Some Notes on the Linearization of Matrix Polynomials in Standard and Tschebyscheff Basis

Heike Faßbender[a], Philip Saltenberger[a,*]

[a] *Institut* Computational Mathematics/ *AG Numerik, TU Braunschweig, Pockelsstr. 14, 38106 Braunschweig, Germany*

Abstract

Eigenvalue problems of matrix polynomials $P(\lambda)$ in Tschebyscheff basis and suitable linearizations are considered. Following the ideas in [13], a vector space $\mathbb{T}_1(P)$ of potential linearizations is introduced and analyzed. All pencils in $\mathbb{T}_1(P)$ are characterized. An easy to check criterion whether a pencil in $\mathbb{T}_1(P)$ is a (strong) lineariation of $P(\lambda)$ is given. Moreover, a new criterion for determining whether a matrix pencil in the vector space $\mathbb{L}_1(P)$ (of potential linearizations for matrix polynomials in monomial basis) is a strong linearization for $P(\lambda)$ is derived. A structural resemblance between the matrix pencils in $\mathbb{L}_1$ and $\mathbb{T}_1$ is pointed out.

Keywords: matrix polynomial, (strong) linearization, Tschebyscheff basis
2000 MSC: 65F15, 15A18, 15A22, 47J10

1. Introduction

In [4] a 3D Laplace eigenvalue problem with Dirichlet boundary conditions is solved via a polynomial eigenvalue problem $P(\lambda)x = 0$, where P is given in Tschebyscheff basis

$$P(\lambda) = A_k t_k(\lambda) + A_{k-1} t_{k-1}(\lambda) + \cdots + A_1 t_1(\lambda) + A_0 t_0(\lambda) \tag{1}$$

with $n \times n$ matrices $A_0, \ldots, A_k$ and the scalar Tschebyscheff polynomials $t_j : \mathbb{R} \to \mathbb{R}$ defined as

$$t_0(\lambda) = 1, t_1(\lambda) = \lambda, t_n(\lambda) = 2\lambda t_{n-1}(\lambda) - t_{n-2}(\lambda), n \geq 2. \tag{2}$$

*Corresponding author, Email philip.saltenberger@tu-braunschweig.de

The eigenvalue problem $P(\lambda)x = 0$ is solved via the equivalent generalized linear eigenvalue problem $\mathcal{L}_0 y = \lambda \mathcal{L}_1 y$ with

$$\mathcal{L}_0 = \begin{bmatrix} 0_n & I_n & & & \\ I_n & 0_n & I_n & & \\ & \ddots & \ddots & \ddots & \\ & & I_n & 0_n & I_n \\ -A_0 & \cdots & -A_{k-3} & A_k - A_{k-2} & -A_{k-1} \end{bmatrix}, \mathcal{L}_1 = \begin{bmatrix} I_n & & & & \\ & 2I_n & & & \\ & & \ddots & & \\ & & & 2I_n & \\ & & & & 2A_k \end{bmatrix}. \tag{3}$$

This can be seen as a companion form linearization of (1).

Linearization is the classical approach for solving polynomial eigenvalue problems. This has been analyzed extensively for matrix polynomials $P(z)$ of degree k

$$P(\lambda) = A_k\lambda^k + \cdots + A_1\lambda + A_0, \qquad A_0, A_1, \ldots, A_k \in \mathbb{R}^{n\times n}, A_k \neq 0, \tag{4}$$

expressed in the monomial basis $\{1, \lambda, \cdots, \lambda^k\}$. In practice, when the polynomial $P(\lambda)$ is expressed in the monomial basis as in (4), the most used linearization to solve the polynomial eigenvalue problem $P(\lambda)y = 0$ is the Frobenius companion form

$$L(\lambda) = \lambda \begin{bmatrix} A_k & & & \\ & I_n & & \\ & & \ddots & \\ & & & I_n \end{bmatrix} + \begin{bmatrix} A_{k-1} & A_{k-2} & \cdots & A_0 \\ -I_n & 0_n & & \\ & \ddots & \ddots & \\ & & -I_n & 0_n \end{bmatrix}.$$

It is well-known that the conditioning of the Frobenius companion form linearization may be worse than the one of the original problem. Moreover, it usually does not preserve any structure present in $P(\lambda)$ (e.g., for $A_j = A_j^T \in \mathbb{R}^{n\times n}$ we have $P(\lambda)^T = P(\lambda)$, but $L(\lambda)^T \neq L(\lambda)$). Therefore it is of interest to have many classes of strong linearizations from which one can select a linearization with the most favorable properties in terms of, e.g., conditioning and backward errors of eigenvalues, or sparsity patterns. This has motivated a flurry of activity with the goal of finding new linearizations. The following list of references is an incomplete sample of recent papers on this topic [3, 13, 14, 15, 18].

When the polynomial $P(\lambda)$ is expressed in the monomial basis many linearizations are available in the literature [3, 13]. However, it is becoming of interest to solve polynomial eigenvalue problems for polynomials expressed in nonmonomial polynomial bases (see, e.g., [11, 4, 14, 15, 18]). In many such cases it is advisable to avoid reformulating $P(\lambda)$ in monomial basis, since this change of basis can be poorly conditioned, and may introduce numerical errors. Moreover, the instability increases with the degree [10]. Hence, constructing linearizations of matrix polynomials from the coefficients of $P(\lambda)$ in the given basis has become an active topic of research. Several linearizations for different polynomial bases have been proposed in [1], and particularly linearizations in the Tschebyscheff basis [8, 15], Bernstein basis [14, 19], Newton basis [9, 17], Lagrange basis [2, 5, 18], and Hermite basis [18] have been investigated. The backward stability of the polynomial rootfinding problem for polynomials in Tschebyscheff basis solved with colleague matrices is considered in [16].

Our interest here lies with linearizations of matrix polynomials $P(\lambda)$ in Tschebyscheff basis. In line with the derivations in [13] we introduce a vector space $\mathbb{T}_1(P)$ of potential linearizations for matrix polynomials in Tschebyscheff basis. A characterization of all pencils in $\mathbb{T}_1(P)$ is derived as well as an easy to check condition whether a pencil in $\mathbb{T}_1(P)$ is a (strong) linearization of $P(\lambda)$. Moreover, a result on the recovery of eigenvectors for matrix pencils in $\mathbb{T}_1(P)$ is presented. Comparing the results obtained here with those of [13], it turns out that the characterization of the pencils in $\mathbb{T}_1$ is structurally similar to the characterization of the pencils in $\mathbb{L}_1$. In order to see this, we consider the main characterization theorem for the pencils in $\mathbb{L}_1$ in a slightly different form as in the original paper. This helps to derive a new criterion on how to determine whether a matrix pencil in $\mathbb{L}_1$ is a strong linearization for $P(\lambda)$ which is essentially the same as the one for pencils in $\mathbb{T}_1(P)$.

After introducing some basic definitions and notions in the next section, Section 3 reviews some of the results derived in [13] on the ansatz space $\mathbb{L}_1$. Moreover, some new observations are given. In Section 4 we introduce and analyze the vector space $\mathbb{T}_1$ of potential linearizations analogous to $\mathbb{L}_1$.

2. Preliminaries

Matrix polynomials of degree k in standard monomial basis are of the form (4): $P(\lambda) = \sum_{j=0}^{k} A_j \lambda^j$, where $A_0, \ldots, A_k \in \mathbb{R}^{n \times n}$, $A_k \neq 0$, and $k \geq 0$. The set of all scalar polynomials of degree at most k will be denoted by $\mathbb{R}_k[\lambda]$. With $\mathbb{R}_k^{\ell}[\lambda]$ we will denote the vector space of all vectors of length ℓ whose entries are elements of $\mathbb{R}_k[\lambda]$; $\mathbb{R}_k^{\ell \times m}[\lambda]$ is the set of $\ell \times m$ matrices whose entries are elements of $\mathbb{R}_k[\lambda]$. Hence, $P(\lambda)$ may also be considered as an element of $\mathbb{R}_k^{n \times n}[\lambda]$.

