IMA Journal of Numerical Analysis (2015) **35**, 909–930
doi:10.1093/imanum/dru019
Advance Access publication on May 7, 2014

Linearization of Lagrange and Hermite interpolating matrix polynomials

Roel Van Beeumen, Wim Michiels and Karl Meerbergen
Department of Computer Science, KU Leuven, University of Leuven, 3001 Heverlee, Belgium
*Corresponding author: roel.vanbeeumen@cs.kuleuven.be wim.michiels@cs.kuleuven.be
karl.meerbergen@cs.kuleuven.be

[Received on 12 May 2013; revised on 5 March 2014]

This paper considers interpolating matrix polynomials $P(\lambda)$ in Lagrange and Hermite bases. A classical approach to investigating the polynomial eigenvalue problem $P(\lambda)x = 0$ is linearization, by which the polynomial is converted into a larger matrix pencil with the same eigenvalues. Since the current linearizations of degree n Lagrange polynomials consist of matrix pencils with $n + 2$ blocks, they introduce additional eigenvalues at infinity. Therefore, we introduce new linearizations which overcome this. Initially, we restrict to Lagrange and barycentric Lagrange matrix polynomials and give two new and more compact linearizations, resulting in matrix pencils of $n + 1$ and n blocks for polynomials of degree n. For the latter, there is a one-to-one correspondence between the eigenpairs of $P(\lambda)$ and the eigenpairs of the pencil. We also prove that these linearizations are strong. Moreover, we show how to exploit the structure of the proposed matrix pencils in Krylov-type methods, so that in this case we only have to deal with linear system solves of matrices of the original matrix polynomial dimension. Finally, we generalize for Hermite interpolation and introduce new linearizations for Hermite Lagrange and barycentric Hermite matrix polynomials. Again, we can show that the linearizations are strong and that there is a one-to-one correspondence of the eigenpairs.

Keywords: matrix polynomials; matrix pencil; linearization; strong linearization; Lagrange interpolation; Hermite interpolation; barycentric form.

1. Introduction

The original Lagrange form, first introduced by Waring (1779), has certain shortcomings, e.g., increasing the degree of the polynomial by adding a new interpolation point requires computations from scratch and also the computation is numerically unstable (see Berrut & Trefethen, 2004). Therefore, Lagrange interpolation is frequently considered as a bad choice for practical computations and thus mainly an analytic or theoretical tool for proving theorems. Nevertheless, rewriting in the so-called modified Lagrange form and the barycentric Lagrange form overcomes the shortcomings of the original form and makes Lagrange interpolation very suitable for practical use.

In this paper, we consider matrix polynomials in Lagrange and Hermite bases. Generally, for every polynomial basis, an interpolating matrix polynomial $P(\lambda)$ of degree n is uniquely determined by $n + 1$ samples of the function $A_i := P(\sigma_i)$, where $\sigma_i \in \mathbb{C}$, $i = 0, 1, \ldots, n$, are distinct interpolation points. The polynomials $P(\lambda)$ in modified or barycentric Lagrange form can be constructed very easily, since we immediately use these function values A_i in combination with the barycentric weights, which are computed from the interpolation points σ_i. Furthermore, for several point distributions we have explicit formulas for these barycentric weights. This is in contrast to Newton's interpolation, where divided differences have to be computed from the function values A_i. Also, for Chebyshev interpolation coefficient matrices have to be computed.

Polynomial eigenvalue problems (PEPs): $P(\lambda)x = 0$, where $P(\lambda)$ is a complex $s \times s$ matrix polynomial of degree n in λ and $x \in \mathbb{C}^s \backslash \{0\}$, occur in a wide number of applications, e.g., vibration analysis of buildings and machines. The classical and most common approach to solve PEPs is *linearization*, i.e., we mean the conversion of $P(\lambda)x = 0$ into a larger size linear eigenvalue problem $L(\lambda)y = (C_0 - \lambda C_1)y = 0$ with the same eigenvalues. This linear eigenvalue problem can then be solved by standard techniques.

Linearization is also a commonly used technique for solving nonlinear eigenvalue problems. In the last decade, several linearizations for different polynomial bases have been proposed in the literature. See, e.g., Jarlebring *et al.* (2012) for monomial basis; Effenberger & Kressner (2012), Jarlebring *et al.* (2012) for Chebyshev basis; and Van Beeumen *et al.* (2013) and Güttel *et al.* (2013) for Newton basis. However, to the authors' knowledge, Lagrange basis was never used. A possible explanation is that the current linearizations for matrix polynomials in Lagrange basis (Amiraslani, 2006; Amiraslani *et al.*, 2009) and also in Hermite Lagrange basis (Shakoori, 2007) contain more eigenvalues than the original polynomial $P(\lambda)$ by introducing additional eigenvalues at infinity.

Therefore, we will now propose new linearizations for the Lagrange and barycentric Lagrange polynomial which ensure a one-to-one correspondence between the eigenpairs of $P(\lambda)$ and the eigenpairs of the pencil obtained after linearization. We also generalize for Hermite interpolation and introduce new linearizations for the Hermite Lagrange and barycentric Hermite polynomial.

This paper is organized as follows. Section 2 introduces some basic definitions and notation. Section 3 reviews Lagrange interpolation and the derivation of the barycentric Lagrange form. In Section 4, we reformulate the linearization of the Lagrange and barycentric Lagrange polynomial of dimension $(n+2)s$ and introduce two new linearizations of dimension $(n+1)s$ and ns. We also prove for the last one that this is a strong linearization. In Section 5, we illustrate how the structure of the proposed pencil can be exploited in Krylov-type methods. Section 6 reviews the Hermite interpolating Lagrange and barycentric matrix polynomial. In Section 7, we generalize the linearizations of Section 4 for Hermite interpolation. Finally, the main conclusions are summarized in Section 8.

2. Definitions and notation

We study linearizations of matrix polynomials $P(\lambda)$ with $P : \mathbb{C} \to \mathbb{C}^{s\times s}$, where $P(\lambda)$ is *regular*, i.e., $\det P(\lambda)$ does not vanish identically. Linearization is the classical approach for investigating and solving PEPs. In this case, matrix polynomials are transformed into linear matrix pencils with the same eigenvalues. Therefore, *unimodular* matrix polynomials are used, i.e., matrix polynomials $E(\lambda)$ such that $\det E(\lambda)$ is a nonzero constant and independent of λ. We now introduce some basic definitions and notation in order to support the elaboration in the remaining sections.

DEFINITION 2.1 (Weak linearization, see Gohberg *et al.*, 1982) Let $P(\lambda)$ be an $s \times s$ matrix polynomial of degree n with $n \geqslant 1$. A pencil $L(\lambda) = (C_0 - \lambda C_1)$ with $C_0, C_1 \in \mathbb{C}^{ns\times ns}$ is called a linearization of $P(\lambda)$ if there exist unimodular matrix polynomials $E(\lambda)$, $F(\lambda)$ such that

$$E(\lambda)L(\lambda)F(\lambda) = \begin{bmatrix} P(\lambda) & 0 \\ 0 & I_{(n-1)s} \end{bmatrix}.$$

Thus, $L(\lambda)$ is a linearization of $P(\lambda)$ if and only if the finite eigenvalues of $L(\lambda)$, together with their partial multiplicities, coincide with those of $P(\lambda)$. Before introducing the definition of a strong linearization, we define the *extended degree* of a matrix polynomial as follows: $P(\lambda) = \sum_{i=0}^{n} A_i \lambda^i$ has *extended degree* n if the degree is $n_0 < n$, $A_n = A_{n-1} = \cdots = A_{n_0+1} = 0$, and $A_{n_0} \neq 0$. In order to study eigenvalues of $P(\lambda)$ at ∞, we introduce the definition of the reversal of a matrix polynomial.

DEFINITION 2.2 (Reversal of matrix polynomial) For a matrix polynomial $P(\lambda)$ of degree n the reversal of $P(\lambda)$ is the polynomial $P^{\#}(\lambda) := \lambda^n P(\lambda^{-1})$.

Note that the nonzero finite eigenvalues of $P^{\#}(\lambda)$ are the reciprocals of those of $P(\lambda)$ and that an eigenvalue at ∞ of $P(\lambda)$ corresponds to an eigenvalue 0 of the reversal polynomial $P^{\#}(\lambda)$.

DEFINITION 2.3 (Strong linearization; see Gohberg *et al.*, 1988) An $ns \times ns$ linear matrix pencil $C_0 - \lambda C_1$ is a strong linearization of the $s \times s$ regular matrix polynomial $P(\lambda)$ of (possibly extended) degree n if there are unimodular matrix polynomials $E(\lambda)$ and $F(\lambda)$ such that

$$\begin{bmatrix} P(\lambda) & 0 \\ 0 & I_{(n-1)s} \end{bmatrix} = E(\lambda)\,(C_0 - \lambda C_1)\,F(\lambda),$$

and there are unimodular matrix polynomials $H(\lambda)$ and $K(\lambda)$ such that

$$\begin{bmatrix} P^{\#}(\lambda) & 0 \\ 0 & I_{(n-1)s} \end{bmatrix} = H(\lambda)\,(\lambda C_0 - C_1)\,K(\lambda).$$

The following theorem gives conditions for a (strong) linearization, which will be used in the subsequent analysis. It is based on the local Smith form.

THEOREM 2.4 (Lancaster, 2008) Let $P(\lambda)$ be an $s \times s$ regular matrix polynomial of extended degree n and let $C_0 - \lambda C_1$ be an $ns \times ns$ linear matrix function. Assume that, for each distinct finite eigenvalue λ_i, there exist functions $E_i(\lambda)$ and $F_i(\lambda)$ which are unimodular and analytic in a neighbourhood of λ_i and for which

$$\begin{bmatrix} P(\lambda) & 0 \\ 0 & I_{(n-1)s} \end{bmatrix} = E_j(\lambda)\,(C_0 - \lambda C_1)\,F_j(\lambda);$$

then $C_0 - \lambda C_1$ is a (weak) linearization of $P(\lambda)$.

