Linearizations of Matrix Polynomials in Newton Bases

Perovic, Vasilije and Mackey, D. Steven

2017

MIMS EPrint: **2017.23**

ISSN 1749-9097

Linearizations of Matrix Polynomials in Newton Bases

Vasilije Perović[a,*], D. Steven Mackey[b]

[a]*Department of Mathematics, University of Rhode Island, Kingston, RI 02881, USA*
[b]*Department of Mathematics, Western Michigan University, Kalamazoo, MI 49008, USA*

Abstract

We discuss matrix polynomials expressed in a Newton basis, and the associated polynomial eigenvalue problems. Properties of the generalized ansatz spaces associated with such polynomials are proved directly by utilizing a novel representation of pencils in these spaces. Also, we show how the family of Fiedler pencils can be adapted to matrix polynomials expressed in a Newton basis. These new Newton-Fiedler pencils are shown to be strong linearizations, and some computational aspects related to them are discussed.

Keywords: matrix polynomial, Newton bases, strong linearization, Newton-Fiedler pencil, ansatz space, updating.
AMS classification: 65F15, 65D05, 41A10, 47J10, 15A18, 15A22.

---

1. Introduction

Consider a nonlinear eigenvalue problem $P(\lambda)x = 0, x \neq 0$, where $P(\lambda)$ is a matrix polynomial of the form

$$P(\lambda) = \sum_{i=0}^{k} A_i \phi_i(\lambda)\,, \quad A_0, A_1, \ldots A_k \in \mathbb{F}^{n\times n}\,. \tag{1.1}$$

Here $\mathcal{B} = \{\phi_i(\lambda)\}_{i=0}^{k}$ is a polynomial basis for the space of univariate scalar polynomials of degree at most k; classical examples of such bases include Chebyshev, Newton, Hermite, Lagrange, Bernstein, etc. Matrix polynomials expressed in those bases arise either directly from applications, or as approximations when solving more general nonlinear eigenvalue problems, see for example [12, 17, 29, 30] and the references therein.

The classical and most widely used approach to solving the polynomial eigenvalue problem $P(\lambda)x = 0$ is to first *linearize* P, i.e., convert P into a matrix pencil L with the same spectral structure, and then compute with L. Since the 1950's this approach has been extensively developed for matrix polynomials expressed in the *standard* basis. It was not until recently, though, that polynomial eigenvalue problems expressed in other bases have been seriously considered, see for example [1, 5, 12, 25, 27, 29, 30]. On the one hand, it is tempting to simply convert P from (1.1) to the standard basis, and then leverage the existing body of knowledge about linearizations. However, it is important to avoid reformulating P into the standard basis, since a change of basis has the potential to introduce numerical errors not present in the original statement of the problem.

In this paper we focus on *matrix polynomials expressed in a Newton basis*, i.e., polynomials of the form

$$P(\lambda) = \sum_{i=0}^{k} A_i n_i(\lambda)\,, \quad A_0, A_1, \ldots A_k \in \mathbb{F}^{n\times n}\,, \tag{1.2}$$

where $n_i(\lambda)$ is the i^{th} degree Newton polynomial associated with an ordered list of nodes $(\alpha_1, \ldots, \alpha_k)$. Our main goal is to generate large new families of linearizations for P by working *directly* with the coefficients A_i from (1.2), while avoiding additions, subtractions, multiplications, or inverses of those coefficients.

*Corresponding author
Email addresses: `perovic@uri.edu` (Vasilije Perović), `steve.mackey@wmich.edu` (D. Steven Mackey)

The first effort to study strong linearizations for matrix polynomials expressed in a Newton basis was made in [1], where a single example generalizing the first Frobenius companion form was provided. Further results came in [27], where the authors generalized the results from [18, 22] and considered spaces of potential linearizations for *regular* matrix polynomials expressed in arbitrary degree-graded polynomial bases.

In this work we also aim to generalize the results of [22], but in contrast to [27] where such generalizations are obtained via a bivariate polynomial approach with emphasis on the *double* ansatz space, our focus lies primarily on generalizations of the ansatz spaces themselves. Additionally, the analysis included here is valid for both regular *and singular* square matrix polynomials. Using a novel representation for pencils, we show how elements of these generalized ansatz spaces associated with a Newton basis can be *easily* constructed. Indeed, this new type of pencil representation is instrumental both for constructing these pencils as well as for analyzing their basic properties, and is one of the main contributions of this paper. We also show that *almost all* of the pencils in these generalized ansatz spaces are strong linearizations, no matter whether the original polynomial is regular or singular.

By adapting the construction of Fiedler pencils for matrix polynomials expressed in the standard basis from [8], we show how to generate a new family of *Newton-Fiedler* pencils associated with matrix polynomials expressed in a Newton basis. This construction employs the same novel representation of matrix pencils as used for the generalized ansatz spaces, but these Newton-Fiedler pencils are *always* strong linearizations for any regular or square singular matrix polynomial in Newton basis, not just almost always. The construction given here can also be extended using the results from [9] to produce strong linearizations for *rectangular* matrix polynomials. Further, these Newton-Fiedler pencils can be easily updated if additional nodes are introduced, making them a suitable candidate for linearizations of matrix polynomials that arise as a result of *adaptive* approximation of more general matrix-valued functions. Finally, it is important to emphasize that the results in this paper are not just of theoretical importance, but can be a starting point for the development of new eigenvalue algorithms for matrix polynomials expressed in Newton basis.

We begin in Section 2 by reviewing the necessary background on matrix polynomials and linearizations, while in Section 3 vector spaces of pencils for matrix polynomials in Newton basis are studied. The main result regarding the family of Fiedler-like pencils for matrix polynomials expressed in Newton basis is given in Section 4, and in Section 5 some computational issues related to Newton-Fiedler pencils are discussed.

2. Background and Notation

Throughout this paper $\mathbb{N}$ denotes the set of non-negative integers, and $\mathbb{F}$ is an arbitrary field, unless otherwise stated. The ring of all univariate polynomials with coefficients from $\mathbb{F}$ and the field of rational functions over $\mathbb{F}$ are denoted by $\mathbb{F}[\lambda]$ and $\mathbb{F}(\lambda)$, respectively. The $\mathbb{F}$-vector space of univariate scalar polynomials of degree at most k is denoted by $\mathcal{P}_k$.

In this paper we focus our attention on *square* matrix polynomials that, when expressed in standard basis, are of the form

$$P(\lambda) = \sum_{i=0}^{k} A_i \lambda^i \ , \tag{2.1}$$

where $A_0, \ldots, A_k \in \mathbb{F}^{n\times n}$ and $k \geq 0$. A matrix polynomial of the form (2.1) is said to have *grade* k, and A_k is referred to as the *leading coefficient*, no matter whether A_k is the zero matrix or not. The *degree* of a nonzero matrix polynomial retains its usual meaning as the largest integer j such that coefficient of λ^j in $P(\lambda)$ is nonzero. Declaring that a polynomial $P(\lambda)$ has grade k indicates that $P(\lambda)$ 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 degree and grade are equal if the leading coefficient is a nonzero matrix, otherwise grade is strictly larger than the degree. The notion of grade plays a key role when defining and studying a variety of notions, including elementary divisors at infinity, Möbius transformations of matrix polynomials [24], and matrix polynomials expressed in non-degree graded bases [1, 25].

A matrix polynomial P is said to be *regular* if it is square and invertible over the field $\mathbb{F}(\lambda)$, equivalently if $\det(P) \not\equiv 0$; otherwise it is said to be *singular*.

Definition 2.1. (Finite eigenvalue)
A scalar $\lambda_0 \in \mathbb{F}$ is said to be a (finite) eigenvalue *of an $m \times n$ matrix polynomial P if*

$$\operatorname{rank}\big(P(\lambda_0)\big) < \operatorname{nrank}\big(P(\lambda)\big),$$

where $\operatorname{rank}\big(P(\lambda_0)\big)$ *is the rank of the matrix $P(\lambda_0)$ over $\mathbb{F}$, and* $\operatorname{nrank}(P(\lambda))$ *is the rank of $P(\lambda)$ as a matrix over the field $\mathbb{F}(\lambda)$. A nonzero vector $x \in \mathbb{F}^n$ is said to be a* right eigenvector *of P corresponding to λ_0 if $P(\lambda_0)x = 0$ holds.*

If P is a regular $n \times n$ matrix polynomial, then $\operatorname{nrank}(P) = n$, and Definition 2.1 is equivalent to saying that λ_0 is an eigenvalue of P if and only if $\det(P(\lambda_0)) = 0$.

It is often useful to consider ∞ as a possible eigenvalue of P, but in order to define this notion we first need the following concept.

Definition 2.2. (Reversal of matrix polynomials)
For a matrix polynomial P of grade $k \geq 0$ as in (2.1), the reversal of P with respect to grade k *is the matrix polynomial*

$$(\operatorname{rev}_k P)(\lambda) := \lambda^k P(1/\lambda) = \sum_{i=0}^{k} \lambda^i A_{k-i}\,. \tag{2.2}$$

Note that the matrix polynomial P in Definition 2.2 is allowed to have grade zero, i.e., P may be just a constant matrix A_0, and in that case we have $\operatorname{rev}_0 P = P$.

The next proposition summarizes some of the fundamental properties of reversal; proofs can be found in [24].

Proposition 2.3. *Let P and Q be $m \times n$ and $n \times t$ matrix polynomials of grades k and ℓ, respectively, with $k, \ell \geq 0$. Then the following properties hold:*

(a) $\operatorname{rev}_{k+\ell}(P \cdot Q) = \operatorname{rev}_k P \cdot \operatorname{rev}_\ell Q$,

(b) $\operatorname{rev}_{k+\ell}(P \otimes Q) = \operatorname{rev}_k P \otimes \operatorname{rev}_\ell Q$,

(c) $\operatorname{rev}_k\big(\operatorname{rev}_k P\big) = P$,

(d) $(\operatorname{rev}_k P)_{ij} = \operatorname{rev}_k(P_{ij})$,

(e) $\operatorname{nrank}(P) = \operatorname{nrank}(\operatorname{rev}_k P)$,

(f) *Let $\{x_i(\lambda)\}_{i=1}^m$ be a set of m scalar polynomials in $\mathcal{P}_k$. Then $\{x_i(\lambda)\}_{i=1}^m$ is linearly independent* if and only if *the set $\{(\operatorname{rev}_k x_i)(\lambda)\}_{i=1}^m \subseteq \mathcal{P}_k$ is linearly independent.*

Note that the nonzero finite eigenvalues of $\operatorname{rev}_k P$ are the reciprocals of those of P. Thus it is sensible to view 0 and ∞ as reciprocals, and use this to define eigenvalues at ∞.

Definition 2.4. (Eigenvalue at ∞)
Let $P(\lambda)$ be a regular matrix polynomial of grade $k \geq 1$. Then $P(\lambda)$ is said to have an eigenvalue at ∞ *with eigenvector x if $\operatorname{rev}_k P(\lambda)$ has the eigenvalue 0 with eigenvector x.*

The classical approach to computing eigenvalues of a matrix polynomial P is to first *linearize* P, i.e., convert P into a matrix pencil[1] L with the same spectral information, and then compute with L. This approach has been extensively studied when P is expressed in the standard basis as in (2.1), and relevant portions of related work are reviewed in later subsections as they are needed.

Before stating the definition of *linearization*, let us recall two common and extremely useful equivalence relations of matrix polynomials.

[1]Throughout this paper the term *pencil* is reserved for matrix polynomials of grade one.

Definition 2.5. (Strict and Unimodular Equivalence)
Suppose P and Q are two matrix polynomials of the same size.

(a) *P and Q are said to be* unimodularly equivalent *if there exist* unimodular *matrices* (*i.e., polynomial matrices with nonzero constant determinant*) *$U(\lambda)$ and $V(\lambda)$ such that $Q(\lambda) = U(\lambda) \cdot P(\lambda) \cdot V(\lambda)$.*

(b) *P and Q are said to be* strictly equivalent *if there exist* constant and invertible *matrices S and T such that $Q(\lambda) = S \cdot P(\lambda) \cdot T$.*

Remark 2.6. One of the most important features of these equivalence relations is their effect on elementary divisors of P. More specifically, it is well known that unimodular equivalence preserves all finite elementary divisors of P, while strict equivalence preserves both finite and infinite elementary divisors of P [14]. The aspect that is relevant to this paper, and concerns the concepts in Definition 2.7, is that any pencil that is unimodularly (resp., *strictly*) equivalent to a linearization (resp., *strong* linearization) for P is also a linearization (resp., *strong* linearization) for P.

Definition 2.7. (Linearization of matrix polynomials)
An $nk \times nk$ matrix pencil $L(\lambda)$ is a linearization *for an $n \times n$ matrix polynomial $P(\lambda)$ of grade k if $L(\lambda)$ and* $\operatorname{diag}\left[P(\lambda), I_{(k-1)n}\right]$ *are unimodularly equivalent. A linearization $L(\lambda)$ is called a* strong linearization *if* $\operatorname{rev}_1 L(\lambda)$ *is also a linearization of* $\operatorname{rev}_k P(\lambda)$.

Definition 2.7 says that in order to show a matrix pencil L is a linearization for P, we must find unimodular matrix polynomials $U(\lambda)$ and $V(\lambda)$ such that

$$U(\lambda)L(\lambda)V(\lambda) = \begin{bmatrix} P(\lambda) & 0 \\ 0 & I_{(k-1)n} \end{bmatrix}. \tag{2.3}$$

But that is not the only way. An alternative criterion for determining (using *rational* transformations) whether a pencil is a linearization for a *regular* matrix polynomial over $\mathbb{F} = \mathbb{C}$ was introduced in [20], and is based on the *local Smith form*. In Theorem 2.12 we introduce an extension of this criterion to allow the use of rational transformations in a way that is valid for both regular *and* singular matrix polynomials, over an *arbitrary* field $\mathbb{F}$ [26]. In order to clearly describe this extension we need a few preliminary definitions. Note that the technique described in Theorem 2.12 will not be used until Section 4.

Definition 2.8. *Let $r(\lambda)$ and $t(\lambda)$ be scalar polynomials in $\mathbb{F}[\lambda]$. Then:*

(a) *$r(\lambda)$ and $t(\lambda)$ are said to be* coprime (*or* relatively prime) *if* $\gcd(r(\lambda), t(\lambda)) = 1$.

(b) *A rational expression $\frac{r(\lambda)}{t(\lambda)}$ (with $t(\lambda) \not\equiv 0$) for an element in $\mathbb{F}(\lambda)$ is said to be in* reduced form *if $r(\lambda)$ and $t(\lambda)$ are coprime.*

Definition 2.9. (q - prime matrices)
Let $q(\lambda)$ be a non-constant $\mathbb{F}$-irreducible scalar polynomial over an arbitrary field $\mathbb{F}$. An $n \times n$ rational (*resp., polynomial*) *matrix $S(\lambda)$ with entries $\big(S(\lambda)\big)_{ij} = \frac{r_{ij}(\lambda)}{t_{ij}(\lambda)}$ is said to be a* q - prime rational (*resp.,* q - prime polynomial) matrix *if:*

(a) *$S(\lambda)$ is regular (i.e.,* $\det S(\lambda) \not\equiv 0$),

(b) *for* each *entry of $S(\lambda)$, expressed in reduced form $\frac{r_{ij}(\lambda)}{t_{ij}(\lambda)}$, the* denominator $t_{ij}(\lambda)$ *is coprime to $q(\lambda)$, and*

(c) *if* $\det S(\lambda)$ *is expressed in reduced form $\frac{r(\lambda)}{t(\lambda)} \not\equiv 0$, then* both the numerator $r(\lambda)$ and denominator $t(\lambda)$ *are coprime to $q(\lambda)$.*

Remark 2.10. It is easy to see that any *invertible* constant matrix or *unimodular* polynomial matrix is also a q - prime matrix, for *every* $\mathbb{F}$-irreducible scalar polynomial $q(\lambda)$.

The following proposition, which can be easily verified, concerns the closure of q-prime matrices under finite products.

Proposition 2.11. *Let $q(\lambda)$ be an arbitrary non-constant $\mathbb{F}$-irreducible scalar polynomial. Then any finite product of q-prime rational (resp., q-prime polynomial) matrices is q-prime rational (resp., q-prime polynomial).*

We have now established all the background needed to state the theorem that provides an essential tool for proving the results in Section 4.

Theorem 2.12. *Let $P(\lambda)$ be an $n \times n$ regular or singular matrix polynomial of grade k over an arbitrary field $\mathbb{F}$, and let $L(\lambda)$ be an $nk \times nk$ matrix pencil over $\mathbb{F}$. Then $L(\lambda)$ is a* linearization for $P(\lambda)$ *if and only if for each $\mathbb{F}$-irreducible scalar polynomial $q(\lambda)$ there exist q-prime (*rational or polynomial*) matrices $E_q(\lambda)$ and $F_q(\lambda)$ such that*

$$E_q(\lambda) \cdot L(\lambda) \cdot F_q(\lambda) = \begin{bmatrix} P(\lambda) & 0 \\ 0 & I_{(k-1)n} \end{bmatrix}.$$

A linearization $L(\lambda)$ of $P(\lambda)$ is a strong linearization *if and only if there exist λ-prime (*rational or polynomial*) matrices $E_\lambda(\lambda)$ and $F_\lambda(\lambda)$ such that*

$$E_\lambda(\lambda) \cdot \mathrm{rev}_1 L(\lambda) \cdot F_\lambda(\lambda) = \begin{bmatrix} \mathrm{rev}_k P(\lambda) & 0 \\ 0 & I_{(k-1)n} \end{bmatrix}.$$

Remark 2.13. Theorem 2.12 can be viewed as a completely algebraic analog of the result for analytic matrix functions in [20, Thm. 3.2]. More specifically, the "*unimodular* matrix functions that are *analytic* on a neighborhood of λ_0", as stated in [20, Thm. 3.2], are replaced in Theorem 2.12 by the algebraically defined class of q-prime matrices, where $q(\lambda) = \lambda - \lambda_0$.

