A sparse linearization for Hermite interpolation matrix polynomials

Heike, Fassbender and Javier, Perez Alvaro and Nikta, Shayanfar

2015

MIMS EPrint: **2015.98**

ISSN 1749-9097

A sparse linearization for Hermite interpolation matrix polynomials[☆]

Heike Faßbender[a], Javier Pérez[b], Nikta Shayanfar[a,*]

[a] *Institut* Computational Mathematics*, AG Numerik, Technische Universität Braunschweig, 38092 Braunschweig, Germany*

[b] *School of Mathematics, The University of Manchester, Manchester, England, M13 9PL*

Abstract

The polynomial eigenvalue problem for Hermite interpolation matrix polynomials is discussed. The standard approach to solve a polynomial eigenvalue problem is via linearization. In this work we introduce a new linearization for Hermite interpolation matrix polynomials expressed in the first barycentric form that is more sparse than the ones known so far. In addition, we show that this linearization is a strong linearization, and that eigenvectors of the polynomial and those of the linearization are related in simple ways. Finally, the backward errors of computed eigenpairs of the original and the linearized problem are compared as well as eigenvalue condition numbers.

Keywords: polynomial eigenvalue problem, linearization, eigenvalue problem, matrix polynomial, Hermite basis, interpolation polynomial
2000 MSC: 65F15, 15A18

---

1. Introduction

The simplest but still most important among nonlinear eigenvalue problems are the polynomial eigenvalue problems (PEP). Given a *regular* $m \times m$ matrix polynomial $P(z)$, the associated PEP consists in finding a scalar λ and nonzero vectors x and y

[☆] The work of second author was partially supported by the Engineering and Physical Sciences Research Council grant EP/I005293, and Alexander von Humboldt Foundation has funded the work of third author.

[*] Corresponding author

Email addresses: `h.fassbender@tu-braunschweig.de` (Heike Faßbender), `javier.perezalvaro@manchester.ac.uk` (Javier Pérez), `n.shayanfar@tu-braunschweig.de` (Nikta Shayanfar)

such that

$$P(\lambda)x = 0, \quad \text{and} \quad y^*P(\lambda) = 0,$$

where the vectors x and y are called, respectively, right and left eigenvectors of $P(z)$ associated with the eigenvalue λ. By regular we mean that the determinant of $P(z)$ is not identically equal to the zero polynomial. Polynomial eigenproblems arise directly from applications, from finite element discretizations of continuous models in mechanics, control theory, computer-aided graphic design and differential algebraic equations [3] or as approximations of nonlinear eigenvalue problems [33]. We will assume throughout the work the regularity of the matrix polynomial $P(z)$.

The standard form to express a matrix polynomial $P(z)$ of *grade* d is

$$P(z) = P_d z^d + \cdots + P_1 z + P_0, \qquad P_0, P_1, \ldots, P_d \in \mathbb{C}^{m\times m}, \tag{1}$$

that is, expanding $P(z)$ in the monomial basis $\{1, z, \ldots, z^d\}$, where we allow any of the coefficients matrices P_i with $i = 0, \ldots, d$ to be the zero matrix. It is worth mentioning that the *degree* of the matrix polynomial (1) is referred to the maximum integer k such that P_k is a nonzero matrix. In other words, a polynomial of degree k can be considered as a polynomial of grade higher than or equal to k. Throughout this paper, when the grade of a polynomial is not explicitly stated, we consider its grade is the same as the degree of the polynomial.

One of the main objective for regular matrix polynomials $P(z)$ is to determine the eigenvalues, as well as their multiplicities (see [10, Definition 2.8] for the precise definition of the multiplicity of an eigenvalue of a matrix polynomial), and the associated eigenvectors of $P(z)$. The finite eigenvalues of $P(z)$ are the roots of the scalar polynomial $\det P(z)$, while its infinite eigenvalues are the zero eigenvalues of the *reversal polynomial* of $P(z)$ defined as $\mathrm{rev}P(z) := z^d P(z^{-1})$. Notice that the definition of the reversal polynomial depends on the choice for the grade of $P(z)$, which should be clear from the context. For more about the finite and infinite eigenstructure of a matrix polynomial we refer the reader to [10] and the references therein.

The standard computational approach for solving polynomial eigenvalue problems is via linearizations. A *linearization* is a matrix pencil $L(z) = zC_1 - C_0$ such that there exist unimodular (i.e. with nonzero constant determinant) matrix polynomials $E(z)$ and $F(z)$ satisfying

$$E(z)L(z)F(z) = \begin{bmatrix} I_{(d-1)m} & \\ & P(z) \end{bmatrix}.$$

In addition, if $\mathrm{rev}L(z) = C_1 - zC_0$ is a linearization of $\mathrm{rev}P(z)$, then $L(z)$ is said to be a *strong linearization* of $P(z)$. The key property of any strong linearization of

$L(z)$ of a regular matrix polynomial $P(z)$ is that it preserves the finite and infinite eigenstructure (eigenvalues and their multiplicities) of $P(z)$. Then, the eigenstructure of $P(z)$ may be computed using any of the well-known algorithms for solving the linear eigenvalue problem.

In practice, when the polynomial $P(z)$ is expressed in the monomial basis as in (1), the most used linearization to solve a PEP is the *Frobenius companion form*

$$L(z) = z \begin{bmatrix} P_d & & & \\ & I_m & & \\ & & \ddots & \\ & & & I_m \end{bmatrix} + \begin{bmatrix} P_{d-1} & P_{d-2} & \cdots & P_0 \\ -I_m & 0_m & & \\ & \ddots & \ddots & \\ & & -I_m & 0_m \end{bmatrix},$$

where 0_m denotes the $m \times m$ zero matrix. 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 $P(z)$ (e.g., for $P_j = P_j^T \in \mathbb{R}^{m\times m}$ we have $P(z)^T = P(z)$, but $L(z)^T \neq L(z)$). 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 [9, 24, 25, 26, 35].