If $P^{\#}(\lambda)$ has an eigenvalue at zero, assume also that there are functions $E_0(\lambda)$ and $F_0(\lambda)$ which are unimodular and analytic in a neighbourhood of $\lambda = 0$ and for which

$$\begin{bmatrix} P^{\#}(\lambda) & 0 \\ 0 & I_{(n-1)s} \end{bmatrix} = E_0(\lambda)\,(\lambda C_0 - C_1)\,F_0(\lambda);$$

then $C_0 - \lambda C_1$ is a strong linearization of $P(\lambda)$.

3. Lagrange interpolation

In this section, we review the interpolating Lagrange matrix polynomial. We start with the original form, followed by the modified form and end with the barycentric form.

3.1 *Original Lagrange form*

Suppose an $s \times s$ matrix function $A(\lambda)$ is sampled at $n+1$ distinct interpolation points (nodes) σ_i, $i = 0, \ldots, n$, with corresponding values $A_i := A(\sigma_i)$. The Lagrange interpolation problem addressed here is that of finding the $s \times s$ matrix polynomial $P(\lambda)$, of degree at most n, such that P interpolates A at the points σ_i, i.e.,

$$P(\sigma_i) = A_i, \quad i = 0, \dots, n.$$

This problem is well-posed and the solution can be written in *Lagrange form* (Lagrange, 1877):

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

where the *Lagrange polynomials* $\ell_i(\lambda)$ are defined as

$$\ell_i(\lambda) = \frac{\prod_{k=0,k\neq i}^{n}(\lambda - \sigma_k)}{\prod_{k=0,k\neq i}^{n}(\sigma_i - \sigma_k)}, \quad i = 0, \dots, n, \tag{3.2}$$

with the following property at the nodes

$$\ell_i(\sigma_k) = \begin{cases} 1, & i = k, \\ 0, & \text{otherwise}, \end{cases} \quad i, k = 0, \dots, n.$$

3.2 *Modified Lagrange form*

The original Lagrange formula (3.1) can be rewritten in such a way that it can be evaluated and updated in $O(n)$ operations (see Berrut & Trefethen, 2004). Therefore, note that the numerator of $\ell_i(\lambda)$ in (3.2) can be written as the polynomial

$$\ell(\lambda) = (\lambda - \sigma_0)(\lambda - \sigma_1) \cdots (\lambda - \sigma_n) \tag{3.3}$$

divided by $\lambda - \sigma_i$. Defining the nonzero *barycentric weights* by

$$w_i = \frac{1}{\prod_{k\neq i}(\sigma_i - \sigma_k)}, \quad i = 0, \dots, n, \tag{3.4}$$

that is, $w_i = 1/\ell'(\sigma_i)$, allows us to write $\ell_i(\lambda)$ as

$$\ell_i(\lambda) = \ell(\lambda)\frac{w_i}{\lambda - \sigma_i}, \quad i = 0, \dots, n. \tag{3.5}$$

Now, note that all terms of the sum in (3.1) contain the factor $\ell(\lambda)$, which is independent of i. Bringing this factor in front of the sum yields the *modified Lagrange form* (see Berrut & Trefethen, 2004):

$$P(\lambda) = \ell(\lambda)\sum_{i=0}^{n} A_i \frac{w_i}{\lambda - \sigma_i}. \tag{3.6}$$

This modified Lagrange form (3.6) is shown to be backward stable (see Higham, 2004).

3.3 *Barycentric Lagrange form*

The modified Lagrange form (3.6) can still be modified to an even more elegant form. Therefore, we start from

$$1 = \sum_{i=0}^{n} \ell_i(\lambda) = \ell(\lambda) \sum_{i=0}^{n} \frac{w_i}{\lambda - \sigma_i}. \tag{3.7}$$

Dividing the modified Lagrange form for $P(\lambda)$ (3.6) by (3.7) and cancelling out the common factor $\ell(\lambda)$, we obtain the *barycentric form* (see Berrut & Trefethen, 2004):

$$P(\lambda) = \frac{\sum_{i=0}^{n} A_i (w_i/(\lambda - \sigma_i))}{\sum_{i=0}^{n} (w_i/(\lambda - \sigma_i))} = \sum_{i=0}^{n} A_i b_i(\lambda), \tag{3.8}$$

where

$$b_i(\lambda) = \frac{1}{b(\lambda)} \cdot \frac{w_i}{\lambda - \sigma_i}, \quad i = 0, \ldots, n, \tag{3.9}$$

with

$$b(\lambda) = \sum_{i=0}^{n} \frac{w_i}{\lambda - \sigma_i}.$$

The barycentric form is a Lagrange form, but one with a special symmetry. The weights w_i, still defined by (3.4), appear in the denominator exactly as in the numerator, except without the data factors A_i. Therefore, any common factor in all the weights w_i can be cancelled without affecting the value P.

Like the modified Lagrange form, the barycentric one also takes advantage of updating the weights w_i in $O(n)$ flops to incorporate a new data pair (σ_{n+1}, A_{n+1}). In Higham (2004), it is proved that the barycentric Lagrange interpolation form is forward stable for any set of interpolating points with a small Lebesgue constant. Finally, note that even if other weights w_i than (3.4) would be chosen in (3.8), the resulting rational function would still interpolate at the nodes σ_i in the sense that $P(\sigma_i) = A_i$.

4. Linearization of the Lagrange and barycentric Lagrange polynomial

To the authors' knowledge, a linearization in a companion pencil of the Lagrange polynomial was first introduced by Corless (2004). First in Section 4.1, we review this linearization of dimension $(n+2)s$ in a slightly different form and extend it also to the barycentric Lagrange polynomial. Next, we introduce two new and more compact linearizations of dimension $(n+1)s$ and ns in Sections 4.2 and 4.3, respectively. Note that, for the remainder, we drop the subscripts to indicate the dimensions of the identity matrices where possible.

4.1 *Linearization of dimension* $(n+2)s$

The companion pencil of the Lagrange polynomial (Corless, 2004; Amiraslani, 2006; Amiraslani *et al.*, 2009) can be extended to the barycentric Lagrange polynomial. We use a slightly different form which allows for an easy extension of the linearization matrices with one column at the right and one row at the bottom when a new data pair is added, while the other part remains unchanged. We review the arrowhead linearization by the following theorem.

THEOREM 4.1 Let $P(\lambda)$ be an $s \times s$ matrix polynomial of degree n in modified Lagrange form (3.6) or in barycentric Lagrange form (3.8). Then, the $(n+2)s \times (n+2)s$ linear companion pencil

$$L(\lambda) = C_0 - \lambda C_1,$$

where

$$C_0 = \begin{bmatrix} 0 & A_0 & A_1 & \cdots & A_n \\ w_0 I & \sigma_0 I & & & \\ w_1 I & & \sigma_1 I & & \\ \vdots & & & \ddots & \\ w_n I & & & & \sigma_n I \end{bmatrix}, \quad C_1 = \begin{bmatrix} 0 & & & & \\ & I & & & \\ & & I & & \\ & & & \ddots & \\ & & & & I \end{bmatrix} \tag{4.1}$$

is a linearization of $P(\lambda)$.

In Amiraslani *et al.* (2009), it is proved that the pencil $C_0 - \lambda C_1$, as defined in (4.1), is a strong linearization of

$$\underline{\hat{P}}(\lambda) := \lambda^{n+2}\, 0_s + \lambda^{n+1}\, 0_s + P(\lambda). \tag{4.2}$$

PROPOSITION 4.2 Suppose that $(\lambda_\star, x)$ is an eigenpair of $P(\lambda)$ and that $L(\lambda) = C_0 - \lambda C_1$ is defined by Theorem 4.1. Then, $\lambda_\star$ is also an eigenvalue of $L(\lambda)$ with the corresponding structured eigenvector $\underline{\Lambda}(\lambda_\star) \otimes x$, where

$$\underline{\Lambda}(\lambda) := \begin{bmatrix} \ell(\lambda) \\ \ell_0(\lambda) \\ \ell_1(\lambda) \\ \vdots \\ \ell_n(\lambda) \end{bmatrix}.$$

Proof. We first show that if $\lambda_\star$ is an eigenvalue of $P(\lambda)$, then $\lambda_\star$ is also an eigenvalue of $L(\lambda)$. Next, we prove that the corresponding eigenvector $\underline{\Lambda}(\lambda_\star) \otimes x \neq 0$. Following the notation of Mackey *et al.* (2006), we have, for the Lagrange polynomial (3.6),

$$(C_0 - \lambda C_1)(\underline{\Lambda}(\lambda) \otimes I) = e_1 \otimes P(\lambda), \tag{4.3}$$

where the product of the first block row of $C_0 - \lambda C_1$ with $\underline{\Lambda}(\lambda) \otimes I$ is the matrix polynomial $P(\lambda)$. The remaining products simply reproduce the relations (3.5). Evaluating (4.3) at $\lambda_\star$ and multiplying to the right by x yields

$$L(\lambda_\star) \cdot (\underline{\Lambda}(\lambda_\star) \otimes x) = 0.$$

Thus, $\lambda_\star$ is also an eigenvalue of $L(\lambda) = C_0 - \lambda C_1$ with the corresponding structured eigenvector $\underline{\Lambda}(\lambda_\star) \otimes x$. For proving that $\underline{\Lambda}(\lambda_\star) \otimes x \neq 0$, suppose first that $\lambda_\star$ is not an interpolation point, i.e., $\lambda_\star \neq \sigma_i$, $i = 0, \ldots, n$. Then, $\ell(\lambda_\star) \neq 0$ and $\ell_i(\lambda_\star) \neq 0$, $i = 0, \ldots, n$, and thus also $\underline{\Lambda}(\lambda_\star) \neq 0$. Next, suppose that $\lambda_\star$ is an interpolation point, i.e., $\lambda_\star = \sigma_k$. Then, only $\ell_k(\lambda_\star) \neq 0$ and again $\underline{\Lambda}(\lambda_\star) \neq 0$. Since $(\lambda_\star, x)$ is an eigenpair of $P(\lambda)$ this yields $\underline{\Lambda}(\lambda_\star) \otimes x \neq 0$. This completes the proof. □

4.2 *Linearization of dimension* $(n+1)s$

We now propose a new linearization for the Lagrange and barycentric Lagrange polynomial which consists of a companion pencil of dimensions $(n+1)s \times (n+1)s$ instead of $(n+2)s \times (n+2)s$ in Amiraslani *et al.* (2009). We start with the following lemma.