On the other hand, a more general version of Theorem 2.12 that holds for rational matrices is given in [2]. In the language of [2], for a fixed non-constant $\mathbb{F}$-irreducible scalar polynomial $q(\lambda)$, the q-prime matrices from Definition 2.9 are elements of $GL_n(\mathbb{F}_q(\lambda))$. With this understanding, Theorem 2.12 can then be seen to be a special case of [2, Thm. 4.2].

The most essential feature of Theorem 2.12 that is relevant to this paper is that it enables the comparison of the spectral structures of P and L one $\mathbb{F}$-irreducible at a time, while allowing a much larger family of transformations to be employed. This added flexibility has already been successfully exploited to study isolated examples of linearizations for *regular* matrix polynomials over $\mathbb{C}$ expressed in non-standard bases [1], and will be crucial to us in Section 4.

2.1. Vector spaces of potential linearizations

In this section we review vector spaces of pencils associated with a matrix polynomial P expressed in the standard basis as in (2.1). For many decades there were two particular linearizations for solving polynomial eigenvalue problems $P(\lambda)x = 0$ that were in predominant use: either $C_1(\lambda) = \lambda X_1 + Y_1$ or $C_2(\lambda) = \lambda X_2 + Y_2$, where

$$X_1 = X_2 = \mathrm{diag}(A_k, I_{(k-1)n}),$$

$$Y_1 = \begin{bmatrix} A_{k-1} & A_{k-2} & \cdots & A_0 \\ -I_n & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & -I_n & 0 \end{bmatrix}, \quad \text{and} \quad Y_2 = \begin{bmatrix} A_{k-1} & -I_n & \cdots & 0 \\ A_{k-2} & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & -I_n \\ A_0 & \cdots & 0 & 0 \end{bmatrix}, \tag{2.4}$$

or simple block permutations of C_1 or C_2. These linearizations are now often referred to as the *Frobenius companion forms* of $P(\lambda)$, or more simply as the first and second companion forms [15] of $P(\lambda)$.

Only in the last 14 years have larger classes of linearizations been studied in a more systematic fashion; the first progress can be found in [3] and [22]. We now recall the definitions of the vector spaces introduced in [22], and remind the reader that *almost all* of the pencils in these spaces are strong linearizations [22].

Definition 2.14. *Let P be an $n\times n$ matrix polynomial of grade k, and define the following spaces of $nk\times nk$* pencils *associated with P:*

$$\begin{aligned}
\mathbb{L}_1(P) &:= \left\{L(\lambda) : L(\lambda)(\Lambda\otimes I_n) = v\otimes P(\lambda),\ \text{for some } v\in\mathbb{F}^k\right\},\\
\mathbb{L}_2(P) &:= \left\{L(\lambda) : (\Lambda^T\otimes I_n)L(\lambda) = w^T\otimes P(\lambda),\ \text{for some } w\in\mathbb{F}^k\right\},\\
\mathbb{DL}(P) &:= \mathbb{L}_1(P)\cap\mathbb{L}_2(P),
\end{aligned}\tag{2.5}$$

where $\Lambda := \left[\lambda^{k-1},\cdots,\lambda,1\right]^T \in \mathbb{F}[\lambda]^{k\times 1}$.

Note that the phrase "v is the right ansatz vector for $L(\lambda)$" is often used when $L(\lambda)$ is in $\mathbb{L}_1(P)$ with $L(\lambda)(\Lambda\otimes I) = v\otimes P(\lambda)$. Analogously, the phrase "w is the left ansatz vector for $K(\lambda)$" is used when $K(\lambda)$ is in $\mathbb{L}_2(P)$ with $(\Lambda^T\otimes I)K(\lambda) = w^T\otimes P(\lambda)$. The spaces $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ have sometimes been referred to as the "ansatz spaces" associated with P, while $\mathbb{DL}(P)$ is the "double ansatz space" associated to P [22]. In Section 3.1 generalizations of these ansatz spaces will be introduced.

One of the fundamental properties of the $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ spaces is that they are related via *block transposition*, see [18, Thm. 2.2].

Definition 2.15. (Block transpose)
If $A = (A_{ij})$ is a block $k\times\ell$ matrix with $m\times n$ blocks A_{ij}, the block transpose of A *is the block $\ell\times k$ matrix $A^{\mathcal{B}}$ with $m\times n$ blocks defined by $(A^{\mathcal{B}})_{ij} = A_{ji}$.*

The next proposition describes one of the main properties of block transpose.

Proposition 2.16. ([21], Lemma 3.1.4) *Let A and C be $k\times\ell$ and $\ell\times m$ block matrices with $n\times n$ blocks A_{ij} and C_{gh}, respectively. If each A_{ij} commutes with each C_{gh}, then $(AC)^{\mathcal{B}} = C^{\mathcal{B}}A^{\mathcal{B}}$.*

2.2. Fiedler companion pencils

Another important source of strong linearizations for a matrix polynomial is the family of *Fiedler pencils.* Based on the ideas in [13], these pencils were introduced in [3] for regular matrix polynomials over $\mathbb{C}$, and then further developed for general matrix polynomials over arbitrary fields in [8] and [9]. For an $n\times n$ matrix polynomial

$$P(\lambda) = \sum_{i=0}^{k} A_i\lambda^i \tag{2.6}$$

of grade k, the *Fiedler factors* of $P(\lambda)$ are defined as the $nk\times nk$ matrices

$$M_k := \begin{bmatrix} A_k & \\ & I_{(k-1)n}\end{bmatrix}, \quad M_0 := \begin{bmatrix} I_{(k-1)n} & \\ & -A_0\end{bmatrix}, \tag{2.7}$$

$$\text{and}\qquad M_i := \begin{bmatrix} I_{(k-i-1)n} & & & \\ & -A_i & I_n & \\ & I_n & 0 & \\ & & & I_{(i-1)n}\end{bmatrix}, \quad i=1,\ldots,k-1. \tag{2.8}$$

Remark 2.17. Note that the M_0 and M_k factors might or might not be invertible, depending on the coefficient matrices A_0 and A_k, but the M_i matrices for $i=1,\ldots,k-1$ in (2.8) are *always* invertible, no matter what the A_i's are like.

Definition 2.18. (Fiedler pencils)
Let $P(\lambda)$ be the matrix polynomial in (2.6), and let M_i for $i=0,\ldots,k$ be the matrices defined in (2.7) and (2.8). Given any bijection $\sigma : \{0,1,\ldots,k-1\}\to\{1,\ldots,k\}$, the Fiedler pencil for $P(\lambda)$ associated with σ *is the $nk\times nk$ matrix pencil*

$$F_\sigma(\lambda) := \lambda M_k - M_{\sigma^{-1}(1)}\cdot M_{\sigma^{-1}(2)}\cdots M_{\sigma^{-1}(k)}. \tag{2.9}$$

Note that $\sigma(i)$ describes the position of the factor M_i in the product $M_{\sigma^{-1}(1)} \cdots M_{\sigma^{-1}(k)}$ defining the zero-degree term in (2.9); i.e., $\sigma(i) = j$ means that M_i is the j^{th} factor in the product. For brevity, we denote this product by

$$M_\sigma \;:=\; M_{\sigma^{-1}(1)} \cdot M_{\sigma^{-1}(2)} \cdots M_{\sigma^{-1}(k)}\,, \tag{2.10}$$

so that $F_\sigma(\lambda) = \lambda M_k - M_\sigma$.

Throughout the rest of the paper the bijection σ will often be written using the standard *array* notation $\sigma := \big(\sigma(0), \sigma(1), \ldots, \sigma(k-1)\big)$.

Theorem 2.19. [8, Thm. 4.6] *Let P be an $n \times n$ matrix polynomial of grade k, regular or singular, over an arbitrary field $\mathbb{F}$. Then any Fiedler pencil associated with P is a strong linearization for P.*

Remark 2.20. There are a number of important features that distinguish Fiedler pencils from the pencils in the $\mathbb{L}_1$ and $\mathbb{L}_2$ spaces, including:

1. *All* Fiedler pencils are *strong* linearizations. In particular, *no* additional "linearization condition" needs to be imposed on a Fiedler pencil to guarantee that it is a strong linearization, as is required for pencils in $\mathbb{L}_1$ and $\mathbb{L}_2$ (see [7] and [22]).
2. *All* Fiedler pencils are strong linearizations for regular *and* singular (square) matrix polynomials, over any field.
3. No linear combinations of the matrix coefficients A_i are required when constructing Fiedler pencils, as is the case for many pencils in $\mathbb{L}_1$ and $\mathbb{L}_2$.
4. Fiedler pencils can be extended in a natural way to provide strong linearizations for *rectangular* matrix polynomials.

3. Matrix polynomials expressed in Newton basis

Let $\mathcal{A} = (\alpha_1, \ldots, \alpha_k)$ be an *ordered list* of elements from $\mathbb{F}$, where the α_i's need *not* be distinct, or numerically ordered in any way. With any such list $\mathcal{A}$ we associate the *scalar* polynomials defined as

$$n_0(\lambda) \equiv 1\,, \quad n_i(\lambda) := \prod_{j=1}^{i} (\lambda - \alpha_j) \quad \text{for each } \; i = 1, \ldots, k. \tag{3.1}$$

If $\mathcal{A}$ consists of *distinct* elements, the polynomials in (3.1) are classically referred to as the *Newton polynomials* associated with $\mathcal{A}$ [6, Ch. 2]. When $\mathcal{A}$ contains *non-distinct* elements, one can think of the polynomials in (3.1) as *generalized Newton polynomials.* If the list $\mathcal{A}$ is ordered so that repeated elements are in contiguous blocks, i.e., as

$$\mathcal{A} = \big(\beta_1, \ldots, \beta_1\,, \beta_2, \ldots, \beta_2\,, \ldots\,, \beta_\ell, \ldots, \beta_\ell\big)$$

where $\beta_i \neq \beta_j$ when $i \neq j$, then the polynomials in (3.1) arise in the problem of *Hermite interpolation.* Since *none* of the results in this paper depend on elements in $\mathcal{A}$ being distinct, for the sake of brevity we refer to the polynomials in (3.1) simply as Newton polynomials and the elements of $\mathcal{A}$ as *nodes.*

We also associate two sets of grade one scalar polynomials with $\mathcal{A}$, namely $\{\gamma_i(\lambda)\}_{i=1}^k$ and $\{\widetilde{\gamma}_i(\lambda)\}_{i=1}^k$, defined by

$$\gamma_i(\lambda) := \lambda - \alpha_i \quad \text{and} \quad \widetilde{\gamma}_i(\lambda) := \mathrm{rev}_1\gamma_i(\lambda) = \big(1 - \alpha_i\lambda\big)\,, \quad \text{for all } \; i = 1, \ldots, k\,. \tag{3.2}$$

Further, for each $i = 1, \ldots, k$ we consider the diagonal $ni \times ni$ matrices in (3.3) and (3.4), that will play a key role when studying matrix polynomials expressed in Newton basis:

$$\Gamma_i(\lambda) \;:=\; \mathrm{diag}\big(\gamma_i(\lambda), \gamma_{i-1}(\lambda), \ldots, \gamma_1(\lambda)\big) \otimes I_n\,, \tag{3.3}$$

$$\widetilde{\Gamma}_i(\lambda) \;:=\; \mathrm{rev}_1\Gamma_i(\lambda) = \mathrm{diag}\big(\widetilde{\gamma}_i(\lambda), \widetilde{\gamma}_{i-1}(\lambda), \ldots, \widetilde{\gamma}_1(\lambda)\big) \otimes I_n\,. \tag{3.4}$$

Note that in light of (3.2) the Newton polynomials can be alternatively defined in terms of the γ_i's via the multiplicative recurrence relation

$$n_0(\lambda) \equiv 1 \quad \text{and} \quad n_i(\lambda) = \prod_{j=1}^{i} \gamma_j(\lambda) = n_{i-1}(\lambda) \cdot \gamma_i(\lambda)\,, \quad i = 1, \ldots, k\,. \tag{3.5}$$

We now turn our attention to the fundamental objects of our study, matrix polynomials of the form

$$P(\lambda) = \sum_{i=0}^{k} A_i n_i(\lambda)\,, \quad \text{where} \quad A_0, \ldots, A_k \in \mathbb{F}^{n\times n}\,, \tag{3.6}$$

i.e., matrix polynomials expressed in a Newton basis. Recall that the aim of this work is to find ways to easily generate families of strong linearizations for matrix polynomials of the form (3.6) by using the matrix coefficients A_i *directly*, without first converting $P(\lambda)$ in (3.6) into the standard basis. For the remainder of Section 3, we develop one approach to achieving this aim by studying generalizations of the ansatz spaces in Definition 2.14 that are better adapted to Newton bases. In Section 4 we then develop a completely independent approach to achieving this aim that adapts Fiedler pencils [8] to matrix polynomials expressed in a Newton basis.

3.1. *Generalized ansatz vector spaces*

Let $P(\lambda)$ be a square matrix polynomial of grade k, and recall the vector spaces of pencils defined in (2.5). We consider analogous vector spaces obtained by generalizing the ansatz relation $L(\lambda)(\Lambda \otimes I_n) = v \otimes P(\lambda)$.

Definition 3.1. *Let P be an $n \times n$ matrix polynomial of grade k, and define the following spaces of $nk \times nk$* pencils *associated with P:*

$$\begin{aligned}
\mathcal{N}_1(P) &:= \left\{L(\lambda) : L(\lambda)(N \otimes I_n) = v \otimes P(\lambda),\ \text{for some } v \in \mathbb{F}^k\right\}, \\
\mathcal{N}_2(P) &:= \left\{L(\lambda) : (N^T \otimes I_n)L(\lambda) = w^T \otimes P(\lambda),\ \text{for some } w \in \mathbb{F}^k\right\}, \\
\mathcal{DN}(P) &:= \mathcal{N}_1(P) \cap \mathcal{N}_2(P)\,,
\end{aligned} \tag{3.7}$$

where $N := [n_{k-1}(\lambda), \cdots, n_1(\lambda), n_0(\lambda)]^T \in \mathbb{F}[\lambda]^{k\times 1}$ is a column of Newton polynomials as in (3.5).

Convention: Throughout the rest of this paper we assume that the ordered list of nodes $\mathcal{A} = (\alpha_1, \ldots, \alpha_k)$ is arbitrary but fixed, and that $\{n_i(\lambda)\}_{i=0}^k$ are the associated (generalized) Newton polynomials.

There are several important observations regarding Definitions 2.14 and 3.1. First, it should be emphasized that neither of these two definitions depend on the basis used to express either P or L. On the other hand, we will see that the *ease of construction* of pencils satisfying either (2.5) or (3.7), as well as the simplicity of the relationship between the matrix coefficients of L and P, *can* be greatly affected by the choice of basis for L and P. Second, the similarity of Definitions 2.14 and 3.1 suggests the possibility of a close relationship between the spaces $\mathbb{L}_1(P)$ and $\mathcal{N}_1(P)$; such a relationship is described in Proposition 3.2. An analogous relation exists between $\mathbb{L}_2(P)$ and $\mathcal{N}_2(P)$ spaces, but for the sake of brevity we omit it.

Proposition 3.2. *Let P be an $n \times n$ matrix polynomial of grade k. Then $\mathbb{L}_1(P)$ and $\mathcal{N}_1(P)$ are isomorphic as vector spaces.*

Proof. The fact that $\mathbb{L}_1(P)$ and $\mathcal{N}_1(P)$ are vector spaces follows directly from their definitions and the properties of Kronecker product. Further, the fact that $\left\{\lambda^i\right\}_{i=0}^{k-1}$ and $\{n_i(\lambda)\}_{i=0}^{k-1}$ are both bases for $\mathcal{P}_{k-1}$ implies the existence of a *nonsingular, constant* change-of-basis matrix S such that $S \cdot \Lambda(\lambda) = N(\lambda)$, where $\Lambda(\lambda)$ and $N(\lambda)$ are defined in Definitions 2.14 and 3.1, respectively. Now consider the mapping

$$\begin{aligned}
\mathbb{L}_1(P) &\xrightarrow{\mathfrak{S}} \mathcal{N}_1(P) \\
K(\lambda) &\mapsto K(\lambda) \cdot (S^{-1} \otimes I_n)\,.
\end{aligned}$$

We start by showing that $\mathfrak{S}$ is a well-defined map. Suppose $K(\lambda) \in \mathbb{L}_1(P)$ with right ansatz vector v. Then

$$\begin{aligned} K(\lambda)\cdot(\Lambda \otimes I_n) = v \otimes P(\lambda) \quad &\Longleftrightarrow \quad K(\lambda)\cdot(S^{-1} \otimes I_n)\cdot(S \otimes I_n)\cdot(\Lambda(\lambda) \otimes I_n) = v \otimes P(\lambda) \\ &\Longleftrightarrow \quad K(\lambda)\cdot(S^{-1} \otimes I_n)\cdot(N(\lambda) \otimes I_n) = v \otimes P(\lambda)\,. \end{aligned}$$

This last equation implies that $K(\lambda)\cdot(S^{-1} \otimes I_n) \in \mathcal{N}_1(P)$ with right ansatz vector v, showing that $\mathfrak{S}$ is a well-defined map from $\mathbb{L}_1(P)$ to $\mathcal{N}_1(P)$. It is now straightforward to check that $\mathfrak{S}$ is also a linear map.