When the polynomial $P(z)$ is expressed in the monomial basis many linearization are available in the literature [9, 24]. However, it is becoming of interest to solve PEPs for polynomials expressed in nonmonomial polynomial bases (see for example [21, 11, 25, 26, 35]). In many such cases it is advisable to avoid reformulating $P(z)$ in monomial basis, since this change of basis can be poorly conditioned, and may introduce numerical errors. Moreover, the instability increases with the degree [15]. Hence, constructing linearizations of matrix polynomials from the coefficients of $P(z)$ in the given basis has become an active topic of research. Several linearizations for different polynomial bases have been proposed in [1], and particularly linearization in the Chebyshev basis [13, 19, 26, 28], Bernstein basis [20, 25, 37], Newton basis [14, 34], Lagrange basis [5, 6, 12, 35], and Hermite basis [35] has been investigated.

In this paper, we consider matrix polynomials obtained from Hermite interpolation problems. Generally, an interpolating matrix polynomial of degree d can be uniquely determined by $d + 1$ samples of the polynomial and its derivatives. More precisely, for the sample nodes $z_0, \ldots, z_n$, the Hermite interpolation polynomial $P(z)$ is a polynomial of degree d which matches the predetermined values $P(z_i)$ of the polynomial and its first $s_i - 1$ predetermined derivatives at the point z_i, where $\sum_{i=0}^{n} s_i = d + 1$. In this work, we will construct a new strong linearization with a sparse structure for the Hermite interpolation polynomial. Besides, we will show how eigenvectors of $P(z)$ are related with those of the new linearization.

The rest of the paper is organized as follows. In Section 2, we review the Hermite interpolation problem as well as the first barycentric form to express the Hermite interpolation polynomial. A new strong linearization for the first barycentric form of the Hermite interpolation matrix polynomial is introduced in Section 3. Moreover, a similar linearization proposed in [30] (see also [7, 35]) is shown to be strong as well. In Section 4 we show that eigenvectors of the Hermite interpolation polynomial are related in a simply fashion to those of the proposed linearization. Furthermore, we show in Section 5 how one may obtain an a posteriori upper bound of the forward error of the solution of a PEP problem solved via the presented linearization. Finally, in Section 6 we draw some conclusions.

2. The Hermite interpolation problem

This section gives a brief description of the Hermite interpolation problem. We will follow [2, 27, 29], where the scalar and matrix polynomial cases are discussed.

Consider a sufficiently smooth matrix function $F : \mathbb{C} \to \mathbb{C}^{m\times m}$. Assume that the points $z_0, \dots, z_n$ are some known interpolation nodes in $\mathbb{C}$, and that at the node z_i the value of F and its derivatives up to order $s_i - 1$ are available. The jth derivative of $F(z)$ at the node z_i, $i = 0, \dots, n$, is denoted by $F_{i,j}$, where $j = 0, \dots, s_i - 1$. The total number of samples of $F(z)$ is

$$d + 1 = \sum_{i=0}^{n} s_i.$$

Then, the Hermite interpolation problem consists in constructing an $m \times m$ matrix polynomial $P(z)$ of degree less than or equal to d such that

$$\frac{d^j P(z)}{dz^j}|_{z=z_i} = F_{i,j}, \qquad i = 0, \dots, n,\ j = 0, \dots, s_i - 1. \tag{2}$$

The polynomial $P(z)$ solving this problem is called the *Hermite interpolation polynomial*, and its existence and uniqueness are proved in [8]. Hermite interpolation polynomials can be represented in several different basis, e.g., in Lagrange [2, 27] or Hermite basis [8]. Here we will concentrate on the barycentric Hermite interpolation, to be precise, on the first barycentric form of the Hermite interpolation polynomial. This form allows for an easy update of the interpolation polynomial in case additional information needs to be added, while many other interpolation polynomials require a complete new construction of the interpolation polynomial. Please note, that there is also a second barycentric form of the Hermite interpolation polynomial [27, 29].

With the polynomial $\omega(z)$ of degree $d+1$

$$\omega(z) = \prod_{i=0}^{n} (z - z_i)^{s_i}, \tag{3}$$

the first barycentric form of the Hermite interpolation polynomial $P(z)$ is given by

$$P(z) = \omega(z) \sum_{i=0}^{n} \sum_{j=0}^{s_i-1} \frac{F_{i,j}}{j!} \sum_{k=0}^{s_i-j-1} \frac{w_{i,k}}{(z - z_i)^{s_i-j-k}}, \tag{4}$$

where $w_{i,k}$ are the so called *generalized barycentric weights* (see for example [27]). To the authors knowledge, there does not exist an explicit closed formula for computing generalized barycentric weights. However, algorithmic approaches to their computation are presented in [2, 27, 29]. The algorithm in [29] uses confluent divided differences and although it is fast, it is an unstable algorithm [30]. In [2], the barycentric weights are determined using divided differences in $\mathcal{O}(n^2)$ operations using contour integration and the manipulation of infinite series. A more efficient method for the computation of the barycentric weights is suggested in [27]. This approach has good numerical stability even when derivatives of high order are involved; for a detailed discussion see [27]. Its main advantage is the possibility of updating the barycentric coefficients using only $\mathcal{O}(n)$ operations. Thus we proceed with the algorithm of computing generalized barycentric form which has been presented in [27]. First, we define the sequence of $\{\mathcal{I}_r\}_{r\geq 1}$ as follows:

$$\mathcal{I}_0 := 1$$

$$k\mathcal{I}_k := \sum_{j=0}^{k-1} \mathcal{I}_j \mathcal{P}_{k-j}, \quad k \geqslant 1$$

where

$$\mathcal{P}_r := \sum_{j\neq k} \frac{s_j}{(z_j - z_k)^r}, \qquad r \geq 1.$$

Then, the generalized barycentric weights are

$$w_{i,k} = w_i \mathcal{I}_k, \tag{5}$$

where w_i, for $i = 0, \ldots, n$, is the ith barycentric weight given by

$$w_i := \prod_{\substack{k=0,\\ k\neq i}}^{n} \frac{1}{(z_i - z_k)^{s_k}}. \tag{6}$$

It is evident from inspection that $P(z)$, as represented in (4), is a polynomial of grade d. Moreover, choosing the generalized barycentric weights as in (5), the polynomial $P(z)$ satisfies the interpolation conditions (2). For the interested reader, we refer to [27], where it is shown how to update the weights $w_{i,k}$ when a new node z_{n+1} or an additional derivative F_{i,s_i} for some i is added to the interpolation problem.