LEMMA 4.3 Suppose that $\ell_i(\lambda)$ and $b_i(\lambda)$ are defined by (3.5) and (3.9), respectively. Let $p_i(\lambda)$ be $\ell_i(\lambda)$ or $b_i(\lambda)$; then

$$(\lambda - \sigma_{i-1})\, p_{i-1}(\lambda) = \frac{w_{i-1}}{w_i}(\lambda - \sigma_i)\, p_i(\lambda)$$

for $i = 1, \ldots, n$.

Proof. The relations between $p_{i-1}(\lambda)$ and $p_i(\lambda)$ follow immediately from the definitions of $\ell_i(\lambda)$ and $b_i(\lambda)$. □

THEOREM 4.4 Let $P(\lambda)$ be an $s \times s$ matrix polynomial of degree n in modified Lagrange form (3.6) or in barycentric Lagrange form (3.8). Then, the $(n+1)s \times (n+1)s$ linear companion pencil

$$L(\lambda) = C_0 - \lambda C_1,$$

where

$$C_0 = \begin{bmatrix} A_0 & A_1 & A_2 & \cdots & A_n \\ \sigma_0 I & -\sigma_1\theta_1 I & & & \\ & \sigma_1 I & -\sigma_2\theta_2 I & & \\ & & \ddots & \ddots & \\ & & & \sigma_{n-1} I & -\sigma_n\theta_n I \end{bmatrix} \tag{4.4}$$

and

$$C_1 = \begin{bmatrix} 0 & & & & \\ I & -\theta_1 I & & & \\ & I & -\theta_2 I & & \\ & & \ddots & \ddots & \\ & & & I & -\theta_n I \end{bmatrix}, \tag{4.5}$$

with $\theta_i = w_{i-1}/w_i$ for $i = 1, \ldots, n$ is a linearization of $P(\lambda)$.

PROPOSITION 4.5 Let C_0 and C_1 be defined by (4.4) and (4.5), respectively. Then, $C_0 - \lambda C_1$ is a strong linearization of

$$\hat{P}(\lambda) := \lambda^{n+1}\, 0_s + P(\lambda). \tag{4.6}$$

We will not prove this proposition, since the proof is similar to the one in Amiraslani *et al.* (2009) and the proof of the linearization in the next section.

PROPOSITION 4.6 Suppose that $(\lambda_\star, x)$ is an eigenpair of $P(\lambda)$ and that $L(\lambda) = C_0 - \lambda C_1$ is defined by Theorem 4.4. Then, $\lambda_\star$ is also an eigenvalue of $L(\lambda)$ with the corresponding structured eigenvector $\Lambda(\lambda_\star) \otimes x$, where

$$\Lambda(\lambda) := \begin{bmatrix} \ell_0(\lambda) \\ \ell_1(\lambda) \\ \vdots \\ \ell_n(\lambda) \end{bmatrix}.$$

Proof. We first show that if $\lambda_\star$ is an eigenvalue of $P(\lambda)$, then $\lambda_\star$ is also an eigenvalue of $L(\lambda)$. Next, we prove that the corresponding eigenvector $\Lambda(\lambda_\star) \otimes x \neq 0$. For the Lagrange polynomial (3.6), we have

$$(C_0 - \lambda C_1)(\Lambda(\lambda) \otimes I) = e_1 \otimes P(\lambda), \tag{4.7}$$

where the product of the first block row of $C_0 - \lambda C_1$ with $\Lambda(\lambda) \otimes I$ is the matrix polynomial $P(\lambda)$. The remaining products simply result in the relations of Lemma 4.3. Evaluating (4.7) at $\lambda_\star$ and multiplying to the right by x yields

$$L(\lambda_\star) \cdot (\Lambda(\lambda_\star) \otimes x) = 0.$$

Thus, $\lambda_\star$ is also an eigenvalue of $L(\lambda) = C_0 - \lambda C_1$ with the corresponding structured eigenvector $\Lambda(\lambda_\star) \otimes x$. For proving that $\Lambda(\lambda_\star) \otimes x \neq 0$, suppose first that $\lambda_\star$ is not an interpolation point, i.e., $\lambda_\star \neq \sigma_i$, $i = 0, \ldots, n$. Then, $\ell_i(\lambda_\star) \neq 0$, $i = 0, \ldots, n$ and thus also $\Lambda(\lambda_\star) \neq 0$. Next, suppose that $\lambda_\star$ is an interpolation point, i.e., $\lambda_\star = \sigma_k$. Then, only $\ell_k(\lambda_\star) \neq 0$ and again $\Lambda(\lambda_\star) \neq 0$. Since $(\lambda_\star, x)$ is an eigenpair of $P(\lambda)$ this yields $\Lambda(\lambda_\star) \otimes x \neq 0$. This completes the proof. □

4.3 *Linearization of dimension ns*

The linearizations from Sections 4.1 and 4.2 are not linearizations of $P(\lambda)$ but of $\underline{\hat{P}}(\lambda)$ (4.2) and $\hat{P}(\lambda)$ (4.6), respectively. Consequently, they contain, besides all the eigenvalues of $P(\lambda)$, also extra eigenvalues at infinity. Here, we introduce a new linearization of dimension *ns* which results in a one-to-one mapping between the eigenstructure of the original matrix polynomial $P(\lambda)$ and the pencil $C_0 - \lambda C_1$, corresponding to both finite eigenvalues and the eigenvalue at infinity.

We start by defining

$$\tilde{\ell}_i(\lambda) := -\frac{\ell_i(\lambda)}{\lambda - \sigma_{i+1}} = -\ell(\lambda)\frac{w_i}{(\lambda - \sigma_i)(\lambda - \sigma_{i+1})}, \quad i = 0, \ldots, n-1. \tag{4.8}$$

Next, using (4.8) for $i = n-1$, we can rewrite $\ell_n(\lambda)$ as follows:

$$\ell_n(\lambda) = \frac{w_n}{w_{n-1}}(\sigma_{n-1} - \lambda)\, \tilde{\ell}_{n-1}(\lambda). \tag{4.9}$$

Then, combining (4.8) and (4.9) yields

$$\begin{aligned}
P(\lambda) &= \sum_{i=0}^{n} A_i \ell_i(\lambda), \\
&= \sum_{i=0}^{n-1} A_i \ell_i(\lambda) + A_n \ell_n(\lambda), \\
&= \sum_{i=0}^{n-1} A_i (\sigma_{i+1} - \lambda)\, \tilde{\ell}_i(\lambda) + A_n \frac{w_n}{w_{n-1}} (\sigma_{n-1} - \lambda)\, \tilde{\ell}_{n-1}(\lambda), \\
&= \sum_{i=0}^{n-2} A_i (\sigma_{i+1} - \lambda)\, \tilde{\ell}_i(\lambda) + \left[A_{n-1}(\sigma_n - \lambda) + A_n \frac{w_n}{w_{n-1}} (\sigma_{n-1} - \lambda) \right] \tilde{\ell}_{n-1}(\lambda).
\end{aligned}$$

Similarly, for the barycentric Lagrange polynomial we define

$$\tilde{b}_i(\lambda) := -\frac{b_i(\lambda)}{\lambda - \sigma_{i+1}} = -\frac{1}{b(\lambda)} \cdot \frac{w_i}{(\lambda - \sigma_i)(\lambda - \sigma_{i+1})}, \quad i = 0, \ldots, n-1, \tag{4.10}$$

and using (4.10) for $i = n-1$, we rewrite $b_n(\lambda)$ as follows:

$$b_n(\lambda) = \frac{w_n}{w_{n-1}}(\sigma_{n-1} - \lambda)\,\tilde{b}_{n-1}(\lambda). \tag{4.11}$$

Combining (4.10) and (4.11) results in

$$P(\lambda) = \sum_{i=0}^{n-2} A_i(\sigma_{i+1} - \lambda)\,\tilde{b}_i(\lambda) + \left[A_{n-1}(\sigma_n - \lambda) + A_n \frac{w_n}{w_{n-1}}(\sigma_{n-1} - \lambda)\right]\tilde{b}_{n-1}(\lambda).$$

Before presenting the linearization, we formulate in the following lemma the relations between successive $\tilde{\ell}_i(\lambda)$ and $\tilde{b}_i(\lambda)$, respectively.

LEMMA 4.7 Suppose $\tilde{\ell}_i(\lambda)$ and $\tilde{b}_i(\lambda)$ are defined by (4.8) and (4.10), respectively. Let $\tilde{p}_i(\lambda)$ be $\tilde{\ell}_i(\lambda)$ or $\tilde{b}_i(\lambda)$; then

$$(\lambda - \sigma_{i-1})\,\tilde{p}_{i-1}(\lambda) = \frac{w_{i-1}}{w_i}(\lambda - \sigma_{i+1})\,\tilde{p}_i(\lambda)$$

for $i = 1, \ldots, n-1$.

Proof. The relations between $\tilde{p}_{i-1}(\lambda)$ and $\tilde{p}_i(\lambda)$ follow immediately from the definitions of $\tilde{\ell}_i(\lambda)$ and $\tilde{b}_i(\lambda)$. □

THEOREM 4.8 Let $P(\lambda)$ be an $s \times s$ matrix polynomial of degree n in modified Lagrange form (3.6) or in barycentric Lagrange form (3.8). Then, the $ns \times ns$ linear companion pencil

$$L(\lambda) = C_0 - \lambda C_1,$$

where

$$C_0 = \begin{bmatrix} \sigma_1 A_0 & \sigma_2 A_1 & \cdots & \sigma_{n-1}A_{n-2} & \sigma_n A_{n-1} + \sigma_{n-1}\theta_n^{-1}A_n \\ \sigma_0 I & -\sigma_2\theta_1 I & & & \\ & \ddots & \ddots & & \\ & & \sigma_{n-3}I & -\sigma_{n-1}\theta_{n-2}I & \\ & & & \sigma_{n-2}I & -\sigma_n\theta_{n-1}I \end{bmatrix} \tag{4.12}$$

and

$$C_1 = \begin{bmatrix} A_0 & A_1 & \cdots & A_{n-2} & A_{n-1} + \theta_n^{-1}A_n \\ I & -\theta_1 I & & & \\ & \ddots & \ddots & & \\ & & I & -\theta_{n-2}I & \\ & & & I & -\theta_{n-1}I \end{bmatrix}, \tag{4.13}$$

with $\theta_i = w_{i-1}/w_i$ for $i = 1, \ldots, n$, is a strong linearization of $P(\lambda)$.