A matrix polynomial $P(\lambda) \in \mathbb{R}_k^{n\times n}[\lambda]$ is said to be regular if $\det(P(\lambda)) \not\equiv 0$, otherwise it is said to be singular.

If $\lambda \in \mathbb{C}$ and $x \in \mathbb{C}^n \backslash \{0\}$ satisfy $P(\lambda)x = 0$, then x is said to be a (right) eigenvector of P corresponding to the (finite) eigenvalue λ. In case A_k is allowed to be singular, it is necessary to consider ∞ as a possible eigenvalue of $P(\lambda)$. In order to do so, the concept of the reversal of a polynomial is needed.

Definition 1 (Reversal of matrix polynomial). *For a matrix polynomial $P(\lambda) = \sum_{j=0}^{k} A_j\lambda^j \in \mathbb{R}_k^{n\times n}[\lambda]$ with $A_k \neq 0$ and $k \geq 1$, the reversal of $P(\lambda)$ is the polynomial*

$$\operatorname{rev}(P(\lambda)) := \lambda^k P(\frac{1}{\lambda}) = \sum_{j=0}^{k} A_{k-j}\lambda^j.$$

The nonzero finite eigenvalues of $\operatorname{rev}(P(\lambda))$ are the reciprocals of those of $P(\lambda)$.

Definition 2 (Eigenvalue ∞). *$P(\lambda)$ is said to have an eigenvalue ∞ with eigenvector x if $\operatorname{rev}(P(\lambda))$ has the eigenvalue 0 with eigenvector x.*

Typically polynomial eigenvalue problems are solved via linearization. That is, the polynomial $P(\lambda) \in \mathbb{R}_k^{n\times n}[\lambda]$ is transformed into a linear matrix pencil $L(\lambda) = L_0 + \lambda L_1$ with $L_0, L_1 \in \mathbb{R}^{kn\times kn}$ (which is the same as to say that $L(\lambda) \in \mathbb{R}_1^{kn\times kn}[\lambda]$) with the same eigenvalues, see, e.g., [7]. In order to deal with eigenvalues at ∞, the notion of a strong linearization was introduced, see, e.g. [6, 12].

Definition 3 ((Strong) Linearization). *A $kn \times kn$ matrix pencil $L(\lambda)$ is called (weak) linearization of the $n \times n$ matrix polynomial $P(\lambda)$ of degree $k \geq 1$, if there exist $kn \times kn$ unimodular matrix polynomials $U(\lambda)$ and $V(\lambda)$ so that*

$$U(\lambda)L(\lambda)V(\lambda) = \left[\begin{array}{c|ccc} P(\lambda) & 0_n & \cdots & 0_n \\ \hline 0_n & & & \\ \vdots & & I_{(k-1)n} & \\ 0_n & & & \end{array}\right]. \tag{5}$$

The linearization is called a strong linearization, if $L(\lambda)$ is a linearization of $P(\lambda)$ and $\operatorname{rev}(L(\lambda))$ is a linearization of $\operatorname{rev}(P(\lambda))$.
An unimodular matrix polynomial is a regular square matrix $E(\lambda)$ with a determinant independent of λ, that is, $\det(E(\lambda))$ is a nonzero constant.

In case $L(\lambda)$ is a strong linearization of the regular matrix polynomial $P(\lambda)$, $L(\lambda)$ and $P(\lambda)$ share the same eigenvalues with the same algebraic and geometric multiplicities.

Other notation used throughout the paper: $I = I_\ell$ is the $\ell \times \ell$ identity matrix, $0 = 0_\ell$ is the $\ell \times \ell$ matrix of all zeros, and R_ℓ denotes the $\ell \times \ell$ reverse identity, $R_\ell = \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix} \in \mathbb{R}^{\ell \times \ell}$. In case, a $p \times q$ rectangular part of 0 is used, we sometimes use $0_{p \times q}$ to specify the dimension. The jth column of I_n will be denoted by e_j, and E_j denotes $E_j := (I_k \otimes e_j) = \begin{bmatrix} e_j & e_{n+j} & e_{2n+j} & \cdots & e_{(k-1)n+j} \end{bmatrix} \in \mathbb{R}^{kn \times k}$. The Kronecker product is denoted by $\otimes$ as usual.

3. Vector space $\mathbb{L}_1$ of potential linearizations (for matrix polynomials in monomial basis)

This section briefly reviews the main results on linearizations for matrix polynomials expressed in standard basis as given in [13, Section 3., 4., 4.1.]. Moreover, a new criterion for determining whether a matrix pencil in the vector space $\mathbb{L}_1(P)$ is a strong linearization for $P(\lambda)$ is derived.

The matrix polynomial $P(\lambda) = \sum_{j=0}^{k} A_j \lambda^j \in \mathbb{R}_k^{n \times n}[\lambda]$ with $A_k \neq 0$, and $k \geq 1$ can be written as

$$\begin{aligned} P(\lambda) &= \begin{bmatrix} \lambda A_k + A_{k-1} & A_{k-2} & \cdots & A_1 & A_0 \end{bmatrix} \begin{bmatrix} \lambda^{k-1} I_n \\ \lambda^{k-2} I_n \\ \vdots \\ \lambda I_n \\ I_n \end{bmatrix} & (6) \\ &=: \mathbf{M}(\lambda)^T (\mu(\lambda) \otimes I_n) & (7) \end{aligned}$$

with $\mathbf{M}(\lambda) \in \mathbb{R}_1^{kn \times n}[\lambda]$ and $\mu(\lambda) = \begin{bmatrix} \lambda^{k-1} & \lambda^{k-2} & \cdots & \lambda & 1 \end{bmatrix}^T \in \mathbb{R}_{k-1}^{k}[\lambda]$.

Definition 4 (Ansatz space $\mathbb{L}_1(P)$). *The set of all $kn \times kn$ matrix pencils $L(\lambda)$ satisfying the equation*

$$L(\lambda)\big(\mu(\lambda) \otimes I_n\big) = v \otimes P(\lambda) \tag{8}$$

for any $v \in \mathbb{R}^k$ is called ansatz space for $P(\lambda)$ as in (6) and is denoted $\mathbb{L}_1(P)$. The vector v in (8) is called ansatz vector for $L(\lambda)$.

It is stated in [13, Corollary 3.6] that $\mathbb{L}_1(P)$ is an $\mathbb{R}$-vector subspace of $\mathbb{R}_1^{kn\times kn}[\lambda]$ and that $\dim(\mathbb{L}_1(P)) = k(k-1)n^2 + k$ for any $n \times n$ matrix polynomial $P(\lambda)$ of degree $k \in \mathbb{N}$. The matrix pencils $L(\lambda)$ satisfying (8) can be characterized nicely. The following theorem states Theorem 3.5 from [13] in a slightly different way which will be more convenient for our discussion.

Theorem 1. *Let $P(\lambda) = \sum_{j=0}^{k} A_j\lambda^j \in \mathbb{R}_k^{n\times n}[\lambda]$ with $A_k \neq 0$ and $k \geq 1$, $v \in \mathbb{R}^k$ and $L(\lambda) \in \mathbb{R}_1^{kn\times kn}[\lambda]$. Then $L(\lambda)$ satisfies $L(\lambda)(\mu(\lambda)\otimes I_n) = v\otimes P(\lambda)$ if and only if it may be expressed as*

$$L(\lambda) = v \otimes \mathbf{M}(\lambda)^T + B\widehat{M}(\lambda)^T \tag{9}$$

with a particular matrix $B \in \mathbb{R}^{kn\times(k-1)n}$ and

$$\widehat{M}(\lambda)^T := \widehat{M}_\star(\lambda)^T \otimes I_n \in \mathbb{R}_1^{(k-1)n\times kn}[\lambda]$$

and

$$\widehat{M}_\star(\lambda)^T := \begin{bmatrix} -1 & \lambda & & & \\ & -1 & \lambda & & \\ & & \ddots & \ddots & \\ & & & -1 & \lambda \end{bmatrix} \in \mathbb{R}_1^{(k-1)\times k}[\lambda].$$

The expression in (9) is unique, that is, every $L(\lambda) \in \mathbb{L}_1(P)$ may be uniquely identified through v and B.