3. A sparse linearization for the Hermite interpolation polynomial

In this section we introduce a new strong linearization of the Hermite interpolation polynomial. One of the main drawbacks of solving a PEP via a linearization is the increase of the size of the problem. For this reason, it is of fundamental importance to construct linearizations with a structure that may be exploited by eigenvalue algorithms to reduce the cost of the computation. As we will see at the end of this section, the proposed strong linearization has such structure.

We start by rewriting the barycentric form (4) of the Hermite interpolation polynomial.

Proposition 3.1. *Let $P(z)$ be the $m \times m$ Hermite interpolation polynomial expressed in the first barycentric form (4). Then, $P(z)$ can be rewritten as*

$$P(z) = \omega(z) \sum_{i=0}^{n} \sum_{j=0}^{s_i-1} \frac{M_{i,j} w_i}{(z - z_i)^{s_i - j}}, \tag{7}$$

where $\omega(z)$ and w_i are defined in (3) *and* (6)*, respectively, and*

$$M_{i,j} := \sum_{k=0}^{j} \frac{F_{i,k}}{k!} \mathcal{I}_{j-k} \in \mathbb{C}^{m \times m}. \tag{8}$$

Proof. The polynomial presentation of the Hermite interpolant (4) can be written in terms of different powers of $(z - z_i), i = 0, \ldots, n$ as follows:

$$P(z) = \omega(z) \sum_{i=0}^{n} \sum_{j=0}^{s_i-1} \sum_{k=0}^{j} \frac{F_{i,k}}{k!} w_{i,j-k} \frac{1}{(z - z_i)^{s_i - j}}.$$

Considering (5) and defining (8) complete the proof. □

We refer to the interpolation polynomial of the Proposition 3.1 as the *modified barycentric form*. It is the basis of our construction of a sparse linearization $L(z) = zC_1 - C_0$ of the Hermite interpolation polynomial $P(z)$. To this purpose, let $P(z)$

be the Hermite interpolation polynomial expressed in the modified barycentric form (7). Then, let us define the matrices C_1 and C_0 of size $m(d+2) \times m(d+2)$ as

$$C_1 = \operatorname{diag}(\underbrace{I_m, \dots, I_m}_{(d+1) \text{ times}}, 0_m), \tag{9}$$

and

$$C_0 = \begin{bmatrix} J_0 & & & & M_0 \\ & J_1 & & & M_1 \\ & & \ddots & & \vdots \\ & & & J_n & M_n \\ -W_0 & -W_1 & \dots & -W_n & 0_m \end{bmatrix}, \tag{10}$$

where

$$J_i = \begin{bmatrix} z_i & 1 & & \\ & z_i & 1 & \\ & & \ddots & 1 \\ & & & z_i \end{bmatrix} \otimes I_m \in \mathbb{C}^{ms_i \times ms_i}, \qquad i = 0, \dots, n,$$

$$W_i = \begin{bmatrix} w_i & 0 & \cdots & 0 \end{bmatrix} \otimes I_m \in \mathbb{C}^{m \times ms_i} \qquad i = 0, \dots, n,$$

with the weight w_i as in (6), and

$$M_i = \begin{bmatrix} M_{i,s_i-1} & \cdots & M_{i,1} & M_{i,0} \end{bmatrix}^T \in \mathbb{C}^{ms_i \times m} \qquad i = 0, \dots, n,$$

where the matrices $M_{i,j}$ are defined in (8). Here, $\otimes$ denotes the standard Kronecker product of two matrices.

To prove that $zC_1 - C_0$ ((9)-(10)) is indeed a strong linearization of the Hermite interpolation polynomial $P(z)$, we will make use of [22, Theorem 3.2], that we restate here as a lemma.

Lemma 3.2. *Let $P(z)$ be an $m \times m$ matrix polynomial of grade d and let $zC_1 - C_0$ be a $dm \times dm$ matrix pencil. Assume that, for each distinct finite eigenvalue λ_j, there exist unimodular matrices $E_j(z)$ and $F_j(z)$ that are analytic on a neighborhood of λ_j such that*

$$E_j(z)(zC_1 - C_0)F_j(z) = \begin{bmatrix} I_{(d-1)m} & \\ & P(z) \end{bmatrix},$$

then $zC_1 - C_0$ is a linearization of $P(z)$. If, in addition, there are unimodular matrices $E_0(z)$ and $F_0(z)$ which are analytic on a neighborhood of $z = 0$ such that

$$E_0(z)(C_1 - zC_0)F_0(z) = \begin{bmatrix} I_{(d-1)m} & \\ & \operatorname{rev}P(z) \end{bmatrix},$$

then $zC_1 - C_0$ is a strong linearization of $P(z)$.

Now we are in a position to prove one of the main contribution of this work, namely, that the pencil $zC_1 - C_0$ is a strong linearization of the Hermite interpolation polynomial $P(z)$ in (4), that is, Theorem 3.3. To be able to apply Lemma 3.2 we need to assume that the interpolation nodes do not coincide with the eigenvalues of $P(z)$.

Theorem 3.3. *Let $P(z)$ be the Hermite interpolation polynomial expressed in the modified barycentric form* (7), *and let $L(z) = zC_1 - C_0$ be the linearization of $P(z)$, where C_0 and C_1 are the matrices in* (10) *and* (9), *respectively. Assume that none of the interpolation nodes coincides with an eigenvalue of $P(z)$. Then, the pencil $L(z)$ is a strong linearization of $P(z)$ considered as a matrix polynomial of grade $d+2$.*

Proof. The proof of this theorem follows closely the ideas in [1]. We will find unimodular matrices $U(z)$ and $V(z)$ such that $E(z)(zC_1 - C_0)F(z) = I_{m(d+1)} \oplus P(z)$. To this purpose, we start factorizing $L(z) = zC_1 - C_0$ via a block LU factorization

$$zC_1 - C_0 = \mathfrak{L}(z)\mathfrak{U}(z),$$

where $\mathfrak{L}(z)$ and $\mathfrak{U}(z)$ are lower and upper block triangular matrices, respectively.

