of pencil space *restricts* to a bijection between the set of all strong linearizations of $\widehat{P}$ and the set of all strong linearizations of P.

Also associated to a square polynomial $\widehat{P}$ of grade k in monomial basis are several well-studied vector spaces of pencils $\widehat{L}$; these spaces have proven to be fertile sources of strong linearizations [9, 27], and are defined by

$$\begin{aligned} \mathbb{L}_1(\widehat{P}) &:= \Big\{\widehat{L}(\lambda) \,:\, \widehat{L}(\lambda)(\Lambda\otimes I) = v\otimes\widehat{P}(\lambda) \ \text{ for some } v\in\mathbb{F}^k\Big\}, \\ \mathbb{L}_2(\widehat{P}) &:= \Big\{\widehat{L}(\lambda) \,:\, (\Lambda^T\otimes I)\widehat{L}(\lambda) = w^T\otimes\widehat{P}(\lambda), \ \text{ for some } w\in\mathbb{F}^k\Big\}, \\ \text{and}\quad \mathbb{DL}(\widehat{P}) &:= \mathbb{L}_1(\widehat{P})\cap\mathbb{L}_2(\widehat{P})\,, \end{aligned} \tag{4.3}$$

where $\Lambda := \begin{bmatrix} \lambda^{k-1} & \dots & \lambda & 1 \end{bmatrix}^T$ is of grade $k-1$. In this section we introduce some analogous pencil spaces that are naturally associated to the partner polynomial P as in (4.2), and investigate the relationships among these spaces using a different restriction of the Möbius transformation $\mathbf{M}_B$.

Letting Φ denote the matrix polynomial of grade $k-1$ given by

$$\Phi := \begin{bmatrix} \phi_{k-1,k-1}(\lambda) \\ \vdots \\ \phi_{1,k-1}(\lambda) \\ \phi_{0,k-1}(\lambda) \end{bmatrix},$$

we define the following spaces of pencils L associated with a matrix polynomial $P(\lambda)$ of grade k expressed in scaled Bernstein basis as in (4.2):

$$\begin{aligned} \mathbb{B}_1(P) &:= \left\{ L(\lambda) : \, L(\lambda)(\Phi \otimes I) \; = \; v \otimes P(\lambda) \;\; \text{for some } v \in \mathbb{F}^k \right\}, \\ \mathbb{B}_2(P) &:= \left\{ L(\lambda) : \, (\Phi^T \otimes I)L(\lambda) \; = \; w^T \otimes P(\lambda) \;\; \text{for some } w \in \mathbb{F}^k \right\}, \\ \text{and} \quad \mathbb{DB}(P) &:= \mathbb{B}_1(P) \cap \mathbb{B}_2(P). \end{aligned} \tag{4.4}$$

Note that the only change from (4.3) to (4.4) is that Λ has been replaced by Φ.

From basic properties of Kronecker product it is easy to see that $\mathbb{B}_1(P)$, $\mathbb{B}_2(P)$, and $\mathbb{DB}(P)$ are all $\mathbb{F}$-vector spaces. Observe that while the definitions in (4.3) are well-adapted to having both $\widehat{L}$ and $\widehat{P}$ expressed in monomial basis, the definitions in (4.4) are more readily understood if L and P are both expressed in scaled Bernstein bases. The simple relationship between the spaces in (4.3) and those in (4.4) is revealed by the following theorem.

Theorem 4.1. *Let $P(\lambda) = \sum_{i=0}^{k} A_i \beta_{i,k}(\lambda; a, b) = \sum_{i=0}^{k} \widehat{A}_i \phi_{i,k}(\lambda; a, b)$ be an $n \times n$ matrix polynomial in Bernstein basis of grade k, and consider the partner polynomial $\widehat{P}(\lambda) := \sum_{i=0}^{k} \widehat{A}_i \lambda^i$. If $B = \left[\begin{smallmatrix} 1 & -a \\ -1 & b \end{smallmatrix}\right]$, then*

(a) *the restriction of the Möbius transformation $\mathbf{M}_B$ to $\mathbb{L}_1(\widehat{P})$, i.e.,*

$$\begin{aligned} \mathbf{M}_B : \mathbb{L}_1(\widehat{P}) \quad &\longrightarrow \quad \mathbb{B}_1(P) \\ \widehat{L}(\lambda) = \lambda X + Y \quad &\longmapsto \quad \phi_{1,1}(\lambda)X + \phi_{0,1}(\lambda)Y = L(\lambda) \end{aligned} \tag{4.5}$$

is a vector space isomorphism.

(b) *the restriction of the Möbius transformation $\mathbf{M}_B$ to $\mathbb{L}_2(\widehat{P})$ is a vector space isomorphism between $\mathbb{L}_2(\widehat{P})$ and $\mathbb{B}_2(P)$.*

(c) *the restriction of the Möbius transformation $\mathbf{M}_B$ to $\mathbb{DL}(\widehat{P})$ is a vector space isomorphism between $\mathbb{DL}(\widehat{P})$ and $\mathbb{DB}(P)$.*

Proof. (a) That $\mathbf{M}_B$ is linear and injective when restricted to $\mathbb{L}_1(\widehat{P})$ follows immediately from Proposition 3.10(a) and (b). What remains is to see why this restriction is well-defined as a map into $\mathbb{B}_1(P)$, and why this map is an isomorphism.

Let $\widehat{L}(\lambda) \in \mathbb{L}_1(\widehat{P})$, and define $L(\lambda) := \mathbf{M}_B(\widehat{L}(\lambda))$. Then by (4.3) there exists a right ansatz vector $v \in \mathbb{F}^k$ such that $\widehat{L}(\lambda)\,(\Lambda \otimes I) = v \otimes \widehat{P}(\lambda)$. Applying $\mathbf{M}_B$ to each side of this equality and using properties of Möbius transformations from Proposition 3.10(e) gives

$$\begin{aligned} \mathbf{M}_B\big(\widehat{L}(\lambda)(\Lambda \otimes I_n)\big) \; &= \; \mathbf{M}_B\big(\widehat{L}(\lambda)\big)\mathbf{M}_B\big(\Lambda \otimes I_n\big) \\ &= \; L(\lambda)\big(\mathbf{M}_B(\Lambda) \otimes I_n\big) \; = \; L(\lambda)\,(\Phi \otimes I) \end{aligned} \tag{4.6}$$

and

$$\mathbf{M}_B\big(v \otimes \widehat{P}(\lambda)\big) \; = \; v \otimes \mathbf{M}_B\big(\widehat{P}(\lambda)\big) \; = \; v \otimes P(\lambda)\,. \tag{4.7}$$

Note that in (4.6) we have used the easily checked fact that

$$\mathbf{M}_B(\Lambda) \; = \; \Phi\,. \tag{4.8}$$

The equality of $\widehat{L}(\lambda)(\Lambda \otimes I_n)$ and $v \otimes \widehat{P}(\lambda)$ now implies that $L(\lambda)(\Phi \otimes I_n) \; = \; v \otimes P(\lambda)$, so $L(\lambda) = \mathbf{M}_B(\widehat{L})(\lambda) \in \mathbb{B}_1(P)$, i.e., the restricted transformation $\mathbf{M}_B$ is well defined as a map from $\mathbb{L}_1(\widehat{P})$ into $\mathbb{B}_1(P)$.

Next observe that if we apply $(\mathbf{M}_B)^{-1} = \mathbf{M}_{B^{-1}}$ to the ends of (4.6) and (4.7) we obtain

$$\mathbf{M}_{B^{-1}}\big(L(\lambda)(\Phi \otimes I)\big) \;=\; \widehat{L}(\lambda)\,(\Lambda \otimes I) \tag{4.9}$$

and

$$\mathbf{M}_{B^{-1}}\big(v \otimes P(\lambda)\big) \;=\; v \otimes \widehat{P}(\lambda)\,. \tag{4.10}$$

Thus if $L(\lambda) \in \mathbb{B}_1(P)$, then $L(\lambda)(\Phi \otimes I) = v \otimes P(\lambda)$, which implies that $\widehat{L}(\lambda) = \mathbf{M}_{B^{-1}}(L)$ satisfies the relation $\widehat{L}(\lambda)\,(\Lambda \otimes I) = v \otimes \widehat{P}(\lambda)$, that is, $\widehat{L}(\lambda) \in \mathbb{L}_1(\widehat{P})$. Consequently, we see that the restriction of $\mathbf{M}_{B^{-1}}$ to $\mathbb{B}_1(P)$ gives a well-defined linear map from $\mathbb{B}_1(P)$ back to $\mathbb{L}_1(\widehat{P})$; this provides an inverse for the map in (4.5), showing that it is indeed an isomorphism.