Note that $\widehat{M}_\star(\lambda)^T\mu(\lambda) = 0 \in \mathbb{R}^{k-1}$, so we immediately obtain $\widehat{M}(\lambda)^T(\mu(\lambda)\otimes I_n) = 0 \in \mathbb{R}^{(k-1)n\times n}$. In [13, Theorem 3.5], the matrix pencils $L(\lambda)$ satisfiying the ansatz equation (8) are characterized in polynomial form as

$$L(\lambda) = \lambda \begin{bmatrix} v \otimes A_k & B \end{bmatrix} + \begin{bmatrix} -B + (v \otimes [\, A_{k-1} \ \cdots \ A_1 \,]) & v \otimes A_0 \end{bmatrix} \tag{10}$$

with $B \in \mathbb{R}^{kn\times(k-1)n}$. It is easy to see that the expressions (9) and (10) are the same.

Not all matrix pencils in $\mathbb{L}_1(P)$ are linearizations for P. The following theorem (see Theorem 4.1 and Section 4.1 in [13]) states a criterion for deciding whether special structured pencils from $\mathbb{L}_1(P)$ are strong linearizations or not.

Theorem 2. *Suppose that $P(\lambda) = \sum_{j=0}^{k} A_j \lambda^j \in \mathbb{R}_k^{n\times n}[\lambda]$ with $A_k \neq 0$ and $k \geq 1$, and $L(\lambda) = \lambda X + Y \in \mathbb{L}_1(P)$ has nonzero right ansatz vector $v = \alpha e_1 \in \mathbb{R}^k$, so that*

$$L(\lambda) \cdot (\mu(\lambda) \otimes I_n) = \alpha e_1 \otimes P(\lambda).$$

In this case, X and Y can be partitioned such that

$$L(\lambda) = \lambda X + Y = \lambda \left[\begin{array}{c|c} \alpha A_k & X_{12} \\ \hline 0 & Z \end{array}\right] + \left[\begin{array}{c|c} Y_{11} & \alpha A_0 \\ \hline -Z & 0 \end{array}\right], \tag{11}$$

where $Z \in \mathbb{R}^{(k-1)n\times(k-1)n}$. Then $L(\lambda)$ is a strong linearization of $P(\lambda)$ if $\operatorname{rank}(Z) = (k-1)n$.

Observe that for the special $L(\lambda)$ in (11) the matrix B in (9) is just $\left[\begin{smallmatrix} X_{12} \\ Z \end{smallmatrix}\right]$ as can be seen from (10). In the case $P(\lambda)$ is a polynomial of degree one, (11) actually reduces to $\alpha A_1 \lambda + \alpha A_0$.

Now suppose that $P(\lambda) = \sum_{j=1}^{k} A_j \lambda^j$ is an $n \times n$ regular matrix polynomial, and $L(\lambda) = \lambda X + Y \in \mathbb{L}_1(P)$ has nonzero right ansatz vector v, so that

$$L(\lambda) \cdot (\mu(\lambda) \otimes I_n) = v \otimes P(\lambda).$$

There always exists a nonsingular matrix $H \in \mathbb{R}^{k\times k}$ such that $Hv = \alpha e_1$. Applying $(H \otimes I_n)$ to $L(\lambda)$ generates $\breve{L}(\lambda) = (H \otimes I_n)L(\lambda)$ which must be of the form (11) since it has ansatz vector αe_1 (this may be seen easily from (9) or (10)). It is clear that $\breve{L}(\lambda)$ is a linearization for $P(\lambda)$ if and only if $L(\lambda)$ is a linearization for $P(\lambda)$, so, extracting Z from $\breve{L}(\lambda)$, the latter may now be checked according to Theorem 2.

Actually it is not necessary to transform a matrix pencil $L(\lambda) \in \mathbb{L}_1(P)$ into the form (11) in order to determine whether $L(\lambda)$ is a strong linearization. This can be directly determined from $L(\lambda) = v \otimes \mathbf{M}(\lambda)^T + B\widehat{M}(\lambda)^T$.

Theorem 3. *Let $P(\lambda) = \sum_{j=0}^{k} A_j \lambda^j \in \mathbb{R}_k^{n\times n}[\lambda]$ with $A_k \neq 0$ and $k \geq 1$, and $L(\lambda) = v \otimes \mathbf{M}(\lambda)^T + B\widehat{M}(\lambda)^T$ with (nonzero) ansatz vector $v \in \mathbb{R}^k$ and arbitrary $B \in \mathbb{R}^{kn\times(k-1)n}$. Then $L(\lambda)$ is a strong linearization for $P(\lambda)$ if* $\operatorname{rank}([(v \otimes I_n) \;\; B]) = kn$.

In order to prove this theorem, let us first observe an equivalent condition for Z in (11) being nonsingular.

Lemma 1. *Let* $\breve{B} = \left[\begin{smallmatrix} X_{12} \\ Z \end{smallmatrix}\right] \in \mathbb{R}^{kn\times(k-1)n}$ *be the last* $(k-1)n$ *columns of* X *in (11), that is,* $Z \in \mathbb{R}^{(k-1)n\times(k-1)n}$ *and* $X_{12} \in \mathbb{R}^{n\times(k-1)n}$*. Moreover, let* α *in (11) be nonzero. Then* $\operatorname{rank}(Z) = (k-1)n$ *if and only if* $\operatorname{rank}([\,(\alpha e_1 \otimes I_n)\;\; \breve{B}\,]) = kn$*.*

Proof. $\Rightarrow$ Let $\operatorname{rank}(Z) = (k-1)n$ and suppose $\operatorname{rank}([\,(\alpha e_1 \otimes I_n)\;\; \breve{B}\,]) < kn$. Then there exists a vector $u = [u_1 \;\cdots\; u_{kn}]^T \in \mathbb{R}^{kn}$ with $[\,(\alpha e_1 \otimes I_n)\; \breve{B}\,]u = 0_{kn\times 1}$. This implies

$$\breve{B}\begin{bmatrix} u_{n+1} & \cdots & u_{kn} \end{bmatrix}^T = \begin{bmatrix} \breve{w}^T & 0_{1\times(k-1)n} \end{bmatrix}^T$$

with some $\breve{w} \in \mathbb{R}^n$. Therefore $Z[u_{n+1} \;\cdots\; u_{kn}]^T = 0_{(k-1)n\times 1}$ which contradicts $\operatorname{rank}(Z) = (k-1)n$. Hence we have $\operatorname{rank}([\,(\alpha e_1 \otimes I_n)\;\; \breve{B}\,]) = kn$.
$\Leftarrow$ Now let $\operatorname{rank}([\,(\alpha e_1 \otimes I_n)\;\; \breve{B}\,]) = kn$ and assume $\operatorname{rank}(Z) < (k-1)n$. Then there exists a vector $u = [u_1 \;\cdots\; u_{(k-1)n}]^T \in \mathbb{R}^{(k-1)n}$ with $Zu = 0_{(k-1)n\times 1}$ and $\breve{B}u = [\breve{u}^T\, 0_{1\times(k-1)n}]^T$ for a particular $\breve{u} = [\breve{u}_1 \;\cdots\; \breve{u}_n]^T \in \mathbb{R}^n$. We immediately obtain

$$\left[\,(\alpha e_1 \otimes I_n)\; \breve{B}\,\right]\begin{bmatrix} -\dfrac{\breve{u}_1}{\alpha} & -\dfrac{\breve{u}_2}{\alpha} & \cdots & -\dfrac{\breve{u}_n}{\alpha} & u^T \end{bmatrix}^T = 0_{kn\times 1},$$

which contradicts $\operatorname{rank}([\,(\alpha e_1 \otimes I_n)\;\; \breve{B}\,]) = kn$. Thus we must have $\operatorname{rank}(Z) = (k-1)n$. This completes the proof. $\square$