Proof. The proof consists of three parts and is similar to the one in Amiraslani *et al.* (2009). Weak linearization is established in parts (a) and (b). Part (a) concerns eigenvalues of $P(\lambda)$ which are not equal to an interpolation point σ_i for any i. Part (b) concerns eigenvalues, which happen to coincide with an interpolation point, and completes the proof of the weak linearization property. Part (c) shows, based on Theorem 2.4, that the linearization is strong.

Part (a). We first introduce the $ns \times ns$ block permutation matrix

$$S := \begin{bmatrix} 0 & I & & & \\ 0 & 0 & I & & \\ \vdots & \vdots & \ddots & \ddots & \\ 0 & 0 & \cdots & 0 & I \\ I & 0 & \cdots & 0 & 0 \end{bmatrix},$$

and note that $C_0 - \lambda C_1$ is a strong linearization if and only if the same is true for $S(C_0 - \lambda C_1)$. Define the λ-dependent block LU decomposition of $S(C_0 - \lambda C_1) = L(\lambda)U(\lambda)$, where

$$L(\lambda) = \begin{bmatrix} I & & & \\ & \ddots & & \\ & & I & \\ L_{n,1}(\lambda) & \cdots & L_{n,n-1}(\lambda) & I \end{bmatrix}, \tag{4.14}$$

with

$$L_{n,i}(\lambda) = \sum_{k=0}^{i-1} \frac{w_k}{w_{i-1}} \frac{\lambda - \sigma_i}{\lambda - \sigma_k} A_k, \quad i = 1, \ldots, n-1 \tag{4.15}$$

and

$$U(\lambda) = \begin{bmatrix} -(\lambda - \sigma_0)I & (\lambda - \sigma_2)\theta_1 I & & \\ & \ddots & \ddots & \\ & & -(\lambda - \sigma_{n-2})I & (\lambda - \sigma_n)\theta_{n-1} I \\ & & & U_{n,n}(\lambda) \end{bmatrix}, \tag{4.16}$$

with

$$U_{n,n}(\lambda) = -\frac{P(\lambda)}{w_{n-1}(\lambda - \sigma_0)\cdots(\lambda - \sigma_{n-2})}.$$

Note that $L(\lambda)$ is well-defined and nonsingular for all $\lambda \neq \sigma_i$ and $\det L(\lambda) \equiv \pm 1$. However, $U(\lambda)$ is singular at the eigenvalues of $P(\lambda)$ and since we supposed $\lambda \neq \sigma_i$ in this part of the proof, all these eigenvalues are associated with $U_{n,n}(\lambda)$. Therefore, we define $\tilde{U}(\lambda)$ to be the same as $U(\lambda)$ except for the last block entry which is replaced by

$$\tilde{U}_{n,n}(\lambda) = -\frac{I}{w_{n-1}(\lambda - \sigma_0)\cdots(\lambda - \sigma_{n-2})};$$

then we also have $\det \tilde{U}(\lambda)$ is a nonzero constant and

$$S(C_0 - \lambda C_1) = L(\lambda)U(\lambda) = L(\lambda) \begin{bmatrix} I_{(n-1)s} & 0 \\ 0 & P(\lambda) \end{bmatrix} \tilde{U}(\lambda).$$

Thus, it follows that

$$\begin{bmatrix} I_{(n-1)s} & 0 \\ 0 & P(\lambda) \end{bmatrix} = E(\lambda)\,(C_0 - \lambda C_1)\,F(\lambda),$$

where $E(\lambda) := L^{-1}(\lambda)S$ and $F(\lambda) := \tilde{U}^{-1}(\lambda)$ are unimodular. For completeness, we now give the explicit forms of $E(\lambda)$ and $F(\lambda)$, respectively:

$$E(\lambda) = \begin{bmatrix} 0 & I & & & \\ & 0 & I & & \\ & & \ddots & \ddots & \\ & & & 0 & I \\ I & -L_{n,1}(\lambda) & -L_{n,2}(\lambda) & \cdots & -L_{n,n-1}(\lambda) \end{bmatrix}, \tag{4.17}$$

$$F_{i,j}(\lambda) = \begin{cases} -\dfrac{1}{\lambda - \sigma_{i-1}} I, & i = 1, \dots, n-1 \\ & j = i, \\ -\dfrac{w_{i-1}(\lambda - \sigma_j)}{w_{j-1}(\lambda - \sigma_{i-1})(\lambda - \sigma_i)} I, & i = 1, \dots, n-2, \\ & j = i+1, \dots, n-1, \\ -w_{i-1} \dfrac{\prod_{k=0}^{n}(\lambda - \sigma_k)}{(\lambda - \sigma_{i-1})(\lambda - \sigma_i)} I, & i = 1, \dots, n, \\ & j = n, \\ 0, & \text{otherwise}, \end{cases} \tag{4.18}$$

where $L_{n,i}(\lambda)$, $i = 1, \dots, n-1$, is defined by (4.15).

Part (*b*). By using the construction of equivalence transformations which are well-defined everywhere except at the nodes, part (a) of the proof shows that the partial multiplicities of all finite eigenvalues of $P(\lambda)$, with the possible exception of an eigenvalue at an interpolation point σ_i, $i = 0, \dots, n$, are reproduced in $C_0 - \lambda C_1$. Now, suppose that σ_i is an eigenvalue of $P(\lambda)$ and also an interpolation point. Without loss of generality, we can reorder the nodes so that this node becomes σ_n.

We will show that the partial multiplicities of σ_n in $P(\lambda)$ and $C_0 - \lambda C_1$ are the same. Therefore, we return to the λ-dependent block LU decomposition of $S(C_0 - \lambda C_1) = L(\lambda)U(\lambda)$ and, once again, we define $E(\lambda)$ and $F(\lambda)$ as in (4.17) and (4.18), respectively. Now, we observe that they are unimodular and analytic in some neighbourhood of σ_n. Hence, the partial multiplicities of an eigenvalue, which is also an interpolation point, are the same for $P(\lambda)$ and $C_0 - \lambda C_1$. Consequently, all finite eigenvalues of $P(\lambda)$ reappear in $C_0 - \lambda C_1$, together with their partial multiplicities. Hence, together with part (a), this concludes the proof for $C_0 - \lambda C_1$ to be a weak linearization of $P(\lambda)$.

Part (*c*). In order to prove the linearization is strong, we consider the reverse polynomial $P^{\#}(\lambda)$ and using Theorem 2.4, we need to show that there exist matrix functions $H(\lambda)$ and $K(\lambda)$ which are unimodular and analytic on a neighbourhood of $\lambda = 0$ and for which

$$\begin{bmatrix} I_{(n-1)s} & 0 \\ 0 & P^{\#}(\lambda) \end{bmatrix} = H(\lambda)\,(\lambda C_0 - C_1)\,K(\lambda).$$

First, we return to the LU decomposition of $S(C_0 - \lambda C_1)$ and apply the transformation $\lambda \to \lambda^{-1}$, which yields

$$S(\lambda C_0 - C_1) = \lambda L(\lambda^{-1})U(\lambda^{-1}).$$

Thus, we obtain the LU factors for the reverse pencil: $S(\lambda C_0 - C_1) = L_{\text{rev}}(\lambda)U_{\text{rev}}(\lambda)$, where

$$L_{\text{rev}}(\lambda) := L(\lambda^{-1}), \quad U_{\text{rev}}(\lambda) := \lambda U(\lambda^{-1}),$$

with $L(\lambda)$ and $U(\lambda)$ defined by (4.14) and (4.16), respectively. Note that $L_{\text{rev}}(\lambda)$ is well-defined and nonsingular for all $\lambda \neq 1/\sigma_i$ and $\det L_{\text{rev}}(\lambda) \equiv \pm 1$. Similarly to part (a), we define $\tilde{U}_{\text{rev}}(\lambda)$ to be the same as $U_{\text{rev}}(\lambda)$ except for the last block entry which is replaced by

$$\tilde{U}_{\text{rev}\ n,n}(\lambda) = -\frac{I}{w_{n-1}(1-\lambda\sigma_0)\cdots(1-\lambda\sigma_{n-2})}.$$

Thus, $\det \tilde{U}_{\text{rev}}(\lambda)$ is a nonzero constant and

$$S(\lambda C_0 - C_1) = L_{\text{rev}}(\lambda)U_{\text{rev}}(\lambda) = L_{\text{rev}}(\lambda)\begin{bmatrix} I_{(n-1)s} & 0 \\ 0 & P^{\#}(\lambda) \end{bmatrix}\tilde{U}_{\text{rev}}(\lambda),$$

and we define $H(\lambda) := [L_{\text{rev}}(\lambda)]^{-1}S$ and $K(\lambda) := [\tilde{U}_{\text{rev}}(\lambda)]^{-1}$. Note that $H(\lambda) = E(\lambda^{-1})$ and

$$K_{i,j}(\lambda) = \begin{cases} -\dfrac{1}{1-\lambda\sigma_{i-1}}I, & i = 1,\ldots,n-1, \\ & j = i, \\ -\dfrac{w_{i-1}(1-\lambda\sigma_j)}{w_{j-1}(1-\lambda\sigma_{i-1})(1-\lambda\sigma_i)}I, & i = 1,\ldots,n-2, \\ & j = i+1,\ldots,n-1, \\ -w_{i-1}\dfrac{\prod_{k=0}^{n}(1-\lambda\sigma_k)}{(1-\lambda\sigma_{i-1})(1-\lambda\sigma_i)}I, & i = 1,\ldots,n, \\ & j = n, \\ 0, & \text{otherwise.} \end{cases} \tag{4.19}$$