(b) The argument for this part is completely analogous to that given for part (a), and so is omitted.

(c) This follows immediately from parts (a) and (b). □

Remark 4.2. Theorem 4.1 has some very useful consequences. It allows us to effectively work with matrix polynomials expressed in Bernstein basis via pencils from the $\mathbb{B}_1$, $\mathbb{B}_2$, and $\mathbb{DB}$ vector spaces with essentially the same ease as using pencils from the $\mathbb{L}_1$, $\mathbb{L}_2$, and $\mathbb{DL}$ vector spaces to handle matrix polynomials in the monomial basis. In particular, it enables the immediate transfer of results already proved for the $\mathbb{L}_1$, $\mathbb{L}_2$, and $\mathbb{DL}$ spaces over to the setting of the $\mathbb{B}_1$, $\mathbb{B}_2$, and $\mathbb{DB}$ spaces. For example, it is known [9, 27] that almost every[1] pencil $\widehat{L}(\lambda) = \lambda X + Y \in \mathbb{L}_1(\widehat{P})$ is a strong linearization for $\widehat{P}(\lambda)$. As a consequence of Theorems 3.3 and 4.1 we can now conclude that almost every pencil $L(\lambda) = \phi_{1,1}(\lambda)X + \phi_{0,1}(\lambda)Y \in \mathbb{B}_1(P)$ is a strong linearization for $P(\lambda)$.

The following example illustrates the fact that the $\mathbb{B}_1$, $\mathbb{B}_2$, and $\mathbb{DB}$ vector spaces are, in general, nice sources of linearizations for matrix polynomials in Bernstein basis.

Example 4.3. Let $P(\lambda) = \sum_{i=0}^{3} A_i\beta_{i,3}(\lambda) = \sum_{i=0}^{3} \widehat{A}_i\phi_{i,3}(\lambda)$ be an $n \times n$ matrix polynomial of grade 3 expressed in Bernstein basis, and consider the partner matrix polynomial $\widehat{P}(\lambda) := \sum_{i=0}^{3} \widehat{A}_i\lambda^i$. It is easy to check [27] that the pencil $\widehat{L}(\lambda) = \lambda X + Y$, where

$$X = \begin{bmatrix} 0 & \widehat{A}_3 & 0 \\ \widehat{A}_3 & \widehat{A}_2 & 0 \\ 0 & 0 & -\widehat{A}_0 \end{bmatrix} \quad \text{and} \quad Y = \begin{bmatrix} -\widehat{A}_3 & 0 & 0 \\ 0 & \widehat{A}_1 & \widehat{A}_0 \\ 0 & \widehat{A}_0 & 0 \end{bmatrix}, \tag{4.11}$$

is in $\mathbb{DL}(\widehat{P})$ with the ansatz vector $e_2 = [0, 1, 0]^T$.

It was shown in [27] that a necessary and sufficient condition for $\widehat{L}(\lambda)$ to be a strong linearization for a *regular* $\widehat{P}$ is for $\widehat{A}_3$ and $\widehat{A}_0$ to both be nonsingular. Now by Proposition 3.10(d) we know that $P(\lambda) = \mathbf{M}_B(\widehat{P}(\lambda))$ is regular if and only if $\widehat{P}$ is regular. Thus we see from Theorem 4.1 that $L(\lambda) = \phi_{1,1}(\lambda)X + \phi_{0,1}(\lambda)Y$ is in $\mathbb{DB}(P)$, and from Theorem 3.3 that $L(\lambda)$ is a strong linearization for a regular $P(\lambda)$ if and only if $\widehat{A}_3$ and $\widehat{A}_0$, equivalently A_3 and A_0, are both nonsingular.

Remark 4.4. In contrast to Example 4.3, when $P(\lambda)$ is *singular* the space $\mathbb{DB}(P)$ is not a good source of strong linearizations. It was shown in [9] that *no pencil in* $\mathbb{DL}(\widehat{P})$ can ever be a strong linearization when $\widehat{P}$ is singular. But $\widehat{P}$ is singular if and only if P is singular by Proposition 3.10(d), and Theorem 4.1 implies that a strong linearization in $\mathbb{DB}(P)$ can only arise from some strong linearization in $\mathbb{DL}(\widehat{P})$. Thus there can not be any strong linearization in $\mathbb{DB}(P)$ whenever P is singular.

[1] Here by "almost every" we mean for all but a closed, nowhere dense set of measure zero in $\mathbb{L}_1(\widehat{P})$.

4.1 Eigenvalues

In this section we exploit the Möbius connection between partner polynomials P and $\widehat{P}$ as in (4.2) to study the relationship between their eigenvalues. This simple relationship makes it possible to obtain an analog of the Eigenvalue Exclusion Theorem [27] that is adapted to matrix polynomials expressed in Bernstein basis; it also aids the study in Section 5 of matrix polynomials that are structured with respect to Bernstein basis.

The following theorem describes the fundamental connection between the partial multiplicity sequences of two matrix polynomials related by a Möbius transformation.

Theorem 4.5. (Partial multiplicity sequences of Möbius transforms, [32])
Let $P(\lambda)$ be an $m \times n$ matrix polynomial over $\mathbb{F}$ with grade k and $\operatorname{rank} P(\lambda) = r$. *Suppose A is any nonsingular 2×2 matrix over $\mathbb{F}$, with associated Möbius transformation $\mathbf{M}_A$ and Möbius function m_A. Then for any $\mu_0 \in \overline{\mathbb{F}}_\infty$,*

$$\mathcal{J}\big(\mathbf{M}_A(P), \mu_0\big) \;\equiv\; \mathcal{J}\big(P, \mathsf{m}_A(\mu_0)\big)\,.$$

From this theorem we immediately obtain a basic result about the nodes a and b as potential eigenvalues of a matrix polynomial expressed in Bernstein basis.

Corollary 4.6. *Let P and $\widehat{P}$ be $n \times n$ matrix polynomials of grade k as in* Theorem 4.1. *Then*

$$\mathcal{J}\big(P, a\big) \;\equiv\; \mathcal{J}\big(\widehat{P}, 0\big) \qquad \text{and} \qquad \mathcal{J}\big(P, b\big) \;\equiv\; \mathcal{J}\big(\widehat{P}, \infty\big).$$

Proof. Recall from (3.14) that with $B := \left[\begin{smallmatrix} 1 & -a \\ -1 & b \end{smallmatrix}\right]$, we have $\mathbf{M}_B(\widehat{P}) = P$. Then from Theorem 4.5 the following equivalences are immediate:

$$\mathcal{J}\big(P, a\big) \;\equiv\; \mathcal{J}\big(\mathbf{M}_B(\widehat{P}), a\big) \;\equiv\; \mathcal{J}\big(\widehat{P}, \mathsf{m}_B(a)\big) \;\equiv\; \mathcal{J}\big(\widehat{P}, 0\big), \tag{4.12}$$

and

$$\mathcal{J}\big(P, b\big) \;\equiv\; \mathcal{J}\big(\mathbf{M}_B(\widehat{P}), b\big) \;\equiv\; \mathcal{J}\big(\widehat{P}, \mathsf{m}_B(b)\big) \;\equiv\; \mathcal{J}\big(\widehat{P}, \infty\big). \tag{4.13}$$

□

Remark 4.7. From Corollary 4.6 it is clear that the numbers a and b play the same role for P as 0 and ∞ play for $\widehat{P}$. Consequently, we see that b is an eigenvalue for P if and only if $\operatorname{rank} A_k < \operatorname{rank} P$; similarly, a is an eigenvalue of P if and only if $\operatorname{rank} A_0 < \operatorname{rank} P$.

Next we consider the question of determining when a pencil in $\mathbb{DB}(P)$ is a strong linearization for a given matrix polynomial P expressed in Bernstein basis. Because of the isomorphism between $\mathbb{DB}(P)$ and $\mathbb{DL}(\widehat{P})$, it is reasonable to expect that known results about $\mathbb{DL}$ spaces can be transferred over to analogous results about $\mathbb{DB}$ spaces. In order to do this, we begin by briefly recalling the fundamental theorem about pencils in $\mathbb{DL}$ spaces.