Next observe that by a completely analogous argument, the map

$$\begin{array}{ccc} \mathcal{N}_1(P) & \xrightarrow{\mathfrak{T}} & \mathbb{L}_1(P) \\ H(\lambda) & \mapsto & H(\lambda)\cdot(S \otimes I_n)\,, \end{array}$$

is also well defined and linear. Finally, it is easy to see that $\mathfrak{T}$ is the inverse mapping of $\mathfrak{S}$, showing that $\mathfrak{S}$ is a linear isomorphism between $\mathbb{L}_1(P)$ and $\mathcal{N}_1(P)$, as desired. $\square$

Given that $\mathbb{L}_1$ and $\mathcal{N}_1$ are isomorphic as vector spaces, one might wonder why it is worth bothering with the $\mathcal{N}_1$ space at all, since so much is already known about $\mathbb{L}_1$. But when P is expressed in Newton basis as in (3.6), it turns out to be more natural to look for strong linearizations in the $\mathcal{N}_1(P)$ and $\mathcal{N}_2(P)$ spaces, rather than in either $\mathbb{L}_1(P)$ or $\mathbb{L}_2(P)$. In particular, pencils in $\mathcal{N}_1$ and $\mathcal{N}_2$ are much easier to construct from the matrix coefficients of P than are the pencils in the $\mathbb{L}_1$ and $\mathbb{L}_2$ spaces, especially if the pencils do not need to be block-symmetric.

We now establish some technical results needed to effectively study the $\mathcal{N}_1$ and $\mathcal{N}_2$ spaces.

Lemma 3.3. *Let $\mathcal{B} := \{\phi_i(\lambda)\}_{i=0}^{\ell-1}$ be any set of ℓ* linearly independent *scalar polynomials and define $\Phi := [\phi_{\ell-1}(\lambda), \cdots, \phi_1(\lambda), \phi_0(\lambda)]^T$. If C is a constant $1 \times \ell n$ matrix, then*

$$C\cdot(\Phi \otimes I_n) = 0_{1\times n} \quad \Longleftrightarrow \quad C = 0_{1\times \ell n}\,.$$

Proof. ($\Leftarrow$) This direction is obvious.
($\Rightarrow$) Assume $C\cdot(\Phi \otimes I_n) = 0_{1\times n}$ and let $1 \le j \le n$. The $(1,j)$ entry of $C\cdot(\Phi \otimes I_n)$ is given by

$$0 = \Big(C\cdot(\Phi \otimes I_n)\Big)_{1,j} = C\cdot(\Phi \otimes I_n)_{:,j} = \sum_{h=0}^{\ell-1} C_{1,j+(\ell-1-h)\cdot n}\cdot\phi_h(\lambda)\,. \tag{3.8}$$

Since $\mathcal{B}$ is a linearly independent set by assumption, (3.8) implies that $C_{1,j+(\ell-1-h)\cdot n} = 0$ for all $h = 0,\ldots,\ell-1$. But j was also arbitrary, so it follows that the matrix C is the $1\times \ell n$ zero matrix as desired. $\square$

Before stating the next result we consider two new column vectors obtained by appending an additional element to $\Lambda(\lambda)$ and $N(\lambda)$ from Definitions 2.14 and 3.1, respectively. In particular, we define the $(k+1)\times 1$ column vectors

$$\breve{\Lambda}(\lambda) := \left[\begin{array}{c} \lambda^k \\ \hline \Lambda(\lambda) \end{array}\right] = \begin{bmatrix} \lambda^k \\ \lambda^{k-1} \\ \vdots \\ \lambda \\ 1 \end{bmatrix} \quad \text{and} \quad \breve{N}(\lambda) := \left[\begin{array}{c} n_k(\lambda) \\ \hline N(\lambda) \end{array}\right] = \begin{bmatrix} n_k(\lambda) \\ n_{k-1}(\lambda) \\ \vdots \\ n_1(\lambda) \\ n_0(\lambda) \end{bmatrix}. \tag{3.9}$$

Using Γ_k from (3.3) with $n = 1$, the multiplicative recurrence relations (3.5) can now be expressed as

$$\Gamma_k(\lambda)\cdot N(\lambda) = \begin{bmatrix} \gamma_k(\lambda) & & & \\ & \gamma_{k-1}(\lambda) & & \\ & & \ddots & \\ & & & \gamma_1(\lambda) \end{bmatrix} \cdot \begin{bmatrix} n_{k-1}(\lambda) \\ n_{k-2}(\lambda) \\ \vdots \\ n_0(\lambda) \end{bmatrix} = \begin{bmatrix} n_k(\lambda) \\ n_{k-1}(\lambda) \\ \vdots \\ n_1(\lambda) \end{bmatrix}, \tag{3.10}$$

where the right-hand side of (3.10) comprises the *top* k rows of $\breve{N}(\lambda)$. The relation (3.10) can be viewed as the "Newton analog" of the equation $\lambda I \cdot \Lambda(\lambda) = \left[\lambda^k, \lambda^{k-1}, \dots, \lambda\right]^T$, and is a *crucial* fact for the following Lemma 3.4.

Lemma 3.4. *Let $P(\lambda) = \sum_{i=0}^{k} A_i n_i(\lambda)$ be an $n \times n$ matrix polynomial of grade k in a Newton basis, and define the* partner polynomial *$\widehat{P}(\lambda) := \sum_{i=0}^{k} A_i \lambda^i$, using the same coefficients A_i as in $P(\lambda)$. Then for matrices $X, Y \in \mathbb{F}^{nk \times nk}$,*

$$\lambda X + Y \in \mathbb{L}_1(\widehat{P}) \quad \Longleftrightarrow \quad X \cdot \Gamma_k + Y \in \mathcal{N}_1(P)\,,$$

where Γ_k is as in (3.3). Moreover, the pencils $\lambda X + Y$ and $X \cdot \Gamma_k + Y$ share the same ansatz vector v.

Proof. Let $\lambda X + Y \in \mathbb{L}_1(\widehat{P})$. Then there exists a right ansatz vector $v \in \mathbb{F}^k$ such that $(\lambda X + Y)(\Lambda \otimes I_n) = v \otimes \widehat{P}(\lambda)$. The properties of block matrix multiplication and Kronecker product then give the following chain of equalities:

$$\begin{aligned}
(\lambda X + Y) \cdot (\Lambda \otimes I_n) \;&=\; v \otimes \widehat{P}(\lambda) \qquad (3.11)\\
\lambda X \cdot (\Lambda \otimes I_n) \;+\; Y \cdot (\Lambda \otimes I_n) \;&=\; (v \cdot 1) \otimes \Big(\big[\, A_k\, A_{k-1} \,\cdots\, A_0 \,\big] \cdot (\breve{\Lambda} \otimes I_n) \Big)\\
\Big(\big[\, X \,|\, 0_{kn \times n} \,\big] \;+\; \big[\, 0_{kn \times n} \,|\, Y \,\big] \Big) \cdot (\breve{\Lambda} \otimes I_n) \;&=\; \Big(v \otimes \big[\, A_k\, A_{k-1} \,\cdots\, A_0 \,\big] \Big) \cdot (\breve{\Lambda} \otimes I_n)\,.
\end{aligned}$$

From Lemma 3.3 it now follows that

$$\big[\, X \,|\, 0_{kn \times n} \,\big] \;+\; \big[\, 0_{kn \times n} \,|\, Y \,\big] \;=\; v \otimes \big[\, A_k\, A_{k-1} \,\cdots\, A_0 \,\big]\,, \qquad (3.12)$$

and so

$$\Big(\big[\, X \,|\, 0_{kn \times n} \,\big] + \big[\, 0_{kn \times n} \,|\, Y \,\big] \Big) \cdot (\breve{N} \otimes I_n) \;=\; \Big(v \otimes \big[A_k\, A_{k-1} \,\cdots\, A_0 \big] \Big) \cdot (\breve{N} \otimes I_n)\,.$$

Then

$$\big(X \cdot \Gamma_k + Y \big) \cdot (N \otimes I_n) \;=\; v \otimes P(\lambda) \qquad (3.13)$$

follows from (3.10), showing that $X \cdot \Gamma_k + Y \in \mathcal{N}_1(P)$. It is easy to see that the argument is reversible, leading back from the defining equation (3.13) for $X \cdot \Gamma_k + Y \in \mathcal{N}_1(P)$ through (3.12), and then on to the defining equation (3.11) for $\lambda X + Y \in \mathbb{L}_1(\widehat{P})$. ☐

Observe that Lemma 3.4 enables us to easily construct pencils in $\mathcal{N}_1(P)$ by leveraging the knowledge of how to construct pencils in $\mathbb{L}_1(\widehat{P})$. Now the remaining question is to determine *if* and *when* pencils from $\mathcal{N}_1(P)$ are strong linearizations, given a P in Newton basis as in (3.6). Before tackling this question we introduce a new concept, analogous to a notion first used in [7].

Definition 3.5. ($\mathcal{Z}$-rank of pencils in $\mathcal{N}_1$ or $\mathbb{L}_1$ with ansatz vector e_1)
Let $Q(\lambda)$ be an $n \times n$ matrix polynomial of grade k, and let $V(\lambda)$ be an $nk \times nk$ pencil in either $\mathcal{N}_1(Q)$ or $\mathbb{L}_1(Q)$, with ansatz vector $e_1 \in \mathbb{F}^k$. Partition $V(\lambda)$ as

$$V(\lambda) = \left[\begin{array}{c} V_T(\lambda) \\ \hline V_B(\lambda) \end{array} \right], \quad \text{with } V_T(\lambda) \in \mathbb{F}^{n \times nk}[\lambda] \quad \text{and} \quad V_B(\lambda) \in \mathbb{F}^{n(k-1) \times nk}[\lambda]\,.$$

*The $\mathcal{Z}$-*rank *of $V(\lambda)$, denoted by $\mathcal{Z}(V(\lambda))$, is defined to be*

$$\mathcal{Z}(V(\lambda)) := \operatorname{nrank}\Big(V_B(\lambda) \Big). \qquad (3.14)$$

The pencil $V(\lambda)$ is said to have full $\mathcal{Z}$-rank *if $\mathcal{Z}\big(V(\lambda)\big) = n(k-1)$.*

Before continuing, we briefly explain the choice of notation. When the pencil $V(\lambda)$ from Definition 3.5 is partitioned, the subscript T is used to denote the *top* n rows of $V(\lambda)$, whereas the subscript B is used to denote the *bottom* $n(k-1)$ rows of $V(\lambda)$.

Theorem 3.6. *Let $P(\lambda) = \sum_{i=0}^{k} A_i n_i(\lambda)$ be an $n\times n$ matrix polynomial of grade k, and consider the partner polynomial defined as in* Lemma 3.4*, i.e., $\widehat{P}(\lambda) := \sum_{i=0}^{k} A_i\lambda^i$. If $\widehat{L}(\lambda) = \lambda X + Y \in \mathbb{L}_1(\widehat{P})$ with ansatz vector e_1, then:*

(a) *the matrices $X, Y \in \mathbb{F}^{nk\times nk}$ are of the form*

$$X = \left[\begin{array}{c|c} A_k & X_{12} \\ \hline 0 & -Z \end{array}\right] \quad \text{and} \quad Y = \left[\begin{array}{c|c} Y_{11} & A_0 \\ \hline Z & 0 \end{array}\right],$$

where Z is a $(k-1)n \times (k-1)n$ constant *matrix.*

(b) *the pencil $L(\lambda) := X\cdot\Gamma_k + Y$ is in $\mathcal{N}_1(P)$ with ansatz vector e_1.*

(c) *$\widehat{L}(\lambda)$ and $L(\lambda)$ are unimodularly equivalent to* $\left[\begin{array}{c|c} A(\lambda) & \widehat{P}(\lambda) \\ \hline Z & 0 \end{array}\right]$ and $\left[\begin{array}{c|c} B(\lambda) & P(\lambda) \\ \hline Z & 0 \end{array}\right]$, *respectively, with the* same *$(2,1)$-block Z.*

(d) $\mathcal{Z}\Big(\widehat{L}(\lambda)\Big) = \operatorname{rank}\big(Z\big) = \mathcal{Z}\Big(L(\lambda)\Big)$.

Proof. Part (a) follows from [22, Cor. 3.7]; Lemma 3.4 with $v = e_1$ implies (b).
To see why (c) is true consider the unimodular matrices

$$\widehat{G}(\lambda) := \begin{bmatrix} I_n & & & & \lambda^{k-1}I_n \\ & I_n & & & \lambda^{k-2}I_n \\ & & \ddots & & \vdots \\ & & & I_n & \lambda I_n \\ & & & & I_n \end{bmatrix}, \quad G(\lambda) := \begin{bmatrix} I_n & & & & n_{k-1}I_n \\ & I_n & & & n_{k-2}I_n \\ & & \ddots & & \vdots \\ & & & I_n & n_1 I_n \\ & & & & I_n \end{bmatrix}, \tag{3.15}$$

$$\text{and} \quad H_i\big(\zeta(\lambda)\big) := \big(I_k + \zeta(\lambda)\cdot E_{i,i+1}\big)\otimes I_n\,, \tag{3.16}$$

where $\zeta(\lambda) \in \mathbb{F}[\lambda]$ and $E_{i,j}$ is the $k\times k$ zero matrix with (i,j)-entry set equal to 1. It is worth emphasizing that the last block columns of matrices $\widehat{G}(\lambda)$ and $G(\lambda)$ are just $(\Lambda\otimes I_n)$ and $(N\otimes I_n)$, respectively. Thus, post-multiplication of $L(\lambda)$ by $G(\lambda)$ (resp., $\widehat{L}(\lambda)$ by $\widehat{G}(\lambda)$) simply recovers the $e_1\otimes P$ (resp., $e_1\otimes\widehat{P}$) part of the corresponding ansatz relation in the last block column of the product. On the other hand, post-multiplication of an arbitrary matrix $R(\lambda)$ by $H_i(\zeta(\lambda))$ is equivalent to the ith column of R being scaled by $\zeta(\lambda)$ and then added to the $(i+1)$th column of R.

Now define matrices $\widehat{T}(\lambda)$ and $T(\lambda)$ by

$$\begin{aligned} \widehat{T}(\lambda) &:= \widehat{G}(\lambda)\cdot H_1(\lambda)\cdot H_2(\lambda)\cdots H_{k-2}(\lambda)\,, \\ T(\lambda) &:= G(\lambda)\cdot H_1(\gamma_{k-1})\cdot H_2(\gamma_{k-2})\cdots H_{k-2}(\gamma_2)\,, \end{aligned}$$

and observe that both are unimodular, since they are the product of unimodular matrices. Then using parts (a) and (b) it is straightforward to verify that

$$\widehat{L}(\lambda)\widehat{T}(\lambda) = \left[\begin{array}{c|c} A(\lambda) & \widehat{P}(\lambda) \\ \hline Z & 0 \end{array}\right] \quad \text{and} \quad L(\lambda)T(\lambda) = \left[\begin{array}{c|c} B(\lambda) & P(\lambda) \\ \hline Z & 0 \end{array}\right], \tag{3.17}$$

where $A(\lambda)$ and $B(\lambda)$ are some matrix polynomials. This proves part (c).

Finally, to verify (d), we first partition $\widehat{L}(\lambda) = \begin{bmatrix} \widehat{L}_T(\lambda) \\ \widehat{L}_B(\lambda) \end{bmatrix}$, where $\widehat{L}_T(\lambda) \in \mathbb{F}[\lambda]^{n\times kn}$. Then (3.17) gives

$$\widehat{L}(\lambda)\cdot\widehat{T}(\lambda) = \left[\begin{array}{c} \widehat{L}_T(\lambda)\cdot\widehat{T}(\lambda) \\ \hline \widehat{L}_B(\lambda)\cdot\widehat{T}(\lambda) \end{array}\right] = \left[\begin{array}{c|c} A(\lambda) & \widehat{P}(\lambda) \\ \hline Z & 0 \end{array}\right]. \tag{3.18}$$

From the definition of $\mathcal{Z}$-rank for pencils with ansatz vector e_1, we then have

$$\mathcal{Z}\Big(\widehat{L}(\lambda)\Big) = \operatorname{nrank}\Big(\widehat{L}_B(\lambda)\Big) = \operatorname{nrank}\Big(\widehat{L}_B(\lambda)\cdot\widehat{T}(\lambda)\Big) = \operatorname{nrank}\Big(\big[Z\,|\,0\big]\Big) = \operatorname{rank}(Z)\,,$$

where the second equality is a consequence of the fact that normal rank is invariant under unimodular transformations; this proves the first part of (d). To complete the proof, also partition $L(\lambda) = \begin{bmatrix} L_T(\lambda) \\ L_B(\lambda) \end{bmatrix}$ with $L_T(\lambda) \in \mathbb{F}[\lambda]^{n\times nk}$, and observe from (3.17) that

$$L(\lambda)\cdot T(\lambda) = \left[\begin{array}{c} L_T(\lambda)\cdot T(\lambda) \\ \hline L_B(\lambda)\cdot T(\lambda) \end{array}\right] = \left[\begin{array}{c|c} B(\lambda) & P(\lambda) \\ \hline Z & 0 \end{array}\right]. \tag{3.19}$$

Then $\mathcal{Z}\big(L(\lambda)\big) = \operatorname{rank}\big(Z\big)$ by an argument completely analogous to the one just used to find $\mathcal{Z}\big(\widehat{L}(\lambda)\big)$. $\square$

The next result provides a sufficient condition for a pencil in $\mathcal{N}_1(P)$ with ansatz vector e_1 to be a strong linearization for a square matrix polynomial P.