In order to examine the behaviour of $S(\lambda C_0 - C_1)$ near 0, we consider the properties of $H(0)$ and $K(0)$. Since $\lim_{\lambda\to\infty} L_{n,i}(\lambda)$ exists, with $L_{n,i}(\lambda)$ as defined by (4.15), it follows from (4.14) that $\lim_{\lambda\to 0} L(\lambda^{-1})$ exists too and hence $\lim_{\lambda\to 0} L^{-1}(\lambda^{-1})$ exists. By definition, it follows that $H(\lambda)$ is unimodular and analytic and invertible at $\lambda = 0$. From (4.19), we observe that $\lim_{\lambda\to 0} K(\lambda)$ is a constant upper-triangular matrix with nonzero diagonal entries. Consequently, $K(\lambda)$ is also unimodular and analytic and invertible at $\lambda = 0$. This completes the proof. □

PROPOSITION 4.9 Suppose that $(\lambda_\star, x)$ is an eigenpair of $P(\lambda)$ and that $L(\lambda) = C_0 - \lambda C_1$ is defined by Theorem 4.8. Then, $\lambda_\star$ is also an eigenvalue of $L(\lambda)$ with the corresponding structured eigenvector $\tilde{\Lambda}(\lambda_\star) \otimes x$, where

$$\tilde{\Lambda}(\lambda) := \begin{bmatrix} \tilde{\ell}_0(\lambda) \\ \tilde{\ell}_1(\lambda) \\ \vdots \\ \tilde{\ell}_{n-1}(\lambda) \end{bmatrix}.$$

Proof. We first show that if $\lambda_\star$ is an eigenvalue of $P(\lambda)$; then $\lambda_\star$ is also an eigenvalue of $L(\lambda)$. Next, we prove that the corresponding eigenvector $\tilde{\Lambda}(\lambda_\star) \otimes x \neq 0$. For the Lagrange polynomial (3.6) we have

$$(C_0 - \lambda C_1)(\tilde{\Lambda}(\lambda) \otimes I) = e_1 \otimes P(\lambda), \tag{4.20}$$

where the product of the first block row of $C_0 - \lambda C_1$ with $\tilde{\Lambda}(\lambda) \otimes I$ is the matrix polynomial $P(\lambda)$. The remaining products simply reproduce the relations of Lemma 4.7. Evaluating (4.20) at $\lambda_\star$ and multiplying to the right by x yields

$$L(\lambda_\star) \cdot (\tilde{\Lambda}(\lambda_\star) \otimes x) = 0.$$

Thus, $\lambda_\star$ is also an eigenvalue of $L(\lambda) = C_0 - \lambda C_1$ with the corresponding structured eigenvector $\tilde{\Lambda}(\lambda_\star) \otimes x$. For proving that $\tilde{\Lambda}(\lambda_\star) \otimes x \neq 0$, suppose first that $\lambda_\star$ is not an interpolation point, i.e., $\lambda_\star \neq \sigma_i, i = 0, \ldots, n$. Then, $\tilde{\ell}_i(\lambda_\star) \neq 0, i = 0, \ldots, n-1$ and thus also $\tilde{\Lambda}(\lambda_\star) \neq 0$. Next, suppose that $\lambda_\star$ is an interpolation point, i.e., $\lambda_\star = \sigma_k, k = 0, \ldots, n-1$. Then, only $\tilde{\ell}_k(\lambda_\star) \neq 0$ and $\tilde{\ell}_{k+1}(\lambda_\star) \neq 0$, and again $\tilde{\Lambda}(\lambda_\star) \neq 0$. Finally, suppose that $\lambda_\star$ is the last interpolation point, $\lambda_\star = \sigma_n$. Then, only $\tilde{\ell}_{n-1}(\lambda_\star) \neq 0$ and again $\tilde{\Lambda}(\lambda_\star) \neq 0$. Since $(\lambda_\star, x)$ is an eigenpair of $P(\lambda)$ this yields $\tilde{\Lambda}(\lambda_\star) \otimes x \neq 0$. This completes the proof. □

Note that the arrowhead linearization of size $(n+2)s$ is the only linearization which treats all nodes equally since the second till the last block rows of the linearization matrices C_0 and C_1 correspond to the relations between $\ell(\lambda)$ and $\ell_i(\lambda)$ for $i = 0, 1, \ldots, n$. This kind of *'symmetry'* is lost in the linearizations of size $(n+1)s$ and ns, where the second till the last block rows express, respectively, the relations between $\ell_0(\lambda)$ and $\ell_1(\lambda)$, $\ell_1(\lambda)$ and $\ell_2(\lambda)$, ... This implies some ordering and some *'graph'*, where an edge denotes a relation between the functions $\ell_i(\lambda)$ which appear in the linearization. We could, for instance, also express a connection between $\ell_0(\lambda)$ and $\ell_n(\lambda)$, $\ell_1(\lambda)$ and $\ell_n(\lambda)$, ... This corresponds to a star-shaped graph with centre $\ell_n(\lambda)$.

As a consequence, the linearization proposed in Theorem 4.8 exhibits an asymmetry due to a special role of the interpolation point σ_n (which can be freely chosen). We believe this asymmetry is hard to avoid and the price to pay for getting a linearization of minimal dimensions. Note that a similar special role of the interpolation point σ_n is used in the Newton basis for obtaining a strong linearization (Amiraslani *et al.*, 2009, Section 3.3).

5. Exploitation of pencil structure

The companion-type matrices of the pencil $L(\lambda) = C_0 - \lambda C_1$ from Theorem 4.8 are of dimension ns. Thus, as a consequence of linearization, the problem dimension is multiplied by n. However, in Krylov-type methods we can exploit the structure of C_0 and C_1 (4.12–4.13) such that we only have to deal with matrices of the original polynomial dimension s.

THEOREM 5.1 Let C_0 and C_1 be defined by (4.12) and (4.13), respectively. Then, the linear system

$$(C_0 - \lambda C_1)x = y, \tag{5.1}$$

with $\lambda \in \mathbb{C}$ can be efficiently solved by using only n matrix–vector products of dimension s and one linear system solve of dimension s.

Proof. Let

$$x := [x_1^* \quad x_2^* \quad \cdots \quad x_n^*]^* \quad \text{and} \quad y := [y_1^* \quad y_2^* \quad \cdots \quad y_n^*]^*,$$

where $x_i, y_i \in \mathbb{C}^s$ for $i = 1, \ldots, n$. Then, the first block row of (5.1) results in

$$\sum_{k=1}^{n} (\sigma_k - \lambda) A_{k-1} x_k + \frac{w_n}{w_{n-1}} (\sigma_{n-1} - \lambda) A_n x_n = y_1, \tag{5.2}$$

and the next block rows can be written as

$$x_i = \frac{w_{i-1}}{w_{i-2}} \cdot \frac{1}{\lambda - \sigma_i} y_i + \frac{w_{i-1}}{w_{i-2}} \cdot \frac{\lambda - \sigma_{i-2}}{\lambda - \sigma_i} x_{i-1}, \qquad i = 2, \ldots, n. \tag{5.3}$$

Now, substituting the relations in (5.3) into (5.2) yields

$$-\frac{1}{w_0} (\lambda - \sigma_0)(\lambda - \sigma_1) \left(\sum_{k=0}^{n} \frac{w_k}{\lambda - \sigma_k} A_k \right) x_1 = y_1 + \sum_{i=1}^{n} A_i \sum_{j=1}^{i} \frac{w_i}{w_{j-1}} \cdot \frac{\lambda - \sigma_j}{\lambda - \sigma_i} y_{j+1},$$

with $y_{n+1} := 0$ and which can be rewritten as follows:

$$P(\lambda) x_1 = \frac{-w_0 \ell(\lambda)}{(\lambda - \sigma_0)(\lambda - \sigma_1)} \left(y_1 + \sum_{i=1}^{n} A_i \sum_{j=1}^{i} \frac{w_i}{w_{j-1}} \cdot \frac{\lambda - \sigma_j}{\lambda - \sigma_i} y_{j+1} \right) \tag{5.4}$$

or

$$P(\lambda) x_1 = \frac{-w_0}{(\lambda - \sigma_0)(\lambda - \sigma_1) b(\lambda)} \left(y_1 + \sum_{i=1}^{n} A_i \sum_{j=1}^{i} \frac{w_i}{w_{j-1}} \cdot \frac{\lambda - \sigma_j}{\lambda - \sigma_i} y_{j+1} \right). \tag{5.5}$$

Note that, taking into account the definition of $\ell(\lambda)$ and $b(\lambda)$, the right-hand sides of (5.4) and (5.5), respectively, have polynomial dependence on λ. Thus, from (5.4) or (5.5) we can compute x_1 with only one linear system solve with $P(\lambda)$ and n matrix–vector products for computing the right-hand side of (5.4) or (5.5). Next, $x_2, \ldots, x_n$ can be computed from (5.3), which completes the proof. □

Note that, in case λ is equal to one of the interpolation points σ_i, the linear system (5.1) has the following sparsity pattern:

$$\left[\begin{array}{cccc|c|c|cccc}
\star & \star & \cdots & \star & 0 & \star & \star & \star & \cdots & \star \\ \hline
\star & \star & & & & & & & & \\
& \ddots & \ddots & & & & & & & \\
& & \star & \star & & & & & & \\
& & & \star & 0 & & & & & \\ \hline
& & & & \star & \star & & & & \\ \cline{5-10}
& & & & & 0 & \star & & & \\
& & & & & & \star & \star & & \\
& & & & & & & \ddots & \ddots & \\
& & & & & & & & \star & \star
\end{array}\right]
\left[\begin{array}{c}
x_1 \\ x_2 \\ \vdots \\ x_{i-1} \\ \hline x_i \\ \hline x_{i+1} \\ \hline x_{i+2} \\ x_{i+3} \\ \vdots \\ x_n
\end{array}\right]
=
\left[\begin{array}{c}
y_1 \\ \hline y_2 \\ \vdots \\ y_{i-1} \\ y_i \\ \hline y_{i+1} \\ \hline y_{i+2} \\ y_{i+3} \\ \vdots \\ y_n
\end{array}\right],$$

which yields a decoupling. The system can now be solved by using forward and backward substitution. Again, we only need n matrix–vector products and one linear system solve, both of dimension s.

6. Hermite interpolation

In the previous sections, we considered interpolating matrix polynomials in distinct points. From here on, we also allow Hermite interpolation. We start with reviewing the Hermite interpolating Lagrange and barycentric matrix polynomial. The next section discusses its corresponding linearizations.