Definition 4.8. (v-polynomial, [27])
Let $v = [v_1, v_2, \ldots, v_k]^T$ be a vector in $\mathbb{F}^k$. The scalar polynomial

$$\mathsf{p}(x; v) = v_1 x^{k-1} + v_2 x^{k-2} + \cdots + v_{k-1} x + v_k\,,$$

is referred to as the "v-polynomial" of the vector v. We adopt the convention that $\mathsf{p}(x; v)$ has grade $k-1$, and hence that $\mathsf{p}(x; v)$ has a root at ∞ whenever $v_1 = 0$.

Theorem 4.9. (Eigenvalue Exclusion Theorem [27])
Suppose that $Q(\lambda)$ is a regular *matrix polynomial expressed in the standard basis and $L(\lambda) \in \mathbb{DL}(Q)$ with nonzero ansatz vector v. Then $L(\lambda)$ is a (strong) linearization for $Q(\lambda)$ if and only if no root of the v-polynomial $\mathsf{p}(x; v)$ is an eigenvalue of $Q(\lambda)$. (Note that this statement includes ∞ as one of the possible roots of $\mathsf{p}(x; v)$ or possible eigenvalues of $Q(\lambda)$.)*

As a consequence of Theorem 3.3, an analog of Theorem 4.9 for matrix polynomials expressed in Bernstein basis can now be obtained.

Theorem 4.10. (Eigenvalue Exclusion Theorem for Polynomials in Bernstein Basis) *Suppose $P(\lambda) = \sum_{i=0}^{k} A_i\beta_{i,k}(\lambda) = \sum_{i=0}^{k} \widehat{A}_i\phi_{i,k}(\lambda)$ is a* regular *matrix polynomial of grade k expressed in Bernstein basis. For a given nonzero vector $v \in \mathbb{F}^k$, let $L(\lambda) = \phi_{1,1}(\lambda)X + \phi_{0,1}(\lambda)Y$ be the unique pencil in $\mathbb{DB}(P)$ with ansatz vector v. Then $L(\lambda)$ is a (strong) linearization for $P(\lambda)$ if and only if no root of $\mathbf{M}_B\big(\mathsf{p}(x;v)\big)$ is an eigenvalue of $P(\lambda)$.*

Proof. Consider the partner matrix polynomial $\widehat{P}(\lambda) = \sum_{i=0}^{k} \widehat{A}_i\lambda^i$, so that $\mathbf{M}_B(\widehat{P}) = P(\lambda)$. Letting $\widehat{L}(\lambda) = \lambda X + Y$ be the unique pencil in $\mathbb{DL}(\widehat{P})$ with ansatz vector v, we also have $\mathbf{M}_B(\widehat{L}) = L(\lambda)$. Then Theorem 3.3 implies that $L(\lambda)$ is a strong linearization for $P(\lambda)$ if and only if $\widehat{L}(\lambda)$ is a strong linearization for $\widehat{P}(\lambda)$. By the Eigenvalue Exclusion Theorem, that is equivalent to saying that no root of $\mathsf{p}(x;v)$ is an eigenvalue of $\widehat{P}(\lambda)$. But this in turn is equivalent by Theorem 4.5 to saying that no root of $\mathbf{M}_B\big(\mathsf{p}(x;v)\big)$ is an eigenvalue of $\mathbf{M}_B(\widehat{P}) = P(\lambda)$, which completes the proof. $\square$

4.2 Eigenvector recovery

There are a number of ways to obtain eigenvectors of a regular matrix polynomial P expressed in Bernstein basis; in this section we describe two such methods. The first is based on the following result on eigenpairs of Möbius-related matrix polynomials.

Proposition 4.11. ([32, Corollary of Theorem 6.11]) *Let $P(\lambda) = \sum_{i=0}^{k} \widehat{A}_i\phi_{i,k}(\lambda)$ be a* regular *matrix polynomial of grade k, and consider the partner polynomial $\widehat{P}(\lambda) := \sum_{i=0}^{k} \widehat{A}_i\lambda^i$. If (x, λ_0) is an eigenpair for $\widehat{P}$, then $(x, \mathsf{m}_{B^{-1}}(\lambda_0))$ is an eigenpair of $P = \mathbf{M}_B(\widehat{P})$.*

This proposition enables us to leverage our knowledge of eigenvector recovery for matrix polynomials $\widehat{P}(\lambda)$ in *standard* basis and apply it to matrix polynomials $P(\lambda)$ expressed in Bernstein basis. Choose any strong linearization $\widehat{L}(\lambda) = \lambda X + Y$ for $\widehat{P}(\lambda)$ for which there is a known (preferably simple) way to recover eigenvectors of $\widehat{P}$ from eigenvectors of $\widehat{L}$; e.g., one could employ any Fiedler pencil [10], or any strong linearization from $\mathbb{L}_1(\widehat{P})$, $\mathbb{L}_2(\widehat{P})$, or $\mathbb{DL}(\widehat{P})$ [27]. Then by Proposition 4.11 any eigenpair (x, λ_0) for $\widehat{P}$ that is recovered from $\widehat{L}$ will induce a corresponding eigenpair $(x, \mathsf{m}_{B^{-1}}(\lambda_0))$ for P.

A second method is based on Theorem 3.1 in [23], which concerns any linearization $L(\lambda)$ of a polynomial $P(\lambda)$ that satisfies an identity of the form

$$L(\lambda)F(\lambda) \;=\; G(\lambda)P(\lambda)$$

for some matrix functions $F(\lambda)$ and $G(\lambda)$ of full rank. Observe that for $P(\lambda)$ in Bernstein basis, any strong linearization $L(\lambda) = \phi_{1,1}X + \phi_{0,1}Y$ in the space $\mathbb{B}_1(P)$ satisfies just such an identity, namely

$$L(\lambda)(\Phi \otimes I_n) \;=\; (v \otimes I_n)P(\lambda)\,,$$

a small variation of the defining equation (4.4). Applying Theorem 3.1 from [23] to this scenario then gives the following eigenvector recovery result.

Proposition 4.12. *Let $P(\lambda) = \sum_{i=0}^{k} \widehat{A}_i\phi_{i,k}(\lambda)$ be a regular $n \times n$ matrix polynomial in Bernstein basis over $\mathbb{C}$, and let $L(\lambda) = \phi_{1,1}X + \phi_{0,1}Y \in \mathbb{B}_1(P)$ be a strong linearization with nonzero ansatz vector $g \in \mathbb{C}^k$. Then the following holds:*

(a) *x is a right eigenvector of P with eigenvalue λ_0 if and only if $\Phi(\lambda) \otimes x$ is a right eigenvector of L with eigenvalue λ_0.*

(b) *If $y \in \mathbb{C}^n$ is a left eigenvector of L with eigenvalue μ, then $w = (g \otimes I_n)^* y$ is a left eigenvector of P with eigenvalue μ, provided that w is nonzero.*

5 Structured Matrix Polynomials in Bernstein Basis

The matrix polynomials considered so far have been quite general in the sense that the only relevant distinguishing features have been regularity/singularity and size (square or rectangular). However, matrix polynomials that arise from applications often have additional structure, at least when expressed in the monomial basis. In this section we extend classical definitions of structure to matrix polynomials expressed in Bernstein basis, and study the impact of those structures on eigenvalue pairings and the existence of structured linearizations.