The matrices $\mathfrak{L}(z)$ and $\mathfrak{U}(z)$ are given by

$$\mathfrak{L}(z) = \begin{bmatrix} I_{s_0} & & & & \\ & I_{s_1} & & & \\ & & \ddots & & \\ & & & I_{s_n} & \\ L_0(z) & L_1(z) & \dots & L_n(z) & 1 \end{bmatrix} \otimes I_m,$$

where

$$L_i(z) = \begin{bmatrix} \frac{w_i}{(z-z_i)} & \cdots & \frac{w_i}{(z-z_i)^{s_i}} \end{bmatrix} \in \mathbb{C}^{1\times s_i} \qquad i = 0, \dots, n, \tag{11}$$

and

$$\mathfrak{U}(z) = \begin{bmatrix} zI_{ms_0} - J_0 & & & & -M_0 \\ & zI_{ms_1} - J_1 & & & -M_1 \\ & & \ddots & & \vdots \\ & & & zI_{ms_n} - J_n & -M_n \\ & & & & \frac{1}{\omega(z)}P(z) \end{bmatrix}.$$

Then, by direct matrix multiplication it may be checked that $zC_1 - C_0 = \mathfrak{L}(z)\mathfrak{U}(z)$.

Now, define $\widetilde{\mathfrak{U}}(z)$ by replacing the bottom-right element of $\mathfrak{U}(z)$ by $I_m/\omega(z)$, with $\omega(z)$ as in (3). Then, let $E(z) := \mathfrak{L}(z)^{-1}$ and $F(z) := \widetilde{\mathfrak{U}}(z)^{-1}$. The explicit forms of $E(z)$ and $F(z)$ are as follows:

$$E(z) = \begin{bmatrix} I_{s_0} & & & & \\ & I_{s_1} & & & \\ & & \ddots & & \\ & & & I_{s_n} & \\ -L_0(z) & -L_1(z) & \dots & -L_n(z) & 1 \end{bmatrix} \otimes I_m,$$

and

$$F(z) = \begin{bmatrix} F_0(z) & & & & \widehat{F}_0(z) \\ & F_1(z) & & & \widehat{F}_1(z) \\ & & \ddots & & \vdots \\ & & & F_n(z) & \widehat{F}_n(z) \\ & & & & \omega(z) I_m \end{bmatrix},$$

where, for $k = 0, \dots, n$, the matrix $F_k(z) = [F_k^{(i,j)}(z)] \in \mathbb{C}^{ms_k \times ms_k}$ is an upper block triangular matrix. Its (i, j)th block entry of size $m \times m$ is given by

$$F_k^{(i,j)}(z) = \frac{1}{(z - z_k)^{j-i+1}} I_m, \qquad \text{if } 1 \leqslant i \leqslant j \leqslant s_k,$$

and $F_k^{(i,j)}(z) = 0$ otherwise. Moreover, for $k = 0, \dots, n$, the matrix

$$\widehat{F}_k(z) = [\widehat{F}_k^{(i)}(z)] \in \mathbb{C}^{ms_k \times m}$$

is a column block vector. Its ith block entry of size $m \times m$ is equal to

$$\widehat{F}_k^{(i)}(z) = \frac{\omega(z)}{(z - z_k)^{s_k - i + 1}} \sum_{j=0}^{s_k - i} M_{k,j} (z - z_k)^j, \qquad 1 \leqslant i \leqslant s_k.$$

Then, it is immediate to check that

$$\begin{bmatrix} I_{m(d+1)} & 0 \\ 0 & P(z) \end{bmatrix} = E(z) L(z) F(z).$$

Notice that the matrices $E(z)$ and $F(z)$ are unimodular, since $\det E(z) = \det F(z) = 1$, and that they are analytic for any $z \notin \{z_0, z_1, \dots, z_n\}$. Therefore, by Lemma 3.2, the pencil $zC_1 - C_0$ is a linearization of $P(z)$.

We now prove that $\text{rev}L(z) = -zC_0 + C_1$ is a linearization of $\text{rev}P(z) = z^{d+2} P(z^{-1})$. We use again the block LU factorization approach. The LU factors of $zC_0 - C_1$ can be computed from the LU factorization of $zC_1 - C_0$. Indeed, since

$$zC_0 - C_1 = -z\mathfrak{L}(z^{-1})\mathfrak{U}(z^{-1}),$$

then $\mathfrak{L}_1(z) := -\mathfrak{L}(z^{-1})$ and $\mathfrak{U}_1(z) := z\mathfrak{U}(z^{-1})$ are the LU factors of $zC_0 - C_1$. From here, the proof goes as the proof to show that $L(z)$ is a linearization of $P(z)$, so we invite the reader to fill in the omitted details. □

Notice that the linearization $L(z) = zC_1 - C_0$ is a strong linearization of the Hermite interpolation polynomial $P(z)$ considered as a matrix polynomial of grade $d+2$, that is, it is a strong linearization of $0_m z^{d+2} + 0_m z^{d-1} + P(z)$. Since the degree of $P(z)$ is upper bounded by d, this implies that the linearization is introducing $2m$ spurious infinite eigenvalues. This fact is a common drawback for linearizations of matrix polynomials in the Lagrange or Hermite interpolation bases [23, 35].

The arrowhead form of the linearization $zC_1 - C_0$ in the previous theorem is a common phenomenon in companion form type linearizations for polynomials in Lagrange or Hermite interpolation basis. This structure of $zC_1 - C_0$ can be exploited in Krylov-type algorithms in order to compute approximate eigenvalues and eigenvectors since for a given vector v the vector $\widetilde{v} = (zC_1 - C_0)v$ may be computed cheaply. Indeed, let us partition the vector $v \in \mathbb{C}^{m(d+2)}$ into $n+2$ blocks as $v = [v_0^T \ v_1^T \ \ldots \ v_n^T \ v_{n+1}^T]$ with $v_i \in \mathbb{C}^{ms_i \times 1}, i = 0, \ldots, n$, and $v_{n+1} \in \mathbb{C}^{m \times 1}$, and let us partition $\widetilde{v}$ in the same way as v. Then, we have $\widetilde{v}_i = (zI_{ms_i} - J_i)v_i - M_i v_{n+1}, i = 0, \ldots, n$, and $\widetilde{v}_{n+1} = zv_{n+1} + W_0 v_0 + \ldots + W_n v_n$. Thus, one does not have to deal with the large $m(d+2) \times m(d+2)$ matrix pencil, all computations can be done working with matrices of smaller size. Taking the special structure of J_i and W_i into account, only n matrix-vector multiplications $M_i v_i$ are required for computing $\widetilde{v}$.