Theorem 3.7. *Let $P(\lambda) = \sum_{i=0}^{k} A_i n_i(\lambda)$ be an $n\times n$ matrix polynomial of grade k, and let $L(\lambda) \in \mathcal{N}_1(P)$ with ansatz vector e_1. If $L(\lambda)$ has full $\mathcal{Z}$-rank, then $L(\lambda)$ is a strong linearization for $P(\lambda)$.*

Proof. We first show that a pencil $L(\lambda)$ with the given properties is a linearization for $P(\lambda)$, and then see why this linearization is strong. From Theorem 3.6(c) we know there exists an $nk\times nk$ unimodular matrix $T(\lambda)$ and a constant matrix $W \in \mathbb{F}^{n(k-1)\times n(k-1)}$ such that

$$L(\lambda)T(\lambda) = \left[\begin{array}{c|c} B(\lambda) & P(\lambda) \\ \hline W & 0 \end{array}\right].$$

Further, the hypothesis that L has full $\mathcal{Z}$-rank together with Theorem 3.6(d) implies that

$$\mathcal{Z}\big(L(\lambda)\big) = \operatorname{rank}(W) = (k-1)n\,, \tag{3.20}$$

and hence that W is *nonsingular*. Consequently, the unimodular equivalence

$$\left[\begin{array}{c|c} I_n & -B(\lambda)\cdot W^{-1} \\ \hline & W^{-1} \end{array}\right]\cdot L(\lambda)\cdot T(\lambda)\cdot\left[\begin{array}{c|c} & I_{(k-1)n} \\ \hline I_n & \end{array}\right] = \left[\begin{array}{c|c} P(\lambda) & \\ \hline & I_{(k-1)n} \end{array}\right]$$

shows that $L(\lambda)$ is a linearization for $P(\lambda)$.

To see why $L(\lambda)$ is a *strong* linearization for $P(\lambda)$, all that remains is to show that $\operatorname{rev}_1 L(\lambda)$ is a linearization for $\operatorname{rev}_k P(\lambda)$. Applying rev_k to each side of the ansatz relation $L(\lambda)\big(N(\lambda)\otimes I\big) = e_1\otimes P(\lambda)$, and using properties of reversal from Proposition 2.3 gives on the left-hand side

$$\begin{aligned}\operatorname{rev}_k\Big(L(\lambda)\cdot(N(\lambda)\otimes I_n)\Big) &= \operatorname{rev}_1 L(\lambda)\cdot\Big(\operatorname{rev}_{k-1}\big[N(\lambda)\otimes I_n\big]\Big) \\ &= \operatorname{rev}_1 L(\lambda)\cdot\Big(\big[\operatorname{rev}_{k-1}N(\lambda)\big]\otimes I_n\Big)\,,\end{aligned} \tag{3.21}$$

and on the right-hand side

$$\operatorname{rev}_k\big(e_1\otimes P(\lambda)\big) = e_1\otimes\operatorname{rev}_k P(\lambda)\,, \tag{3.22}$$

so that

$$\mathrm{rev}_1 L(\lambda) \cdot \Big(\mathrm{rev}_{k-1} N(\lambda) \otimes I_n\Big) = e_1 \otimes \mathrm{rev}_k P(\lambda)\,. \tag{3.23}$$

On the other hand, Proposition 2.3(f) implies that $\{\mathrm{rev}_{k-1}(n_i(\lambda))\}_{i=0}^{k-1}$ is a set of k linearly independent polynomials, each with degree less than or equal to $k-1$, and therefore a basis for $\mathcal{P}_{k-1}$. Consequently, there exists a *nonsingular, constant* change-of-basis matrix U such that $U \cdot N(\lambda) = \mathrm{rev}_{k-1} N(\lambda)$, so that (3.23) becomes

$$\mathrm{rev}_1 L(\lambda) \cdot \Big(\big[U \cdot N(\lambda)\big] \otimes I_n\Big) = \mathrm{rev}_1 L(\lambda) \cdot \Big(U \otimes I_n\Big) \cdot \Big(N(\lambda) \otimes I_n\Big) = e_1 \otimes \mathrm{rev}_k P(\lambda)\,. \tag{3.24}$$

Now defining the pencil

$$K(\lambda) := \mathrm{rev}_1 L(\lambda) \cdot (U \otimes I)\,, \tag{3.25}$$

we see from (3.24) that $K(\lambda) \in \mathcal{N}_1(\mathrm{rev}_k P(\lambda))$ with ansatz vector e_1. Since $\mathrm{rev}_1 L(\lambda)$ and $K(\lambda)$ are strictly equivalent, the pencil $\mathrm{rev}_1 L(\lambda)$ is a linearization for $\mathrm{rev}_k P(\lambda)$ if and only if $K(\lambda)$ is a linearization for $\mathrm{rev}_k P(\lambda)$ (see Remark 2.6). But we also have

$$\begin{aligned}
\mathcal{Z}\Big(L(\lambda)\Big) &\overset{(1)}{=} \mathrm{nrank}\Big(L_B(\lambda)\Big) \overset{(2)}{=} \mathrm{nrank}\Big(\mathrm{rev}_1\big[L_B(\lambda)\big]\Big) \overset{(3)}{=} \mathrm{nrank}\Big(\big[\mathrm{rev}_1 L(\lambda)\big]_B\Big) \\
&\overset{(4)}{=} \mathrm{nrank}\Big(\big[\mathrm{rev}_1 L(\lambda)\big]_B \cdot (U \otimes I_n)\Big) \\
&\overset{(5)}{=} \mathrm{nrank}\Big(\big[\mathrm{rev}_1 L(\lambda) \cdot (U \otimes I_n)\big]_B\Big) \overset{(6)}{=} \mathcal{Z}\Big(K(\lambda)\Big)\,.
\end{aligned} \tag{3.26}$$

The equalities (1) and (6) in (3.26) follow from Definition 3.5, while (2) and (3) are consequences of Proposition 2.3(e) and (d), respectively. The fact that rank is invariant under multiplication by a nonsingular matrix implies equality (4), while (5) is just a property of matrix multiplication.

Thus we see from (3.26) that full $\mathcal{Z}$-rank of $K(\lambda)$ follows from the full $\mathcal{Z}$-rank of $L(\lambda)$. Since $K(\lambda) \in \mathcal{N}_1(\mathrm{rev}_k P(\lambda))$ with ansatz vector e_1 and has full $\mathcal{Z}$-rank, we have by the result of the first part of this proof that $K(\lambda)$ is a linearization for $\mathrm{rev}_k P(\lambda)$, as desired. □

Example 3.8. Let $P(\lambda)$ and $\widehat{P}(\lambda)$ be $n \times n$ matrix polynomials of grade k defined by

$$P(\lambda) = \sum_{i=0}^{k} A_i n_i(\lambda) \quad \text{and} \quad \widehat{P}(\lambda) := \sum_{i=0}^{k} A_i \lambda^i\,,$$

with the same coefficients A_i. From [22] we know that the Frobenius companion pencil $\widehat{L}_1(\lambda) := \lambda X_1 + Y_1$ with X_1 and Y_1 given in (2.4) is in $\mathbb{L}_1(\widehat{P})$ with ansatz vector e_1. Theorem 3.6(b) then implies that $L_1(\lambda) := X_1 \cdot \Gamma_k + Y_1$ is in $\mathcal{N}_1(P)$ with ansatz vector e_1. Now note that both $\widehat{L}_1(\lambda)$ and $L_1(\lambda)$ have full $\mathcal{Z}$-rank, so Theorem 3.7 implies that $L_1(\lambda)$ is a strong linearization for any regular or singular matrix polynomial $P(\lambda)$ over an arbitrary field $\mathbb{F}$. A pencil very similar to $L_1(\lambda)$, indeed strictly equivalent to $L_1(\lambda)$, was shown in [1] to be a strong linearization when P is regular and $\mathbb{F} = \mathbb{C}$, but the argument provided in [1] is completely different from the one given here.

Remark 3.9. If a matrix polynomial is singular, then the converse of Theorem 3.7 need not be true; Example 2 from [7] can be adapted to see this. However, the converse does hold in the regular case.

Corollary 3.10. *Let $P(\lambda)$ be a* regular *matrix polynomial, and let $L(\lambda) \in \mathcal{N}_1(P)$ with ansatz vector e_1. Then the pencil $L(\lambda)$ is a strong linearization for $P(\lambda)$* if and only if *$L(\lambda)$ has full $\mathcal{Z}$-rank.*

Proof. ($\Leftarrow$) This direction is Theorem 3.7.
($\Rightarrow$) If L is a (strong) linearization for P, then regularity of P and Definition 2.7 imply that L is also regular, which in turn implies that L has full $\mathcal{Z}$-rank. □

So far we have seen that the notion of $\mathcal{Z}$-rank plays a key role in determining whether a pencil in $\mathcal{N}_1(P)$ with ansatz vector e_1 is a strong linearization for P. But now this begs the question of whether the notion of $\mathcal{Z}$-rank can be extended in a meaningful (and useful) way to pencils in $\mathcal{N}_1(P)$ with *any* ansatz vector.

Definition 3.11. *Let $Q(\lambda)$ be an $n \times n$ matrix polynomial of grade k, and let $V(\lambda) \in \mathcal{N}_1(Q)$ with an arbitrary nonzero ansatz vector $v \in \mathbb{F}^k$. Further, let M be a* nonsingular, constant *matrix such that $Mv = e_1$. The* $\mathcal{Z}$-rank of $V(\lambda)$ *is defined by*

$$\mathcal{Z}\Big(V(\lambda)\Big) := \mathcal{Z}\Big((M \otimes I_n) \cdot V(\lambda)\Big). \tag{3.27}$$

Note that if $V(\lambda)$ is in $\mathcal{N}_1(Q)$ with ansatz vector v, then the pencil $(M \otimes I_n) \cdot V(\lambda)$ is also in $\mathcal{N}_1(Q)$ but with ansatz vector e_1. This shows that the right hand side of (3.27) is defined. However, it is not immediately clear that (3.27) provides a notion of the $\mathcal{Z}$-rank that is *well-defined*, i.e., that (3.27) is *independent* of the choice of the matrix M. The following lemma addresses this issue.

Lemma 3.12. *Let $P(\lambda) = \sum_{i=0}^{k} A_i n_i(\lambda)$ be an $n \times n$ matrix polynomial of grade k, and let $L(\lambda) = X \cdot \Gamma_k + Y$ be a pencil in $\mathcal{N}_1(P)$ with a nonzero ansatz vector $v \in \mathbb{F}^k$. Then the $\mathcal{Z}$-rank of L is* independent *of the choice of the matrix M in* Definition 3.11.

Proof. We start by defining the partner polynomial $\widehat{P}(\lambda) := \sum_{i=0}^{k} A_i \lambda^i$ as in Lemma 3.4 and Theorem 3.6. Then Lemma 3.4 implies that $\widehat{L}(\lambda) = \lambda X + Y \in \mathbb{L}_1(\widehat{P})$, also with ansatz vector v, that is, $\widehat{L}(\lambda)(\Lambda \otimes I) = v \otimes \widehat{P}(\lambda)$.

Now let $M_1 \neq M_2$ be any two nonsingular $k \times k$ matrices such that $M_1 v = M_2 v = e_1$, and define new pencils

$$\begin{aligned} \widehat{L}_1(\lambda) &:= (M_1 \otimes I) \cdot \widehat{L}(\lambda) = \lambda \cdot \{(M_1 \otimes I)X\} + \{(M_1 \otimes I)Y\}, \\ \widehat{L}_2(\lambda) &:= (M_2 \otimes I) \cdot \widehat{L}(\lambda) = \lambda \cdot \{(M_2 \otimes I)X\} + \{(M_2 \otimes I)Y\}. \end{aligned} \tag{3.28}$$

It is easy to verify that

$$\widehat{L}_1(\lambda)(\Lambda \otimes I) = e_1 \otimes \widehat{P}(\lambda) = \widehat{L}_2(\lambda)(\Lambda \otimes I),$$

i.e., $\widehat{L}_1(\lambda)$ and $\widehat{L}_2(\lambda)$ are both pencils in $\mathbb{L}_1(\widehat{P})$ with ansatz vector e_1. In [7, Lemma 4.2] it is shown that two pencils in $\mathbb{L}_1(\widehat{P})$ such as those in (3.28) have the same $\mathcal{Z}$-rank, i.e.,

$$\mathcal{Z}\Big(\widehat{L}_1(\lambda)\Big) = \mathcal{Z}\Big(\widehat{L}_2(\lambda)\Big). \tag{3.29}$$

Recalling that the pencil $L(\lambda) := X \cdot \Gamma_k + Y$ is by assumption in $\mathcal{N}_1(P)$ with ansatz vector v, define new pencils

$$\begin{aligned} L_1(\lambda) &:= (M_1 \otimes I) \cdot L(\lambda) = \{(M_1 \otimes I)X\} \cdot \Gamma_k + \{(M_1 \otimes I)Y\}, \\ L_2(\lambda) &:= (M_2 \otimes I) \cdot L(\lambda) = \{(M_2 \otimes I)X\} \cdot \Gamma_k + \{(M_2 \otimes I)Y\}. \end{aligned}$$

Applying Theorem 3.6(b) to $\widehat{L}_1(\lambda)$ and $\widehat{L}_2(\lambda)$ in (3.28) shows that $L_1(\lambda)$ and $L_2(\lambda)$ are pencils in $\mathcal{N}_1(P)$ with ansatz vector e_1. But even more is true, in particular,

$$\mathcal{Z}\Big(L_1(\lambda)\Big) = \mathcal{Z}\Big(\widehat{L}_1(\lambda)\Big) = \mathcal{Z}\Big(\widehat{L}_2(\lambda)\Big) = \mathcal{Z}\Big(L_2(\lambda)\Big). \tag{3.30}$$

The first and third equalities in (3.30) follow from Theorem 3.6(d), while the second equality is just (3.29). Hence the $\mathcal{Z}$-rank of $L(\lambda)$ is independent of the choice of M. $\square$

The next result now extends Theorem 3.7 to include pencils in $\mathcal{N}_1(P)$ with *any* nonzero ansatz vector v, not just ones with $v = e_1$.

Theorem 3.13. *Let $P(\lambda) = \sum_{i=0}^{k} A_i n_i(\lambda)$ be an $n \times n$ matrix polynomial of grade k, and let $L(\lambda) \in \mathcal{N}_1(P)$ with any nonzero ansatz vector $v \in \mathbb{F}^k$. If $L(\lambda)$ has full $\mathcal{Z}$-rank, then $L(\lambda)$ is a strong linearization for P.*

Proof. Let M be any nonsingular, constant matrix such $Mv = e_1$. Then the pencil $K(\lambda) := (M \otimes I_n) \cdot L(\lambda)$ is in $\mathcal{N}_1(P)$ with ansatz vector e_1, and has full $\mathcal{Z}$-rank by assumption. But now Theorem 3.7 implies that $K(\lambda)$ is a strong linearization for P. Since $L(\lambda)$ is strictly equivalent to $K(\lambda)$, it must also be a strong linearization for P (see Remark 2.6). $\square$

Analogous to Corollary 3.10, the converse of Theorem 3.13 is true for regular matrix polynomials. The proof is omitted since it is nearly identical to the proof of Corollary 3.10.

Corollary 3.14. *Let $P(\lambda)$ be a* regular *matrix polynomial, and let $L(\lambda) \in \mathcal{N}_1(P)$ with arbitrary ansatz vector $v \neq 0$. Then the pencil $L(\lambda)$ is a strong linearization for $P(\lambda)$* if and only if *$L(\lambda)$ has full $\mathcal{Z}$-rank.*

We conclude this section by combining several of the preceding results together into a single theorem, that can then be used to show that *almost every* pencil in $\mathcal{N}_1(P)$ is a strong linearization for a regular or singular matrix polynomial P.

Theorem 3.15. *Let P and $\widehat{P}$ be $n \times n$ matrix polynomials of grade k of the form*

$$P(\lambda) = \sum_{i=0}^{k} A_i n_i(\lambda) \qquad \text{and} \qquad \widehat{P}(\lambda) := \sum_{i=0}^{k} A_i \lambda^i \,,$$

both defined using the same coefficients A_i. Then the map

$$\begin{aligned} \mathcal{G} : \mathbb{L}_1(\widehat{P}) &\to \mathcal{N}_1(P) \\ \lambda X + Y &\mapsto X \cdot \Gamma_k + Y \end{aligned}$$

*is a $\mathcal{Z}$-*rank-preserving *linear isomorphism.*

Proof. That $\mathcal{G}$ is well-defined and a bijection follows from Lemma 3.4, while the linearity of $\mathcal{G}$ is easy to check. The fact that $\mathcal{G}$ preserves $\mathcal{Z}$-rank is a consequence of Definition 3.11 and Theorem 3.6(d). $\square$

Recall now a result from [7] about the full $\mathcal{Z}$-rank property of pencils in the classical ansatz space $\mathbb{L}_1$.

Theorem 3.16. [7, Cor. 4.5] *For any (regular or singular) $n \times n$ matrix polynomial $P(\lambda)$ of grade k, almost every pencil in $\mathbb{L}_1(P)$ has full $\mathcal{Z}$-rank. (Here by "almost every" we mean for all but a closed, nowhere dense set of measure zero in $\mathbb{L}_1(P)$.)*

Before we state an analogous result for pencils in the space $\mathcal{N}_1(P)$, we remind the reader that Theorem 3.16 was proved in [7] by showing that the set of all pencils in $\mathbb{L}_1(P)$ that do *not* have full $\mathcal{Z}$-rank is a proper algebraic subset of $\mathbb{L}_1(P)$, and then using the fact that proper algebraic sets are closed, nowhere dense sets of measure zero.

Theorem 3.17. *For any (regular or singular) $n \times n$ matrix polynomial $P(\lambda)$ of grade k, almost every pencil in $\mathcal{N}_1(P)$ has full $\mathcal{Z}$-rank.*

Proof. Let $\mathcal{R}_{\widehat{P}} \subseteq \mathbb{L}_1(\widehat{P})$ denote the subset of all $\mathcal{Z}$-rank-deficient pencils in $\mathbb{L}_1(\widehat{P})$, i.e., all pencils that do *not* have full $\mathcal{Z}$-rank. Then by Theorem 3.15 the $\mathcal{G}$-image $\mathcal{G}(\mathcal{R}_{\widehat{P}}) \subseteq \mathcal{N}_1(P)$ is exactly the subset of all $\mathcal{Z}$-rank-deficient pencils in $\mathcal{N}_1(P)$. This set $\mathcal{G}(\mathcal{R}_{\widehat{P}})$ is a proper algebraic subset of $\mathcal{N}_1(P)$, since the image of a proper algebraic set under any linear isomorphism is again a proper algebraic set. Thus almost every pencil in $\mathcal{N}_1(P)$ has full $\mathcal{Z}$-rank. $\square$

Together with Theorem 3.13, Theorem 3.17 now implies the following corollary.

