



GENERALIZED STANDARD TRIPLES FOR ALGEBRAIC LINEARIZATIONS OF MATRIX POLYNOMIALS*

EUNICE Y. S. CHAN[†], ROBERT M. CORLESS[‡], AND LEILI RAFIEE SEVYERI[§]

Abstract. We define *generalized standard triples* \mathbf{X} , \mathbf{Y} , and $L(z) = z\mathbf{C}_1 - \mathbf{C}_0$, where $L(z)$ is a linearization of a regular matrix polynomial $\mathbf{P}(z) \in \mathbb{C}^{n \times n}[z]$, in order to use the representation $\mathbf{X}(z\mathbf{C}_1 - \mathbf{C}_0)^{-1}\mathbf{Y} = \mathbf{P}^{-1}(z)$ which holds except when z is an eigenvalue of \mathbf{P} . This representation can be used in constructing so-called *algebraic linearizations* for matrix polynomials of the form $\mathbf{H}(z) = z\mathbf{A}(z)\mathbf{B}(z) + \mathbf{C} \in \mathbb{C}^{n \times n}[z]$ from generalized standard triples of $\mathbf{A}(z)$ and $\mathbf{B}(z)$. This can be done even if $\mathbf{A}(z)$ and $\mathbf{B}(z)$ are expressed in differing polynomial bases. Our main theorem is that \mathbf{X} can be expressed using the coefficients of the expression $1 = \sum_{k=0}^{\ell} e_k \phi_k(z)$ in terms of the relevant polynomial basis. For convenience, we tabulate generalized standard triples for orthogonal polynomial bases, the monomial basis, and Newton interpolational bases; for the Bernstein basis; for Lagrange interpolational bases; and for Hermite interpolational bases.

Key words. Standard triple, Regular matrix polynomial, Polynomial bases, Companion matrix, Colleague matrix, Comrade matrix, Algebraic linearization, Linearization of matrix polynomials.

AMS subject classifications. 65F15, 15A22, 65D05.

1. Introduction. A *matrix polynomial* $\mathbf{P}(z) \in \mathbb{F}^{m \times n}[z]$ is a polynomial in the variable z with coefficients that are m by n matrices with entries from the field \mathbb{F} . We will use $\mathbb{F} = \mathbb{C}$, the field of complex numbers, in this paper. Typically, an expression in the monomial basis $\phi_k(z) = z^k$ is given for $\mathbf{P}(z)$, and often only *regular* matrix polynomials are considered, that is, with $m = n$ (we will use n for the dimension) and where $\det \mathbf{P}(z)$ is not identically zero. Matrix polynomials have many applications, and their study is of both classic and ongoing interest. See the classic work [20] and the surveys [21] and [29] for theory and applications.

In this paper, as is done in [2], we consider the case when *any* polynomial basis $\phi_k(z)$ is used. We require that the set $\{\phi_k(z)\}$ for $0 \leq k \leq \ell$ forms a basis for polynomials of grade ℓ . The word ‘grade’ is short for ‘degree at most.’ Thus, we write our regular matrix polynomial as

$$(1.1) \quad \mathbf{P}(z) = \sum_{k=0}^{\ell} \mathbf{P}_k \phi_k(z),$$

where the matrices $\mathbf{P}_k \in \mathbb{F}^{n \times n}$ are square, and the *degree*¹ of $\mathbf{P}(z)$ is at most ℓ ($\mathbf{P}(z)$ has grade ℓ).

*Received by the editors on May 29, 2019. Accepted for publication on September 26, 2021. Handling Editor: Froilán Dopico. Corresponding Author: Eunice Y. S. Chan

[†]Centre for Medical Evidence, Decision Integrity and Clinical Impact (MEDICI Centre), Department of Anesthesia and Perioperative Medicine, Schulich School of Medicine and Dentistry, Western University, London, ON, Canada (echan295@uwo.ca)

[‡]Ontario Research Centre for Computer Algebra, School of Mathematical and Statistical Sciences, Western University, London, ON, Canada (rcorless@uwo.ca)

[§]David R. Cheriton School of Computer Science, University of Waterloo, Waterloo, ON, Canada (lrafiees@uwaterloo.ca)

¹The *degree* of a matrix polynomial $\mathbf{P}(z)$ is defined as follows. If $\mathbf{P}(z)$ is identically zero, the degree is $-\infty$. Otherwise, if $\mathbf{P}(z) = \sum_{k=0}^{\ell} \mathbf{A}_k z^k$ expressed in the monomial basis and \mathbf{A}_{ℓ} is not the zero matrix, then the degree of $\mathbf{P}(z)$ is ℓ .

The notion of ‘grade’ is useful even for the monomial basis, but it is especially useful if the basis is an interpolational basis or the Bernstein basis $\phi_k(z) = B_k^n(z) = \binom{n}{k} z^k (1-z)^{n-k}$, when the degree of the polynomial may not be clear from the data.