6.1 *Lagrange Hermite form*

We still suppose that σ_i, $i = 0, \ldots, n$, are $n+1$ distinct interpolation points, but now with corresponding multiplicities m_i, with

$$m_0 + \cdots + m_n = N + 1,$$

where N is the degree of the corresponding interpolating polynomial $P(\lambda)$. The Lagrange form can now be generalized to Hermite interpolation by

$$P(\lambda) = \sum_{i=0}^{n} \sum_{j=0}^{m_i-1} \frac{A_i^{(j)}}{j!} \ell_{i,j}(\lambda),$$

where $A_i^{(j)} := A^{(j)}(\sigma_i)$ denotes the jth derivative of A evaluated at σ_i and

$$\ell_{i,j}(\lambda) = \ell(\lambda) \sum_{k=j}^{m_i-1} \frac{w_{i,k}}{(\lambda - \sigma_i)^{k-j+1}} \tag{6.1}$$

is the generalization of (3.5) for Hermite interpolation with

$$\ell(\lambda) = (\lambda - \sigma_0)^{m_0} (\lambda - \sigma_1)^{m_1} \cdots (\lambda - \sigma_n)^{m_n},$$

the generalization of (3.3). The constants $w_{i,j}$ are called the *generalized barycentric weights*. For the computation of these $w_{i,k}$ we refer the reader to Butcher *et al.* (2011) and Sadiq & Viswanath (2013). Similar to (3.6), we can bring the factor $\ell(\lambda)$ in front of the sums, yielding

$$P(\lambda) = \ell(\lambda) \sum_{i=0}^{n} \sum_{j=0}^{m_i-1} \frac{A_i^{(j)}}{j!} \sum_{k=j}^{m_i-1} \frac{w_{i,k}}{(\lambda - \sigma_i)^{k-j+1}}. \tag{6.2}$$

6.2 *Barycentric Hermite form*

The barycentric interpolating matrix polynomial for Hermite interpolation can be obtained in a similar way as in Section 3.3. Again, we start from

$$1 = \ell(\lambda) \sum_{i=0}^{n} \sum_{j=0}^{m_i-1} \frac{w_{i,j}}{(\lambda - \sigma_i)^{j+1}}. \tag{6.3}$$

Dividing the Lagrange Hermite form (6.2) by (6.3) and cancelling out the common factor $\ell(\lambda)$, we obtain the *barycentric Hermite form* (see Schneider & Werner, 1991):

$$P(\lambda) = \frac{\sum_{i=0}^{n} \sum_{j=0}^{m_i-1} (A_i^{(j)}/j!) \sum_{k=j}^{m_i-1} (w_{i,k}/(\lambda - \sigma_i)^{k-j+1})}{\sum_{i=0}^{n} \sum_{j=0}^{m_i-1} (w_{i,j}/(\lambda - \sigma_i)^{j+1})} = \sum_{i=0}^{n} \sum_{j=0}^{m_i-1} \frac{A_i^{(j)}}{j!} b_{i,j}(\lambda), \tag{6.4}$$

where

$$b_{i,j}(\lambda) = \frac{1}{b(\lambda)} \sum_{k=j}^{m_i-1} \frac{w_{i,k}}{(\lambda - \sigma_i)^{k-j+1}}, \quad i = 0, \ldots, n, \quad j = 0, \ldots, m_i - 1, \tag{6.5}$$

with

$$b(\lambda) = \sum_{i=0}^{n} \sum_{j=0}^{m_i-1} \frac{w_{i,j}}{(\lambda - \sigma_i)^{j+1}}.$$

7. Linearization of the Lagrange and barycentric Hermite polynomial

The linearization of the Lagrange polynomial by Amiraslani (2006) was generalized for Hermite interpolation by Shakoori (2007). We review this linearization for the Hermite Lagrange and barycentric Hermite matrix polynomial. Next, we generalize our new linearization of Section 4.3 for which there is again a one-to-one correspondence between the eigenvalues of the original matrix polynomial $P(\lambda)$ and the ones of the companion pencil $C_0 - \lambda C_1$.

7.1 *Linearization of dimension* $(N+2)s$

The companion pencil of the barycentric Hermite polynomial was introduced in Shakoori (2007) and Corless *et al.* (2008). Here, we review this linearization for matrix polynomials in a similar form as in Theorem 4.1 by the following theorem.

THEOREM 7.1 Let $P(\lambda)$ be an $s \times s$ matrix polynomial of degree N in Lagrange Hermite form (6.2) or in barycentric Hermite form (6.4). Then, the $(N+2)s \times (N+2)s$ linear companion pencil

$$L(\lambda) = \mathbf{C}_0 - \lambda \mathbf{C}_1,$$

where

$$\mathbf{C}_0 = \begin{bmatrix} 0 & \mathbf{A}_0 & \mathbf{A}_1 & \cdots & \mathbf{A}_n \\ \mathbf{W}_0 & \mathbf{J}_0 & & & \\ \mathbf{W}_1 & & \mathbf{J}_1 & & \\ \vdots & & & \ddots & \\ \mathbf{W}_n & & & & \mathbf{J}_n \end{bmatrix}, \quad \mathbf{C}_1 = \begin{bmatrix} 0 & & & & \\ & I & & & \\ & & I & & \\ & & & \ddots & \\ & & & & I \end{bmatrix}, \tag{7.1}$$

with

$$\mathbf{A}_i = \begin{bmatrix} \frac{A_i^{(0)}}{0!} & \frac{A_i^{(1)}}{1!} & \cdots & \frac{A_i^{(m_i-1)}}{(m_i-1)!} \end{bmatrix} \in \mathbb{C}^{s \times m_i s}, \quad i = 0, \ldots, n, \tag{7.2}$$

$$\mathbf{W}_i = \begin{bmatrix} w_{i,0}I \\ w_{i,1}I \\ \vdots \\ w_{i,m_i-1}I \end{bmatrix} \in \mathbb{C}^{m_i s \times s}, \quad i = 0, \ldots, n, \tag{7.3}$$

$$\mathbf{J}_i = \begin{bmatrix} \sigma_i I & I & & \\ & \ddots & \ddots & \\ & & \sigma_i I & I \\ & & & \sigma_i I \end{bmatrix} \in \mathbb{C}^{m_i s \times m_i s}, \quad i = 0, \ldots, n \tag{7.4}$$

is a linearization of $P(\lambda)$.

PROPOSITION 7.2 Suppose that $(\lambda_\star, x)$ is an eigenpair of $P(\lambda)$ and that $L(\lambda) = \mathbf{C}_0 - \lambda\mathbf{C}_1$ is defined by Theorem 7.1. Then, $\lambda_\star$ is also an eigenvalue of $L(\lambda)$ with the corresponding structured eigenvector $\underline{\Lambda}(\lambda_\star) \otimes x$, where

$$\underline{\Lambda}(\lambda) := \begin{bmatrix} \ell(\lambda) \\ \boldsymbol{\ell}_0(\lambda) \\ \boldsymbol{\ell}_1(\lambda) \\ \vdots \\ \boldsymbol{\ell}_n(\lambda) \end{bmatrix}, \quad \text{with } \boldsymbol{\ell}_i(\lambda) := \begin{bmatrix} \ell_{i,0}(\lambda) \\ \ell_{i,1}(\lambda) \\ \vdots \\ \ell_{i,m_i-1}(\lambda) \end{bmatrix}, \quad i = 0, \ldots, n.$$

Proof. We first show that if $\lambda_\star$ is an eigenvalue of $P(\lambda)$, then $\lambda_\star$ is also an eigenvalue of $L(\lambda)$. Next, we prove that the corresponding eigenvector $\underline{\Lambda}(\lambda_\star) \otimes x \neq 0$. Following the notation of Mackey *et al.* (2006), we have, for the Hermite Lagrange form (6.2),

$$(\mathbf{C}_0 - \lambda\mathbf{C}_1)(\underline{\Lambda}(\lambda) \otimes I) = e_1 \otimes P(\lambda), \tag{7.5}$$

where the product of the first s rows of $\mathbf{C}_0 - \lambda\mathbf{C}_1$ with $\underline{\Lambda}(\lambda) \otimes I$ is the matrix polynomial $P(\lambda)$ and the remaining products simply reproduce the relations (6.1). Evaluating (7.5) at $\lambda_\star$ and multiplying to the right by x yields

$$L(\lambda_\star) \cdot (\underline{\Lambda}(\lambda_\star) \otimes x) = 0.$$

Thus, $\lambda_\star$ is also an eigenvalue of $L(\lambda) = \mathbf{C}_0 - \lambda\mathbf{C}_1$ with the corresponding structured eigenvector $\underline{\Lambda}(\lambda_\star) \otimes x$. For proving that $\underline{\Lambda}(\lambda_\star) \otimes x \neq 0$, suppose first that $\lambda_\star$ is not an interpolation point, i.e., $\lambda_\star \neq \sigma_i$, $i = 0, \ldots, n$. Then, $\ell(\lambda_\star) \neq 0$ and $\ell_{i,j}(\lambda_\star) \neq 0$, $i = 0, \ldots, n$, $j = 0, \ldots, m_i - 1$ and thus also $\underline{\Lambda}(\lambda_\star) \neq 0$. Next, suppose that $\lambda_\star$ is an interpolation point, i.e., $\lambda_\star = \sigma_k$. Then, only $\ell_{k,m_k-1}(\lambda_\star) \neq 0$ and again $\underline{\Lambda}(\lambda_\star) \neq 0$. Since $(\lambda_\star, x)$ is an eigenpair of $P(\lambda)$ this yields $\underline{\Lambda}(\lambda_\star) \otimes x \neq 0$. This completes the proof. □

7.2 *Linearization of dimension Ns*

We now propose a new linearization for the Hermite Lagrange and barycentric Hermite polynomial which consists of a companion pencil of dimensions $Ns \times Ns$ instead of $(N+2)s \times (N+2)s$.