Next we observe that for any nonsingular $H \in \mathbb{R}^{k\times k}$ such that $Hv = \alpha e_1$ for some $\alpha \in \mathbb{R}\backslash\{0\}$, $(H \otimes I_n)L(\lambda)$ is of the form (11) which is the same as

$$(H \otimes I_n)L(\lambda) = \alpha e_1 \otimes \mathbf{M}(\lambda)^T + \breve{B}\widehat{M}(\lambda)^T$$

with $\breve{B} = (H \otimes I_n)B = \begin{bmatrix} \breve{B}_1 \\ \breve{B}_2 \end{bmatrix}$ and $\breve{B}_2 \in \mathbb{R}^{(k-1)n\times(k-1)n}$. From Lemma 1 we have the equivalence of $\operatorname{rank}(\breve{B}_2) = (k-1)n$ and $\operatorname{rank}([\,(\alpha e_1 \otimes I_n)\, \breve{B}\,]) = kn$. Moreover,

$$\begin{aligned}
\big(H \otimes I_n\big)\left[\,(v \otimes I_n)\;\; B\,\right] &= \left[(H \otimes I_n)(v \otimes I_n)\;\; (H \otimes I_n)B\,\right] \\
&= \left[(Hv \otimes I_n)\;\; (H \otimes I_n)B\,\right] \\
&= \left[(\alpha e_1 \otimes I_n)\;\; \breve{B}\,\right]
\end{aligned}$$

and thus, since $H \otimes I_n$ is regular, $[(v \otimes I_n)\;\; B]$ has full rank if and only if $[(\alpha e_1 \otimes I_n)\;\; \breve{B}]$ has full rank. Hence $\operatorname{rank}(\breve{B}_2) = (k-1)n$ if and only if $\operatorname{rank}([\,(v \otimes I_n)\;\; B\,]) = kn$. This immediately gives Theorem 3.

The converse of Theorem 3 (and Theorem 2) is true only for regular $P(\lambda)$. This is because a linearization $L(\lambda) = v \otimes \mathbf{M}(\lambda)^T + B\widehat{M}(\lambda)^T$ of a regular $P(\lambda)$ needs to be regular, too, but $\operatorname{rank}(\begin{bmatrix}(v \otimes I_n) & B\end{bmatrix}) < kn$ always implies the existence of a vector $w \in \mathbb{C}^{kn}$ with $w^T \begin{bmatrix}(v \otimes I_n) & B\end{bmatrix} = 0_{1\times kn}$. Therefore it follows that

$$w^T L(\lambda) = w^T \left((v \otimes I_n)\mathbf{M}(\lambda)^T + B\widehat{M}(\lambda)^T \right) = 0_{1\times kn}$$

for all $\lambda \in \mathbb{C}$. In this case, the matrix pencil $L(\lambda)$ is singular which contradicts the fact that $L(\lambda)$ was assumed to be a linearization of a regular $P(\lambda)$. The following theorem from [13] (complemented by our observation in Theorem 3) reveals the connection between linearizations and regular pencils in $\mathbb{L}_1(P)$.

Theorem 4. *Let $P(\lambda)$ be a regular $n \times n$ matrix polynomial of degree $k \geq 1$ and $L(\lambda) = v \otimes \mathbf{M}(\lambda)^T + B\widehat{M}(\lambda)^T \in \mathbb{L}_1(P)$. Then the following statements are equivalent:*

1. *$L(\lambda)$ is a linearization for $P(\lambda)$.*
2. *$L(\lambda)$ is a regular matrix pencil.*
3. *$L(\lambda)$ is a strong linearization for $P(\lambda)$.*
4. *$\operatorname{rank}([(v \otimes I_n) \;\; B]) = kn$.*

Finally, we state a result on the recovery of eigenvectors for matrix pencils in $\mathbb{L}_1(P)$, see [13, Theorem 3.8, 4.4].

Theorem 5. *Let $P(\lambda)$ be a regular $n \times n$ matrix polynomial of degree $k \in \mathbb{N}$ and $L(\lambda) \in \mathbb{L}_1(P)$ a linearization for $P(\lambda)$.*

1. *Let $\alpha \in \mathbb{C}$ be a finite eigenvalue of $P(\lambda)$. Then $u \in \mathbb{C}^{kn}$ is an eigenvector of $L(\alpha)$ if and only if $u = \mu(\alpha) \otimes w$ for an eigenvector w of $P(\alpha)$.*
2. *Let $\alpha = \infty$. Then $u \in \mathbb{C}^{kn}$ is eigenvector for L with eigenvalue ∞ if and only if $u = e_1 \otimes w$ with an eigenvector w of P with eigenvalue ∞.*

4. Vector space $\mathbb{T}_1$ of potential linearizations (for matrix polynomials in Tschebyscheff basis)

In this section $n \times n$ matrix polynomials of degree $k \geq 1$ in Tschebyscheff basis

$$P(\lambda) = A_k t_k(\lambda) + A_{k-1} t_{k-1}(\lambda) + \cdots + A_1 t_1(\lambda) + A_0 t_0(\lambda) \in \mathbb{R}_k^{n\times n}[\lambda] \quad (12)$$

with $A_k \neq 0$ are considered. Following the idea of (7), $P(\lambda)$ can be written as

$$\begin{aligned} P(\lambda) &= \begin{bmatrix} 2\lambda A_k + A_{k-1} & A_{k-2} - A_k & A_{k-3} & \cdots & A_1 & A_0 \end{bmatrix} \begin{bmatrix} t_{k-1}(\lambda) I_n \\ t_{k-2}(\lambda) I_n \\ \vdots \\ t_1(\lambda) I_n \\ t_0(\lambda) I_n \end{bmatrix} \\ &=: \mathbf{T}(\lambda)^T (\tau(\lambda) \otimes I_n), \end{aligned} \tag{13}$$

where $\mathbf{T}(\lambda) \in \mathbb{R}_1^{kn \times n}[\lambda]$ and $\tau(\lambda) = \begin{bmatrix} t_{k-1}(\lambda) & t_{k-2}(\lambda) & \cdots & t_1(\lambda) & t_0(\lambda) \end{bmatrix}^T \in \mathbb{R}_{k-1}^k[\lambda]$.

Analogous to the definition of $\mathbb{L}_1(P)$ for a matrix polynomial in standard basis we define the ansatz space $\mathbb{T}_1(P)$ for matrix polynomials (12) in Tschebyscheff basis.

Definition 5 (Ansatz space $\mathbb{T}_1(P)$). *The set of all $kn \times kn$ matrix pencils $L(\lambda)$ satisfying the equation*

$$L(\lambda)\big(\tau(\lambda) \otimes I_n\big) = v \otimes P(\lambda) \tag{14}$$

for any $v \in \mathbb{R}^k$ is called ansatz space for $P(\lambda)$ as in (12) and is denoted $\mathbb{T}_1(P)$. The vector v in (14) is called ansatz vector for $L(\lambda)$.

Our main goal is to derive a characterization of $\mathbb{T}_1(P)$ similar to Theorem 1 for $\mathbb{L}_1(P)$. Since $\mathbb{T}_1(P) = \{cP(\lambda) \,|\, c \in \mathbb{R}\}$ whenever $P(\lambda)$ is of degree one, every element in $\mathbb{T}_1(P)$ is a linearization for $P(\lambda)$ (set $U(\lambda) = V(\lambda) = I_n$ in (5)). Therefore, we opt out of discussing this case in detail and restrict our attention to matrix polynomials of degree $k \geq 2$ from now on.

Theorem 6. *Let $P(\lambda) = \sum_{j=0}^k A_j t_j(\lambda) \in \mathbb{R}_k^{n \times n}[\lambda]$ with $A_k \neq 0$ and $k \geq 2$, $L(\lambda) \in \mathbb{R}_1^{kn \times kn}[\lambda]$, and $v \in \mathbb{R}^k$ arbitrary. Then $L(\lambda)$ satisfies (14) if and only if it may be expressed as*

$$L(\lambda) = v \otimes \mathbf{T}(\lambda)^T + B\widehat{T}(\lambda)^T \tag{15}$$

with a particular matrix $B \in \mathbb{R}^{kn \times (k-1)n}$ where $\tau(\lambda)$ and $\mathbf{T}(\lambda)$ are as in (13),

$$\widehat{T}(\lambda)^T := \widehat{T}_\star(\lambda)^T \otimes I_n \in \mathbb{R}_1^{(k-1)n \times kn}$$

and

$$\widehat{T}_\star(\lambda) := \begin{bmatrix} -1 & & & & \\ 2\lambda & -1 & & & \\ -1 & 2\lambda & \ddots & & \\ & -1 & \ddots & -1 & \\ & & \ddots & 2\lambda & -1 \\ & & & -1 & \lambda \end{bmatrix} \in \mathbb{R}_1^{k\times k-1}[\lambda]. \tag{16}$$

The expression (15) uniquely characterizes every $L(\lambda) \in \mathbb{T}_1(P)$ through v and B.