Definition 5.1. ($\star$−Adjoint)
Let $\Pi = \{\pi_0, \pi_1, \ldots, \pi_k\}$ be an ordered basis for $\mathcal{P}_k$, and let $P(\lambda) = \sum_{i=0}^{k} A_i \pi_i(\lambda)$ be an $n \times n$ matrix polynomial of grade k over $\mathbb{F}$, expressed in the Π-basis. Then

$$P^\star(\lambda) := \sum_{i=0}^{k} A_i^\star \pi_i^\star(\lambda) \tag{5.1}$$

defines the $\star$-adjoint *$P^\star(\lambda)$. Here $\star$ is used as an abbreviation for transpose T when $\mathbb{F}$ is arbitrary, and either T or conjugate transpose $*$ when $\mathbb{F} = \mathbb{C}$. For a scalar polynomial $\pi(\lambda) = \sum_{j=0}^{k} \alpha_j \lambda^j$ of grade k over $\mathbb{F}$, we define*

$$\pi^\star(\lambda) := \begin{cases} \pi(\lambda) & \text{when } \star = T\,, \\ \overline{\pi}(\lambda) := \sum_{j=0}^{k} \overline{\alpha}_j \lambda^j & \text{when } \mathbb{F} = \mathbb{C} \text{ and } \star = *\,. \end{cases} \tag{5.2}$$

Remark 5.2. Note that Definition 5.1 also applies to scalars, so that for any $t \in \mathbb{F}_\infty$ we have

$$t^\star = \begin{cases} t & \text{when } \star = T \\ \bar{t} & \text{when } \mathbb{F} = \mathbb{C} \text{ and } \star = *\,, \end{cases} \tag{5.3}$$

where $\overline{\infty} := \infty$. Also when $\mathbb{F} = \mathbb{C}$ and $\star = *$ we have

$$P^\star(\lambda) = \overline{P}^T(\lambda) \qquad \text{where} \qquad \overline{P}(\lambda) := \sum_{i=0}^{k} \overline{A}_i \overline{\pi}_i(\lambda)\,. \tag{5.4}$$

With this definition of $\overline{P}(\lambda)$ we have the property $\overline{P(\lambda)} = \overline{P}(\overline{\lambda})$. This property together with (2.4), (5.4), and Lemma 4.5(c) from [32] then imply that

$$\mathcal{J}(P, \lambda_0) = \mathcal{J}(\overline{P}, \overline{\lambda}_0) = \mathcal{J}(\overline{P}^T, \overline{\lambda}_0) = \mathcal{J}(P^\star, \lambda_0^\star) \tag{5.5}$$

for any $\lambda_0 \in \mathbb{C}_\infty$ with $\star = *$. Thus we see that $\mathcal{J}(P, \lambda_0) = \mathcal{J}(P^\star, \lambda_0^\star)$ holds for any λ_0 in any extended field $\mathbb{F}_\infty$, with any choice for $\star$.

Definition 5.3. *Let $\Pi = \{\pi_0, \pi_1, \ldots, \pi_k\}$ be an ordered basis for $\mathcal{P}_k$, and consider an $n \times n$ matrix polynomial $P(\lambda) = \sum_{i=0}^{k} A_i \pi_i(\lambda)$ of grade k over $\mathbb{F}$. Then $P(\lambda)$ is said to be*

(a) Π-Hermitian *if $\mathbb{F} = \mathbb{C}$, $\star = *$, $A_i^* = A_i$, and $\pi_i^* = \overline{\pi}_i = \pi_i$ for $i = 0, \ldots, k$, (the condition on the π_i polynomials just says that Π is a* real basis *for $\mathcal{P}_k$.)*

(b) Π-palindromic *if $A_i^\star = \varepsilon \cdot A_{k-i}$ for $i = 0, \ldots, k$ and $\varepsilon = \pm 1$,*
(and when $\mathbb{F} = \mathbb{C}$ and $\star = *$, then also $\pi_i^* = \overline{\pi}_i = \pi_i$ for $i = 0, \ldots, k$)

(c) Π-alternating *if* $A_i^\star = \varepsilon \cdot (-1)^i A_i$ *for* $i = 0, \dots, k$ *and* $\varepsilon = \pm 1$.
(*and when* $\mathbb{F} = \mathbb{C}$ *and* $\star = *$, *then also* $\pi_i^* = \overline{\pi}_i = \pi_i$ *for* $i = 0, \dots, k$)

Remark 5.4. Throughout the rest of this paper matrix polynomials with any of the structures described in Definition 5.3 are collectively referred to as Π-structured. Now consider a polynomial P_1 expressed in the Π_1-basis, and a P_2 expressed in the Π_2-basis. When we say that P_1 is Π_1-structured and P_2 is Π_2-structured *of the same type*, then this is to be understood to mean that both are Hermitian, or both are palindromic, or both are alternating, each with respect to their own basis.

In particular, we investigate matrix polynomials that are structured with respect to either a Bernstein basis $\mathcal{B} = \{\beta_{0,k}(\lambda), \beta_{1,k}(\lambda), \dots, \beta_{k,k}(\lambda)\}$, a scaled Bernstein basis $\mathcal{S} = \{\phi_{0,k}(\lambda), \phi_{1,k}(\lambda), \dots, \phi_{k,k}(\lambda)\}$, or the monomial basis $\mathcal{M} = \{1, \lambda, \dots, \lambda^k\}$ for $\mathcal{P}_k$.

A few things are worth observing. First, a matrix polynomial might be structured when expressed in one basis but need not be when expressed in another basis. For example, the pencil

$$L(\lambda) = (\lambda - a)\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + (b - \lambda)\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \lambda \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} -a & b \\ -b & -a \end{bmatrix}$$

is $\mathcal{B}$-alternating but not $\mathcal{M}$-alternating. On the other hand, it is not hard to see that every $\mathcal{B}$-palindromic matrix polynomial is also $\mathcal{S}$-palindromic, and vice versa. To see why, consider an $n \times n$ matrix polynomial of grade k

$$P(\lambda) = \sum_{i=0}^{k} C_i \beta_{i,k}(\lambda) = \sum_{i=0}^{k} \widehat{C}_i \phi_{i,k}(\lambda), \quad \text{where} \quad \widehat{C}_i = \frac{1}{(b-a)^k}\binom{k}{i} C_i .$$

(Note that if $\mathbb{F} = \mathbb{C}$ and $\star = *$, then we will also assume $a, b \in \mathbb{R}$ so that $\mathcal{B}$ and $\mathcal{S}$ are *real bases*, as required by Definition 5.3.) Then

$$\begin{aligned} C_i^\star &= \varepsilon \cdot C_{k-i} \iff \\ \frac{1}{(b-a)^k}\binom{k}{i} C_i^\star &= \varepsilon \cdot \frac{1}{(b-a)^k}\binom{k}{i} C_{k-i} = \varepsilon \cdot \frac{1}{(b-a)^k}\binom{k}{k-i} C_{k-i} \\ \iff \quad \widehat{C}_i^\star &= \varepsilon \cdot \widehat{C}_{k-i} , \end{aligned} \tag{5.6}$$

where the second line uses the fact that $\binom{k}{i} = \binom{k}{k-i}$ for all $i = 0, 1, \dots, k$. The equivalence (5.6) shows that $P(\lambda)$ is $\mathcal{B}$-palindromic if and only if $P(\lambda)$ is $\mathcal{S}$-palindromic. Similar relations hold for alternating and Hermitian structures.

Proposition 5.5. *A matrix polynomial* P *is* $\mathcal{B}$*-structured if and only if it is also* $\mathcal{S}$*-structured of the same type.*

In order to more effectively study matrix polynomials that are structured with respect to an ordered basis $\Pi = \{\pi_0, \pi_1, \dots, \pi_k\}$ for $\mathcal{P}_k$, it is useful to introduce two mappings $\Psi_1, \Psi_2 : V \to V$ on the space V of all $n \times n$ matrix polynomials of grade k. These mappings are defined by

$$\Psi_1\left(\sum_{i=0}^{k} A_i \pi_i(\lambda)\right) := \sum_{i=0}^{k} A_i \pi_{k-i}(\lambda)$$

$$\text{and} \qquad \Psi_2\left(\sum_{i=0}^{k} A_i \pi_i(\lambda)\right) := \sum_{i=0}^{k} A_i (-1)^i \pi_i(\lambda) . \tag{5.7}$$

Then it is easy to see from Definition 5.3 that a matrix polynomial P is Π-palindromic if and only if $\Psi_1(P) = \varepsilon P^\star$, and $\Pi-$alternating if and only if $\Psi_2(P) = \varepsilon P^\star$. Note that the definition and meaning of Ψ_1 and Ψ_2 are heavily dependent on the choice of the basis Π.