A similar companion form linearization has been proposed in [30] (see also [7, 35]):

$$zC_1 - \widetilde{C}_0 \tag{12}$$

with C_1 introduced in (9), and

$$\widetilde{C}_0 = \begin{bmatrix} J_0^T & & & & \widetilde{M}_0 \\ & J_1^T & & & \widetilde{M}_1 \\ & & \ddots & & \vdots \\ & & & J_n^T & \widetilde{M}_n \\ -\widetilde{W}_0 & -\widetilde{W}_1 & \ldots & -\widetilde{W}_n & 0_m \end{bmatrix}, \tag{13}$$

with J_j as above,

$$\widetilde{W}_i = \begin{bmatrix} w_{i,s_i-1} & \dots & w_{i,1} & w_{i,0} \end{bmatrix} \otimes I_m \in \mathbb{C}^{m \times ms_i} \qquad i = 0, \dots, n,$$

with the weight $w_{i,j}$ in (4), and

$$\widetilde{M}_i = \begin{bmatrix} \frac{F_{i,0}}{0!} & \frac{F_{i,1}}{1!} & \cdots & \frac{F_{i,s_i-1}}{(s_i-1)!} \end{bmatrix}^T \in \mathbb{C}^{ms_i \times m} \qquad i = 0, \dots, n.$$

A similar argument to the proof of Theorem 3.3 leads to the following proposition.

Proposition 3.4. *Let $P(z)$ be an $m \times m$ matrix polynomial of grade d in the form* (4)*, and assume that none of the interpolation nodes coincides with an eigenvalue of $P(z)$. The pencil* (12)*, with the matrices $\widetilde{C}_0$ and C_1 in* (13) *and* (9)*, respectively, is a strong linearization of $P(z)$ considered as a matrix polynomial of grade $d+2$.*

Similar as for the linearization $zC_1 - C_0$, the structure of the linearization $zC_1 - \widetilde{C}_0$ can be exploited by algorithms that only use matrix-vector multiplications such as Krylov-based methods (see [35, Section 5] for a discussion).

A comparison of the two linearizations of $P(z)$ yields that the linearization (12) makes use of the given data $F_{i,j}$ and the generalized barycentric weights (5), while the linearization proposed here makes use of the $M_{i,j}$ (8) and the barycentric weights (6). Notice that, but for large values of s_i, the linearization proposed here is much sparser than (12) due to the form of the W_j. This allows for a faster matrix-vector multiplication. Although sparsity is a favorable advantage, we want to emphasize that in practice it may be of interest to have more than one linearization available for the same problem; see, e.g., [38], where the authors propose a backward stable method to solve quadratic eigenvalue problems that requires the use of two different linearizations.

4. Eigenvector formulas and efficient eigenvector recovery procedures

In this section we show that the eigenvectors of $P(z)$ and those of the linearization $L(z) = zC_1 - C_0$ are related in a simple way. This, in turn, will allow us to recover easily and efficiently the eigenvectors of $P(z)$ from those of $L(z)$. The existence of an eigenvector recovery procedure is essential for a linearization to be relevant for applications.

First, notice that the LU factorization of $L(z) = zC_1 - C_0 = \mathfrak{L}(z)\mathfrak{U}(z)$ is easy and cheap to compute, as we have seen in the proof of Theorem 3.3. This implies that once an eigenvalue $z = \lambda$ of $L(z)$ has been computed, the associated right eigenvector may be easily computed solving the triangular linear systems $\mathfrak{L}(\lambda)y = 0$ and $\mathfrak{U}(\lambda)z = y$. In the following we show how to do it. Initially, let us partition the vector $y \in \mathbb{C}^{m(d+2)}$ into blocks conformable to the blocks of $L(z)$. Then it follows from the special structure of $\mathfrak{L}$ that $y_i = 0$ for $i = 0, \dots, n$ and therefore $y_{n+1} = -\sum_{j=0}^{n} L_j(z) I_{ms_j} v_j = 0$. Hence, y comes for free. Next, we need to solve the linear system $\mathfrak{U}(\lambda)z = y = 0$. To this end, let $z \in \mathbb{C}^{m(d+2)}$ be blocked as the vector y above. Then the $m \times m$ linear system

$$\frac{1}{w(\lambda)} P(\lambda) z_{n+1} = 0 \tag{14}$$

has to be solved, since the other z_j can be essentially read off as $(\lambda I_{ms_j} - J_j)z_j = M_j z_{n+1}$ (taking the special structure of J_j into account). Therefore, after solving one $m \times m$ linear system, only n matrix-vector multiplications are needed to compute an eigenvector of $zC_1 - C_0$. A similar result holds for left eigenvectors. For a similar observation concerning the linearization (12) see [35].