Based on this characterization of $\mathbb{T}_1(P)$, one obtains the same criterion whether $L(\lambda) \in \mathbb{T}_1(P)$ is a (strong) linearization for P as in Theorem 4.

Theorem 7. *Let $P(\lambda)$ be a regular $n \times n$ matrix polynomial of degree $k \geq 2$ and $L(\lambda) = v \otimes \mathbf{T}(\lambda)^T + B\widehat{T}(\lambda)^T \in \mathbb{T}_1(P)$. Then the following statements are equivalent:*

1. $\operatorname{rank}([(v \otimes I_n) \;\; B]) = kn$.
2. *$L(\lambda)$ is a strong linearization for $P(\lambda)$.*
3. *$L(\lambda)$ is a linearization for $P(\lambda)$.*
4. *$L(\lambda)$ is a regular matrix pencil.*

It is fairly easy to see that $\mathbb{L}_1(P)$ and $\mathbb{T}_1(P)$ are isomorphic and therefore that $\dim(\mathbb{T}_1(P)) = k + k(k-1)n^2$, which is just the sum of the dimension of v and the dimension of B. Moreover, as almost all matrix pencils in $\mathbb{L}_1(P)$ are strong linearizations for $P(\lambda)$ (see Theorem 4.7 in [13]), almost all matrix pencils in $\mathbb{T}_1(P)$ are strong linearizations.

Before we discuss how to prove Theorems 6 and 7, let us mention that the result on the recovery of eigenvectors for matrix pencils in $\mathbb{T}_1(P)$ is essentially the same as for the recovery of eigenvectors for matrix pencils in $\mathbb{L}_1(P)$. As the proof is similar to that of Theorem 3.8 in [13] no proof is given here.

Theorem 8. *Let $P(\lambda)$ be a regular $n \times n$ matrix polynomial of degree $k \in \mathbb{N}$ and $L(\lambda) \in \mathbb{T}_1(P)$ a linearization for $P(\lambda)$.*

1. *Let $\alpha \in \mathbb{C}$ be a finite eigenvalue of $P(\lambda)$. Then $u \in \mathbb{C}^{kn}$ is an eigenvector of $L(\alpha)$ if and only if $u = \tau(\alpha) \otimes w$ for an eigenvector w of $P(\alpha)$.*
2. *Let $\alpha = \infty$. Then $u \in \mathbb{C}^{kn}$ is eigenvector for L with eigenvalue ∞ if and only if $u = e_1 \otimes w$ with an eigenvector w of P with eigenvalue ∞.*

4.1. Proof of Theorem 6

Now we will prove Theorem 6.

First recall that $P(\lambda) = \sum_{s=0}^{k} A_s t_s(\lambda) \in \mathbb{R}_k^{n\times n}[\lambda]$ may be interpreted as an $n \times n$ matrix whose entries are scalar polynomials $p_{ij}(\lambda) \in \mathbb{R}_k^1[\lambda]$,

$$[P(\lambda)]_{ij} = p_{ij}(\lambda) = \sum_{s=0}^{k} a_{ij}^{(s)} t_s(\lambda) \quad \text{with} \quad [A_s]_{ij} = a_{ij}^{(s)}$$

for $1 \leq i,j \leq n$. Whenever $L(\lambda)$ satisfies $L(\lambda)(\tau(\lambda) \otimes I_n) = v \otimes P(\lambda)$, we obtain

$$\widetilde{L}_{ij}(\lambda)\tau(\lambda) = v \otimes p_{ij}(\lambda), \tag{17}$$

for every $\widetilde{L}_{ij}(\lambda) = E_i^T L(\lambda) E_j$, $1 \leq i,j \leq n$ where $E_i = (I_k \otimes e_i)$. This is seen easily since

$$\begin{aligned}
\widetilde{L}_{ij}(\lambda)\tau(\lambda) &= (I_k \otimes e_i^T)L(\lambda)(I_k \otimes e_j)\tau(\lambda) \\
&= (I_k \otimes e_i^T)L(\lambda)(I_k \otimes e_j)(\tau(\lambda) \otimes 1) \\
&= (I_k \otimes e_i^T)L(\lambda)(\tau(\lambda) \otimes e_j) \\
&= (I_k \otimes e_i^T)L(\lambda)(\tau(\lambda) \otimes I_n)(1 \otimes e_j) \\
&= (I_k \otimes e_i^T)(v \otimes P(\lambda))(1 \otimes e_j) \\
&= (I_k \otimes e_i^T)(v \otimes P(\lambda)e_j) \\
&= v \otimes e_i^T P(\lambda) e_j \\
&= v \otimes p_{ij}(\lambda).
\end{aligned}$$

Hence, the ansatz equation $L(\lambda)(\tau(\lambda) \otimes I_n) = v \otimes P(\lambda)$ is composed of n^2 subproblems of the form

$$\widetilde{L}_{ij}(\lambda)\tau(\lambda) = v \otimes p_{ij}(\lambda) \qquad \big(= vp(\lambda)\big), \qquad i,j = 1,\ldots,n \tag{18}$$

for scalar polynomials $p_{ij}(\lambda)$ and $\widetilde{L}_{ij}(\lambda) \in \mathbb{R}_1^{k\times k}[\lambda]$. In order to easily identify this ansatz space for $n = 1$, the set of all $k \times k$ matrix pencils $\widetilde{L}(\lambda)$ satisfying $\widetilde{L}(\lambda)\tau(\lambda) = v \otimes p(\lambda)$ for any $v \in \mathbb{R}^k$ will be called scalar Tschebyscheff ansatz space for $p(\lambda) = \sum_{s=0}^{k} a_{ij} t_s(\lambda)$ and is denoted by $\mathbb{T}_\star(p)$.

In particular, the following scalar version of Theorem 6 holds. Notice that not all polynomials $p_{ij}(\lambda)$ appearing in (18) need to have degree k, so we need to formulate the scalar version of Theorem 6 in a quite general fashion. Theorem 6 then follows from the above considerations easily.

Theorem 9. *Let $k \geq 2$ and $p(\lambda) = \sum_{j=0}^{k} a_j t_j(\lambda) \in \mathbb{R}_k[\lambda]$ with degree $s, 0 \leq s \leq k$ be given. Assume $\widetilde{L}(\lambda) \in \mathbb{R}_1^{k\times k}[\lambda]$ and $v \in \mathbb{R}^k$ to be arbitrary. Then $\widetilde{L}(\lambda)$ satisfies $\widetilde{L}(\lambda)\tau(\lambda) = vp(\lambda)$ if and only if $\widetilde{L}(\lambda)$ may be expressed as*

$$\widetilde{L}(\lambda) = v \otimes \mathbf{T}_\star(\lambda)^T + \widetilde{B}\widehat{T}_\star(\lambda)^T \tag{19}$$

with a particular matrix $\widetilde{B} \in \mathbb{R}^{k\times k-1}$ where $\tau(\lambda)$ is as in (13), $\widehat{T}_\star(\lambda)$ as in (16), and

$$\mathbf{T}_\star(\lambda)^T = [2\lambda a_k + a_{k-1} \quad a_{k-2} - a_k \quad a_{k-3} \quad a_{k-2} \quad \cdots \quad a_1 \quad a_0] \in \mathbb{R}_1^k[\lambda]$$

is the scalar version of $\mathbf{T}(\lambda)^T$ as in (13). Moreover, the expression in (19) is unique, i.e. every element in $\mathbb{T}_\star(p)$ is uniquely determined by v and $\widetilde{B}$.