For more information about matrix polynomials, consult [27]. See also [29, 21] and consult the seminal book [20]. Linearizations using different polynomial bases were first systematically studied in [2]; for a more up-to-date treatment see [5]. Some other recent papers of interest include [4, 28, 12, 17], and [32]; this is a very active area. See also [1]. In that paper, standard triples for structured matrices are studied. In [10], several proofs are given of strict equivalence of various linearizations to the standard second linearization for the monomial basis, which we will use without much further comment in this paper.

1.1. Organization of the paper. In Section 1.2, we establish notation, give the definition of algebraic linearization, and define generalized standard triples. We define this last in Definition 1.1 with reference to the representation in equation (1.7). In Section 1.3, we show how to use a generalized standard triple in the construction of algebraic linearizations. We also give a proof, by construction of the necessary equivalence matrices $\mathbf{E}(z)$ and $\mathbf{F}(z)$, that algebraic linearizations are *local* linearizations (a very useful notion from [13]) in the same sense that its components are. In particular, if the components \mathbf{A} and \mathbf{B} are local linearizations in the sets $\Sigma_{\mathbf{A}}$ and $\Sigma_{\mathbf{B}}$, respectively, then the ‘algebraic linearization’ is a local linearization in the intersection $\Sigma_{\mathbf{A}} \cap \Sigma_{\mathbf{B}}$.

In Section 2, we prove our main result, giving a universal expression for generalized standard triples for the linearizations using the polynomial bases that appear in Sections 3.1, 3.2, and 3.3. In Section 3, we introduce all the polynomial bases by scalar examples. We sum up in the final section.

1.2. Notation and definitions. Two matrix polynomials $\mathbf{P}_1(z)$ and $\mathbf{P}_2(z)$ are called *unimodularly equivalent* if there exist unimodular matrix polynomials (that is, matrix polynomials with constant nonzero determinant) $\mathbf{E}(z)$ and $\mathbf{F}(z)$ with $\mathbf{P}_1(z) = \mathbf{E}(z)\mathbf{P}_2(z)\mathbf{F}(z)$. A matrix pencil $\mathbf{L}(z) := z\mathbf{C}_1 - \mathbf{C}_0$ is called a *linearization* of the matrix polynomial $\mathbf{P}(z)$ if both \mathbf{C}_1 and \mathbf{C}_0 are of dimension $N \geq n$ and $\mathbf{L}(z)$ is unimodularly equivalent to the block diagonal matrix $\text{diag}(\mathbf{P}(z), \mathbf{I}_{N-n})$. Two linearizations $\mathbf{L}_m(z)$ and $\mathbf{L}_\phi(z)$ are called *strictly equivalent* if the corresponding matrices are equivalent in the following stronger sense: $\mathbf{C}_{1,m} = \mathbf{E}(z)\mathbf{C}_{1,\phi}\mathbf{F}(z)$ and $\mathbf{C}_{0,m} = \mathbf{E}(z)\mathbf{C}_{0,\phi}\mathbf{F}(z)$, with the same unimodular matrix polynomials $\mathbf{E}(z)$ and $\mathbf{F}(z)$.

In the paper [13], we find the powerful notion of a *local* linearization. In the notation of their Definition 2.1, two *rational* matrices $\mathbf{G}_1(z)$ and $\mathbf{G}_2(z)$ are said to be equivalent in a nonempty set $\Sigma \subset \mathbb{F}$ if there exist rational matrices $\mathbf{R}_1(z)$ and $\mathbf{R}_2(z)$ each nonsingular in Σ with $\mathbf{R}_1(z)\mathbf{G}_1(z)\mathbf{R}_2(z) = \mathbf{G}_2(z)$. This notion, intended for use with matrix rational functions, is also useful for matrix polynomials, as we will see.

There is a related idea to linearization, that of ‘companion pencil’, sometimes denoted by an ordered pair of matrices (\mathbf{A}, \mathbf{B}) , where the only requirement is that $\det \mathbf{P}(z) = K \det(z\mathbf{B} - \mathbf{A})$ for some nonzero constant K . A companion pencil, therefore, has the same generalized eigenvalues as the matrix polynomial. Companion pencils that are not linearizations do not necessarily preserve eigenvectors or elementary divisors and are less useful than linearizations.

We will frequently use the notation (\mathbf{A}, \mathbf{B}) to indicate the pencil $z\mathbf{B} - \mathbf{A}$.

The usual reversal² of a matrix polynomial of grade ℓ is the polynomial $\text{rev } \mathbf{P}(z) = z^\ell \mathbf{P}(z^{-1})$. A linearization $\mathbf{L}(z) = z\mathbf{C}_1 - \mathbf{C}_0$ of \mathbf{P} is called a *strong* linearization if $\text{rev } \mathbf{L}(z) = \mathbf{C}_1 - z\mathbf{C}_0$ is also a linearization of $\text{rev } \mathbf{P}(z)$.