Corollary 3.18. *For any (regular or singular) $n \times n$ matrix polynomial $P(\lambda)$ of grade k, almost every pencil in $\mathcal{N}_1(P)$ is a strong linearization for P.*

Remark 3.19. In Remark 3.9 it was noted that for singular polynomials P, being a strong linearization in $\mathcal{N}_1(P)$ or $\mathbb{L}_1(P)$ is not equivalent to having full $\mathcal{Z}$-rank. Consequently, in the singular case it is not guaranteed a priori that a linear isomorphism that preserves $\mathcal{Z}$-rank (like $\mathcal{G}$ in Theorem 3.15) will also preserve the property of being a strong linearization. This remains an open question.

3.2. The spaces $\mathcal{N}_2(P)$ and $\mathcal{DN}(P)$

For the sake of brevity, we will not include details about the space $\mathcal{N}_2(P)$. All of the results for $\mathcal{N}_1(P)$ from the previous section can be re-stated for pencils in $\mathcal{N}_2(P)$ as a consequence of the following very simple relation between these spaces; note that Theorem 3.20 is completely analogous to a similar relation [18, Thm. 2.2] between the spaces $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$.

Theorem 3.20. *Let $P(\lambda)$ be an $n \times n$ matrix polynomial of grade k. The* block transpose *map*

$$\begin{array}{ccc} \mathcal{N}_1(P) & \xrightarrow{\mathfrak{B}} & \mathcal{N}_2(P) \\ L(\lambda) & \mapsto & L(\lambda)^{\mathcal{B}} \end{array}$$

is a linear isomorphism *between $\mathcal{N}_1(P)$ and $\mathcal{N}_2(P)$, where pencils in $\mathcal{N}_1(P)$ and $\mathcal{N}_2(P)$ are to be viewed as block $k \times k$ matrices with $n \times n$ blocks, and block transpose is as in* Definition 2.15.

We omit the proof of Theorem 3.20, but emphasize that the key relation needed to show that $\mathfrak{B}$ is a well-defined map is the following consequence of Proposition 2.16:

$$\Big(L(\lambda)(N \otimes I_n)\Big)^{\mathcal{B}} = \Big(v \otimes P(\lambda)\Big)^{\mathcal{B}} \quad \Longleftrightarrow \quad (N^T \otimes I_n) L(\lambda)^{\mathcal{B}} = v^T \otimes P(\lambda)\,.$$

Using Theorem 3.20 and properties of block transposition, results from Section 3.1 are now easily extended to pencils in $\mathcal{N}_2(P)$. For example, adapting Lemma 3.4 to the $\mathcal{N}_2(P)$ space gives, for matrices $X, Y \in \mathbb{F}^{nk \times nk}$, the equivalence

$$\lambda X + Y \in \mathbb{L}_2(\widehat{P}) \quad \Longleftrightarrow \quad \Gamma_k \cdot X + Y \in \mathcal{N}_2(P)\,.$$

Observe that the Γ_k matrix appears on the *left* of X in pencils in $\mathcal{N}_2(P)$, in contrast to being on the right side of X for pencils in $\mathcal{N}_1(P)$.

Earlier we have seen that representing pencils in the space $\mathcal{N}_1(P)$ as $X \cdot \Gamma_k + Y$ was crucial both for their construction and for studying their properties; now Theorem 3.20 allows us to study pencils in $\mathcal{N}_2(P)$ via an analogous representation $\Gamma_k \cdot X + Y$. Surprisingly, neither one of these representations is helpful when trying to construct pencils that are in *both* spaces $\mathcal{N}_1(P)$ and $\mathcal{N}_2(P)$, i.e., pencils in the space $\mathcal{DN}(P)$ (see Definition 3.1). This somewhat unexpected behavior can be attributed to the fact that the linear isomorphism $\mathcal{G}$ in Theorem 3.15 does *not* restrict to a map between the double ansatz spaces $\mathbb{DL}(P)$ and $\mathcal{DN}(P)$. We illustrate this point with an example.

Example 3.21. Let $\mathcal{A} = (\alpha_1, \alpha_2)$ be an ordered list of two *distinct* nodes and consider an $n \times n$ quadratic matrix polynomial $P(\lambda)$ expressed in the corresponding Newton basis, together with its partner polynomial $\widehat{P}(\lambda)$, i.e.,

$$P(\lambda) = A_2 n_2(\lambda) + A_1 n_1(\lambda) + A_0 \qquad \text{and} \qquad \widehat{P}(\lambda) = A_2\lambda^2 + A_1\lambda + A_0\,.$$

Defining X and Y to be the $2n \times 2n$ matrices

$$X := \begin{bmatrix} 0 & A_2 \\ A_2 & A_1 \end{bmatrix} \qquad \text{and} \qquad Y := \begin{bmatrix} -A_2 & 0 \\ 0 & A_0 \end{bmatrix}, \tag{3.31}$$

it is easy to see that $\widehat{L}(\lambda) := \lambda X + Y$ is in the space $\mathbb{DL}(\widehat{P}) \subset \mathbb{L}_1(\widehat{P})$ with ansatz vector $e_2 = [0, 1]^T$. However, the matrix pencil $L(\lambda)$

$$L(\lambda) := \mathcal{G}\Big(\widehat{L}(\lambda)\Big) = X \cdot \Gamma_2 + Y$$

is in $\mathcal{N}_1(P)$ but not in $\mathcal{N}_2(P)$, and hence not in $\mathcal{DN}(P)$; this can be verified by a direct computation.

Example 3.21 suggests that even though the use of Γ_k is an effective way to work with pencils in the $\mathcal{N}_1(P)$ and $\mathcal{N}_2(P)$ spaces, it is not well-adapted to the study of pencils in $\mathcal{DN}(P)$. It is important to mention, though, that a completely different approach based on bivariate polynomials has been used in [27] to effectively investigate the space $\mathcal{DN}(P)$, albeit only for regular polynomials P.

3.3. Eigenvector Recovery

In Section 3.1 we have seen how to construct pencils in the space $\mathcal{N}_1(P)$, and developed a criterion to determine when they are strong linearizations for P. Whenever a pencil L is a strong linearization for P, then all the finite and infinite eigenvalues of L and P are the same, with the same multiplicities. On the other hand, eigenvectors of P and L can never be the same, due to the difference in the sizes of P and L. In Theorem 3.22 we establish several results on how eigenvectors of P can be easily recovered from eigenvectors of any strong linearization belonging to $\mathcal{N}_1(P)$, which in turn are just analogs of known eigenvector recovery results for the space $\mathbb{L}_1(P)$. We note that eigenvector recovery results can also be derived for strong linearizations in $\mathcal{N}_2(P)$, but we omit them here in light of the discussion in Section 3.2.

Theorem 3.22. *Let $P(\lambda)$ be a* regular *$n\times n$ matrix polynomial of grade k over the field $\mathbb{F}$, and define $N(\lambda) = [n_{k-1}(\lambda), \dots, n_0(\lambda)]^T$. Further, let $L(\lambda)$ in $\mathcal{N}_1(P)$ with nonzero ansatz vector v be a strong linearization for $P(\lambda)$. Then*

(a) *For a finite eigenvalue $\lambda_0 \in \mathbb{F}$, $x \in \mathbb{F}^n$ is a right eigenvector of P* if and only if *$N(\lambda_0) \otimes x$ is a right eigenvector of L.*

(b) *$x \in \mathbb{F}^n$ is a right eigenvector of P at ∞* if and only if *$e_1 \otimes x$ is a right eigenvector of L at ∞.*

(c) *If $w \in \mathbb{F}^{kn}$ is a left eigenvector of L with eigenvalue $\lambda_0 \in \mathbb{F}$, then*

$$y := (v^T \otimes I_n)w = \begin{bmatrix} v_1 I_n , \dots , v_k I_n \end{bmatrix} \cdot w$$

is a left eigenvector of P with eigenvalue λ_0, as long as $y \neq 0$.

Proof. Assume $L(\lambda) \in \mathcal{N}_1(P)$ with ansatz vector $v \neq 0$.

(a) Post-multiplication of the ansatz relation $L(\lambda)(N \otimes I_n) = v \otimes P(\lambda)$ by $(1 \otimes x)$ gives the polynomial identity

$$L(\lambda)(N \otimes x) = v \otimes P(\lambda)x \,. \tag{3.32}$$

It follows that x is a right eigenvector of P at a finite eigenvalue λ_0 if and only if $N(\lambda_0) \otimes x$ is a right eigenvector of L at λ_0.

(b) Taking the reversal with respect to grade k of both sides of the ansatz relation $L(\lambda)(N \otimes I_n) = v \otimes P(\lambda)$, together with Proposition 2.3 gives

$$\mathrm{rev}_1 L(\lambda)\big(\mathrm{rev}_{k-1} N(\lambda) \otimes I_n\big) = v \otimes \mathrm{rev}_k P(\lambda) \,. \tag{3.33}$$

Now evaluating (3.33) at $\lambda = 0$ and multiplying both sides on the right by $1 \otimes x$ gives

$$\mathrm{rev}_1 L\big|_{\lambda=0} \cdot (e_1 \otimes x) = v \otimes \big(\mathrm{rev}_k P\big|_{\lambda=0} \cdot x\big) \,. \tag{3.34}$$

It follows that $e_1 \otimes x$ is an eigenvector of $\mathrm{rev}_1 L$ with eigenvalue $\lambda = 0$ if and only if x is an eigenvector of $\mathrm{rev}_k P$ at $\lambda = 0$. Equivalently, $e_1 \otimes x$ is an eigenvector of L with eigenvalue ∞ if and only if x is an eigenvector of P at ∞.

(c) This is a special case of Theorem 3.1(a) in [16]. Note that in this case the matrix $G(\lambda)$ in [16, Thm. 3.1] is $G(\lambda) = v \otimes I_n$.

□

4. Newton-Fiedler pencils

In the last 14 years Fiedler pencils for scalar and matrix polynomials expressed in the standard basis have been extensively studied [3, 4, 8, 9, 10, 11, 13, 19, 31]; this includes work [10, 11, 19] concerned especially with the numerical properties of Fiedler pencils, and with algorithms for solving polynomial eigenvalue problems based on these pencils. In this section we extend the notion of Fiedler pencils and show how to adapt them to polynomials expressed in a Newton basis. In particular, if P is an $n \times n$ matrix polynomial of grade k expressed in Newton basis

$$P(\lambda) = A_k n_k(\lambda) + \cdots + A_1 n_1(\lambda) + A_0 n_0(\lambda)\,, \tag{4.1}$$

we show how to construct Fiedler-like pencils for P using the matrix coefficients A_i *directly* from (4.1).

Definition 4.1. (Newton-Fiedler pencils)
Let $P(\lambda)$ be the matrix polynomial in (4.1), expressed in the Newton basis $\{n_i(\lambda)\}_{i=0}^{k}$ associated with an ordered node list $\mathcal{A}$ as in (3.1). Also let M_i for $i = 0, \dots, k$ be the matrices defined in (2.7) and (2.8). Then given any bijection $\sigma : \{0, 1, \dots, k-1\} \to \{1, \dots, k\}$, the Newton-Fiedler pencil for $P(\lambda)$ *associated with σ and the Newton basis $\{n_i(\lambda)\}_{i=0}^{k}$ is the $nk \times nk$ matrix pencil*

$$F_\sigma(\lambda) := \Gamma_k(\lambda) \cdot M_k - M_{\sigma^{-1}(1)} \cdot M_{\sigma^{-1}(2)} \cdot \ldots \cdot M_{\sigma^{-1}(k)}\ , \tag{4.2}$$

where $\Gamma_k(\lambda)$ is the $nk \times nk$ diagonal matrix defined in (3.3). As in (2.10) of Definition 2.18*, we also define*

$$M_\sigma := M_{\sigma^{-1}(1)} \cdot M_{\sigma^{-1}(2)} \cdot \ldots \cdot M_{\sigma^{-1}(k)}\,, \tag{4.3}$$

so that $F_\sigma(\lambda) = \Gamma_k(\lambda) \cdot M_k - M_\sigma$.

Remark 4.2. Comparing Definitions 2.18 and 4.1 shows that Fiedler and Newton-Fiedler pencils are closely related; in particular, the term $\lambda \cdot M_k$ in Fiedler pencils is just replaced by $\Gamma_k(\lambda) \cdot M_k$ in Newton-Fiedler pencils. Also note that M_k and Γ_k are block-diagonal matrices whose blocks commute; thus M_k and Γ_k *commute*, i.e., $\Gamma_k(\lambda) \cdot M_k = M_k \cdot \Gamma_k(\lambda)$.

In Section 4.2 we show that Definition 4.1 provides a genuine extension of Fiedler pencils to polynomials expressed in a Newton basis, both in the sense that ordinary Fiedler pencils are recovered as a special case of Newton-Fiedler pencils when the node list $\mathcal{A}$ consists of all zeros, as well as because the most important properties of Fiedler pencils are retained by the new Newton-Fiedler pencils. In particular, we will show that for *any* square matrix polynomial expressed in a Newton basis, *all* if its associated Newton-Fiedler pencils are always strong linearizations.

Before we can prove anything we first establish some necessary background. One of the most important observations concerns the commutativity of many pairs of Fiedler factors, i.e.,

$$M_i M_j = M_j M_i \quad \text{for } \ |i-j| \neq 1\,. \tag{4.4}$$

Relation (4.4) implies that Newton-Fiedler pencils associated with different bijections σ may sometimes be identical. This feature was also observed in [8], and led the authors to introduce the following concepts, describing the relative positions of *non*-commuting pairs of Fiedler factors in the product M_σ.

Definition 4.3. [8, Defn. 3.3] *Let $\sigma : \{0, 1, \dots, k-1\} \to \{1, \dots, k\}$ be a bijection. For $i = 0, \dots, k-2$, we say that σ has a* consecution at i *if $\sigma(i) < \sigma(i+1)$, and that σ has an* inversion at i *if $\sigma(i) > \sigma(i+1)$.*

Remark 4.4. Note that σ has a consecution at i if and only if M_i is to the left of M_{i+1} in the product M_σ in (4.3), while σ has an inversion at i if and only if M_i is to the right of M_{i+1} in M_σ.

4.1. Auxiliary matrices and equivalences

This section introduces the infrastructure needed to show that Newton-Fiedler pencils are strong linearizations for matrix polynomials expressed in a Newton basis. As expected, the γ_i's and Γ_i's associated with the given ordered node list $\mathcal{A} = (\alpha_1, \dots, \alpha_k)$ play an important role, so for the convenience of the reader we recall their definitions:

$$\begin{aligned} &\gamma_i(\lambda) := (\lambda - \alpha_i)\,, \quad \widetilde{\gamma}_i(\lambda) := \mathrm{rev}_1 \gamma_i(\lambda) = (1 - \alpha_i \lambda)\,, \quad \text{and} \\ &\Gamma_i(\lambda) := \mathrm{diag}\left[\gamma_i(\lambda), \gamma_{i-1}(\lambda), \dots, \gamma_1(\lambda)\right] \otimes I_n\,, \quad i = 1, \dots, k\,. \end{aligned} \tag{4.5}$$

We also make use of the reversal and inverse reversal of the Γ_k matrix in (4.5), which we abbreviate with the notation

$$\widetilde{\Gamma}_k(\lambda) := \mathrm{rev}_1\Gamma_k(\lambda) \qquad \text{and} \qquad \widetilde{\Gamma}_k^{-1} := \left(\widetilde{\Gamma}_k(\lambda)\right)^{-1}. \tag{4.6}$$

Most of the rest of Section 4.1 is a mild generalization of techniques presented in [8], and therefore proofs are omitted.

Definition 4.5. *Let $P(\lambda) = \sum_{i=0}^{k} A_i n_i(\lambda)$ be a matrix polynomial of grade k. For $d = 0, \dots, k$ define the matrix polynomials $P_d(\lambda)$ recursively:*

$$\begin{array}{rcll} P_0(\lambda) & = & A_k\,, & \\ P_{d+1}(\lambda) & = & \gamma_{k-d}P_d(\lambda) + A_{k-(d+1)} & \text{for } 0 \le d \le k-1\,, \\ P_k(\lambda) & = & P(\lambda)\,. & \end{array} \tag{4.7}$$

Note that the polynomials $P_d(\lambda)$ in Definition 4.5 are analogs of the Horner polynomials associated with a polynomial expressed in the standard basis.

We now introduce some auxiliary matrices that appear repeatedly throughout the rest of Section 4.

Definition 4.6. *For an $n \times n$ matrix polynomial $P(\lambda) = \sum_{i=0}^{k} A_i n_i(\lambda)$, let $P_i(\lambda)$ be the matrix polynomials as in* Definition 4.5*. For $i = 1, \dots, k-1$, define the following $nk \times nk$ matrix polynomials:*

$$\begin{array}{rcl} Q_i(\lambda) & := & \mathrm{diag}\left(I_{(i-1)n}, \begin{bmatrix} I_n & \gamma_{k-i}I_n \\ 0_n & I_n \end{bmatrix}, I_{(k-i-1)n}\right), \\ R_i(\lambda) & := & \mathrm{diag}\left(I_{(i-1)n}, \begin{bmatrix} 0_n & I_n \\ I_n & P_i(\lambda) \end{bmatrix}, I_{(k-i-1)n}\right) \;=\; R_i^{\mathcal{B}}(\lambda)\,, \\ T_i(\lambda) & := & \mathrm{diag}\left(0_{(i-1)n}, \begin{bmatrix} 0_n & \gamma_{k-(i-1)}P_{i-1} \\ \gamma_{k-i}I_n & \gamma_{k-i}\gamma_{k-(i-1)}P_{i-1} \end{bmatrix}, 0_{(k-i-1)n}\right), \\ D_i(\lambda) & := & \mathrm{diag}\left(0_{(i-1)n}, \gamma_{k-(i-1)}P_{i-1}, \Gamma_{k-i}\right), \\ \text{and} \quad D_k(\lambda) & := & \mathrm{diag}\left[0_{(k-1)n}, \gamma_1 P_{k-1}(\lambda)\right]. \end{array}$$

For the sake of brevity, we will sometimes omit the dependence on λ in these matrices and write just Q_i, R_i, T_i or D_i. Note that $D_1(\lambda) = \Gamma_k(\lambda) \cdot M_k$, and that $Q_i(\lambda)$ and $R_i(\lambda)$ are unimodular matrix polynomials for all $i = 1, \dots, k-1$.