The next step is to show how to recover the eigenvectors of $P(z)$ from those of $L(z)$ and vice versa. The recovery procedures are possible thanks to the existence of matrix polynomials $G(z)$ and $H(z)$ of degree $d + 1$ and sizes $m \times m(d + 2)$ and $m(d + 2) \times m$, respectively, such that

$$(zC_1 - C_0)H(z) = e_{d+2} \otimes P(z) \quad \text{and} \quad G(z)(zC_1 - C_0) = e_{d+2}^T \otimes P(z). \tag{15}$$

The matrix polynomials $H(z)$ and $G(z)$ in the equation above are constructed as follows. First, let us define the matrix polynomial $H_i(z)$ as

$$H_i(z) = \begin{bmatrix} h_{i,0}(z) & h_{i,1}(z) & \cdots & h_{i,s_i-1}(z) \end{bmatrix}^T \in \mathbb{C}^{ms_i \times m}, \qquad i = 0, \dots, n,$$

where

$$h_{i,j}(z) = \omega(z) \sum_{k=0}^{s_i-j-1} \frac{M_{i,k}}{(z - z_i)^{s_i-j-k}}, \quad i = 0, \dots, n, \ j = 0, \dots, s_i - 1.$$

Let us define also the matrix polynomial

$$G_i(z) := \begin{bmatrix} g_{i,0}(z) & g_{i,1}(z) & \cdots & g_{i,s_i-1}(z) \end{bmatrix} \in \mathbb{C}^{m \times ms_i} \qquad i = 0, \dots, n,$$

where

$$g_{i,j}(z) := \frac{w(z)w_i}{(z - z_i)^{j+1}} I_m, \quad i = 0, \dots, n, \ j = 0, \dots, s_i - 1.$$

Then, the matrix polynomials $G(z)$ and $H(z)$ in (15) are given by

$$G(z) = \begin{bmatrix} G_0(z) & G_1(z) & \cdots & G_n(z) & w(z)I_m \end{bmatrix}, \quad \text{and} \quad H(z) = \begin{bmatrix} H_0(z) \\ H_1(z) \\ \vdots \\ H_n(z) \\ w(z)I_m \end{bmatrix}. \qquad (16)$$

Equations in (15) can be checked easily by direct matrix multiplication.

The following theorem shows how eigenvectors of $P(z)$ and those of the linearization $L(z)$ are related.

Theorem 4.1. *Let $P(z)$ be the Hermite interpolation polynomial expressed in the modified barycentric form* (7)*. Let $L(z) = zC_1 - C_0$ be the linearization of $P(z)$, where C_0 and C_1 are the matrices in* (10) *and* (9)*, respectively, and let $G(z)$ and $H(z)$ be the matrix polynomials in* (16)*.*

(a1) *Let x be a right eigenvector of $P(z)$ with finite eigenvalue λ, that is, $P(\lambda)x = 0$. Then, $H(\lambda)x$ is a right eigenvector of $L(z)$ with finite eigenvalue λ.*

(a2) *Let z be a right eigenvector of $L(z)$, partitioned into blocks conformable to the blocks of $L(z)$, with finite eigenvalue λ, that is $L(\lambda)z = 0$. Then, the $(d+2)$th block of z is a right eigenvector of $P(z)$ with finite eigenvalue λ.*

(b1) *Let y be a left eigenvector of $P(z)$ with finite eigenvalue λ, that is, $y^*P(\lambda) = 0$. Then, $y^*G(\lambda)$ is a left eigenvector of $L(z)$ with finite eigenvalue λ.*

(b2) *Let w be a left eigenvector of $L(z)$, partitioned into blocks conformable to the blocks of $L(z)$, with finite eigenvalue λ, that is $w^*L(\lambda) = 0$. Then, the $(d+2)$th block of w is a left eigenvector of $P(z)$ with finite eigenvalue λ.*

Proof. The results follow immediately from the equations in (15) and from the fact that the $(d+2)$th block entry of the matrix polynomials $H(z)$ and $G(z)$ is proportional to the identity matrix I_m. □

Besides their intrinsic matrix theoretical interest, formulas for the eigenvectors of the linearization $zC_1 - C_0$, in parts (a2) and (b2) in Theorem 4.1, find applications in numerical analysis, e.g. for conditioning and backward error analysis [17, 18, 32]. This is the subject of the following section.

5. Conditioning and backward errors

Throughout this section we consider the Hermite interpolation polynomial in (4) as a matrix polynomial of grade $d + 2$.

To guide anyone in the choice of linearization is key to study the influence of the linearization process on the accuracy and stability of the computed eigenvalues and eigenvectors. This is done via the study of *eigenvalue condition numbers* and *backward errors*. Both quantities are needed to bound the forward error (the difference between the exact and the computed solution of a problem) of the computed eigenvalues. A basic inequality that relates the forward error with the condition number and the backward error is

$$\text{forward error} \lesssim \text{condition number} \times \text{backward error}$$

(see for example [16]). Thus when solving a PEP via a linearization, it is important to compare the conditioning and the backward error of both problems.

The backward error of a computed eigenpair (λ, x) measure how far the problem has to be perturbed so the eigenpair (λ, x) is an exact solution of the perturbed problem. To be more precise, the backward error of a right eigenpair (λ, x) of a matrix polynomial $Q(z)$, denoted by $\eta_P(\lambda, x)$, is defined as

$$\eta_Q(\lambda, x) := \min\{\epsilon \,:\, (Q(\lambda) + \Delta Q(\lambda))x = 0, \quad \text{where } \|\Delta Q\|_2 \leq \epsilon \|Q\|_2\},$$

and the backward error of a left eigenpair (λ, y^*), denoted by $\eta_Q(\lambda, y)$, is defined as

$$\eta_Q(\lambda, y^*) := \min\{\epsilon \,:\, y^*(Q(\lambda) + \Delta Q(\lambda)) = 0, \quad \text{where } \|\Delta Q\|_2 \leq \epsilon \|Q\|_2\},$$

where $\|\cdot\|_2$ is a matrix polynomial norm defined as

$$\|Q\|_2^2 := \sum_{k=0}^{d} \|Q_i\|_2^2,$$

where Q_i are the coefficients of $Q(\lambda)$ in the monomial basis.