In the special case of the monomial basis $\Pi = \mathcal{M}$ the mappings Ψ_1 and Ψ_2 are well known; Ψ_1 is just the reversal mapping $P \mapsto \operatorname{rev} P$, while Ψ_2 is the map $P(\lambda) \mapsto P(-\lambda)$. Even more significant in our context is the fact that these two mappings are Möbius transformations. In particular, $\Psi_1(P) = \operatorname{rev} P$ is $\mathbf{M}_R(P)$ and $\Psi_2(P)$ is $\mathbf{M}_S(P)$, where R and S are the nonsingular matrices

$$R = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad \text{and} \qquad S = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{5.8}$$

Thus we see that a polynomial P is $\mathcal{M}$-palindromic if and only if $\mathbf{M}_R(P) = \varepsilon P^\star$, and $\mathcal{M}$-alternating if and only if $\mathbf{M}_S(P) = \varepsilon P^\star$.

The key observation for this paper is that for either a Bernstein basis $\mathcal{B}$ or a scaled Bernstein basis $\mathcal{S}$, the maps Ψ_1 and Ψ_2 can still be realized by Möbius transformations. This is the essential content of the next result and its corollary.

Theorem 5.6. *With distinct $a, b \in \mathbb{F}$, let $\mathcal{B} = \{\beta_{0,k}(\lambda; a, b), \ldots, \beta_{k,k}(\lambda; a, b)\}$ be an ordered Bernstein basis and $\mathcal{S} = \{\phi_{0,k}(\lambda; a, b), \ldots, \phi_{k,k}(\lambda; a, b)\}$ be an ordered scaled Bernstein basis for $\mathcal{P}_k$. Define*

$$K := \begin{bmatrix} -1 & a+b \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad A := \frac{1}{a-b} \begin{bmatrix} a+b & -2ab \\ 2 & -(a+b) \end{bmatrix} \tag{5.9}$$

in $GL(2, \mathbb{F})$. Then

$$\mathbf{M}_K\big(\beta_{i,k}(\lambda)\big) = \beta_{k-i,k}(\lambda) \quad \text{and} \quad \mathbf{M}_K\big(\phi_{i,k}(\lambda)\big) = \phi_{k-i,k}(\lambda)\,, \tag{5.10}$$

while

$$\mathbf{M}_A\big(\beta_{i,k}(\lambda)\big) = (-1)^i \beta_{i,k}(\lambda) \quad \text{and} \quad \mathbf{M}_A\big(\phi_{i,k}(\lambda)\big) = (-1)^i \phi_{i,k}(\lambda) \tag{5.11}$$

for all $i = 0, 1, \ldots, k$.

Proof. The claims in (5.10) and (5.11) are easily established by straightforward computations starting from Definition 3.7. For example, it is immediate from the definition that $\mathbf{M}_K\big(\beta_{i,k}(\lambda)\big) = \beta_{i,k}(a+b-\lambda)$. But then $\beta_{i,k}(a+b-\lambda) = \beta_{k-i,k}(\lambda)$ by Proposition 2.11(b), which completes the verification of the first claim in (5.10). The results in (5.11) require nothing more than Definition 3.7 followed by some straightforward simplifications. $\square$

Remark 5.7. The idea that the matrix K in (5.9) will induce a Möbius transformation $\mathbf{M}_K$ that satisfies (5.10) is more or less obvious in light of the mirror image property $\beta_{i,k}(a + b - \lambda) = \beta_{k-i,k}(\lambda)$. On the other hand, it is not so obvious at all that there is any Möbius transformation satisfying (5.11), and if there is what the underlying matrix A might be. This question can be resolved by drawing an analogy to the case of the monomial basis $\mathcal{M}$, where the corresponding Möbius transformation is $\mathbf{M}_S$ with S as in (5.8), and the associated Möbius function is $\mathsf{m}_S(\lambda) = -\lambda$. For this function m_S the numbers 1 and -1 are interchanged, while 0 and ∞ are fixed. But recall from (4.8) that the monomial and scaled Bernstein bases are related via the Möbius transformations induced by the matrix $B = \left[\begin{smallmatrix} 1 & -a \\ -1 & b \end{smallmatrix}\right]$ and its inverse. Now by Definition 3.7 and Remark 3.8 we have

$$\mathsf{m}_{B^{-1}}(1) = \frac{a+b}{2}, \quad \mathsf{m}_{B^{-1}}(-1) = \infty, \quad \mathsf{m}_{B^{-1}}(0) = a, \quad \text{and} \quad \mathsf{m}_{B^{-1}}(\infty) = b\,. \tag{5.12}$$

Thus by analogy with m_S we should search for a Möbius function m_A that interchanges $\frac{a+b}{2}$ and ∞, leaves a and b fixed, and is an involution. Straightforward computations aimed at satisfying these requirements now lead to the Möbius function induced by the matrix A in (5.9); note that A is indeed an involution, as expected.

Based on Theorem 5.6, we now have the following characterizations of palindromic and alternating structures with respect to the $\mathcal{B}$ and $\mathcal{S}$ bases.

Corollary 5.8. *Let $P(\lambda)$ be a square matrix polynomial over $\mathbb{F}$ of grade k. Then*

(a) *P is $\mathcal{B}$-palindromic (and $\mathcal{S}$-palindromic) if and only if $\mathbf{M}_K(P)(\lambda) = \varepsilon P^\star(\lambda)$,*

(b) *P is $\mathcal{B}$-alternating (and $\mathcal{S}$-alternating) if and only if $\mathbf{M}_A(P)(\lambda) = \varepsilon P^\star(\lambda)$,*

where K and A are given by (5.9).

Proof. Both parts follow immediately from (5.7), Theorem 5.6, and the linearity of Möbius transformations of matrix polynomials. $\square$

5.1 Spectral symmetry

It is well known that additional structure in a matrix polynomial often results in certain restrictions on its spectrum, such as eigenvalue pairings. This phenomenon has been previously investigated for several types of $\mathcal{M}$-structured matrix polynomials in [28, 30, 31]. In this section we extend this investigation to $\mathcal{B}$-structured and $\mathcal{S}$-structured matrix polynomials, and discuss the underlying geometry of the resulting eigenvalue pairings.

Proposition 5.9. *Let $P(\lambda) = \sum_{i=0}^{k} A_i \beta_{i,k}(\lambda; a, b)$ with $a, b \in \mathbb{R}$ be a square matrix polynomial over $\mathbb{C}$ of grade k that is $\mathcal{B}$-Hermitian (or equivalently $\mathcal{S}$-Hermitian). Then*

$$\mathcal{J}(P, \lambda) \equiv \mathcal{J}\left(P, \overline{\lambda}\right). \tag{5.13}$$

Proof. Since the bases $\mathcal{B}$ and $\mathcal{S}$ are real, from Definitions 5.1 and 5.3(a) we see that $\mathcal{B}$-Hermitianness (or $\mathcal{S}$-Hermitianness) of $P(\lambda)$ is equivalent to the condition $P^*(\lambda) = P(\lambda)$. Thus (5.5) implies that

$$\mathcal{J}(P, \lambda) \equiv \mathcal{J}(P^*, \lambda^*) \equiv \mathcal{J}(P, \overline{\lambda}), \tag{5.14}$$

as desired. $\square$

Remark 5.10. Observe that (5.13) is a statement about *eigenvalue pairing*; whenever $\lambda_0 \in \mathbb{C}\setminus\mathbb{R}$ is an eigenvalue of a $\mathcal{B}$-Hermitian P, then so is $\overline{\lambda}_0$, with exactly the same algebraic, geometric, and partial multiplicities. In other words, eigenvalues of $\mathcal{B}$-Hermitian polynomials come in $(\lambda_0, \overline{\lambda}_0)$ pairs. But this is the same kind of eigenvalue pairing that is well known for $\mathcal{M}$-Hermitian matrix polynomials. This is not a coincidence, since the same argument used to prove Proposition 5.9 actually applies to Π-Hermitian matrix polynomials for *any* real basis Π. A polynomial $P(\lambda)$ being Π-Hermitian is equivalent to the condition $P^*(\lambda) = P(\lambda)$, and the chain of equalities in (5.14) then follows.

A different kind of eigenvalue pairing holds for $\mathcal{B}$-palindromic matrix polynomials.