Similarly to Section 4.3, we use $\tilde{p}_{i,j}(\lambda)$ to denote

$$\tilde{p}_{i,j}(\lambda) := -\frac{p_{i,j}(\lambda)}{\lambda - \sigma_{i+1}}, \qquad \begin{aligned} i &= 0, \dots, n-1, \\ j &= 0, \dots, m_i - 1, \end{aligned} \tag{7.6}$$

$$\tilde{p}_{n,j}(\lambda) := -\frac{p_{n,j}(\lambda)}{\lambda - \sigma_{n-1}}, \qquad j = 0, \dots, m_i - 2, \tag{7.7}$$

where $p_{i,j}(\lambda)$ is $\ell_{i,j}(\lambda)$ or $b_{i,j}(\lambda)$. Next, using (7.7) we can rewrite $p_{n,m_n-1}(\lambda)$ as follows:

$$p_{n,m_n-1}(\lambda) = \frac{w_{n,m_n-1}}{w_{n-1,m_{n-1}-1}} (\sigma_{n-1} - \lambda)\, \tilde{p}_{n-1,m_{n-1}-1}(\lambda). \tag{7.8}$$

Then, combining (7.7) and (7.8) yields

$$\begin{aligned}
P(\lambda) &= \sum_{i=0}^{n} \sum_{j=0}^{m_i-1} \frac{A_i^{(j)}}{j!}\, p_{i,j}(\lambda), \\
&= \sum_{i=0}^{n-1} \sum_{j=0}^{m_i-1} \frac{A_i^{(j)}}{j!}\, p_{i,j}(\lambda) + \sum_{j=0}^{m_n-2} \frac{A_n^{(j)}}{j!}\, p_{n,j}(\lambda) + \frac{A_n^{(m_n-1)}}{(m_n-1)!}\, p_{n,m_n-1}(\lambda), \\
&= \sum_{i=0}^{n-1} \sum_{j=0}^{m_i-1} \frac{A_i^{(j)}}{j!} (\sigma_{i+1} - \lambda)\, \tilde{p}_{i,j}(\lambda) + \sum_{j=0}^{m_n-2} \frac{A_n^{(j)}}{j!} (\sigma_{n-1} - \lambda)\, \tilde{p}_{n,j}(\lambda) \\
&\quad + \frac{A_n^{(m_n-1)}}{(m_n-1)!} \frac{w_{n,m_n-1}}{w_{n-1,m_{n-1}-1}} (\sigma_{n-1} - \lambda)\, \tilde{p}_{n-1,m_{n-1}-1}(\lambda), \\
&= \sum_{i=0}^{n-2} \sum_{j=0}^{m_i-1} \frac{A_i^{(j)}}{j!} (\sigma_{i+1} - \lambda)\, \tilde{p}_{i,j}(\lambda) + \sum_{j=0}^{m_{n-1}-2} \frac{A_{n-1}^{(j)}}{j!} (\sigma_n - \lambda)\, \tilde{p}_{n-1,j}(\lambda) \\
&\quad + \left[\frac{A_{n-1}^{(m_{n-1}-1)}}{(m_{n-1}-1)!} (\sigma_n - \lambda) + \frac{A_n^{(m_n-1)}}{(m_n-1)!} \frac{w_{n,m_n-1}}{w_{n-1,m_{n-1}-1}} (\sigma_{n-1} - \lambda) \right] \tilde{p}_{n-1,m_{n-1}-1}(\lambda), \\
&\quad + \sum_{j=0}^{m_n-2} \frac{A_n^{(j)}}{j!} (\sigma_{n-1} - \lambda)\, \tilde{p}_{n,j}(\lambda).
\end{aligned}$$

Before presenting the linearization, we formulate the relations between successive $\tilde{\ell}_{i,j}(\lambda)$ and $\tilde{b}_{i,j}(\lambda)$, respectively, in the following lemma.

Lemma 7.3 Suppose that $\tilde{p}_{i,j}(\lambda)$ is $\tilde{\ell}_{i,j}(\lambda)$ or $\tilde{b}_{i,j}(\lambda)$, defined by (7.7). Then, we have the following relations:

$$(\lambda - \sigma_i)\, \tilde{p}_{i,j}(\lambda) = \tilde{p}_{i,j+1}(\lambda) + \frac{w_{i,j}}{w_{i,m_i-1}} (\lambda - \sigma_i)\, \tilde{p}_{i,m_i-1}(\lambda)$$

for $i = 0, \dots, n-1$, $j = 0, \dots, m_i - 2$, and

$$(\lambda - \sigma_{i-1})\, \tilde{p}_{i-1,m_{i-1}-1}(\lambda) = \frac{w_{i-1,m_{i-1}-1}}{w_{i,m_i-1}} (\lambda - \sigma_{i+1})\, \tilde{p}_{i,m_i-1}(\lambda)$$

for $i = 1, \ldots, n-1$. We also have

$$(\lambda - \sigma_n)\,\tilde{p}_{n,j}(\lambda) = \tilde{p}_{n,j+1}(\lambda) + \frac{w_{n,j}}{w_{n-1,m_{n-1}-1}}(\lambda - \sigma_n)\,\tilde{p}_{n-1,m_{n-1}-1}(\lambda)$$

for $j = 0, \ldots, m_i - 3$, and

$$(\lambda - \sigma_n)\,\tilde{p}_{n,m_n-2}(\lambda) = \left(\frac{w_{n,m_n-1}}{w_{n-1,m_{n-1}-1}} + \frac{w_{n,m_n-2}}{w_{n-1,m_{n-1}-1}}(\lambda - \sigma_n)\right)\tilde{p}_{n-1,m_{n-1}-1}(\lambda).$$

Proof. These relations follow immediately from the definitions (7.7) of $\tilde{\ell}_{i,j}(\lambda)$ and $\tilde{b}_{i,j}(\lambda)$ and the relation (7.8). □

In a similar manner as in Section 4.3, the relations between $\tilde{\ell}_{i,j}(\lambda)$ and $\tilde{b}_{i,j}(\lambda)$ of Lemma 7.3 can now be used to construct a linearization of dimensions $Ns \times Ns$ for the Hermite interpolating Lagrange and barycentric matrix polynomial $P(\lambda)$.

THEOREM 7.4 Let $P(\lambda)$ be an $s \times s$ matrix polynomial of degree N in Lagrange Hermite form (6.2) or in barycentric Hermite form (6.4). Then, the $Ns \times Ns$ linear companion pencil

$$L(\lambda) = \mathbf{C}_0 - \lambda\mathbf{C}_1,$$

where

$$\mathbf{C}_0 = \left[\begin{array}{c|c|c|c|c|c}
\begin{matrix}\sigma_1\mathbf{A}_0\\ \Theta_0\end{matrix} & \begin{matrix}\sigma_2\mathbf{A}_1\\ 0\end{matrix} & \begin{matrix}\cdots\\ \cdots\end{matrix} & \begin{matrix}\sigma_{n-1}\mathbf{A}_{n-2}\\ 0\end{matrix} & \begin{matrix}\sigma_n\mathbf{A}_{n-1}+\sigma_{n-1}\tilde{\mathbf{A}}_{n-1}\\ 0\end{matrix} & \begin{matrix}\sigma_{n-1}\tilde{\mathbf{A}}_n\\ 0\end{matrix}\\ \hline
\begin{matrix}\sigma_0\Gamma_0\\ 0\end{matrix} & \begin{matrix}-\sigma_1\Pi_1\\ \Theta_1\end{matrix} & & & & \\ \hline
 & \begin{matrix}\sigma_1\Gamma_1\\ 0\end{matrix} & \begin{matrix}-\sigma_2\Pi_2\\ \Theta_2\end{matrix} & & & \\ \hline
 & & \ddots & \ddots & & \\ \hline
 & & & \begin{matrix}\sigma_{n-2}\Gamma_{n-2}\\ 0\end{matrix} & \begin{matrix}-\sigma_{n-1}\Pi_{n-1}\\ \Theta_{n-1}\end{matrix} & \\ \hline
 & & & & \tilde{\Theta}_{n-1} & \tilde{\Theta}_n
\end{array}\right], \qquad (7.9)$$

$$\mathbf{C}_1 = \left[\begin{array}{c|c|c|c|c|c}
\begin{matrix}\mathbf{A}_0\\ \Delta_0\end{matrix} & \begin{matrix}\mathbf{A}_1\\ 0\end{matrix} & \begin{matrix}\cdots\\ \cdots\end{matrix} & \begin{matrix}\mathbf{A}_{n-2}\\ 0\end{matrix} & \begin{matrix}\mathbf{A}_{n-1}+\tilde{\mathbf{A}}_{n-1}\\ 0\end{matrix} & \begin{matrix}\tilde{\mathbf{A}}_n\\ 0\end{matrix}\\ \hline
\begin{matrix}\Gamma_0\\ 0\end{matrix} & \begin{matrix}-\Pi_1\\ \Delta_1\end{matrix} & & & & \\ \hline
 & \begin{matrix}\Gamma_0\\ 0\end{matrix} & \begin{matrix}-\Pi_1\\ \Delta_1\end{matrix} & & & \\ \hline
 & & \ddots & \ddots & & \\ \hline
 & & & \begin{matrix}\Gamma_{n-2}\\ 0\end{matrix} & \begin{matrix}-\Pi_{n-1}\\ \Delta_{n-1}\end{matrix} & \\ \hline
 & & & & \tilde{\Delta}_{n-1} & \tilde{\Delta}_n
\end{array}\right], \qquad (7.10)$$