If $\mathbf{L}(z) = z\mathbf{C}_1 - \mathbf{C}_0 \in \mathbb{C}^{N \times N}[z]$ is a local linearization of $\mathbf{P}(z)$ in Σ , then as a necessary consequence $\det(\mathbf{P}(z)) = q(z)\det(\mathbf{L}(z)) = \det(z\mathbf{C}_1 - \mathbf{C}_0)$ for some rational $q(z)$ whose zeros and poles lie outside Σ . The eigenvalues of \mathbf{P} in Σ are thus computable from the generalized eigenvalues of \mathbf{L} , which can be computed using well-known methods.

A *standard pair* (\mathbf{X}, \mathbf{T}) for a regular matrix polynomial $\mathbf{P}(z)$ expressed in the monomial basis with coefficients \mathbf{P}_k is defined in [20] or in [27] as having the following properties: \mathbf{X} has dimension $n \times n\ell$, \mathbf{T} has dimension $n\ell \times n\ell$,

$$(1.2) \quad \sum_{k=0}^{\ell} \mathbf{P}_k \mathbf{X} \mathbf{T}^k = \mathbf{0},$$

and the $n\ell \times n\ell$ matrix

$$(1.3) \quad \mathbf{Q} = \begin{bmatrix} \mathbf{X} \\ \mathbf{X}\mathbf{T} \\ \vdots \\ \mathbf{X}\mathbf{T}^{\ell-1} \end{bmatrix},$$

is nonsingular. We can then define a third matrix

$$(1.4) \quad \mathbf{Y} = \mathbf{Q}^{-1} \begin{bmatrix} \mathbf{0}_n \\ \vdots \\ \mathbf{0}_n \\ \mathbf{I}_n \end{bmatrix},$$

and say that the triple $(\mathbf{X}, \mathbf{T}, \mathbf{Y})$ is a *standard triple* for a *monic* $\mathbf{P}(z)$. It is pointed out in [27] that monicity of $\mathbf{P}(z)$ is not required for many of the formulæ to do with standard pairs (but is required for some).

Theorem 12.1.4 of [19] states that if there are matrices \mathbf{X} , \mathbf{T} , and \mathbf{Y} of dimension $n \times n\ell$, $n\ell \times n\ell$, and $n\ell \times n$ for which

$$(1.5) \quad \mathbf{P}^{-1}(z) = \mathbf{X}(z\mathbf{I}_n - \mathbf{T})^{-1}\mathbf{Y},$$

then $(\mathbf{X}, \mathbf{T}, \mathbf{Y})$ is a standard triple for $\mathbf{P}(z)$. This is also reported in Lemma 2 in [1]. There are two other representations of a matrix polynomial given a standard triple: the right canonical form, and the left canonical form. See Theorem 2.4 in [20]. However, we do not need those representations for algebraic linearization: it is the resolvent form above that we seek to generalize in this paper. The reason is that it

²This definition, which is standard, is particularly appropriate for the monomial basis. The coefficients of the reversed matrix polynomial in the monomial basis are simply the same matrices in reverse order. The notion of a reversal, however, is independent of the basis used, and indeed reversals can be done differently. In [9] for instance, we find a slightly different definition of reversal, appropriate for computation in a Lagrange or Hermite interpolational basis, which maps an arbitrary finite point to infinity; this difference allows for greater numerical stability. Something analogous for the Bernstein linearization was introduced in [10].

is this formula that is used in the proof that algebraic linearizations can be performed when the component matrix polynomials are expressed in different bases.

A polynomial basis $\{\phi_k(z)\}_{k=0}^\ell$ for polynomials of grade ℓ is a set of polynomials for which there is a nonsingular matrix Φ relating the polynomials $\phi_k(z)$ to the monomials $1, z, \dots, z^\ell$. We can write this as

$$(1.6) \quad \begin{bmatrix} \phi_\ell(z) \\ \phi_{\ell-1}(z) \\ \vdots \\ \phi_0(z) \end{bmatrix} = \Phi \begin{bmatrix} z^\ell \\ z^{\ell-1} \\ z^{\ell-2} \\ \vdots \\ 1 \end{bmatrix}.$$

Frequently, we want the $\ell \times \ell$ matrix that only goes up to grade $\ell - 1$. This should not cause confusion. This matrix forms the foundation for the proofs in [10] of strict equivalence of various linearizations. The matrix Φ is called the change-of-basis matrix and is usually exponentially ill-conditioned in the dimension. For example, for the Bernstein polynomials $\phi_j^\ell(z) = \binom{\ell}{j} z^j (1 - z)^{\ell-j}$ the change-of-basis matrix has entries $\phi_{i,j} = \binom{j}{i} / \binom{\ell}{j}$ and condition number $K_1 = K_\infty = (\ell + 1) \binom{\ell}{s} 2^{\ell-s}$ where $s = \lceil (\ell - 2)/3 \rceil$ (the notation $\lceil x \rceil$ means the ceiling of x , the least integer not smaller than x). A short computation shows $K_1 \sim 3^{\ell+1} \sqrt{\ell/4\pi}$ as $\ell \rightarrow \infty$ and is thus exponentially growing with the dimension.