Proposition 5.11. *Let $P(\lambda)$ be a square matrix polynomial over $\mathbb{F}$ of grade k that is $\mathcal{B}$–palindromic (or $\mathcal{S}$-palindromic). Then*

$$\mathcal{J}\big(P, \mu\big) \equiv \mathcal{J}\big(P, a+b-\mu^\star\big).$$

Proof. Corollary 5.8(a) implies that $\mathbf{M}_K(P)(\lambda) = \varepsilon P^\star(\lambda)$, where K is given by (5.9). Applying Theorem 4.5 then gives the following chain of equivalences:

$$\begin{aligned}\mathcal{J}(P,\mu) \;\equiv\; \mathcal{J}\left(P^\star,\mu^\star\right) \;&\equiv\; \mathcal{J}\left(\varepsilon P^\star,\mu^\star\right)\\ &\equiv\; \mathcal{J}\left(\mathbf{M}_K(P),\mu^\star\right)\\ &\equiv\; \mathcal{J}\left(P,\mathsf{m}_K(\mu^\star)\right) \;\equiv\; \mathcal{J}\left(P,a+b-\mu^\star\right)\;,\end{aligned}$$

where the first two equivalences follow from (5.5) and (2.4), respectively. □

Remark 5.12. Proposition 5.11 says that if μ is an eigenvalue of a $\mathcal{B}$-palindromic matrix polynomial $P(\lambda)$, then so is $a+b-\mu^\star$, and this pair of eigenvalues have identical algebraic, geometric, and partial multiplicities. When $\mathbb{F}=\mathbb{C}$ this eigenvalue pairing has a simple geometric interpretation, illustrated in Figure 1. If $\star = T$ and μ is an eigenvalue of P, then the eigenvalue $a+b-\mu^\star = a+b-\mu$ can be obtained by *inverting* μ through the midpoint $t=(a+b)/2$. On the other hand, if $\star = *$ then the eigenvalue $a+b-\mu^\star = a+b-\overline{\mu}$ is simply the *reflection* of μ through the vertical line $\mathrm{Re}(z)=t$.

Proposition 5.13. *Let $P(\lambda)$ be a square matrix polynomial over $\mathbb{F}$ of grade k that is $\mathcal{B}$–alternating (or $\mathcal{S}$-alternating). Then*

$$\mathcal{J}(P,\mu) \;\equiv\; \mathcal{J}\left(P,\;\frac{(a+b)\mu^\star - 2ab}{2\mu^\star-(a+b)}\right)\;. \tag{5.15}$$

Proof. Corollary 5.8(b) implies that $\mathbf{M}_A(P)(\lambda)=\varepsilon P^\star(\lambda)$, where A is given by (5.9). Then applying Theorem 4.5, one easily obtains the following chain of equivalences:

$$\begin{aligned}\mathcal{J}(P,\mu) \;\equiv\; \mathcal{J}\left(\varepsilon P^\star,\mu^\star\right) \;&\equiv\; \mathcal{J}\left(\mathbf{M}_A(P),\mu^\star\right)\\ &\equiv\; \mathcal{J}\left(P,\mathsf{m}_A(\mu^\star)\right) \;\equiv\; \mathcal{J}\left(P,\;\frac{(a+b)\mu^\star-2ab}{2\mu^\star-(a+b)}\right)\,,\end{aligned}$$

as desired. □

Remark 5.14. In contrast to the eigenvalue pairing for $\mathcal{B}-$palindromic matrix polynomials, the meaning of the eigenvalue pairing for $\mathcal{B}$-alternating matrix polynomials described by (5.15) seems rather obscure, at least at first glance. However, when $\mathbb{F}=\mathbb{C}$ this eigenvalue pairing also turns out to have a nice geometric interpretation.

Let Γ be the circle with the interval $[a,b]$ as one of its diameters. If $\star = *$ and μ is an eigenvalue of a $\mathcal{B}-$alternating matrix polynomial P, then the eigenvalue $\mathsf{m}_A(\mu^\star)=\mathsf{m}_A(\overline{\mu})$

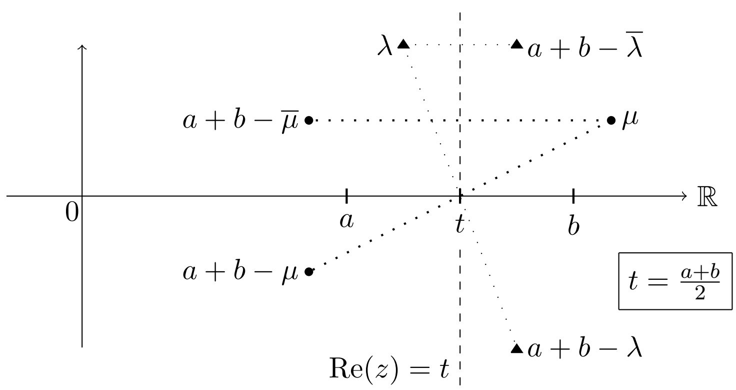

Figure 1: Eigenvalue pairing for $\mathcal{B}-$palindromic matrix polynomials

can be obtained by the classical geometric operation of *inverting μ through the circle* Γ, as illustrated in Figure 2. On the other hand, if $\star = T$ then the eigenvalue $\mathsf{m}_A(\mu^\star) = \mathsf{m}_A(\mu)$ is obtained by first inversion through Γ, followed by conjugation. This second type of eigenvalue pairing from (5.15) is also illustrated in Figure 2.

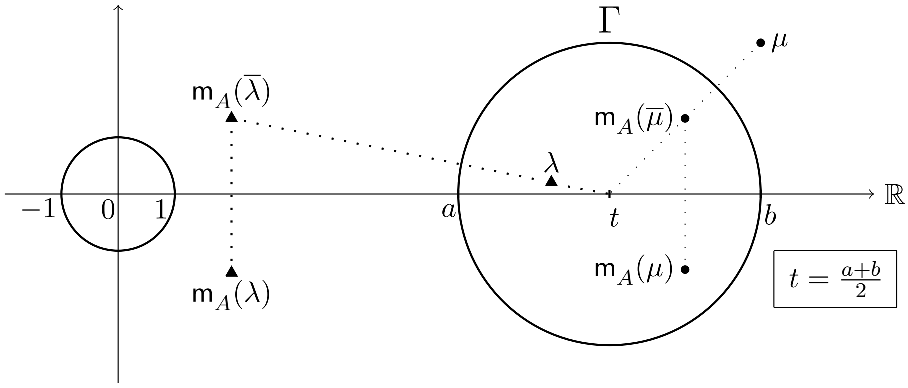

Figure 2: Eigenvalue pairing for $\mathcal{B}-$alternating matrix polynomials

In order to see why the eigenvalue pairings (5.15) truly have the claimed geometry, first observe that the involution A (and its negative $-A$) from (5.9) is similar to the standard involution R; in particular we have $T^{-1}(U^{-1}RU)T = -A$, where

$$T := \begin{bmatrix} 1 & -\frac{a+b}{2} \\ 0 & 1 \end{bmatrix}, \quad U := \begin{bmatrix} \frac{2}{b-a} & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{and} \quad R := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \tag{5.16}$$

Thus the Möbius function m_A can be understood as the composition

$$\mathsf{m}_A \;=\; \mathsf{m}_{-A} \;=\; \mathsf{m}_{T^{-1}} \circ \mathsf{m}_{U^{-1}} \circ \mathsf{m}_R \circ \mathsf{m}_U \circ \mathsf{m}_T\,.$$

Now m_T as a Möbius function on the complex plane simply translates points horizontally so that the center t of Γ maps to the origin, then m_U acts as a scaling transformation centered at the origin that takes the radius $(b-a)/2$ of Γ to 1. Thus the combined effect of $\mathsf{m}_{UT} = \mathsf{m}_U \circ \mathsf{m}_T$ is to map Γ to the unit circle centered at the origin. The function m_R is just the mapping $\lambda \mapsto (1/\lambda)$, which is well known to have the geometric effect of inversion through the unit circle followed by conjugation. Finally, the remaining transformations $\mathsf{m}_{T^{-1}U^{-1}} = \mathsf{m}_{T^{-1}} \circ \mathsf{m}_{U^{-1}}$ simply undo the scaling and translation, returning Γ back to its original position, with the geometric effect of the inversion and conjugation left intact.

5.2 Structured linearizations in Bernstein basis

The classical approach to solving the polynomial eigenvalue problem is via a *(strong) linearization*, i.e., by transformation to a matrix pencil with the same finite (and infinite) elementary divisors. For structured matrix polynomials, as we have seen in the previous section, spectral symmetries are often present, and when computing eigenvalues of such structured eigenproblems in finite precision, it is desirable to preserve those spectral symmetries. One way to ensure that computed solutions have these symmetries is to first find a structured linearization, and then apply a structure-preserving algorithm to that linearization.