Note that the matrices in Definition 4.6 were chosen just so that the results and proofs for Lemma 4.7 are formally identical to [8, Lemma 4.3].

Lemma 4.7. *Let Q_i, R_i, T_i and D_i be the matrices introduced in* Definition 4.6*, and M_j the matrices in* (2.7) *and* (2.8)*. Then the following relations hold for $i = 1, \dots, k-1$.*

(a) $Q_i^{\mathcal{B}}\big(D_i\big)R_i \;=\; D_{i+1} + T_i\,,$ *and* $Q_i^{\mathcal{B}}\big(M_{k-(i+1)}M_{k-i}\big)R_i \;=\; M_{k-(i+1)} + T_i.$

(b) $R_i^{\mathcal{B}}\big(D_i\big)Q_i \;=\; D_{i+1} + T_i^{\mathcal{B}},$ *and* $R_i^{\mathcal{B}}\big(M_{k-i}M_{k-(i+1)}\big)Q_i \;=\; M_{k-(i+1)} + T_i^{\mathcal{B}}.$

(c) $T_iM_j \;=\; M_jT_i \;=\; T_i$ *and* $T_i^{\mathcal{B}}M_j \;=\; M_jT_i^{\mathcal{B}} \;=\; T_i^{\mathcal{B}}$ *for all* $j \le k-i-2.$

The next definition introduces the polynomials which will form the intermediate steps in the unimodular transformation of a Newton-Fiedler pencil $F_\sigma(\lambda)$ into $\mathrm{diag}\big[P(\lambda),\, I_{(k-1)n}\big]$.

Definition 4.8. *Let $P(\lambda) = \sum_{i=0}^{k} A_i n_i(\lambda)$, and let $F_\sigma(\lambda) = \Gamma_k M_k - M_\sigma$ be the Newton-Fiedler pencil for $P(\lambda)$ associated with the bijection σ. For $i = 1, \dots, k$ define*

$$M_\sigma^{(i)} := \prod_{\sigma^{-1}(j)\,\le\, k-i} M_{\sigma^{-1}(j)}\,,$$

where the factors $M_{\sigma^{-1}(j)}$ are in the same relative order *as they are in M_σ. Equivalently, $M_\sigma^{(i)}$ is obtained from M_σ by deleting all factors M_ℓ with index $\ell > k-i$. Note that $M_\sigma^{(1)} = M_\sigma$ and $M_\sigma^{(k)} = M_0$. Also for $i = 1,\ldots,k$ define the $nk \times nk$ polynomials*

$$F_\sigma^{(i)}(\lambda) \;:=\; D_i(\lambda) \;-\; M_\sigma^{(i)}\,.$$

Observe that $F_\sigma^{(1)}(\lambda) = F_\sigma(\lambda)$ and $F_\sigma^{(k)}(\lambda) = \operatorname{diag}\big[-I_{(k-1)n},\, P(\lambda)\big]$.

The final technical lemma shows explicitly how to transform each $F_\sigma^{(i)}(\lambda)$ into $F_\sigma^{(i+1)}(\lambda)$ by unimodular transformations.

Lemma 4.9. *For each $i = 1,\ldots,k-1$, $F_\sigma^{(i)}(\lambda)$ is unimodularly equivalent to $F_\sigma^{(i+1)}(\lambda)$. Specifically, if Q_i and R_i are the unimodular matrices introduced in* Definition 4.6*, then:*

(a) *if σ has a consecution at $k-i-1$, then $F_\sigma^{(i+1)}(\lambda) \;=\; Q_i^{\mathcal{B}}\, F_\sigma^{(i)}(\lambda)\, R_i\,,$*

(b) *if σ has an inversion at $k-i-1$, then $F_\sigma^{(i+1)}(\lambda) \;=\; R_i^{\mathcal{B}}\, F_\sigma^{(i)}(\lambda)\, Q_i\,.$*

Remark 4.10. We emphasize that given the modified Definitions 4.5, 4.6, and 4.8, the proofs of Lemmas 4.7 and 4.9 are formally the same as the corresponding proofs for matrix polynomials in standard basis [8]. This gives another sense in which Newton-Fiedler pencils defined as in (4.2) can be viewed as a generalization of Fiedler pencils for polynomials expressed in the standard basis. In fact, if all nodes in $\mathcal{A}$ are set to zero then the associated Newton basis degenerates to the standard basis, and as expected, Newton-Fiedler pencils degenerate to Fiedler pencils.

4.2. Unimodular reduction of Newton-Fiedler pencils

In this section we state and prove our main result about Newton-Fiedler pencils. We start with the following lemma, which shows that certain analogs of the Q_i and R_i matrices in Definition 4.6 are q-prime.

Lemma 4.11. *Let $q(\lambda)$ be an arbitrary non-constant $\mathbb{F}$-irreducible scalar polynomial, and for $\ell = 1,\ldots,k$ let B_ℓ be in $\mathbb{F}^{n\times n}$. Further, let $\frac{r_\ell(\lambda)}{s_\ell(\lambda)}$ be elements in $\mathbb{F}(\lambda)$ in reduced form such that each denominator $s_\ell(\lambda)$ is coprime to $q(\lambda)$, and define the rational matrix*

$$B(\lambda) \;:=\; \sum_{\ell=1}^{k} \frac{r_\ell(\lambda)}{s_\ell(\lambda)} B_\ell\,.$$

(a) *If $B_{ij}(\lambda) = \frac{t_{ij}(\lambda)}{u_{ij}(\lambda)}$ is any entry of $B(\lambda)$ in* reduced form, *then $u_{ij}(\lambda)$ is coprime to $q(\lambda)$.*

(b) *If $Q_i(\lambda) = \operatorname{diag}\left(I_{(i-1)n}, \begin{bmatrix} I_n & \frac{r_\ell(\lambda)}{s_\ell(\lambda)} I_n \\ 0_n & I_n \end{bmatrix}, I_{(k-i-1)n} \right)$ for any $1 \le i \le k-1$, then $Q_i(\lambda)$ is a q-prime rational matrix.*

(c) *If $R_i(\lambda) = \operatorname{diag}\left(I_{(i-1)n}, \begin{bmatrix} 0_n & I_n \\ I_n & B(\lambda) \end{bmatrix}, I_{(k-i-1)n} \right)$ for any $1 \le i \le k-1$, then $R_i(\lambda)$ is a q-prime rational matrix.*

Proof. For part (a), observe that any entry $B_{ij}(\lambda) = \frac{t_{ij}(\lambda)}{u_{ij}(\lambda)}$ in reduced form is just an $\mathbb{F}$-linear combination of rational expressions with denominators $s_\ell(\lambda)$, each coprime to $q(\lambda)$. Since $\frac{t_{ij}(\lambda)}{u_{ij}(\lambda)}$ is in reduced form, $u_{ij}(\lambda)$ must divide the product of all these denominators $s_\ell(\lambda)$. But this product is coprime to $q(\lambda)$, hence $u_{ij}(\lambda)$ is also coprime to $q(\lambda)$.

To see why the matrices Q_i and R_i are q-prime, first observe that $\det Q_i = \det R_i = \pm 1$ for all $1 \le i \le k-1$, so conditions (a) and (c) from Definition 2.9 are satisfied. It only remains to see that no entry of Q_i or R_i has $q(\lambda)$ as a factor of its denominator. This is clearly the case for Q_i, since any $s_\ell(\lambda)$ is coprime to $q(\lambda)$ by assumption. In R_i the only non-constant denominators come from the block $B(\lambda)$, which are coprime to $q(\lambda)$ by part (a). Hence each Q_i and R_i is a q-prime rational matrix. □

Theorem 4.12. *Let $P(\lambda)$ be an $n\times n$ matrix polynomial (regular or singular) expressed in a Newton basis as in (4.1). Then any Newton-Fiedler pencil $F_\sigma(\lambda)$ for $P(\lambda)$ is a strong linearization for $P(\lambda)$.*

Proof. In order to show that $F_\sigma(\lambda)$ is a linearization for $P(\lambda)$, we exhibit a unimodular equivalence between $F_\sigma(\lambda)$ and $\operatorname{diag}\big[P(\lambda), I_{(k-1)n}\big]$. Since $\operatorname{diag}\big[P(\lambda), I_{(k-1)n}\big]$ and $\operatorname{diag}\big[-I_{(k-1)n}, P(\lambda)\big]$ are strictly equivalent, it suffices to show that $F_\sigma(\lambda)$ is unimodularly equivalent to $\operatorname{diag}\big[-I_{(k-1)n}, P(\lambda)\big]$. But such an equivalence can be explicitly constructed from Lemma 4.9 as the composition of a sequence of $k-1$ unimodular equivalences

$$F_\sigma(\lambda)=F_\sigma^{(1)}(\lambda)\longrightarrow F_\sigma^{(2)}(\lambda)\longrightarrow\cdots\longrightarrow F_\sigma^{(k)}(\lambda)=\operatorname{diag}\big[-I_{(k-1)n}, P(\lambda)\big] \tag{4.8}$$

$$\text{where}\qquad F_\sigma^{(i+1)}(\lambda) \;=\; \begin{cases} Q_i^{\mathcal{B}}\,F_\sigma^{(i)}(\lambda)\,R_i & \text{if } \sigma \text{ has a consecution at } k-i-1\\ R_i^{\mathcal{B}}\,F_\sigma^{(i)}(\lambda)\,Q_i & \text{if } \sigma \text{ has an inversion at } k-i-1\,.\end{cases}$$

This shows that $F_\sigma(\lambda)$ is a linearization for P.

To see why the linearization $F_\sigma(\lambda)$ is strong takes considerably more work. The approach we take consists of finding λ-prime rational matrices E_λ and F_λ such that

$$E_\lambda(\lambda)\cdot\operatorname{rev}_1F_\sigma(\lambda)\cdot F_\lambda(\lambda)=\begin{bmatrix}\operatorname{rev}_kP(\lambda) & \\ & I_{(k-1)n}\end{bmatrix}, \tag{4.9}$$

and then invoking the second half of Theorem 2.12.

Note that given a matrix polynomial P as in (4.1), its reversal with respect to grade k is given by

$$\begin{aligned}\operatorname{rev}_kP(\lambda) \;=\; & A_k\big(\widetilde{\gamma}_k\widetilde{\gamma}_{k-1}\cdots\widetilde{\gamma}_1\big) \;+\; A_{k-1}\big(\lambda\cdot\widetilde{\gamma}_{k-1}\widetilde{\gamma}_{k-2}\cdots\widetilde{\gamma}_1\big)\\ & +\; A_{k-2}\big(\lambda^2\cdot\widetilde{\gamma}_{k-2}\widetilde{\gamma}_{k-3}\cdots\widetilde{\gamma}_1\big) \;+\;\cdots\;+\; A_1\big(\lambda^{k-1}\cdot\widetilde{\gamma}_1\big) \;+\; A_0\big(\lambda^k\big)\,.\end{aligned} \tag{4.10}$$

On the other hand, the reversal of the Newton-Fiedler pencil $F_\sigma(\lambda)$ given by (4.2) is

$$-\operatorname{rev}_1F_\sigma(\lambda) \;=\; \lambda\big(M_{\sigma^{-1}(1)}\cdots M_{\sigma^{-1}(s-1)}M_0M_{\sigma^{-1}(s+1)}\cdots M_{\sigma^{-1}(k)}\big)-\widetilde{\Gamma}_k(\lambda)\cdot M_k\,, \tag{4.11}$$

where $\sigma^{-1}(s)=0$. Further, the relation (4.4) implies that M_0 commutes with all Fiedler factors except possibly with M_1. Hence there exists a bijection τ such that

$$-\operatorname{rev}_1F_\sigma(\lambda) \;=\; \lambda M_\sigma-\widetilde{\Gamma}_k(\lambda)\cdot M_k \;=\; \lambda M_\tau-\widetilde{\Gamma}_k(\lambda)\cdot M_k\,, \tag{4.12}$$

and where M_0 is either at the far right or far left of M_τ, depending on its position relative to M_1 in M_σ. For now we assume that M_0 is to the far left of M_σ. Then pre-multiplying (4.12) with the λ-prime rational matrix $\widetilde{\Gamma}_k^{-1}(\lambda)$ gives

$$\begin{aligned}-\widetilde{\Gamma}_k^{-1}(\lambda)\cdot\operatorname{rev}_1F_\sigma(\lambda) \;&=\; \lambda\widetilde{\Gamma}_k^{-1}(\lambda)\cdot M_\tau-M_k\\ &=\; \lambda\widetilde{\Gamma}_k^{-1}(\lambda)\cdot M_0\big(M_{\tau^{-1}(2)}\cdots M_{\tau^{-1}(k)}\big)-M_k\,.\end{aligned}$$

From Remark 2.17 we know that $M_1,\ldots,M_{k-1}$ are always constant *invertible* matrices, so letting $\widehat{U}:=\big(M^{-1}_{\tau^{-1}(k)}\cdots M^{-1}_{\tau^{-1}(2)}\big)$ we have

$$-\widetilde{\Gamma}_k^{-1}(\lambda)\cdot\operatorname{rev}_1F_\sigma(\lambda)\cdot\widehat{U}=\lambda\widetilde{\Gamma}_k^{-1}M_0-M_k\big(M^{-1}_{\tau^{-1}(k)}\cdots M^{-1}_{\tau^{-1}(2)}\big)\,. \tag{4.13}$$

Let $B:=\begin{bmatrix} & & I_n\\ & ⋰ & \\ I_n & & \end{bmatrix}$ so that $B^2=I_{kn}$, and set $U:=\widehat{U}\cdot B$. Then pre- and post-multiplying (4.13) by B gives

$$-B\cdot\widetilde{\Gamma}_k^{-1}(\lambda)\cdot\operatorname{rev}_1F_\sigma(\lambda)\cdot U \;=\; \lambda(B\widetilde{\Gamma}_k^{-1}B)(BM_0B)-BM_k\big(M^{-1}_{\tau^{-1}(k)}\cdots M^{-1}_{\tau^{-1}(2)}\big)B\,. \tag{4.14}$$

Now define

$$\widehat{M}_0 := BM_0B = \begin{bmatrix} -A_0 & \\ & I_{(k-1)n} \end{bmatrix}, \qquad \widehat{M}_k := BM_kB = \begin{bmatrix} I_{(k-1)n} & \\ & A_k \end{bmatrix},$$

and

$$\widehat{M}_i := BM_i^{-1}B = \begin{bmatrix} I_{(i-1)n} & & & \\ & A_i & I_n & \\ & I_n & 0 & \\ & & & I_{(k-i-1)n} \end{bmatrix}, \quad \text{for } i = 1, \ldots, k-1.$$

In terms of these $\widehat{M}_i$'s, the rational matrix on the right-hand side of (4.14) can be written as

$$\Delta_k(\lambda) \cdot \widehat{M}_0 - \widehat{M}_k \widehat{M}_{\tau^{-1}(k)} \cdots \widehat{M}_{\tau^{-1}(2)} =: \mathcal{F}(\lambda) \tag{4.15}$$

where $\Delta_k(\lambda) := \lambda(B\widetilde{\Gamma}_k^{-1}B)$. Note that $\Delta_k(\lambda)$ is just a diagonal matrix of the form

$$\Delta_k(\lambda) := \Big(\delta_1(\lambda), \delta_2(\lambda), \ldots, \delta_k(\lambda)\Big) \otimes I_n \quad \text{where} \quad \delta_i(\lambda) := \frac{\lambda}{\widetilde{\gamma}_i(\lambda)} = \frac{\lambda}{1 - \alpha_i \lambda}. \tag{4.16}$$

Now here is the crucial point. If we think of $\delta_{k-(i-1)}$ as *symbolically* replacing γ_i in the diagonal matrix Γ_k of (4.2), then the rational matrix $\mathcal{F}(\lambda)$ in (4.15) looks *formally* like a Newton-Fiedler "pencil" for the *rational* matrix defined by

$$-\Big[\big(\delta_1 \cdots \delta_k\big)A_0 + \big(\delta_2 \cdots \delta_k\big)A_1 + \cdots + \big(\delta_k\big)A_{k-1} + A_k\Big] =: \widehat{\mathcal{P}}(\lambda). \tag{4.17}$$

Hence, by a construction analogous to that of the unimodular equivalences built out of the Q_i's and R_i's in (4.8) of the first paragraph of this proof, we can build *rational* matrices $\mathcal{V}(\lambda)$ and $\mathcal{W}(\lambda)$ with nonzero constant determinants such that

$$\mathcal{V}(\lambda) \cdot \mathcal{F}(\lambda) \cdot \mathcal{W}(\lambda) = \begin{bmatrix} \widehat{\mathcal{P}}(\lambda) & 0 \\ 0 & I_{(k-1)n} \end{bmatrix}. \tag{4.18}$$

These matrices $\mathcal{V}(\lambda)$ and $\mathcal{W}(\lambda)$ are just products of matrices Q_i and R_i (and their block transposes) in which each γ_i is *symbolically* replaced by $\delta_{k-(i-1)}$.