The constructions and definitions of standard triple discussed above are apparently tied to the monomial basis because of the powers T^k in equation (1.3). We would like to relax this restriction and extend the notion of standard triple to other bases and also to the non-monic case. In particular, we would like the following extension of Theorem 2.4 in [20] or Theorem 12.1.4 in [19] to be available: If a matrix $X \in \mathbb{C}^{n \times N}$, the linearization $L(z) = zC_1 - C_0 \in \mathbb{C}^{N \times N}[z]$ of a polynomial matrix $P(z)$ expressed in terms of any polynomial basis, and a matrix $Y \in \mathbb{C}^{N \times n}$ satisfy

$$(1.7) \quad P^{-1}(z) = X(zC_1 - C_0)^{-1}Y,$$

for $z \notin z(P)$ (the set of eigenvalues of $P(z)$), then $X, L(z)$, and Y form a *generalized standard triple* for $P(z)$. This obviously requires regularity of P because the formula contains $P^{-1}(z)$.

Indeed, we simply require $L(z)$ to be a linearization (or local linearization, restricted to some set Σ) and take this extension as a *definition*. Such things exist, as we will demonstrate, and are useful, as shown in [7].

A referee pointed out that a simple characterization of linearizations $L(z)$ can be found in [17]; this can be used as a starting point, here.

DEFINITION 1.1. *Matrices $X, zC_1 - C_0$, and Y form a generalized standard triple for the regular matrix polynomial $P(z)$ if $L(z) = zC_1 - C_0$ is a linearization of P and equation (1.7) holds.*

Note that the matrices X and Y do not depend on z , but the linearization $L(z)$ does, albeit only linearly; we could instead have chosen to use the words ‘standard quadruple’ to mean (X, C_1, C_0, Y) where z does not appear of any of these matrices, but this quibble seems to be a matter of aesthetics only; we may use the term ‘triple’ to refer to X , the linearization, and Y .

PROPOSITION 1.2. *If $L(z)$ is a linearization for $P(z)$, then there exists X and Y forming a generalized standard triple with $L(z)$ in the above sense.*

Proof. This is already proved in [20]. If $L(z)$ is a linearization for $P(z)$, then there exist unimodular matrix polynomials $E(z)$ and $F(z)$ with $F^{-1}(z)(zC_1 - C_0)^{-1}E^{-1}(z) = \text{diag}(P^{-1}(z), I_n, \dots, I_n)$. By pre-multiplying by $X_p = [I_n, 0, \dots, 0]$ and postmultiplying by $Y_p = [I_n, 0, \dots, 0]^T$, we may find $X = X_p F^{-1}(z)$ and $Y = E^{-1}(z) Y_p$ so that equation (1.7) holds. \square

REMARK 1.3. Thus, as a referee pointed out, generalized standard triples may be read off from the proof that $L(z)$ is indeed a linearization (we shall see this explicitly, shortly). The matrices $E(z)$ and $F(z)$ above are not unique; given a nonsingular constant matrix B of the right dimension, $\text{diag}(I_n, B)E(z)(zC_1 - C_0)F(z)\text{diag}(I_n, B^{-1})$ is also $\text{diag}(P(z), I_n, \dots, I_n)$. Therefore, generalized standard triples are also not unique.

Without using the rank criterion characterization of [17], say, it is not a priori clear that if equation (1.7) holds then $L(z)$ is necessarily a linearization of $P(z)$. We do have in any case, however, the following:

LEMMA 1.4. *If X, Y , and the matrix pencil $L(z) = zC_1 - C_0$ are such that equation (1.7) holds, then the matrix pencil $L(z)$ is at least a companion pencil for $P(z)$.*

Proof. The norm of the resolvent $\|P^{-1}(z)\|$ will be arbitrarily large if and only if $\|(zC_1 - C_0)^{-1}\|$ is arbitrarily large. Thus, the zeros of $\det P(z)$ are the same as the zeros of $\det L(z)$. \square

For various reasons, we usually do not wish to invert a ‘leading coefficient’ here; for instance, if the polynomial basis is not degree-graded, e.g. for the Bernstein basis, then in order to even look at the true leading coefficient, we have to form a particular linear combination of the existing coefficients. In floating-point arithmetic, rounding errors can disguise the rank of the resulting matrix, hence our interest in the generalization.

If $X, zC_1 - C_0$, and Y form a generalized standard triple according to our definition, then so do $XU, V^{-1}(zC_1 - C_0)U$, and $V^{-1}Y$ for any nonsingular matrices U and V of dimension N by N .

Several similarities are used very frequently. For convenience, we explicitly describe here one of the most common.