From [32, Theorem 1] we have that the backward errors of computed eigenpairs (λ, x) and (λ, y^*) of the Hermite interpolation polynomial $P(z)$ in (4) are

$$\eta_P(\lambda, x) = \frac{\|P(\lambda)x\|_2}{\|P\|_2 \cdot \|x\|_2 \cdot \sum_{k=0}^{d+2} |\lambda|^k}, \quad \text{and} \quad \eta_P(\lambda, y^*) = \frac{\|y^* P(\lambda)\|_2}{\|P\|_2 \cdot \|y\|_2 \cdot \sum_{k=0}^{d+2} |\lambda|^k},$$

while the backward errors of computed eigenpairs (λ, v) and (λ, w^*) of the linearization $L(z) = zC_1 - C_0$ are

$$\eta_L(\lambda, v) = \frac{\|L(\lambda)v\|_2}{\|L\|_2 \cdot \|v\|_2 \cdot (1 + |\lambda|)}, \quad \text{and} \quad \eta_L(\lambda, w^*) = \frac{\|w^* L(\lambda)\|_2}{\|L\|_2 \cdot \|w\|_2 \cdot (1 + |\lambda|)}.$$

Our aim now is to compare $\eta_L(\lambda, v)$ with $\eta_P(\lambda, x)$, and $\eta_L(\lambda, w^*)$ with $\eta_P(\lambda, y^*)$, when the eigenvectors x and y^* of $P(\lambda)$ are recovered from the $(d+2)$th blocks of those of $L(\lambda)$ as it is explained in Theorem 4.1. From (15) we get

$$G(\lambda)L(\lambda)v = (e_{d+2}^T \otimes P(\lambda))v = P(\lambda)(e_{d+2}^T \otimes I_m)v = P(\lambda)x,$$

and

$$w^*L(\lambda)H(\lambda) = w^*(e_{d+2} \otimes P(\lambda)) = w^*(e_{d+2} \otimes I_m)P(\lambda) = y^*P(\lambda).$$

Using the equations above and the explicit formulas for $\eta_P(\lambda, x)$ and $\eta_P(\lambda, y^*)$ yields

$$\eta_P(\lambda, x) \leq \frac{\|G(\lambda)\|_2 \cdot \|L(\lambda)v\|_2}{\|P\|_2 \cdot \|x\|_2 \cdot \sum_{k=0}^{d+2} |\lambda^k|}, \quad \text{and} \quad \eta_P(\lambda, y^*) \leq \frac{\|H(\lambda)\|_2 \cdot \|w^*L(\lambda)\|_2}{\|P\|_2 \cdot \|y\|_2 \cdot \sum_{k=0}^{d+2} |\lambda^k|}.$$

Therefore

$$\frac{\eta_P(\lambda, x)}{\eta_L(\lambda, z)} \leq \frac{\|L\|_2 \cdot (1 + |\lambda|)}{\|P\|_2 \cdot \sum_{k=0}^{d+2} |\lambda|^k} \cdot \frac{\|G(\lambda)\|_2 \cdot \|v\|_2}{\|x\|_2},$$

and

$$\frac{\eta_P(\lambda, y)}{\eta_L(\lambda, w)} \leq \frac{\|L\|_2 \cdot (1 + |\lambda|)}{\|P\|_2 \cdot \sum_{k=0}^{d+2} |\lambda|^k} \cdot \frac{\|H(\lambda)\|_2 \cdot \|w\|_2}{\|y\|_2},$$

which give a posteriori upper bounds for the ratios $\eta_P(\lambda, x)/\eta_L(\lambda, v)$ and $\eta_P(\lambda, y^*)/\eta_L(\lambda, w^*)$.

Eigenvalue condition numbers measure the sensitivity of an eigenvalue to small perturbations. To be more precise, the condition number of a nonzero simple eigenvalue λ of a regular matrix polynomial $Q(z)$, denoted by $\kappa_Q(\lambda)$, is defined as

$$\kappa_Q(\lambda) := \lim_{\epsilon \to 0} \sup \left\{ \frac{|\Delta\lambda|}{\epsilon|\lambda|} : (Q(\lambda + \Delta\lambda) + \Delta Q(\lambda + \Delta Q))(x + \Delta x) = 0, \right.$$
$$\left. \text{where } \|\Delta Q\|_2 \leq \epsilon \|Q\|_2 \right\},$$

where $\|\cdot\|_2$ is the matrix polynomial norm defined before.

From [32, Theorem 5] we have that the condition number of a nonzero simple eigenvalue λ of the regular Hermite interpolation polynomial $P(z)$ in (4) is

$$\kappa_P(\lambda) = \frac{\|P\|_2 \cdot \|x\|_2 \|y\|_2 \sum_{k=0}^{d+2} |\lambda|^k}{|\lambda| \cdot |y^*P'(\lambda)x|},$$

where x and y^* are the associated right and left eigenvectors of $P(z)$. Analogously, the condition number of a nonzero simple eigenvalue λ of the linearization $L(z) = zC_1 - C_0$ is

$$\kappa_L(\lambda) = \frac{\|L\|_2 \cdot \|v\|_2 \|w\|_2 (1+|\lambda|)}{|\lambda| \cdot |w^* L'(\lambda) v|},$$

where v and w^* are the associated right and left eigenvectors of $L(z)$. Our aim now is to compare $\kappa_L(\lambda)$ with $\kappa_P(\lambda)$.

Let x and y^* be right and left eigenvectors of $P(z)$ with eigenvalue λ. Differentiating the first equation in (15) with respect to z gives

$$L'(z)H(z) + L(z)H'(z) = e_{d+2} \otimes P'(z).$$

Evaluating at an eigenvalue λ, pre-multiplying by $w^* = y^* G(\lambda)$, post-multiplying by x and making use of the structure of $G(\lambda)$ (16) gives

$$w^* L'(\lambda) v = y^* G(\lambda)(e_{d+2} \otimes P'(\lambda))x = \omega(\lambda) y^* P'(\lambda) x,$$

where we have used that $v = H(\lambda)x$ and w^* are right and left eigenvectors of $L(z)$ with eigenvalue λ (see Theorem 4.1). Therefore, we have

$$\frac{\kappa_L(\lambda)}{\kappa_P(\lambda)} = \frac{1}{|\omega(\lambda)|} \frac{(1+|\lambda|)\|L\|_2}{\sum_{k=0}^{d+2} |\lambda|^k \|P\|_2} \cdot \frac{\|w\|_2 \|z\|_2}{\|y\|_2 \|x\|_2}.$$

The above expression can be used by any user of the linearization to compare the sizes of $\kappa_L(\lambda)$ and $\kappa_P(\lambda)$.

We finish this section remarking that these results concerning conditioning and backward errors may be used to get a posteriori upper bound of the forward error, which could be used to check in the eigenvalues of $P(z)$ have been computed with the forward errors expected from the sensitivity of the original data, i.e., from $P(z)$.