Hence, for each of the n^2 subproblems

$$\widetilde{L}_{ij}(\lambda)\tau(\lambda) = v \otimes p_{ij}(\lambda)$$

we obtain

$$\widetilde{L}_{ij}(\lambda) = v \otimes \mathbf{T}_{\star ij}(\lambda)^T + \widetilde{B}_{ij}\widehat{T}_\star(\lambda)^T$$

with

$$\mathbf{T}_{\star ij}(\lambda)^T = \left[2\lambda a_{ij}^{(k)} + a_{ij}^{(k-1)} \quad a_{ij}^{(k-2)} - a_{ij}^{(k)} \quad a_{ij}^{(k-3)} \quad a_{ij}^{(k-2)} \quad \cdots \quad a_{ij}^{(1)} \quad a_{ij}^{(0)}\right].$$

Assembling this into $L(\lambda)$ yields the desired expression stated in Theorem 6

$$\begin{aligned}
L(\lambda) &= \sum_{i,j=1}^{n} \widetilde{L}_{ij}(\lambda) \otimes e_i e_j^T \\
&= \sum_{i,j=1}^{n} \left(v \otimes \mathbf{T}_{\star ij}(\lambda)^T + \widetilde{B}_{ij}\widehat{T}_\star(\lambda)^T\right) \otimes e_i e_j^T \\
&= \sum_{i,j=1}^{n} v \otimes \left(\mathbf{T}_{\star ij}(\lambda)^T \otimes e_i e_j^T\right) + \left(\widetilde{B}_{ij}\widehat{T}_\star(\lambda)^T \otimes e_i e_j^T\right) \\
&= v \otimes \mathbf{T}(\lambda)^T + \left(\sum_{i,j=1}^{n} \widetilde{B}_{ij}\widehat{T}_\star(\lambda)^T \otimes e_i e_j^T I_n\right) \\
&= v \otimes \mathbf{T}(\lambda)^T + \left(\sum_{i,j=1}^{n} \widetilde{B}_{ij} \otimes e_i e_j^T\right)\left(\widehat{T}_\star(\lambda)^T \otimes I_n\right) \\
&= v \otimes \mathbf{T}(\lambda)^T + B\widehat{T}(\lambda)^T.
\end{aligned}$$

Thus, it suffices to prove Theorem 9.

Observe that $\widehat{T}_\star(\lambda)^T$ plays a similar role as $\widehat{M}_\star(\lambda)$ in Theorem 1, it essentially encodes the recursion (2)

$$\widehat{T}_\star(\lambda)^T \tau(\lambda) = 0 \in \mathbb{R}^{k-1}.$$

In order to prove Theorem 9, let us first assume that $p(\lambda)$ is of degree k; that is, $\deg(p(\lambda)) = k$. Consider the linear map

$$\begin{aligned} f: \quad & \mathbb{R}_1^{k\times k}[\lambda] \to \mathbb{R}_k^k[\lambda], \\ & \widetilde{L}(\lambda) \mapsto \widetilde{L}(\lambda)\tau(\lambda) \end{aligned}$$

which describes the left hand side of (18). As the Tschebyscheff polynomials $t_{k-1}(\lambda), \ldots, t_1(\lambda), t_0(\lambda)$ are a basis for the vector space $\mathbb{R}_{k-1}[\lambda]$ of scalar polynomials of degree at most $k-1$ and as $\widetilde{L}(\lambda) \in \mathbb{R}_1^{k\times k}[\lambda]$, f is surjective. But, as $t_{j+1}(\lambda) = 2\lambda t_j(\lambda) - t_{j-1}(\lambda)$, f is not injective. Hence the linear map f has to have a nontrivial null space.

Now consider $\mathcal{V} := \{vp(\lambda) | v \in \mathbb{R}^k\}$ which describes the right hand side of (18). Clearly, $\mathcal{V}$ is a subspace of $\mathbb{R}_k^k[\lambda]$ of dimension k and

$$\mathbb{T}_\star(p) = f^{\leftarrow}(\mathcal{V}) = \{\widetilde{L}(\lambda) \in \mathbb{R}_1^{k\times k}[\lambda] \mid f(\widetilde{L}(\lambda)) \in \mathcal{V}\}, \tag{20}$$

that is, $\mathbb{T}_\star(p)$ is the preimage of $\mathcal{V}$ under f. Hence, $\mathbb{T}_\star(p)$ is a subspace of $\mathbb{R}_1^{k\times k}[\lambda]$. Moreover,

$$\mathbb{T}_\star(p) = \mathbb{T}_\star(p) + \mathrm{null}(f)$$

(recall $\mathrm{null}(f) = \{\widetilde{L}(\lambda) \in \mathbb{R}_1^{k\times k}[\lambda] \mid \widetilde{L}(\lambda)\tau(\lambda) = 0\} \subset \mathbb{R}_1^{k\times k}[\lambda]$). Due to the construction of $\mathbb{T}_\star(p)$ and the fundamental homomorphism theorem we have

$$\mathcal{V} \backsimeq \mathbb{T}_\star(p)/\mathrm{null}(f). \tag{21}$$

We already observed that $\widehat{T}_\star(\lambda)^T\tau(\lambda) = 0_{k-1\times 1}$. Therefore, $\widetilde{B}\widehat{T}_\star(\lambda)^T\tau(\lambda) = 0_{k\times 1}$ for any $\widetilde{B} \in \mathbb{R}^{k\times k-1}$ and thus, since $\widetilde{B}\widehat{T}_\star(\lambda)^T \in \mathbb{R}_1^{k\times k}[\lambda]$, $\widetilde{B}\widehat{T}_\star(\lambda)^T \in \mathrm{null}(f)$. Even more, the set of matrices $\widetilde{B}\widehat{T}_\star(\lambda)^T$ already describe the entire kernel $\mathrm{null}(f)$.

Theorem 10. *Let $p(\lambda) = \sum_{j=0}^k a_j t_j(\lambda) \in \mathbb{R}_k[\lambda]$ arbitrary with degree $k \geq 2$. Then*

$$\mathrm{null}(f) = \{\widetilde{B}\widehat{T}_\star(\lambda)^T \mid \widetilde{B} \in \mathbb{R}^{k\times k-1}\}. \tag{22}$$

Proof. We have already seen that $\{\widetilde{B}\widehat{T}_\star(\lambda)^T \mid \widetilde{B} \in \mathbb{R}^{k\times k-1}\} \subseteq \mathrm{null}(f)$. In order to prove $\mathrm{null}(f) \subseteq \{\widetilde{B}\widehat{T}_\star(\lambda)^T \mid \widetilde{B} \in \mathbb{R}^{k\times k-1}\}$ it suffices to show that any $a(\lambda) \in \mathbb{R}_1^k[\lambda]$ with $a(\lambda)^T\tau(\lambda) = 0$ has the form $a(\lambda) = \widehat{T}_\star(\lambda)b$ with a particular vector $b \in \mathbb{R}^{k-1}$. The statement will be proven by contradiction.

Assume

$$a(\lambda) = \begin{bmatrix} a_1(\lambda) & a_2(\lambda) & \cdots & a_{k-1}(\lambda) & a_k(\lambda) \end{bmatrix}^T \in \mathbb{R}_1^k[\lambda]$$

to be given and that $a(\lambda)^T\tau(\lambda) = 0$. Furthermore suppose there is no vector $b \in \mathbb{R}^{k-1}$ so that $a(\lambda)$ satisfies $\widehat{T}_\star(\lambda)b = a(\lambda)$.