LEMMA 1.5 (Flipping). *Let $X, L(z), Y$ be a generalized standard triple for $P(z)$. Further, let J be the $N \times N$ ‘anti-identity’, also called the sip matrix, for standard involutory permutation, $J_{i,j} = 0$ unless $i + j = N + 1$ when $J_{i,N+1-i} = 1$. Consider the ‘flipped’ linearization $L_F(z) = J(zC_1 - C_0)J$. Then $X_F = XJ, L_F(z)$, and $Y_F = JY$ also form a generalized standard triple for $P(z)$.*

Proof. Immediate. \square

REMARK 1.6. Flipping switches both the order of the equations and the order of the variables. It obviously does not change eigenvalues. Flipping, transposition, and flipping-with-transposition give four common equivalent linearizations [34].

1.3. Algebraic linearizations. An algebraic linearization (A_H, B_H) , as referred to in the title of this present note, is defined in [7] as a linearization (A_H, B_H) of a matrix polynomial $H(z) = zA(z)D_0B(z) + C$ constructed recursively from linearizations (A_A, B_A) and (A_B, B_B) of the lower-grade component matrix polynomials $A(z)$ and $B(z)$, together with constant matrices D_0 and C . The paper [7] did not give an explicit unimodular pair $(E_H(z), F_H(z))$ that reduces the linearization to $\text{diag}(zA(z)D_0B(z) + C_0, I_{N-n})$, proving that the construction actually gave a linearization, so we give a method to construct them here. Without loss of generality we take $D_0 = I_n$.

THEOREM 1.7. *Let $\mathbf{E}_A(z)$ and $\mathbf{F}_A(z)$ be rational matrices such that if $z \in \Sigma_A$ then $\mathbf{E}_A(z)$ and $\mathbf{F}_A(z)$ are invertible and $\mathbf{E}_A(z)(z\mathbf{B}_A - \mathbf{A}_A)\mathbf{F}_A(z) = \text{diag}(\mathbf{A}(z), \mathbf{I}_{N_A-n})$, and likewise let $\mathbf{E}_B(z)$ and $\mathbf{F}_B(z)$ be rational matrices such that if $z \in \Sigma_B$ then $\mathbf{E}_B(z)$ and $\mathbf{F}_B(z)$ are invertible and $\mathbf{E}_B(z)(z\mathbf{B}_B - \mathbf{A}_B)\mathbf{F}_B(z) = \text{diag}(\mathbf{B}(z), \mathbf{I}_{N_B-n})$. Then the pencil $z\mathbf{B}_H - \mathbf{A}_H$ is a local linearization of $\mathbf{H}(z) = z\mathbf{A}(z)\mathbf{B}(z) + \mathbf{C}$ for $z \in \Sigma_A \cap \Sigma_B$, where the matrices \mathbf{B}_H and \mathbf{A}_H are given as follows:*

$$(1.8) \quad \mathbf{B}_H = \begin{bmatrix} \mathbf{B}_A & & \\ & \mathbf{I}_n & \\ & & \mathbf{B}_B \end{bmatrix},$$

and

$$(1.9) \quad \mathbf{A}_H = \begin{bmatrix} \mathbf{A}_A & \mathbf{0}_{N_A,n} & -\mathbf{Y}_A\mathbf{c}\mathbf{X}_B \\ -\mathbf{X}_A & \mathbf{0}_n & \mathbf{0}_{n,N_B} \\ \mathbf{0}_{N_B,N_A} & -\mathbf{Y}_B & \mathbf{B}_A \end{bmatrix}.$$

Here $\mathbf{X}_A = [\mathbf{I}_n, 0, \dots, 0]\mathbf{F}_A^{-1}(z)$, $\mathbf{Y}_A = \mathbf{E}_A^{-1}(z)[\mathbf{I}_n, 0, \dots, 0]^T$ and likewise $\mathbf{X}_B = [\mathbf{I}_n, 0, \dots, 0]\mathbf{F}_B^{-1}(z)$ and $\mathbf{Y}_B = \mathbf{E}_B^{-1}(z)[\mathbf{I}_n, 0, \dots, 0]^T$ give the elements of the (generalized) standard triples for $\mathbf{A}(z)$ and $\mathbf{B}(z)$.

Proof. We first construct

$$(1.10) \quad \mathbf{E}_1(z) = \begin{bmatrix} \mathbf{E}_A(z) & & \\ & \mathbf{I}_n & \\ & & \mathbf{E}_B(z) \end{bmatrix},$$

and

$$(1.11) \quad \mathbf{F}_1(z) = \begin{bmatrix} \mathbf{F}_A(z) & & \\ & \mathbf{I}_n & \\ & & \mathbf{F}_B(z) \end{bmatrix}.$$

Applying them we get

$$(1.12) \quad \mathbf{E}_1(z)(z\mathbf{B}_H - \mathbf{A}_H)\mathbf{F}_1(z) = \begin{bmatrix} \mathbf{A}(z) & & \mathbf{E}_A(z)\mathbf{Y}_A\mathbf{c}\mathbf{X}_B\mathbf{F}_B(z) & & \\ & \mathbf{I}_{N_A-n} & & & \\ \mathbf{X}_A\mathbf{F}_A(z) & & z\mathbf{I}_n & & \\ & & \mathbf{E}_B(z)\mathbf{Y}_B & \mathbf{B}(z) & \\ & & & & \mathbf{I}_{N_B-n} \end{bmatrix}.$$