Next observe that $\widehat{\mathcal{P}}(\lambda)$ is closely related to $\mathrm{rev}_k P(\lambda)$ from (4.10); in particular we have

$$\widetilde{\gamma}_1(\lambda)\widetilde{\gamma}_2(\lambda) \cdots \widetilde{\gamma}_k(\lambda) \cdot \widehat{\mathcal{P}}(\lambda) \;=\; \mathrm{rev}_k P(\lambda). \tag{4.19}$$

Now the equation (4.18) can be rewritten as

$$\mathcal{V}(\lambda) \cdot \mathcal{F}(\lambda) \cdot \mathcal{W}(\lambda) = \begin{bmatrix} \frac{1}{\widetilde{\gamma}_1 \cdots \widetilde{\gamma}_k} I_n & 0 \\ 0 & I_{(k-1)n} \end{bmatrix} \cdot \begin{bmatrix} \mathrm{rev}_k P(\lambda) & 0 \\ 0 & I_{(k-1)n} \end{bmatrix},$$

or equivalently using (4.14) as

$$\underbrace{-\begin{bmatrix} \widetilde{\gamma}_1 \cdots \widetilde{\gamma}_k I_n & 0 \\ 0 & -I_{(k-1)n} \end{bmatrix} \cdot \mathcal{V}(\lambda) \cdot B \cdot \widetilde{\Gamma}_k^{-1}}_{=:\mathcal{X}(\lambda)} \cdot \Big(\mathrm{rev}_1 F_\sigma(\lambda)\Big) \cdot \underbrace{U \cdot \mathcal{W}(\lambda)}_{=:\mathcal{Y}(\lambda)}$$

$$= \begin{bmatrix} \mathrm{rev}_k P(\lambda) & 0 \\ 0 & I_{(k-1)n} \end{bmatrix}. \tag{4.20}$$

Several points are now important to emphasize. First, due to the special structure of the δ_j's (whose denominators are coprime to λ) and the results of Lemma 4.11(b) and (c), each of the modified Q_i and R_i matrices (with γ_i symbolically replaced by $\delta_{k-(i-1)}$) in $\mathcal{V}(\lambda)$ and $\mathcal{W}(\lambda)$ is a λ-prime rational matrix. Second, since $\mathcal{V}(\lambda)$ and $\mathcal{W}(\lambda)$ are each just products of λ-prime rational matrices (these modified Q_i's and

R_i's), it follows from Proposition 2.11 that $\mathcal{V}(\lambda)$ and $\mathcal{W}(\lambda)$ are λ-prime rational matrices, too. Third, from Definition 2.9 and Remark 2.10 it is easy to see that the matrices $\operatorname{diag}\big(\widetilde{\gamma}_1\cdots\widetilde{\gamma}_k I_n\,,\,-I_{(k-1)n}\big)$, B, $\widetilde{\Gamma}_k^{-1}$, and U are also λ-prime matrices. Thus the matrices $\mathcal{X}(\lambda)$ and $\mathcal{Y}(\lambda)$ in (4.20) are products of λ-prime matrices, and hence by Proposition 2.11 are λ-prime matrices, too. All together, then, the simplified version of the relation (4.20), i.e.,

$$\mathcal{X}(\lambda)\cdot \mathrm{rev}_1 F_\sigma(\lambda)\cdot \mathcal{Y}(\lambda) \;=\; \operatorname{diag}\big[\mathrm{rev}_k P(\lambda), I_{(k-1)n}\big],$$

together with Theorem 2.12 implies that the linearization F_σ is a strong linearization for P, as desired.

All that remains is to consider the case when M_0 appears on the far *right* of M_τ in (4.12). In this case we start by *post*-multiplying (4.12) by $\widetilde{\Gamma}_k^{-1}(\lambda)$ to obtain

$$\begin{aligned}
-\mathrm{rev}_1 F_\sigma(\lambda)\cdot\widetilde{\Gamma}_k^{-1}(\lambda) \;&=\; \lambda M_\tau\cdot\widetilde{\Gamma}_k^{-1}(\lambda) - M_k \\
&=\; \lambda\big(M_{\tau^{-1}(1)}\cdots M_{\tau^{-1}(k-1)}\big)M_0\cdot\widetilde{\Gamma}_k^{-1}(\lambda) - M_k\,.
\end{aligned}$$

Note that here we use the fact that $\Gamma_k M_k = M_k\Gamma_k$ (see Remark 4.2). Now one continues by *pre*-multiplying in the appropriate order by the inverses of $M_1, M_2, \ldots, M_{k-1}$ to show that $-\mathrm{rev}_1 F_\sigma(\lambda)\cdot\widetilde{\Gamma}_k^{-1}(\lambda)$ is strictly equivalent to

$$\lambda M_0\widetilde{\Gamma}_k^{-1} - \big(M^{-1}_{\tau^{-1}(k-1)}\cdots M^{-1}_{\tau^{-1}(1)}\big)M_k\,.$$

Since $M_0\widetilde{\Gamma}_k^{-1} = \widetilde{\Gamma}_k^{-1}M_0$, we have reached a stage that is essentially equivalent to (4.13). Thus the rest of the proof for this case can be completed in a manner completely analogous to the case already considered. $\square$

By comparing Theorems 4.12 and 2.19 one can see that Newton-Fiedler and Fiedler pencils have much in common. Indeed, Fiedler and Newton-Fiedler pencils share all of the properties outlined in Remark 2.20, a fact that plays an important role in Section 5.

Example 4.13. Let $\{n_i(\lambda)\}_{i=0}^5$ be the Newton polynomials associated with the (not necessarily distinct) nodes $(\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5)$. Further, let $P(\lambda)$ be an $n\times n$ matrix polynomial $P(\lambda) = \sum_{i=0}^5 A_i n_i(\lambda)$, and consider the bijection $\tau = (1,4,2,5,3)$. Then the Newton-Fiedler pencil for P associated with the bijection τ, $F_\tau(P)$, is given by

$$F_\tau(P) \;=\; \Gamma_5(\lambda)\cdot M_5 - M_0M_2M_4M_1M_3\,,$$

where the M_i matrices are defined as in (2.7) and (2.8). More concretely,

$$F_\tau(P) \;=\; \begin{bmatrix} \gamma_5 A_5 & & & & \\ & \gamma_4 I_n & & & \\ & & \gamma_3 I_n & & \\ & & & \gamma_2 I_n & \\ & & & & \gamma_1 I_n \end{bmatrix} - \begin{bmatrix} -A_4 & -A_3 & I_n & & \\ I_n & 0 & 0 & 0 & \\ 0 & -A_2 & 0 & -A_1 & I_n \\ & I_n & 0 & 0 & 0 \\ & & 0 & -A_0 & 0 \end{bmatrix}, \tag{4.21}$$

where $\gamma_i = \gamma_i(\lambda) := \lambda - \alpha_i$. Theorem 4.12 implies that $F_\tau(P)$ is a strong linearization for P.

5. Some Computational Aspects of Newton-Fiedler Pencils

In Sections 3 and 4 we have seen that for an $n\times n$ matrix polynomial P of grade k, *almost all* the pencils in $\mathcal{N}_1(P)$, and *all* the Newton-Fiedler pencils for P are strong linearizations. The common feature of all these pencils is the novel form in which they are represented: $X\cdot\Gamma_k + Y$, where X and Y are constant matrices. This representation was crucial when studying the properties of these pencils, but it is not so obvious how to do numerical computations with this form. In particular, algorithms used to solve a generalized eigenvalue problem typically treat the underlying matrix pencil $\lambda A + B$ as a pair of matrices (A,B). The challenge when computing with a pencil $L(\lambda)$ in the form $X\Gamma_k + Y$ is that strict equivalences performed on $L(\lambda)$ *cannot* be treated as strict equivalences acting on the matrix pair (X,Y); this is because strict equivalence transformations *do not commute* (in general) with the diagonal matrix Γ_k.

An alternative approach to computing with a pencil $L(\lambda) = X\cdot\Gamma_k + Y$ is to first rewrite it as $(\lambda-\beta)X + \widetilde{Y}$, where

$$\widetilde{Y} := Y - X\Big(\operatorname{diag}\left(\alpha_k - \beta, \ldots, \alpha_2 - \beta, \alpha_1 - \beta\right) \otimes I_n\Big), \tag{5.1}$$

and β is an *arbitrary* scalar. The pencil $(\lambda-\beta)X + \widetilde{Y}$ can now be treated as the matrix pair $(X, \widetilde{Y})$, no matter what the choice of $\beta \in \mathbb{F}$, analogous to the way a *standard* pencil $\lambda M + N$ is treated as the matrix pair (M, N). Having the choice of β available may potentially be useful in improving numerical accuracy when forming and then computing with the pair $(X, \widetilde{Y})$. It is also worth noting that the product of X and the diagonal matrix $\big(\alpha_k - \beta, \ldots, \alpha_2 - \beta, \alpha_1 - \beta\big) \otimes I_n$ is not numerically problematic, since it only involves scaling of columns.

Assume $L(\lambda)$ is a pencil of the form $(\lambda-\beta)X + \widetilde{Y}$ where $\widetilde{Y}$ is given by (5.1). Using the QZ-algorithm one can simultaneously reduce X and $\widetilde{Y}$ to upper triangular forms T_X and $T_{\widetilde{Y}}$; then the pencil $T(\lambda) = (\lambda-\beta)T_X + T_{\widetilde{Y}}$ has the same spectrum as $L(\lambda)$. But the eigenvalues of $T(\lambda)$ are now readily computed from the diagonal entries to be

$$\lambda_i = \frac{-(T_{\widetilde{Y}})_{ii}}{(T_X)_{ii}} + \beta\,; \qquad i = 1, \ldots, nk\,.$$

If $(T_X)_{ii} = 0$ and $(T_{\widetilde{Y}})_{ii} \neq 0$ for some i, then of course $\lambda_i = \infty$.

Now recall that our original goal was to find strong linearizations for a matrix polynomial in Newton basis $\sum_{i=0}^{k} A_i n_i(\lambda)$ by working (as much as possible) with the coefficients A_i directly, while avoiding any additions, multiplications, or inverses of those coefficients. The underlying motivation is to avoid possible numerical errors that could arise by performing such preliminary operations on the coefficients A_i. It is important to emphasize that there are *two* sources from which such matrix operations arise:

(a) the initial construction of the X and Y coefficient matrices used to generate strong linearizations of the form $X\Gamma_k + Y$,

(b) the rewriting of the pencil $X\Gamma_k + Y$ into the more “computationally friendly” pencil of the form $(\lambda - \beta)X + \widetilde{Y}$.

It is well known that for many pencils $\lambda X + Y$ in $\mathbb{L}_1(\widehat{P})$, the matrices X and Y already have block entries that involve matrix additions [22]. Consequently, the pencils $X\Gamma_k + Y \in \mathcal{N}_1(P)$ will also contain the same such blocks. In addition, when rewriting the pencil $X\Gamma_k + Y \in \mathcal{N}_1(P)$ into the form $(\lambda-\beta)X + \widetilde{Y}$, the matrix $\widetilde{Y}$ from (5.1) is likely to contain even more block additions than the original Y.

On the other hand, for a Newton-Fiedler pencil $C\cdot\Gamma_k + D$, *none* of the block entries of the coefficients C and D ever involve any matrix operations. In fact, the algorithmic construction of Fiedler pencils for matrix polynomials in standard basis given in [9, Theorem 3.4] also implies that only *one* block addition is needed when converting $C\cdot\Gamma_k + D$ into the form $(\lambda-\beta)C + \widetilde{D}$. More specifically, the $(1,1)$ block entry of $\widetilde{D}$ is $(\beta - \alpha_k)A_k + A_{k-1}$, whereas all other block entries require no further operations on the original matrix coefficients A_i other than scaling. These features suggest that Newton-Fiedler pencils may be more desirable than pencils in $\mathcal{N}_1(P)$ as strong linearizations for matrix polynomials in a Newton basis. The rest of this section studies some other features of Newton-Fiedler pencils that could potentially play a significant role in numerical practice.

5.1. Updating Newton-Fiedler pencils

Matrix polynomials expressed in Newton basis often arise as interpolating polynomials for more general nonlinear matrix functions. As such, it is desirable that those polynomials can be easily updated in case additional nodes are considered. More specifically, we let $(\alpha_1, \ldots, \alpha_k)$ be an arbitrary but fixed list of nodes and consider an $n \times n$ matrix polynomial $P(\lambda) = \sum_{i=0}^{k} A_i n_i(\lambda)$, where the $n_i(\lambda)$'s are the (scalar) Newton polynomials associated with the nodes α_i. Now assume that a new node α_{k+1} is appended to the list, and consider an $n \times n$ matrix polynomial

$$\widetilde{P}(\lambda) := A_{k+1}n_{k+1}(\lambda) + P(\lambda) = \sum_{i=0}^{k+1} A_i n_i(\lambda)$$

expressed in the Newton basis associated to the extended node list $(\alpha_1, \ldots, \alpha_k, \alpha_{k+1})$. The question we address is the following:

> *Given a Newton-Fiedler pencil $F_\tau(\lambda)$ for $P(\lambda)$ associated with a bijection*
> $$\tau : \{0, \ldots, k-1\} \to \{1, \ldots, k\} \,,$$
> *is it possible to use knowledge about $F_\tau(\lambda)$ to produce a strong linearization for $\widetilde{P}(\lambda)$ with little to no additional work?*

The answer to this question is yes. Indeed, the strong linearization for $\widetilde{P}$ can even be chosen to be a Newton-Fiedler pencil for $\widetilde{P}$ that is associated with a bijection $\widetilde{\tau} : \{0, 1, \ldots, k\} \to \{1, 2, \ldots, k+1\}$ that is simply related to the given bijection τ, in the sense that the relative order of the Fiedler factors $M_0, M_1, \ldots, M_{k-1}$ is the same in M_τ and $M_{\widetilde{\tau}}$. The only additional thing that must be decided about $\widetilde{\tau}$ is whether there is to be an inversion or consecution at $k-1$, i.e., whether $M_k(\widetilde{P})$ is to be to the *left* or to the *right* of $M_{k-1}(\widetilde{P})$ in $M_{\widetilde{\tau}}$ (see Definition 4.3 and Remark 4.4). The commutativity relation of Fiedler factors (4.4) now implies that only two cases need to be considered, i.e., $M_k(\widetilde{P})$ is either to the *far* right (Option A) or *far* left (Option B) in the product $M_{\widetilde{\tau}}(\widetilde{P})$.

Note that in the next theorem we use MATLAB notation for submatrices on *block* indices; that is, if A is a matrix partitioned into blocks, then $A(i:j;:)$ indicates the submatrix of A consisting of block rows i through j and $A(:;k:\ell)$ indicates the submatrix of A consisting of block columns k through ℓ. Theorem 5.1, justified by [9, Theorem 3.4], now describes an appropriate procedure for building an "updated" Newton-Fiedler pencil $\widetilde{F}_{\widetilde{\tau}}(\lambda)$ for $\widetilde{P}$.

Theorem 5.1. (Updating Newton-Fiedler pencils)
Let $(\alpha_1, \ldots, \alpha_k, \alpha_{k+1})$ be an ordered list of possibly non-distinct nodes, and let $n_i(\lambda)$ be the (generalized) scalar Newton polynomials associated with those nodes. Further consider the $n \times n$ matrix polynomials defined by

$$P(\lambda) := \sum_{i=0}^{k} A_i n_i(\lambda) \quad \text{and} \quad \widetilde{P}(\lambda) := \sum_{i=0}^{k+1} A_i n_i(\lambda) = A_{k+1}n_{k+1}(\lambda) + P(\lambda)\,, \tag{5.2}$$

where $A_i \in \mathbb{F}^{n\times n}$. Let $\tau : \{0, \ldots, k-1\} \to \{1, \ldots, k\}$ be an arbitrary bijection, so that $F_\tau(\lambda) = \Gamma_k \cdot M_k(P) - M_\tau(P)$ is the Newton-Fiedler pencil for $P(\lambda)$ associated with τ. Further let $\widetilde{\tau} : \{0, \ldots, k\} \to \{1, \ldots, k+1\}$ be another bijection.

$\boxed{\text{Option A}}$: *Let $\widetilde{\tau}(k) = k+1$ and $\widetilde{\tau}(j) = \tau(j)$ for all $j = 0, \ldots, k-1$. Then the Newton-Fiedler pencil $F^A_{\widetilde{\tau}}$ for $\widetilde{P}(\lambda)$ is given by $F^A_{\widetilde{\tau}} = \Gamma_{k+1} \cdot M_{k+1}(\widetilde{P}) - M^A_{\widetilde{\tau}}(\widetilde{P})$, where*

$$M^A_{\widetilde{\tau}}(\widetilde{P}) := \begin{bmatrix} -A_k & I_n & 0 \\ M_\tau(P)(:,1) & 0 & M_\tau(P)(:,2:k) \end{bmatrix}.$$

$\boxed{\text{Option B}}$: *Let $\widetilde{\tau}(k) = 1$ and $\widetilde{\tau}(j) = \tau(j)+1$ for all $j = 0, \ldots, k-1$. Then the Newton-Fiedler pencil $F^B_{\widetilde{\tau}}$ for $\widetilde{P}(\lambda)$ is given by $F^B_{\widetilde{\tau}} = \Gamma_{k+1} \cdot M_{k+1}(\widetilde{P}) - M^B_{\widetilde{\tau}}(\widetilde{P})$, where*

$$M^B_{\widetilde{\tau}}(\widetilde{P}) := \begin{bmatrix} -A_k & M_\tau(P)(1,:) \\ I_n & 0 \\ 0 & M_\tau(P)(2:k,:) \end{bmatrix}.$$

Note that the condition $\widetilde{\tau}(k) = k+1$ in Option A simply says that the factor M_k is on the far right in the product $M_{\widetilde{\tau}}(\widetilde{P})$. Analogously, the condition $\widetilde{\tau}(k) = 1$ in Option B implies that M_k is to appear on the far left in the product $M_{\widetilde{\tau}}(\widetilde{P})$.