In the case of $\mathcal{M}$-structured matrix polynomials, the existence of structured linearizations and the development of structure-preserving algorithms for computing eigenvalues of structured pencils have been extensively studied [25, 26, 28, 29, 30, 31]. Ideally, one would like to leverage all of this existing theory for $\mathcal{M}$-structured matrix polynomials and somehow transfer it over to $\mathcal{B}$-structured ones. Fortunately, this is straightforward to do, using a strategy analogous to that described in Remark 3.2(c) for general pencils. We start by considering a typical structure-preserving algorithm, and see how it can be easily adapted to a different structured setting.

Example 5.15. Let $\widehat{L}(\lambda)$ be an $\mathcal{M}$-palindromic pencil, i.e., $\widehat{L}(\lambda) = \lambda W + W^T$, where W is an arbitrary $n \times n$ matrix over $\mathbb{C}$. It was shown in [29] that there always exists a unitary matrix U such that

$$U^T \widehat{L}(\lambda) U \;=\; U^T(\lambda W + W^T)U \;=\; \lambda A + A^T\,, \tag{5.17}$$

where $A = [w_{i,j}] = U^T W U$ is *anti-triangular*, that is, $A_{i,j} = 0$ whenever $i + j \le n$. Furthermore, several methods for computing this anti-triangular form were investigated in [29]. The eigenvalues of the pencil $\widehat{A}(\lambda) = \lambda A + A^T$ are the same as those of $\widehat{L}(\lambda)$, so once this anti-triangular form is achieved the eigenvalues of a regular $\widehat{L}(\lambda)$ can be readily obtained from the anti-diagonal entries of $\widehat{A}(\lambda)$, simply by solving scalar equations of the type $(\lambda \cdot w_{j,n-j+1} + w_{n-j+1,j}) = 0$ for $j = 1, \dots, n$.

Now consider the $\mathcal{S}$-palindromic pencil $L(\lambda) = \phi_{1,1}(\lambda)W + \phi_{0,1}(\lambda)W^T$. Using exactly the same unitary matrix U as in (5.17) for the $\mathcal{M}$-palindromic pencil $\widehat{L}(\lambda)$ transforms the pencil $L(\lambda)$ into the pencil $A(\lambda) = \phi_{1,1}(\lambda)A + \phi_{0,1}(\lambda)A^T$, which is again $\mathcal{S}$-palindromic, and has the same eigenvalues as $L(\lambda)$. But just as in the $\mathcal{M}$-palindromic case, the eigenvalues of the anti-triangular pencil $A(\lambda)$ are readily obtained from the anti-diagonal entries, this time by solving scalar equations of the type $\big(\phi_{1,1}(\lambda) \cdot w_{j,n-j+1} + \phi_{0,1}(\lambda) \cdot w_{n-j+1,j}\big) = 0\,.$

Example 5.15 highlights two very important points. First, for pencils there is no particular disadvantage in being expressed in a non-monomial basis. In fact, as we will see soon, for the special case of $\mathcal{S}$ and $\mathcal{B}$ bases it allows us to readily find structured linearizations. Secondly, any structure-preserving eigenvalue algorithm developed for $\mathcal{M}$-structured pencils that just works on a *pair of matrices*, such as the algorithms in [25, 26, 29], can be used without change on $\mathcal{S}$ and $\mathcal{B}$-structured pencils as well. Consequently, the problem of finding eigenvalues of a $\mathcal{B}$-structured matrix polynomial essentially boils down to finding a structure-preserving linearization.

Let $P(\lambda)$ be an $n \times n$ matrix polynomial of grade k expressed as

$$P(\lambda) = \sum_{i=0}^{k} A_i \beta_{i,k}(\lambda) = \sum_{i=0}^{k} \widehat{A}_i \phi_{i,k}(\lambda)\,, \quad \text{where} \quad \widehat{A}_i := \frac{1}{(b-a)^k}\binom{k}{i} A_i\,. \tag{5.18}$$

For such a polynomial $P(\lambda)$ we consider (as in Theorem 3.3) the associated partner polynomial $\widehat{P}(\lambda) = \sum_{i=0}^{k} \widehat{A}_i \lambda^i$ in the monomial basis. Then using Proposition 5.5 and Definition 5.3, the following equivalences are easily verified:

$$P \text{ is } \mathcal{B}\text{-structured} \iff P \text{ is } \mathcal{S}\text{-structured} \iff \widehat{P} \text{ is } \mathcal{M}\text{-structured}\,, \tag{5.19}$$

where all of the structures are of the same type.

Now Theorem 3.3 together with (5.19) implies that $\widehat{L}(\lambda) = \lambda X + Y$ is a strong $\mathcal{M}$-structured linearization for an $\mathcal{M}$-structured matrix polynomial $\widehat{P}(\lambda)$ if and only if $L(\lambda) = \phi_{1,1}(\lambda)X + \phi_{0,1}(\lambda)Y$ is a strong $\mathcal{S}$-structured linearization for $\mathcal{S}$-structured $P(\lambda)$, where all the structures are of the same type. Further, Proposition 5.5 implies that finding a strong $\mathcal{B}$-structured linearization for a $\mathcal{B}$-structured polynomial P is equivalent to finding a strong $\mathcal{S}$-structured linearization for $\mathcal{S}$-structured P. Hence the entire problem of finding a strong $\mathcal{B}$-structured linearization for P reduces to that of finding a strong $\mathcal{M}$-structured linearization for $\widehat{P}$, a problem about which much is already known.

We now look at a few examples.

Example 5.16. Let $P(\lambda) = \sum_{i=0}^{5} A_i \beta_{i,5}(\lambda) = \sum_{i=0}^{5} \widehat{A}_i \phi_{i,5}(\lambda)$ be a square $\mathcal{B}$-Hermitian (and $\mathcal{S}$-Hermitian) matrix polynomial, regular or singular, of grade five. Then the pencil given by (3.7) is a strong $\mathcal{S}$-Hermitian linearization for P, and by (5.19) it is also a strong $\mathcal{B}$-Hermitian linearization for P.

Example 5.17. Let $P(\lambda) = \sum_{i=0}^{3} A_i \beta_{i,3}(\lambda) = \sum_{i=0}^{3} \widehat{A}_i \phi_{i,3}(\lambda)$ be a square *regular* $\mathcal{B}$-palindromic (and $\mathcal{S}$-palindromic) matrix polynomial of grade 3, with $\varepsilon = +1$. Then the pencil $L(\lambda) = (R_3 \otimes I)\big[\,\phi_{1,1}(\lambda)X + \phi_{0,1}(\lambda)Y\,\big]$ with X, Y given by (4.11) and R_3 as in (2.1) is a strong $\mathcal{B}$-palindromic (and $\mathcal{S}$-palindromic) linearization for P if and only if A_3 and A_0 are nonsingular.

Examples 5.16 and 5.17 clearly illustrate the simplicity of finding a strong $\mathcal{S}$-structured linearization for an $\mathcal{S}$-structured matrix polynomial. In fact, finding a strong $\mathcal{S}$-structured linearization for an $\mathcal{S}$-structured matrix polynomial is almost identical to the Linearization Procedure described in Section 3.1, except that at step 3 we look for a strong $\mathcal{M}$*-structured* linearization of $\widehat{P}$ of the same structure type. Hence we obtain the following structured version of Theorem 3.3.

Theorem 5.18. *Let $P(\lambda) = \sum_{i=0}^{k} A_i \beta_{i,k}(\lambda) = \sum_{i=0}^{k} \widehat{A}_i \phi_{i,k}(\lambda)$ be a square $\mathcal{S}$-structured matrix polynomial and define the $\mathcal{M}$-structured matrix polynomial $\widehat{P} = \sum_{i=0}^{k} \widehat{A}_i \lambda^i$, of the same structure type. Then $\phi_{1,1}(\lambda)X + \phi_{0,1}(\lambda)Y$ is a strong $\mathcal{S}$-structured linearization for P if and only if $\lambda X + Y$ is a strong $\mathcal{M}$-structured linearization for $\widehat{P}$.*

6 Conclusions

We have shown how to generate large new families of strong linearizations for matrix polynomials (regular or singular) expressed in Bernstein basis, which recover as special cases all of the known examples currently in the literature. In fact, we have shown in principle how to find all such strong linearizations, by establishing a simple bijection between the set of all strong linearizations for a matrix polynomial expressed in Bernstein basis, and the set of all strong linearizations for an associated partner matrix polynomial expressed in the monomial basis. As a consequence of an extensive use of Möbius transformations of matrix polynomials throughout our analysis, we have seen that strong linearizations are most naturally expressed in scaled Bernstein basis. This small shift in emphasis turns out to be the essential ingredient for gaining insight into working with matrix polynomials in Bernstein basis and their strong linearizations.