First notice that $a_1(\lambda)$ must be independent of λ, and hence has to be a real constant. This is because multiplying $t_{k-1}(\lambda)$ with λ produces a term λ^k that can not be eliminated by the subsequent computations (keep in mind that $a_i(\lambda) \in \mathbb{R}_1[\lambda]$). Thus $a(\lambda)$ may be expressed as

$$a(\lambda) = \begin{bmatrix} \alpha_1 \\ \alpha_2 + \beta_2\lambda \\ \vdots \\ \alpha_{k-1} + \beta_{k-1}\lambda \\ \alpha_k + \beta_k\lambda \end{bmatrix}. \tag{23}$$

Next define the vector $\tilde{b} := \frac{1}{2}\,[\,\beta_2 \;\; \beta_3 \;\; \cdots \;\; \beta_{k-1} \;\; \beta_k\,]^T \in \mathbb{R}^{k-1}$ and observe that

$$\tilde{c} := a(\lambda) - \widehat{T}_\star(\lambda)\tilde{b} \tag{24}$$

is now completely independent of λ, i.e. $\tilde{c} \in \mathbb{R}^k$. Therefore, $a(\lambda)$ can be written as $a(\lambda) = \widehat{T}_\star(\lambda)\tilde{b} + \tilde{c}$ with $\tilde{b} \in \mathbb{R}^{k-1}$ and $\tilde{c} \in \mathbb{R}^k$. Since $a(\lambda)$ has to satisfy $a(\lambda)^T\tau(\lambda) = 0$ we have

$$0 = a(\lambda)^T\tau(\lambda) = \big(\widehat{T}_\star(\lambda)\tilde{b} + \tilde{c}\big)^T\tau(\lambda) = \tilde{b}^T\widehat{T}_\star(\lambda)^T\tau(\lambda) + \tilde{c}^T\tau(\lambda)$$

and, since $\tilde{b}^T\widehat{T}_\star(\lambda)^T\tau(\lambda) = 0$ as $\{\widetilde{B}\widehat{T}_\star(\lambda)^T \mid \widetilde{B} \in \mathbb{R}^{k\times k-1}\} \subseteq \mathrm{null}(f)$, we obtain $\tilde{c}^T\tau(\lambda) = 0$.

Because $\tilde{c}$ does not contain any λ terms and $t_0(\lambda), \ldots, t_{k-1}(\lambda)$ are nonzero polynomials with $\deg(t_j(\lambda)) = j$, $\tilde{c}^T\tau(\lambda) = 0$ can only be achieved for $\tilde{c} = 0 \in \mathbb{R}^k$. Thus, $a(\lambda) = \widehat{T}_\star(\lambda)\tilde{b} + \tilde{c} = \widehat{T}_\star(\lambda)\tilde{b}$ which contradicts the assumption and establishes our statement. $\square$

Note that the characterization of $\mathrm{null}(f)$ is completely independent of $p(\lambda)$. As $\mathrm{null}(f)$ is a subspace of $\mathbb{R}_1^{k\times k}[\lambda]$, it follows immediately from Theorem 10 that

$$\dim(\mathrm{null}(f)) = k(k-1),$$

because $k \times (k-1)$ is exactly the dimension of $\tilde{B}$ in (22). Moreover, as $\dim(\mathcal{V}) = k$ and due to (21) we have

$$\dim(\mathbb{T}_\star(p)/\mathrm{null}(f)) = k,$$

and, since $\dim(\mathbb{T}_\star(p)/\mathrm{null}(f)) = \dim(\mathbb{T}_\star(p)) - \dim(\mathrm{null}(f))$,

$$\dim(\mathbb{T}_\star(p)) = k^2.$$

Next, a characterization of $\mathbb{T}_\star(p)/\mathrm{null}(f)$ is derived which will give Theorem 9. Observe that $\mathbf{T}_\star(\lambda)^T\tau(\lambda) = p(\lambda)$ due to the construction of $\mathbf{T}_\star(\lambda)$. So, for any $v \in \mathbb{R}^k$, it holds

$$(v \otimes \mathbf{T}_\star(\lambda)^T)\tau(\lambda) = v \otimes (\mathbf{T}_\star(\lambda)^T\tau(\lambda)) = v \otimes p(\lambda) = vp(\lambda).$$

Hence, since $(v \otimes \mathbf{T}_\star(\lambda)^T)$ is a $k \times k$ matrix pencil, $(v \otimes \mathbf{T}_\star(\lambda)^T) \in \mathbb{T}_\star(p)$. It is easy to check that $\{v \otimes \mathbf{T}_\star(\lambda)^T \mid v \in \mathbb{R}^k\}$ is a k dimensional subspace of $\mathbb{T}_\star(p)$. It turns out that

Theorem 11. *Let $p(\lambda) = \sum_{j=0}^{k} a_j t_j(\lambda) \in \mathbb{R}_k[\lambda]$ with degree $k \geq 2$. Then*

$$\mathbb{T}_\star(p)/\mathrm{null}(f) = \{v \otimes \mathbf{T}_\star(\lambda)^T + \mathrm{null}(f) \mid v \in \mathbb{R}^k\}.$$

Proof. We have already seen that $\{v \otimes \mathbf{T}_\star(\lambda)^T \mid v \in \mathbb{R}^k\} \subseteq \mathbb{T}_\star(p)$. This implies

$$\{v \otimes \mathbf{T}_\star(\lambda)^T + \mathrm{null}(f) \mid v \in \mathbb{R}^k\} \subseteq \mathbb{T}_\star(p)/\mathrm{null}(f).$$

In order to prove $\mathbb{T}_\star(p)/\mathrm{null}(f) \subseteq \{(v \otimes \mathbf{T}_\star(\lambda)^T) + \mathrm{null}(f) \mid v \in \mathbb{R}^k\}$ it suffices to show that any $a(\lambda) \in \mathbb{R}_1^k[\lambda]$ with $a(\lambda)^T\tau(\lambda) = \eta p(\lambda)$ for some $\eta \in \mathbb{R}$ has the form $a(\lambda) = \eta\mathbf{T}_\star(\lambda) + \widehat{T}_\star b$ with a particular vector $b \in \mathbb{R}^{k-1}$. The statement can be proven by contradiction similar to the approach in the proof of Theorem 10. $\square$

Theorem 10 gives a characterization of $\mathrm{null}(f)$, such that we have

$$\mathbb{T}_\star(p) = \{v \otimes \mathbf{T}_\star(\lambda)^T + \widetilde{B}\widehat{T}_\star(\lambda)^T \mid \widetilde{B} \in \mathbb{R}^{k\times k-1}, v \in \mathbb{R}^k\}$$

and hence, Theorem 9 is proven given the case $p(\lambda)$ is of degree k.

Now assume $\deg(p(\lambda)) = s < k$ and $a(\lambda) \in \mathbb{R}_1^k[\lambda]$ satisfies $a(\lambda)\tau(\lambda) = \eta p(\lambda)$ for some nonzero $\eta \in \mathbb{R}$. Since $a(\lambda)$ may once more be expressed in the form (23), $\tilde{c}$ can be calculated from (24) which is completely independent of λ. Hence, $\tilde{c}^T\tau(\lambda) = \eta p(\lambda)$ holds. This implies $\tilde{c}^T$ to be of the form

$$\tilde{c}^T = \eta \begin{bmatrix} 0 & 0 & \cdots & 0 & a_s & a_{s-1} & \cdots & a_1 & a_0 \end{bmatrix}$$

with $k-s$ leading zeros. This is just the vector $\mathbf{T}_\star(\lambda)^T$ from Theorem 9 in the case $a_k = a_{k-1} = \ldots = a_{s+1} = 0$. It is easily seen that the expression (19) is unique in both cases which completes the proof.