Simplifying and using the definitions of the matrices appearing in the standard triples, we get

$$(1.13) \quad \left[\begin{array}{c|c|c} \mathbf{A}(z) & & \mathbf{C} \\ \hline & \mathbf{I}_{N_A-n} & \\ \hline \mathbf{I}_n & z\mathbf{I}_n & \\ \hline & \mathbf{I}_n & \mathbf{B}(z) \\ & & \mathbf{I}_{N_B-n} \end{array} \right],$$

which is permutationally equivalent to

$$(1.14) \quad \left[\begin{array}{c|c|c} \mathbf{A}(z) & \mathbf{C} & \\ \hline & \mathbf{I}_{N_A-n} & \\ \hline & \mathbf{B}(z) & \mathbf{I}_n \\ \hline \mathbf{I}_n & \mathbf{I}_{N_B-n} & z\mathbf{I}_n \end{array} \right].$$

Adding to the third block column, the last block column multiplied by $-B(z)$, we get

$$(1.15) \quad \left[\begin{array}{c|c|c} \mathbf{A}(z) & \mathbf{C} & \\ \hline & \mathbf{I}_{N_A-n} & \\ \hline & 0 & \mathbf{I}_n \\ & & \mathbf{I}_{N_B-n} \\ \hline \mathbf{I}_n & -z\mathbf{B}(z) & z\mathbf{I}_n \end{array} \right],$$

which is permutationally equivalent to

$$(1.16) \quad \left[\begin{array}{c|c|c} \mathbf{A}(z) & \mathbf{C} & \\ \hline \mathbf{I}_n & -z\mathbf{B}(z) & z\mathbf{I}_n \\ \hline & \mathbf{I}_{N_A-n} & \\ & & \mathbf{I}_{N_B-n} \\ \hline & & \mathbf{I}_n \end{array} \right].$$

Adding to the second block row, the last block row multiplied by $-z$, we get

$$(1.17) \quad \left[\begin{array}{c|c|c} \mathbf{A}(z) & \mathbf{C} & \\ \hline \mathbf{I}_n & -z\mathbf{B}(z) & \\ \hline & \mathbf{I}_{N_A-n} & \\ & & \mathbf{I}_{N_B-n} \\ \hline & & \mathbf{I}_n \end{array} \right].$$

Finally, take into account that

$$(1.18) \quad \begin{bmatrix} \mathbf{I}_n & -\mathbf{A}(z) \\ & \mathbf{I}_n \end{bmatrix} \begin{bmatrix} \mathbf{A}(z) & \mathbf{C} \\ \mathbf{I}_n & -z\mathbf{B}(z) \end{bmatrix} \begin{bmatrix} z\mathbf{B}(z) & \mathbf{I}_n \\ \mathbf{I}_n & \end{bmatrix} = \begin{bmatrix} z\mathbf{A}(z)\mathbf{B}(z) + \mathbf{C} & \\ & \mathbf{I}_n \end{bmatrix}.$$

This completes the proof. □

REMARK 1.8. In the case that $\Sigma_{\mathbf{A}} = \Sigma_{\mathbf{A}} = \mathbb{F}$, then the intersection is also \mathbb{F} , and this establishes that ‘algebraic linearizations’ are linearizations if their components are linearizations.

REMARK 1.9. The generalized standard triple for the algebraic linearization given here has $\mathbf{X}_H = [0, 0, \mathbf{X}_B]$ and $\mathbf{Y}_H = [\mathbf{Y}_A^T, 0, 0]^T$. For these linearizations, there is no notion of expressing 1 as a linear combination of anything, because this formulation is independent of particular polynomial bases, and indeed may use different bases for different submatrices.

Algebraic linearizations offer a new class of linearizations. In [7] examples are given where the eigenvalue conditioning of such linearizations is *exponentially* better than that of the Frobenius linearization; this demonstrates that this class of linearizations potentially offers more numerically stable algorithms³ for computing matrix polynomial eigenvalues than standard linearizations do. The extent to which this is possible in general has not yet been explored: at this point, we only know that this *can* happen for some cases.

The recursive construction of algebraic linearizations relies on generalized standard triples of each of the component matrix polynomials, and (as does the unrelated paper [32]) allows different polynomial bases to be used for each component. This present note provides some explicit formulas for generalized standard

³Any algorithm that transforms a well-conditioned problem into an ill-conditioned one as a step along the way is likely to be a numerically unstable algorithm.

triples in various bases, for reference. As one reviewer points out, these formulas *could* simply be obtained by reading the proofs that these linearizations are indeed linearizations; one purpose of this paper is simply convenience.