Matrix polynomials that are structured with respect to an arbitrary polynomial basis have also been defined. In the specific case of Bernstein-structured matrix polynomials, we studied the various spectral symmetries that result from these structures, and showed how to easily generate structured strong linearizations.

Finally, it is worth emphasizing that the entire theory in this paper applies equally well to scalar polynomials expressed in Bernstein basis. In particular, large families of new companion pencils for scalar polynomials in Bernstein basis have been introduced, some of them having promising tridiagonal Fiedler-like structure. The impact that these new companion pencils have on numerical properties is a subject for future investigation.

References

[1] A. Amiraslani, R. M. Corless & P. Lancaster, *Linearization of matrix polynomials expressed in polynomial bases*, IMA Journal of Numerical Analysis, 29 (2009), pp. 141–157.

[2] E. N. Antoniou & S. Vologiannidis, *A new family of companion forms of polynomial matrices*, Electron. J. Linear Algebra, 11 (2004), pp. 78–87.

[3] R. Bartle, *The Elements of Real Analysis*, 2nd Edition, John Wiley & Sons, New York, 1976.

[4] S. N. Bernstein, *Démonstration du théorème de Weierstrass fondée sur le calcul des probabilités*, Communications de la Société Mathématique de Kharkov 2, Series XIII No. 1, 1-2 (1912).

[5] D. A. Bini, L. Gemignani & J. R. Winkler, *Structured matrix methods for CAGD: An application to computing the resultant of polynomials in the Bernstein basis*, Numerical Linear Algebra with Applications, 12 (2004), pp. 685–698.

[6] D. A. Bini & V. Noferini, *Solving polynomial eigenvalue problems by means of the Ehrlich-Aberth method*, Linear Algebra Appl., 439 (2013), pp. 1130–1149.

[7] M. I. Bueno, F. De Terán & F. M. Dopico, *Recovery of eigenvectors and minimal bases of matrix polynomials from generalized Fiedler linearizations*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 463–483.

[8] M.I. Bueno & S. Furtado, *Palindromic linearizations of a matrix polynomial of odd degree obtained from Fiedler pencils with repetition*, Electron. J. Linear Algebra, 23 (2012), pp. 562–577.

[9] F. De Terán, F. M. Dopico & D. S. Mackey, *Linearizations of singular matrix polynomials and the recovery of minimal indices*, Electron. J. Linear Algebra, 18 (2009), pp. 371–402.

[10] F. De Terán, F. M. Dopico & D. S. Mackey, *Fiedler companion linearizations and the recovery of minimal indices*, SIAM J. Matrix Anal. Appl., 31 (2010), pp. 2181–2204.

[11] F. De Terán, F. M. Dopico & D. S. Mackey, *Fiedler companion linearizations for rectangular matrix polynomials*, Linear Algebra Appl., 437 (2012), pp. 957–991.

[12] F. De Terán, F. M. Dopico & D. S. Mackey, *Spectral equivalence of matrix polynomials and the index sum theorem*, To appear in Linear Algebra Appl. Also available as MIMS EPrint 2013.47, Manchester Institute for Mathematical Sciences, The University of Manchester, UK., 2013.

[13] F. De Terán, F. M. Dopico & J. Pérez, *Condition numbers for inversion of Fiedler companion matrices*, Linear Algebra Appl., 439 (2013), pp. 944–981.

[14] J. Delgado & J. M. Peña, *Optimal conditioning of Bernstein collocation matrices*, SIAM J. Matrix Anal. Appl., 31 (2009), pp. 990–996.

[15] D. S. Dummit & R. M. Foote, *Abstract Algebra*, 2nd Edition, John Wiley & Sons, Inc, 2004.

[16] G. Farin, *Curves and Surfaces for Computer-Aided Geometric Design*, Academic Press, San Diego, 1997.

[17] R. T. Farouki, *The Bernstein polynomial basis: a centennial retrospective*, Computer Aided Geometric Design, 29 (2012), pp. 379–419.

[18] R. T. Farouki, T. N. T. Goodman & T. Sauer, *Construction of orthogonal bases for polynomials in Bernstein form on triangular simplex domains*, Computer Aided Geometric Design, 20 (2003), pp. 209–230.

[19] G. Frobenius, *Theorie der linearen Formen mit ganzen Coefficienten*, J. Reine Angew. Math. (Crelle), 86 (1878), pp. 146–208.

[20] J. Garloff & A. P. Smith [eds.], *Special Issue on the Use of Bernstein Polynomials in Reliable Computing: A Centennial Anniversary*, Reliable Computing, 17 (2012), pp. 1–237.

[21] I. Gohberg, M. A. Kaashoek & P. Lancaster, *General theory of regular matrix polynomials and band Toeplitz operators*, Integral Equations Operator Theory, 11(6) (1988), pp. 776–882.

[22] I. Gohberg, P. Lancaster & L. Rodman, *Matrix Polynomials*, 2nd Edition, Academic Press, New York, 2005.

[23] L. Grammont, N. Higham & F. Tisseur, *A framework for analyzing nonlinear eigenproblems and parametrized linear systems*, Linear Algebra Appl., 435: (2011), pp. 623–640.

[24] G. F. Jónsson & S. Vavasis, *Solving polynomials with small leading coefficients*, SIAM J. Matrix Anal. Appl., 26 (2005), pp. 400–414.

[25] D. Kressner, *Numerical Methods for General and Structured Eigenvalue Problems*, Lecture Notes in Computational Science and Engineering, Springer-Verlag, 2005.

[26] D. Kressner, C. Schröder & D. S. Watkins, *Implicit QR algorithms for palindromic and even eigenvalue problems*, Numer. Algorithms, (2009), pp. 209–238.

[27] D. S. Mackey, N. Mackey, C. Mehl & V. Mehrmann, *Vector spaces of linearizations for matrix polynomials*, SIAM J. Matrix Anal. Appl., 27 (2006), pp. 821–850.

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

[29] D. S. Mackey, N. Mackey, C. Mehl & V. Mehrmann, *Numerical methods for palindromic eigenvalue problems: computing the anti-triangular Schur form*, Numerical Linear Algebra Appl., 16 (2009), pp. 63–86.

[30] D. S. Mackey, N. Mackey, C. Mehl & V. Mehrmann, *Jordan structures of alternating matrix polynomials*, Linear Algebra Appl., 432 (2010), pp. 867–891.

[31] D. S. Mackey, N. Mackey, C. Mehl & V. Mehrmann, *Smith forms of palindromic matrix polynomials*, Electron. J. Linear Algebra, 22 (2011), pp. 53–91.

[32] D. S. Mackey, N. Mackey, C. Mehl & V. Mehrmann, *Möbius transformations of matrix polynomials*, to appear in Linear Algebra Appl. Also available as MIMS EPrint, 2014.2, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, 2014.

[33] A. Marco & J-J. Martinez, *A fast and accurate algorithm for solving Bernstein-Vandermonde linear systems*, Linear Algebra Appl., 422 (2007), pp. 616–628.

[34] A. Marco & J-J. Martinez, *Polynomial least squares fitting in the Bernstein basis*, Linear Algebra Appl., 433 (2010), pp. 1254–1264.

[35] M. Osman & S. K. Ghosh, *Nonparametric regression models for right-censored data using Bernstein polynomials*, Computational Statistics and Data Analysis, 56 (2012), pp. 559–573.

[36] G. Simmons, *Introduction to Topology and Modern Analysis*, McGraw-Hill Book Company, Inc., New York, 1963.

[37] S. Vologiannidis & E. N. Antoniou, *A permuted factors approach for the linearization of polynomial matrices*, Math. Control Signals Syst., 22 (2011), pp. 317-342.

[38] J. R. Winkler, *A companion matrix resultant for Bernstein polynomials*, Linear Algebra Appl., 362 (2003), pp. 153–175.