6. Conclusions

In this work, we have introduced a new strong linearization for matrix polynomials which come from Hermite interpolation problems, with a much sparser structure than other linearizations in the literature. We have shown how to recover efficiently the eigenvectors of the matrix polynomials from those of the linearization. In addition, we have obtained explicit formulas for the eigenvectors of the linearization which have allowed us to compare backward errors and eigenvalue condition numbers of the polynomial eigenvalue problem with those of the linearized problem.

References

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

[2] J. C. Butcher, R. M. Corless, L. Gonzalez-Vega, A. Shakoori, *Polynomial algebra for Birkhoff interpolants*, Numer. Algorithms 56(3) (2011) 319–347.

[3] T. Betcke, N. J. Higham, V. Mehrmann, C. Schröder, F. Tisseur, *NLEVP: a collection of nonlinear eigenvalue problems*, ACM Trans. Math. Software 39(2) (2013) 1–28.

[4] J.-P. Berrut, L. N. Trefethen, *Barycentric Lagrange interpolation*, SIAM Rev. 46(3) (2004) 501–517.

[5] R. M. Corless. *Generalized companion matrices in the Lagrange basis*, In L. Gonzalez-Vega, editor, Proceedings of EACA, pages 317–322, Santander, Spain, 2004.

[6] R. M. Corless, *On a generalized companion matrix pencil for matrix polynomials expressed in the Lagrange basis*, In D. Wang and L. Zhi, editors, Proceedings of Symbolic-Numeric Computation, pages 1–15, Xian, China, 2005.

[7] R. M. Corless, Robert, A. Shakoori, D. A. Aruliah, L. Gonzalez-Vega, *Barycentric Hermite interpolants for event location in initial-value problems*, JNAIAM J. Numer. Anal. Ind. Appl. Math. 3(1-2) (2008) 1–16.

[8] P. J. Davis, *Interpolation and Approximation*, Dover Publications, 1975.

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

[10] F. De Terán, F. M. Dopico, D. S. Mackey, *Spectral equivalence of matrix polynomials and the index sum theorem*, Linear Algebra Appl. 459 (2014), 264–333.

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

[12] L. Gemignani, *Structured matrix methods for polynomial root-finding*, In Proceedings of the 2007 International Symposium on Symbolic and Algebraic Computation, pages 175–180, New York, 2007. ACM.

[13] I. J. Good, *The colleague matrix, a Chebyshev analogue of the companion matrix*, Quart. J. Math. Oxford Ser. 12(2) (1961) 61–68.

[14] S. Güttel, R. Van Beeumen, K. Meerbergen, W. Michiels, *NLEIGS:A class of robust fully rational Krylov methods for nonlinear eigenvalue problems*, Technical Report TW633. Department of Computer Science, KU Leuven, 2013.

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

[16] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, SIAM, Philadelphia, 1996.

[17] N. J. Higham, D. S. Mackey, F. Tisseur, *The conditioning of linearizations of matrix polynomials*, SIAM J. Matrix Anal. Appl. 28(4) (2006) 1005–1028.

[18] N. J. Higham, R. -C. Li, F. Tisseur, *Backward error of polynomial eigenproblems solved by linearization*, SIAM J. Matrix Anal. Appl., 29(4) (2007) 1218–1241.

[19] E. Jarlebring, W. Michiels, K. Meerbergen, *A linear eigenvalue algorithm for the nonlinear eigenvalue problem*, Numer. Math. 122(1) (2012) 169–195.

[20] G. F. Jónsson, S. Vavasis, *Solving polynomials with small leading coefficients*, SIAM J. Matrix Anal. Appl. 26(2) (2005) 400–414.

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

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

[23] P. W. Lawrence, R. M. Corless, *Backward error of polynomial eigenvalue problems solved by linearization of Lagrange interpolants*, TW Reports, TW655, Department of Computer Science, KU Leuven, February 2015.

[24] D. S. Mackey, N. Mackey, C. Mehl, V. Mehrmann, *Vector spaces of linearizations for matrix polynomials*, SIAM J. Matrix Anal. Appl. 27(4) (2006) 971–1004.

[25] D. S. Mackey, V. Perovic, *Linearizations of Matrix Polynomials in Bernstein Basis*, MIMS Eprint 2014.29, preprint, 2014.

[26] V. Noferini, J. Pérez, *Chebyshev-Fiedler pencils*, MIMS Eprint 2015.90, preprint, 2015.

[27] B. Sadiq, D. Viswanath, *Barycentric Hermite interpolation*, SIAM J. Sci. Comput. 35(3) (2013) A1254–A1270.

[28] P. Saltenberger, *Linearization of Matrix Polynomials in Tschebyscheff Basis*, Master Thesis TU Braunschweig, Germany, 2015.

[29] C. Schneider, W. Werner, *Hermite interpolation: the barycentric approach*, Computing, 46(1) (1991) 35–51.

[30] A. Shakoori, *Bivariate Polynomial Solver by Values*, Ph.D. Thesis, University of Western Ontario, London, Ontario, 2007.

[31] H. Tal-Ezer, *High degree polynomial interpolation in Newton form*, SIAM J. Sci. Statist. Comput. 12(3) (1991) 648–667.

[32] F. Tisseur, *Backward error and condition of polynomial eigenvalue problems*, Linear Algebra Appl. 309(1-3) (2000) 339–361.

[33] F. Tisseur, K. Meerbergen, *The quadratic eigenvalue problem*, SIAM Rev. 43(2) (2001) 235–286.

[34] R. Van Beeumen, K. Meerbergen, W. Michiels, *A rational Krylov method based on Hermite interpolation for nonlinear eigenvalue problems*, SIAM J. Sci. Comput. 35(1) (2013) A327–A350.

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

[36] E. Waring, *Problems concerning interpolations*, Phil. Trans. R. Soc. Lond. 69 (1779) 59–67.

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

[38] L. Zeng, Y. Su, *A backward stable algorithm for quadratic eigenvalue problems*, SIAM J. Matrix Anal. Appl. 35 (2014) 499–516.