2. Expressing 1 in the basis gives the triple. If the $\phi_k(z)$, $0 \leq k \leq \ell - 1$ form a basis, we may express the polynomial 1 in that basis: then $1 = \sum_{k=0}^{\ell-1} e_k \phi_k(z)$ defines the coefficients e_k uniquely. Putting

$$(2.19) \quad \mathbf{X} = [e_{\ell-1} \ e_{\ell-2} \ \cdots \ e_1 \ e_0] \otimes \mathbf{I}_n,$$

for an appropriate choice of basis always gives our generalized standard triple $\mathbf{P}^{-1}(z) = \mathbf{X}(z\mathbf{C}_1 - \mathbf{C}_0)^{-1}\mathbf{Y}$ with

$$(2.20) \quad \mathbf{Y} = [\mathbf{I}_n \ \mathbf{0}_n \ \mathbf{0}_n \ \cdots \ \mathbf{0}_n]^T = \mathbf{e}_1 \otimes \mathbf{I}_n.$$

[Here the notation \mathbf{e}_1 means the first elementary vector: although printed in bold type, it does not look very different from the scalar e_1 , not bold, printed earlier as a component of \mathbf{X} .] We prove this below for all the elementary linearizations we use in this paper. Moreover, if we replace \mathbf{Y} above with $\mathbf{v} \otimes \mathbf{I}_n$ for a general ansatz column vector \mathbf{v} , then the theorem is also true for all the linearizations of [17] as well.

THEOREM 2.1. *Let $\mathbf{P}(z)$ be a regular matrix polynomial expressed in terms of a polynomial basis $\{\phi_i(z)\}$ i.e. $\mathbf{P}(z) = \sum_{k=0}^{\ell} \mathbf{P}_k \phi_k(z)$. Consider a linearization $\mathbf{L}(z)$ of $\mathbf{P}(z)$ such that*

$$(2.21) \quad \mathbf{L}(z) (\Phi_{\ell}(z) \otimes \mathbf{I}_m) = (\mathbf{e}_1 \otimes \mathbf{I}_m) \mathbf{P}(z),$$

where $\mathbf{e}_1 = [1 \ 0 \ \cdots \ 0]^T \in \mathbb{C}^{\ell}$ and $\Phi_{\ell}(z) = [\phi_{\ell-1}(z) \ \cdots \ \phi_0(z)]^T$. Let x be a vector such that $x\Phi_{\ell}(z) = 1$ and define $\mathbf{X} = x \otimes \mathbf{I}_m$ and $\mathbf{Y} = \mathbf{e}_1 \otimes \mathbf{I}_m$. Then

$$(2.22) \quad \mathbf{P}(z)^{-1} = \mathbf{X}\mathbf{L}(z)^{-1}\mathbf{Y}.$$

Proof. The following proof, which uses an idea of an anonymous referee, is simpler than our original one. For each of the polynomial bases we examine in this paper, the linearization satisfies either

$$(2.23) \quad \mathbf{L}(z) \begin{bmatrix} \phi_{\ell-1}(z)\mathbf{I}_n \\ \phi_{\ell-2}(z)\mathbf{I}_n \\ \vdots \\ \phi_0(z)\mathbf{I}_n \end{bmatrix} = \begin{bmatrix} \mathbf{I}_n \\ \mathbf{0}_n \\ \vdots \\ \mathbf{0}_n \end{bmatrix} \mathbf{P}(z),$$

for degree-graded bases, or similar statements for Bernstein bases and Lagrange and Hermite interpolational bases, as follows. For the Bernstein basis, the polynomial elements in the vector on the left are multiples of $B_j^{\ell-1}(z)$: $[\ell/1 \cdot B_{\ell-1}^{\ell-1}(z), \ell/2 \cdot B_{\ell-1}^{\ell-1}(z), \dots, \ell/\ell \cdot B_0^{\ell-1}(z)]^T$. For the Lagrange basis, the vector on the left is $[w(z), \ell_0(z), \ell_1(z), \dots, \ell_{\ell}(z)]^T$. For the Hermite interpolational basis⁴, it is the same as for the Lagrange but with the Lagrange basis elements replaced with the Hermite interpolational basis elements.

Premultiplying by $\mathbf{L}^{-1}(z)$ and post-multiplying by $\mathbf{P}^{-1}(z)$, we have

$$(2.24) \quad \mathbf{L}^{-1}(z) \begin{bmatrix} \mathbf{I}_n \\ \mathbf{0}_n \\ \vdots \\ \mathbf{0}_n \end{bmatrix} = \begin{bmatrix} \phi_{\ell-1}(z)\mathbf{I}_n \\ \phi_{\ell-2}(z)\mathbf{I}_n \\ \vdots \\ \phi_0(z)\mathbf{I}_n \end{bmatrix} \mathbf{P}^{-1}(z).$$

⁴The paper [10] did not prove that the Hermite companion pencil discussed here is in fact a linearization. We believe that it is, but there is no proof published yet, to our knowledge.