4.2. Proof of Theorem 7

In order to prove Theorem 7 it suffices to show that 1. $\Rightarrow$ 2. $\Rightarrow$ 3. $\Rightarrow$ 4. $\Rightarrow$ 1.. We will start with

1. $\Rightarrow$ 2. Under the condition $\operatorname{rank}([\,(v\otimes I_n)\ \ B\,]) = kn$ the constant $kn\times kn$ matrix $[\,(v\otimes I_n)\ \ B\,]$ is regular, i.e. invertible, so we may premultiply $L(\lambda)$ by $[\,(v\otimes I_n)\ \ B\,]^{-1}$ and calculate $\tilde{L}(\lambda) := [\,(v\otimes I_n)\ \ B\,]^{-1}L(\lambda)$:

$$\begin{aligned}\tilde{L}(\lambda) &= \begin{bmatrix} (v\otimes I_n) & B \end{bmatrix}^{-1}\big(v\otimes \mathbf{T}(\lambda)^T + B\widehat{T}(\lambda)^T\big) \qquad (25)\\ &= \begin{bmatrix} 2A_k & 0_n & \cdots & \cdots & 0_n \\ 0_n & 2I_n & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & 2I_n & 0_n \\ 0_n & \cdots & \cdots & 0_n & I_n \end{bmatrix}\lambda + \begin{bmatrix} A_{k-1} & A_{k-2}-A_k & A_{k-3} & \cdots & A_0 \\ -I_n & 0_n & -I_n & & 0_n \\ 0_n & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & -I_n & 0_n & -I_n \\ 0_n & \cdots & 0_n & -I_n & 0_n \end{bmatrix}.\end{aligned}$$

As $[\,(v\otimes I_n)\ \ B\,]$ is unimodular, $L(\lambda)$ is a strong linearization for $P(\lambda)$ if and only if $\tilde{L}(\lambda)$ is a strong linearization for $P(\lambda)$. Now consider the matrix polynomial $\hat{L}(\lambda) := (R_k\otimes I_n)\tilde{L}(\lambda)(R_k\otimes I_n)$, where $R_k = \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix}$ is the reverse identity. It is seen easily, that $\hat{L}(\lambda)$ coincides with the matrix pencil $\lambda\mathcal{L}_1 - \mathcal{L}_0$ (3). It is stated in [4], Theorem 2, and proven in [1], that $\hat{L}$ is in fact a strong linearization for $P(\lambda)$, so $\tilde{L}(\lambda)$ and $L(\lambda)$ are strong linearizations for $P(\lambda)$ as well.

4. $\Rightarrow$ 1. Now assume $L(\lambda) = (v\otimes I_n)\mathbf{T}(\lambda)^T + B\widehat{T}(\lambda)^T$ to be regular and $\operatorname{rank}(\begin{bmatrix}(v\otimes I_n) & B\end{bmatrix}) < kn$. Because of the rank-deficiency we may choose a vector $w \in \mathbb{C}^{kn}$ with $w^T\begin{bmatrix}(v\otimes I_n) & B\end{bmatrix} = 0_{1\times kn}$ and observe that this

immediately implies $w^T L(\lambda) = 0$ independently of λ (see also the discussion preceding Theorem 4). Therefore, $L(\lambda)$ must be singular which contradicts our assumption. Hence we must have $\text{rank}(\begin{bmatrix}(v \otimes I_n) & B\end{bmatrix}) = kn$.

Since 2. $\Rightarrow$ 3. and 3. $\Rightarrow$ 4. are clear this completes the proof.

5. Concluding Remarks

A vector space $\mathbb{T}_1(P)$ of potential linearizations for matrix polynomials in Tschebyscheff basis has been introduced guided by the ideas discussed in [13] in the context of polynomials in monomial basis. A complete characterization of all strong linearizations in $\mathbb{T}_1(P)$ has been derived. Comparing (9) and (15) a structural resemblance between the elements on $\mathbb{L}_1(P)$ and $\mathbb{T}_1(P)$ is apparent

space	ansatz equation	characterization
$\mathbb{L}_1(P)$	$L(\lambda)(\mu(\lambda) \otimes I_n) = v \otimes P(\lambda)$	$v \otimes \mathbf{M}(\lambda)^T + B\widehat{M}(\lambda)^T$
$\mathbb{T}_1(P)$	$L(\lambda)(\tau(\lambda) \otimes I_n) = v \otimes P(\lambda)$	$v \otimes \mathbf{T}(\lambda)^T + B\widehat{T}(\lambda)^T$

Moreover, in both cases, the criterion $\text{rank}([v \otimes I_n \;\; B]) = kn$ helps to identify strong linearizations.

It seems possible to extend the results presented to matrix polynomials in any orthogonal basis as well as to polynomials in bases which can be expressed by 'easy' to encode recursions such as the Newton basis. Moreover, with the results obtained here, it is straightforward to adapt the results for $\mathbb{L}_2$ and $\mathbb{DL}(P)$ introduced and discussed in [13] to analogue spaces for polynomials in Tschebyscheff basis.

Acknowledgement

Both authors would like to thank Nikta Shayanfar for many enlightening discussions on the topic and many other subjects. Moreover, the second author would like to thank Nikta Shayanfar for providing the opportunity to attend the MATTRAID conference 2015 in Coimbra.

References

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

[2] Corless, R., 2007. On a generalized companion matrix pencil for matrix polynomials expressed in the Lagrange basis. In: Symbolic-numeric computation. Invited and contributed presentations given at the international workshop (SNC 2005), Xi'an, China, July 19–21 (2005). Basel: Birkhäuser, pp. 1–15.

[3] De Terán, F., Dopico, F., Mackey, D., 2009. Linearizations of singular matrix polynomials and the recovery of minimal indices. Electron. J. Linear Algebra 18, 371–402.

[4] Effenberger, C., Kressner, D., 2012. Chebyshev interpolation for nonlinear eigenvalue problems. BIT Numer. Math. 52 (4), 933–951.

[5] Gemignani, L., 2007. Structured matrix methods for polynomial rootfinding. In: Proceedings of the 2007 international symposium on symbolic and algebraic computation, ISSAC 2007, Waterloo, ON, Canada, July 29–August 1, 2007. New York, NY: Association for Computing Machinery (ACM), pp. 175–180.

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

[7] Gohberg, I., Lancaster, P., Rodman, L., 2009. Matrix polynomials. Reprint of the 1982 original., reprint of the 1982 original Edition. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM).

[8] Good, I., 1961. The colleague matrix, a Chebyshev analogue of the companion matrix. Q. J. Math., Oxf. II. Ser. 12, 61–68.

[9] Güttel, S., Van Beeumen, R., Meerbergen, K., Michiels, W., 2013. NLEIGS: A class of robust fully rational Krylov methods for nonlinear eigenvalue problems. MIMS EPrint 2013.49, Manchester Institute for Mathematical Sciences, The University of Manchester, UK.

[10] Hermann, T., 1996. On the stability of polynomial transformations between Taylor, Bernstein and Hermite forms. Numer. Algorithms 13 (3-4), 307–320.

[11] Kressner, D., Roman, J., 2014. Memory-efficient Arnoldi algorithms for linearizations of matrix polynomials in Chebyshev basis. Numer. Linear Algebra Appl. 21 (4), 569–588.

[12] Lancaster, P., 2008. Linearization of regular matrix polynomials. Electron. J. Linear Algebra 17, 21–27.

[13] Mackey, D., Mackey, N., Mehl, C., Mehrmann, V., 2006. Vector spaces of linearizations for matrix polynomials. SIAM J. Matrix Anal. Appl. 28 (4), 971–1004.

[14] Mackey, D., Perović, V., 2014. Linearizations of matrix polynomials in Bernstein basis. MIMS EPrint 2014.29, Manchester Institute for Mathematical Sciences, The University of Manchester, UK.

[15] Noferini, V., Pérez, J., 2015. Chebyshev-Fiedler pencils. MIMS EPrint 2015.90, Manchester Institute for Mathematical Sciences, The University of Manchester, UK.

[16] Noferini, V., Pérez, J., 2015. Chebyshev rootfinding via computing eigenvalues of colleague matrices: when is it stable? MIMS EPrint 2015.24, Manchester Institute for Mathematical Sciences, University of Manchester, UK.

[17] Van Beeumen, R., Meerbergen, K., Michiels, W., 2013. A rational Krylov method based on Hermite interpolation for nonlinear eigenvalue problems. SIAM J. Sci. Comput. 35 (1), a327–a350.

[18] Van Beeumen, R., Michiels, W., Meerbergen, K., 2015. Linearization of Lagrange and Hermite interpolating matrix polynomials. IMA J. Numer. Anal. 35 (2), 909–930.

[19] Winkler, J., 2003. A companion matrix resultant for Bernstein polynomials. Linear Algebra Appl. 362, 153–175.