Example 5.2. Let $(\alpha_1, \ldots, \alpha_6)$ be an ordered list of nodes, and consider an $n \times n$ matrix polynomial $\widetilde{P}(\lambda) := A_6 n_6(\lambda) + P(\lambda) = \sum_{i=0}^{6} A_i n_i(\lambda)$. Now, for the convenience of the reader, we recall that the strong linearization for P from Example 4.13, $F_\tau(P)$, is given by

$$\begin{aligned} F_\tau(P) &= \Gamma_5 M_5(P) - M_0(P)M_2(P)M_4(P)M_1(P)M_3(P) \\ &= \begin{bmatrix} \gamma_5 A_5 & & & & \\ & \gamma_4 I_n & & & \\ & & \gamma_3 I_n & & \\ & & & \gamma_2 I_n & \\ & & & & \gamma_1 I_n \end{bmatrix} - \begin{bmatrix} -A_4 & -A_3 & I_n & 0 & 0 \\ I_n & 0 & 0 & 0 & 0 \\ 0 & -A_2 & 0 & -A_1 & I_n \\ 0 & I_n & 0 & 0 & 0 \\ 0 & 0 & 0 & -A_0 & 0 \end{bmatrix}. \end{aligned} \tag{5.3}$$

Then the Newton-Fiedler pencil F_τ for P can be used to construct the following two Newton-Fiedler pencils for $\widetilde{P}$:

$$\begin{aligned} F^A_{\widetilde{\tau}}(\widetilde{P}) &:= \Gamma_6 \cdot M_6(\widetilde{P}) - M_0(\widetilde{P}) \cdot M_2(\widetilde{P}) \cdot M_4(\widetilde{P}) \cdot M_1(\widetilde{P}) \cdot M_3(\widetilde{P}) \cdot M_5(\widetilde{P}), \\ F^B_{\widetilde{\tau}}(\widetilde{P}) &:= \Gamma_6 \cdot M_6(\widetilde{P}) - M_5(\widetilde{P}) \cdot M_0(\widetilde{P}) \cdot M_2(\widetilde{P}) \cdot M_4(\widetilde{P}) \cdot M_1(\widetilde{P}) \cdot M_3(\widetilde{P}). \end{aligned}$$

Note that $M_j(\widetilde{P}) = \mathrm{diag}(I_n, M_j(P))$ for $j = 0, \ldots, 4$.

Now given that $M_\tau(P)$ is known, we can easily construct $F^A_{\widetilde{\tau}}(\widetilde{P})$ and $F^B_{\widetilde{\tau}}(\widetilde{P})$ using Theorem 5.1. For example, $F^B_{\widetilde{\tau}}(\widetilde{P})$ is given by

$$F^B_{\widetilde{\tau}}(\widetilde{P}) = \begin{bmatrix} \gamma_6 A_6 & & & & & \\ & \gamma_5 I_n & & & & \\ & & \gamma_4 I_n & & & \\ & & & \gamma_3 I_n & & \\ & & & & \gamma_2 I_n & \\ & & & & & \gamma_1 I_n \end{bmatrix} - \begin{bmatrix} -A_5 & -A_4 & -A_3 & I_n & 0 & 0 \\ I_n & 0 & 0 & 0 & 0 & 0 \\ 0 & I_n & 0 & 0 & 0 & 0 \\ 0 & 0 & -A_2 & 0 & -A_1 & I_n \\ 0 & 0 & I_n & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -A_0 & 0 \end{bmatrix}.$$

Notice that the shaded blocks in $F^B_{\widetilde{\tau}}(\widetilde{P})$ are exactly the blocks needed to *update* $F_\tau(P)$ from (5.3) and turn it into the Newton-Fiedler pencil $F^B_{\widetilde{\tau}}(\widetilde{P})$. The unshaded blocks in $F^B_{\widetilde{\tau}}(\widetilde{P})$ are the blocks from $F_\tau(P)$ that are reused.

5.2. Hermite-interpolating matrix polynomials and symmetric linearizations

In this section we consider matrix polynomials that arise in Hermite interpolation of more general matrix functions. Such matrix polynomials can be expressed in a generalized Newton basis corresponding to a special type of ordered node list $\mathcal{A}$ containing *non-distinct* elements. Recall that none of the results from Section 4 depend on whether the elements in $\mathcal{A}$ are distinct or not, so those results can be immediately applied to these matrix polynomials. In this section we also show how it is sometimes possible to exploit additional structure in a generalized Newton basis.

Let $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ be ℓ distinct nodes, and with each node associate a *multiplicity* $m_i > 0$. We define the quantities $s_0 = 0$ and $s_j := \sum_{i=1}^{j} m_i$ for $j = 1, \ldots, \ell$, and the set of degree one scalar polynomials $\{\gamma_i(\lambda)\}_{i=0}^{s_\ell}$ where

$$\gamma_0(\lambda) := 1 \quad \text{and} \quad \gamma_{s_{(i-1)}+1}(\lambda) = \ldots = \gamma_{s_i}(\lambda) := \lambda - \alpha_i, \quad i = 1, \ldots, \ell. \tag{5.4}$$

Further, define the scalar polynomials $h_j(\lambda)$ by

$$h_0(\lambda) \equiv 1 \quad \text{and} \quad h_j(\lambda) := h_{j-1}(\lambda) \cdot \gamma_j(\lambda)\,, \quad j = 1, \ldots, s_\ell\,, \tag{5.5}$$

and note that these $h_j(\lambda)$'s are in fact the generalized Newton polynomials from (3.1) corresponding to the *ordered* node list

$$\mathcal{A} = (\,\underbrace{\alpha_1, \ldots, \alpha_1}_{m_1}, \underbrace{\alpha_2, \ldots, \alpha_2}_{m_2}, \ldots, \underbrace{\alpha_\ell, \ldots, \alpha_\ell}_{m_\ell}\,)\,.$$

Nevertheless, we use the notation $h_i(\lambda)$ here instead of $n_i(\lambda)$ to emphasize that some elements in $\mathcal{A}$ are *non-distinct.*

Now consider an $n \times n$ matrix polynomial $H(\lambda)$ of the form

$$H(\lambda) = \sum_{i=0}^{s_\ell} B_i h_i(\lambda)\,, \tag{5.6}$$

and observe that polynomials arising as a consequence of Hermite interpolation of matrix functions are exactly of the form (5.6).

Corollary 5.3. *Any Newton-Fiedler pencil for a polynomial $H(\lambda)$ as in (5.6), with $\gamma_i(\lambda)$ defined as in (5.4), is a strong linearization for $H(\lambda)$.*

Proof. This follows directly from Theorem 4.12. □

It is worth noting that the polynomial basis used to express $H(\lambda)$ in (5.6) may have additional structure that a general Newton basis doesn't have, which can potentially be exploited in order to find *structured* linearizations for $H(\lambda)$; we illustrate this point with an example.

Example 5.4. Let α_1, α_2, and α_3 be three distinct nodes with multiplicities $m_1 = 2$, $m_2 = 4$, and $m_3 = 1$, so that the ordered node list is $\mathcal{A} = (\alpha_1, \alpha_1, \alpha_2, \alpha_2, \alpha_2, \alpha_2, \alpha_3)$. Further, let $H(\lambda) = \sum_{i=0}^{7} A_i h_i(\lambda)$ be an $n \times n$ matrix polynomial where the $h_i(\lambda)$ are defined as in (5.5). Also suppose that $A_i = A_i^T$, so that $H(\lambda)$ is a *symmetric* matrix polynomial.

Now consider the bijection $\sigma : \{0, 1, \ldots, 6\} \to \{1, 2, \ldots, 7\}$ given in the array notation $\sigma = (1, 5, 2, 6, 3, 7, 4)$. The corresponding Newton-Fiedler pencil is given by

$$F_\sigma(H) = \Gamma_7(\lambda) \cdot M_7 - M_0 M_2 M_4 M_6 M_1 M_3 M_5\,,$$

and by Corollary 5.3 is a strong linearization for $H(\lambda)$.

But any pencil that is strictly equivalent to $F_\sigma(H)$ is also a strong linearization for $H(\lambda)$, in particular,

$$L(\lambda) := \Gamma_7(\lambda) \cdot M_7 \cdot M_5^{-1} \cdot M_3^{-1} \cdot M_1^{-1} - M_0 \cdot M_2 \cdot M_4 \cdot M_6\,.$$

In expanded form the pencil $L(\lambda)$ is given by

$$L(\lambda) = \begin{bmatrix} \gamma_7 A_7 + A_6 & -I_n & & & & & \\ -I_n & 0 & \gamma_6 I_n & & & & \\ & \gamma_5 I_n & \gamma_5 A_5 + A_4 & -I_n & & & \\ & & -I_n & 0 & \gamma_4 I_n & & \\ & & & \gamma_3 I_n & \gamma_3 A_3 + A_2 & -I_n & \\ & & & & -I_n & 0 & \gamma_2 I_n \\ & & & & & \gamma_1 I_n & \gamma_1 A_1 + A_0 \end{bmatrix}, \tag{5.7}$$

which can be viewed as a Newton-adapted variation of a construction first given in [3]. The specific choice of multiplicities m_i means that $\gamma_1 = \gamma_2$ and $\gamma_3 = \gamma_4 = \gamma_5 = \gamma_6$, which implies that the pencil $L(\lambda)$ is symmetric. Thus we have started with a symmetric matrix polynomial expressed in a Hermite basis, and found a symmetric strong linearization for it. Note that if $\mathbb{F} = \mathbb{C}$, then a similar construction with all *real* nodes can produce a Hermitian strong linearization for a Hermitian matrix polynomial expressed in a Hermite basis.

Theorem 5.5. *Let $\alpha_1, \dots, \alpha_\ell$ be ℓ distinct nodes, and for each $i = 1, \dots, \ell$ consider a multiplicity $m_i > 0$. Assume that all m_i are* even *except for the last multiplicity m_ℓ, which is odd, and define $k := \sum_{i=1}^{\ell} m_i$. Further, consider linear polynomials $\gamma_i(\lambda)$ defined as in (5.4), and scalar polynomials $h_j(\lambda)$ defined as in (5.5). If $P(\lambda) = \sum_{i=0}^{k} A_i h_i(\lambda)$ is an $n \times n$ symmetric matrix polynomial, i.e, $A_i^T = A_i$, then there exists a strong linearization that is also symmetric and block tridiagonal.*

Proof. Start by observing that k is an odd positive integer. Then Corollary 5.3 implies that the Newton-Fiedler pencil $F_\sigma(\lambda) = \Gamma_k \cdot M_k - M_0 M_2 \cdots M_{k-1} M_1 M_3 \cdots M_{k-2}$ is a strong linearization for P. Thus the strictly equivalent pencil

$$L(\lambda) := \Gamma_k \cdot M_k M_{k-2}^{-1} \cdots M_3^{-1} M_1^{-1} - M_0 M_2 \cdots M_{k-1}$$

is also a strong linearization for P (see Remark 2.6). To see that $L(\lambda)$ is a symmetric pencil, note that the M_i's, M_j^{-1}'s, and Γ_k are each individually symmetric because the A_i's are assumed to be symmetric. Now observe that each of the terms $\Gamma_k \cdot M_k M_{k-2}^{-1} \cdots M_3^{-1} M_1^{-1}$ and $M_0 M_2 \cdots M_{k-1}$ is a product of *commuting* symmetric matrices. Indeed, the hypothesis on the multiplicities m_i is exactly what makes Γ_k commute with both M_k and M_j^{-1} for $j = 1, 3, 5, \dots, k-2$. Thus each term of $L(\lambda)$ is symmetric. The fact that $L(\lambda)$ is block tridiagonal is just a consequence of (4.4) and properties of block multiplication. □

Although Theorem 5.5 has been expressed in a purely existential form, the most important feature is the explicit construction of a symmetric linearization given in the proof. An alternative argument could have simply converted P into the standard basis and used the results from [3] or [23]. But any symmetric linearization produced by this approach would likely involve numerous matrix additions as a consequence of re-expressing P from the Newton basis to the standard basis. Instead, the pencil obtained in the proof above will, when re-expressed in the form $\lambda B + C$, only require at most $(k+1)/2$ block additions $-\alpha_j A_j + A_{j-1}$ at the specific diagonal-block entries $(k+1-j, k+1-j)$, for $j = 1, 3, 5, \dots, k$; this can be clearly seen in the example in (5.7).

6. Conclusion

We have shown how two novel ways of expressing pencils, i.e., $X \cdot \Gamma_k + Y$ and $\Gamma_k \cdot X + Y$, can be used to effectively study the properties of pencils in the generalized ansatz spaces of both regular *and* singular matrix polynomials expressed in a Newton basis. Eigenvector recovery formulas for linearizations in the generalized ansatz spaces have also been derived.

These same pencil representations also proved invaluable for generalizing Fiedler pencils to matrix polynomials in Newton bases. These Newton-Fiedler pencils now constitute a third successful adaptation of Fiedler pencils to matrix polynomials expressed in a non-standard basis; the first two were to polynomials in Bernstein bases [25], and to polynomials in Chebyshev bases [28]. We have shown that all Newton-Fiedler pencils are strong linearizations for square matrix polynomials over arbitrary fields. Also, since our construction of Newton-Fiedler pencils leveraged the results from [9], one could potentially adapt Algorithm 2 from [9, Theorem 3.6] to construct strong linearizations for *rectangular* matrix polynomials in a Newton basis.

From the practical standpoint, we discussed how Newton-Fiedler pencils can be algorithmically constructed, and exploited that fact when building linearizations for a matrix polynomial in a Newton basis that is being updated one node at a time. Finally, we illustrated how Newton-Fiedler pencils can be used to find a symmetric strong linearization for a special class of symmetric Hermite interpolating matrix polynomials. The numerical properties of Newton-Fiedler pencils and their applications are subjects for future investigation.

Acknowledgements

The work of both authors was supported by National Science Foundation grant DMS-1016224.

References

[1] A. Amiraslani, R. M. Corless, and P. Lancaster, *Linearization of matrix polynomials expressed in polynomial bases*, IMA J. Numer. Anal., 29 (2009), pp. 141–157.

[2] A. Amparan, S. Marcaida, and I. Zaballa *Finite and infinite structures of rational matrices: a local approach*, Electron. J. Linear Algebra, 30 (2015), pp. 196–226.

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

[4] M. I. Bueno, F. De Terán, and 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.

[5] R. M. Corless, *Generalized companion matrices in the Lagrange basis*, in Proceedings EACA, L. Gonzalez-Vega and T. Recio [eds.], 317–322, 2004.

[6] P. J. Davis, *Interpolation and Approximation*, Dover Publications, Inc., New York, 1975.

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

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

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

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

[11] F. De Terán, F. M. Dopico, and J. Pérez, *Backward stability of polynomial root-finding using Fiedler companion matrices*, IMA J. Numer. Anal., (2015), pp. 1–41, doi:10.1093/imanum/dru057.

[12] C. Effenberger and D. Kressner, *Chebyshev interpolation for nonlinear eigenvalue problems*, BIT Numer. Math., 52 (2012), pp. 933–951.

[13] M. Fiedler, *A note on companion matrices*, Linear Algebra Appl., 372 (2003), pp. 325–331.

[14] F. Gantmacher, *The Theory of Matrices*, AMS Chelsea, Providence, RI, 1998.

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

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

[17] S. Güttel, R. Van Beeumen, K. Meerbergen, and W. Michiels, *NLEIGS: A class of robust fully rational Krylov methods for nonlinear eigenvalue problems*, SIAM J. Sci. Comput., 36 (2014), pp. 2842–2864.

[18] N. J. Higham , D. S. Mackey, N. Mackey, and F. Tisseur, *Symmetric linearizations for matrix polynomials*, SIAM J. Matrix Anal., 29 (2006), pp. 143–159.

[19] L. Karlsson and F. Tisseur, *Algorithms for Hessenberg-triangular reduction of Fiedler linearization of matrix polynomials*, SIAM J. Sci. Comput., 37 (2015), pp. 384–414.

[20] P. Lancaster, *Linearization of regular matrix polynomials*, Electron. J. Linear Algebra, 17 (2008), pp. 21–27.

[21] D. S. Mackey, *Structured Linearizations for Matrix Polynomials*, Ph. D. Thesis, University of Manchester, Manchester, UK, 2006. Available as MIMS EPrint 2006.68, posted 27 June 2006.

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

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

[24] D. S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann, *Möbius transformations of matrix polynomials*, Linear Algebra Appl., 470 (2014), pp. 120–184.

[25] D. S. Mackey and V. Perović, *Linearizations of matrix polynomials in Bernstein bases.* Linear Algebra Appl., 501 (2016), pp. 162–197.

[26] D.S. Mackey and V. Perović, *Rational tools for matrix polynomials*, in preparation.

[27] Y. Nakatsukasa, V. Noferini, and A. Townsend, *Vector spaces of linearizations of matrix polynomials: a bivariate polynomial approach*, SIAM J. Matrix Anal., 38 (2017), pp. 1–29.

[28] V. Noferini and J. Pérez, *Fiedler-comrade and Fiedler-Chebyshev pencils*, SIAM J. Matrix Anal., 37 (2016), pp. 1600–1624.

[29] R. Van Beeumen, K. Meerbergen, and W. Michiels, *A rational Krylov method based on Hermite interpolation for nonlinear eigenvalue problems*, SIAM. J. Sci. Comput., 35 (2013), pp. 327–350.

[30] R. Van Beeumen, K. Meerbergen, and W. Michiels, *Linearization of Lagrange and Hermite interpolating matrix polynomials*, IMA J. Numer. Anal., 35 (2015), pp. 909–930.

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