with $\mathbf{A}_i$, $i = 0, 1, \ldots, n-1$ as defined by (7.2) and

$$
\begin{aligned}
\mathbf{A}_i &= \begin{bmatrix} \dfrac{A_i^{(0)}}{0!} & \dfrac{A_i^{(1)}}{1!} & \cdots & \dfrac{A_i^{(m_i-1)}}{(m_i-1)!} \end{bmatrix}, \quad i = 0, \ldots, n-1, \\
\Pi_i &= \begin{bmatrix} 0 & \cdots & 0 & \dfrac{w_{i-1,m_{i-1}-1}}{w_{i,m_i-1}} I \end{bmatrix}, \quad i = 1, \ldots, n-1, \\
\Gamma_i &= \begin{bmatrix} 0 & \cdots & 0 & I \end{bmatrix}, \quad i = 0, \ldots, n-2, \\
\Theta_i &= \begin{bmatrix} \sigma_i I & I & & & -\sigma_i \dfrac{w_{i,0}}{w_{i,m_i-1}} I \\ & \ddots & \ddots & & \vdots \\ & & \sigma_i I & I & -\sigma_i \dfrac{w_{i,m_i-3}}{w_{i,m_i-1}} I \\ & & & \sigma_i I & \left(1 - \sigma_i \dfrac{w_{i,m_i-2}}{w_{i,m_i-1}}\right) I \end{bmatrix}, \quad i = 0, \ldots, n-1, \\
\Delta_i &= \begin{bmatrix} I & & & -\dfrac{w_{i,0}}{w_{i,m_i-1}} I \\ & \ddots & & \vdots \\ & & I & -\dfrac{w_{i,m_i-3}}{w_{i,m_i-1}} I \\ & & & I \;\; -\dfrac{w_{i,m_i-2}}{w_{i,m_i-1}} I \end{bmatrix}, \quad i = 0, \ldots, n-1,
\end{aligned}
$$

where $\mathbf{A}_i, \Gamma_i, \Pi_i \in \mathbb{C}^{s \times m_i s}$ and $\Theta_i, \Delta_i \in \mathbb{C}^{(m_i-1)s \times m_i s}$ and

$$
\begin{aligned}
\tilde{\mathbf{A}}_{n-1} &= \begin{bmatrix} 0 & \cdots & 0 & \dfrac{w_{n,m_n-1}}{w_{n-1,m_{n-1}-1}} \dfrac{A_n^{(m_n-1)}}{(m_n-1)!} \end{bmatrix} \in \mathbb{C}^{s \times m_{n-1} s}, \\
\tilde{\mathbf{A}}_n &= \begin{bmatrix} \dfrac{A_n^{(0)}}{0!} & \dfrac{A_n^{(1)}}{1!} & \cdots & \dfrac{A_n^{(m_n-2)}}{(m_n-2)!} \end{bmatrix} \in \mathbb{C}^{s \times (m_n-1)s}, \\
\tilde{\Theta}_{n-1} &= \begin{bmatrix} 0 & & & -\sigma_n \dfrac{w_{n,0}}{w_{n-1,m_{n-1}-1}} I \\ & \ddots & & \vdots \\ & & 0 & -\sigma_n \dfrac{w_{n,m_n-3}}{w_{n-1,m_{n-1}-1}} I \\ & & & 0 \;\; \left(\dfrac{w_{n,m_n-1}}{w_{n-1,m_{n-1}-1}} - \sigma_n \dfrac{w_{n,m_n-2}}{w_{n-1,m_{n-1}-1}} \right) I \end{bmatrix} \in \mathbb{C}^{(m_n-1)s \times m_n s},
\end{aligned}
$$

$$\tilde{\Theta}_n = \begin{bmatrix} \sigma_n I & I & & \\ & \ddots & \ddots & \\ & & \sigma_n I & I \\ & & & \sigma_n I \end{bmatrix} \in \mathbb{C}^{(m_n-1)s \times (m_n-1)s},$$

$$\tilde{\Delta}_{n-1} = \begin{bmatrix} 0 & & & -\frac{w_{n,0}}{w_{n-1,m_{n-1}-1}} I \\ & \ddots & & \vdots \\ & & 0 & -\frac{w_{n,m_n-3}}{w_{n-1,m_{n-1}-1}} I \\ & & & 0 \quad -\frac{w_{n,m_n-2}}{w_{n-1,m_{n-1}-1}} I \end{bmatrix} \in \mathbb{C}^{(m_n-1)s \times m_n s},$$

$$\tilde{\Delta}_n = I \in \mathbb{C}^{(m_n-1)s \times (m_i-1)s}$$

is a strong linearization of $P(\lambda)$.

Proof. The proof is similar to that of Theorem 4.8. □

PROPOSITION 7.5 Suppose that $(\lambda_\star, x)$ is an eigenpair of $P(\lambda)$ and that $L(\lambda) = \mathbf{C}_0 - \lambda \mathbf{C}_1$ is defined by Theorem 7.4. Then, $\lambda_\star$ is also an eigenvalue of $L(\lambda)$ with the corresponding structured eigenvector $\tilde{\Lambda}(\lambda_\star) \otimes x$, where

$$\tilde{\Lambda}(\lambda) = \begin{bmatrix} \tilde{\mathbf{l}}_0(\lambda) \\ \tilde{\mathbf{l}}_1(\lambda) \\ \vdots \\ \tilde{\mathbf{l}}_n(\lambda) \end{bmatrix}, \quad \text{with } \tilde{\mathbf{l}}_i(\lambda) = \begin{bmatrix} \tilde{\ell}_{i,0}(\lambda) \\ \tilde{\ell}_{i,1}(\lambda) \\ \vdots \\ \tilde{\ell}_{i,k_i}(\lambda) \end{bmatrix},$$

for $i = 0, 1, \ldots, n$ with $k_i = m_i - 1$ for $i = 0, 1, \ldots, n-1$ and $k_n = m_n - 2$.

Proof. The proof is similar to that of Proposition 4.9. □

The companion-type matrices of the pencil $L(\lambda) = \mathbf{C}_0 - \lambda \mathbf{C}_1$ from Theorem 7.4 are of dimension Ns. But similar as in Section 5, the structure of $\mathbf{C}_0$ and $\mathbf{C}_1$ (7.9–7.10) can be exploited in Krylov-type methods such that we only have to deal with matrices of the original polynomial dimension s.

8. Conclusions

In this paper, we introduced two new and more compact linearizations for interpolating Lagrange and barycentric Lagrange matrix polynomials $P(\lambda)$. For the proposed linearization of dimension ns there is a one-to-one correspondence between the eigenpairs of $P(\lambda)$ and the eigenpairs of the pencil such that no extra eigenvalues at infinity are introduced any more. We proved that this linearization is strong. Moreover, the structure of the companion-type matrices can be exploited such that in Krylov-type methods only n matrix–vector products and one linear system solve, both of dimension s, are required. We also generalized for Hermite interpolation and introduced new linearizations for Hermite Lagrange and barycentric Hermite matrix polynomials.

We have implemented the linearizations and the exploitation of the pencil structure in Matlab. The codes can be downloaded from http://twr.cs.kuleuven.be/research/software/nleps/lin-lagr.html.

Funding

This work has been supported by the Programme of Interuniversity Attraction Poles of the Belgian Federal Science Policy Office (IAP P6-DYSCO); by OPTEC, the Optimization in Engineering Center of the KU Leuven; by the projects STRT1-09/33 and OT/10/038 of the KU Leuven Research Council; and the project G.0712.11N of the Research Foundation-Flanders (FWO).

References

AMIRASLANI, A. (2006) New algorithms for matrices, polynomials and matrix polynomials. *Ph.D. Thesis*, University of Western Ontario, London, Ontario.

AMIRASLANI, A., CORLESS, R. M. & LANCASTER, P. (2009) Linearization of matrix polynomials expressed in polynomial bases. *IMA J. Numer. Anal.*, **29**, 141–157.

BERRUT, J.-P. & TREFETHEN, L. N. (2004) Barycentric Lagrange interpolation. *SIAM Rev.*, **46**, 501–517.

BUTCHER, J., CORLESS, R. M., GONZALEZ-VEGA, L. & SHAKOORI, A. (2011) Polynomial algebra for Birkhoff interpolants. *Numer. Algorithms*, **56**, 319–347.

CORLESS, R. M. (2004) Generalized companion matrices in the Lagrange basis. *Proceedings EACA* (L. González Vega & T. Recio eds). Santander, Spain: Universidad de Cantabria, pp. 317–322.

CORLESS, R. M., SHAKOORI, A., ARULIAH, D. A. & GONZALEZ-VEGA, L. (2008) Barycentric Hermite interpolants for event location in initial-value problems. *J. Numer. Anal. Ind. Appl. Math.*, **3**, 1–16.

EFFENBERGER, C. & KRESSNER, D. (2012) Chebyshev interpolation for nonlinear eigenvalue problems. *BIT*, **52**, 933–951.

GOHBERG, I., KAASHOEK, M. A. & LANCASTER, P. (1988) General theory of regular matrix polynomials and band Toeplitz operators. *Integral Equations Operator Theory*, **11**, 776–882.

GOHBERG, I., LANCASTER, P. & RODMAN, L. (1982) *Matrix Polynomials*. New York: Academic Press, p. 409.

GÜTTEL, S., VAN BEEUMEN, R., MEERBERGEN, K. & MICHIELS, W. (2013) NLEIGS: A class of robust fully rational Krylov methods for nonlinear eigenvalue problems. *Technical Report TW633*. Department of Computer Science, KU Leuven.

HIGHAM, N. J. (2004) The numerical stability of barycentric Lagrange interpolation. *IMA J. Numer. Anal.*, **24**, 547–556.

JARLEBRING, E., MICHIELS, W. & MEERBERGEN, K. (2012) A linear eigenvalue algorithm for the nonlinear eigenvalue problem. *Numer. Math.*, **122**, 169–195.

LAGRANGE, J. L. (1877) Leçons élémentaires sur les mathématiques données à l'École Normale en 1795. *Oeuvres de Lagrange VII*. Paris: Gauthier-Villars, pp. 183–287.

LANCASTER, P. (2008) Linearization of regular matrix polynomials. *Electron. J. Linear Algebra*, **17**, 21–27.

MACKEY, D. S., MACKEY, N., MEHL, C. & MEHRMANN, V. (2006) Vector spaces of linearizations for matrix polynomials. *SIAM J. Matrix Anal. Appl.*, **28**, 971–1004.

SADIQ, B. & VISWANATH, D. (2013) Barycentric Hermite interpolation. *SIAM J. Sci. Comput.*, **35**, A1254–A1270.

SCHNEIDER, C. & WERNER, W. (1991) Hermite interpolation: the barycentric approach. *Computing*, **46**, 35–51.

SHAKOORI, A. (2007) Bivariate Polynomial Solver by Values. *Ph.D. Thesis*, University of Western Ontario, London, Ontario.

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**, A327–A350.

WARING, E. (1779) Problems concerning interpolations. *Phil. Trans. R. Soc. Lond.*, **69**, 59–67.