If $1 = \sum_{k=0}^{\ell-1} e_k \phi_k(z)$ is the expression of 1 in that basis, then premultiplying both sides by

$$\mathbf{X} = [e_{\ell-1} \mathbf{I}_n \quad e_{\ell-2} \mathbf{I}_n \quad \dots \quad e_0 \mathbf{I}_n],$$

gives the theorem. Compare also Remark 1.2 which gives another formula for \mathbf{X} and \mathbf{Y} . ‡

REMARK 2.2. This result can be extended slightly using the results of [17], and their use of a general *ansatz* vector \mathbf{v} generalizes our results, in the following way: They consider the set of all *matrix pencils* $\mathbf{L}(z)$ satisfying

$$(2.25) \quad \mathbf{L}(z) (\Phi \otimes \mathbf{I}_n) = \mathbf{v} \otimes \mathbf{P}(z),$$

and later characterize just which of these matrix pencils are linearizations. They did not explicitly consider Lagrange or Hermite interpolational basis polynomials, or the Bernstein basis, but as they point out their proofs go through unchanged for matrix pencils where z appears only on the diagonal of the matrix pencil, as it does for all the cases we consider here except the Bernstein case. Their use of the general *ansatz* vector (all the examples we have considered just use $\mathbf{v} = \mathbf{e}_1$) extends our result to the case $\mathbf{Y} = \mathbf{v} \otimes \mathbf{I}_n$. With the appropriate changes, the proof above goes through.

REMARK 2.3. There are linearizations not explicitly considered in this paper; for instance, a referee has pointed out that when a matrix polynomial is expressed in a basis where the elements satisfy a linear recurrence, then there is an automatic way to build what is called a CORK linearization. See [21] and [35] for details.

In what follows we examine specific cases in detail and supply specific proofs for each basis. Indeed, much of the utility of this paper is simply writing down those details, which will allow easier programming for the uses of these generalized standard triples.

3. Scalar examples of generalized standard triples. In this section, we tabulate generalized standard triples for four classes of linearizations. Note that the generalized standard triples will be given for (scalar) polynomials of degree 5.

In the special case $n = 1$ and when the monomial basis is used, a linearization is usually simplified by dividing by the leading coefficient, making the result monic and the second matrix of the pair just becomes the identity. The remaining matrix is called a ‘companion matrix’ or Frobenius companion⁵. Thus, finding roots of a scalar polynomial can be done by finding eigenvalues of the companion matrix. Kublanovskaya calls these ‘accompanying pencils’ in [25].

Construction of a linearization from a companion matrix is, when possible at all, a simple matter of the Kronecker (tensor) product: given $\mathbf{C}_1, \mathbf{C}_0 \in \mathbb{C}^{\ell \times \ell}$, take $\widetilde{\mathbf{C}}_1 = \mathbf{C}_1 \otimes \mathbf{I}_n$ and then replace each block $p_k \mathbf{I}_n$ with the corresponding matrix coefficient $\mathbf{P}_k \in \mathbb{C}^{n \times n}$ (the first p_k , in $p_k \mathbf{I}_n$, is the symbolic coefficient from $p(z) = \sum_{k=0}^{\ell} p_k \phi_k(z)$; the matrix coefficient $\mathbf{P}_k \in \mathbb{C}^{n \times n}$ is from $\mathbf{P}(z) = \sum_{k=0}^{\ell} \mathbf{P}_k \phi_k(z)$.) This will be clearer by example.

3.1. Bases with three-term recurrence relations. The monomial basis, the shifted monomial basis, the Taylor basis, the Newton interpolational bases, and many common orthogonal polynomial bases all have three-term recurrence relations that, except for initial cases, can be written

⁵The *Frobenius form* of a matrix is related, but different: see for instance [33].

TABLE 1

A short list of three-term recurrence relations for some important polynomial bases discussed in Section 3.1. For a more comprehensive list, see *The Digital Library of Mathematical Functions*. These relations and others are coded in Walter Gautschi's packages *OPQ* and *SOPQ* [18] and in the *MatrixPolynomialObject* implementation package in Maple (see [22]).

$\phi_k(z)$	Name	α_k	β_k	γ_k	ϕ_0	ϕ_1
z^k	Monomial	1	0	0	1	z
$(z - a)^k$	Shifted monomial	1	a	0	1	$z - a$
$(z - a)^k/k!$	Taylor	$n + 1$	a	0	1	$z - a$
$\prod_{j=0}^{k-1} (z - \tau_j)$	Newton interpolational	1	τ_n	0	1	$z - \tau_0$
$T_k(z) = \cos(k \cos^{-1}(z))$	Chebyshev	$1/2$	0	$1/2$	1	z
$P_k(z)$	Legendre	$(k + 1)/(2k + 1)$	0	$k/(2k + 1)$	1	z

$$(3.26) \quad z\phi_k(z) = \alpha_k\phi_{k+1}(z) + \beta_k\phi_k(z) + \gamma_k\phi_{k-1}(z).$$

In all cases, we have $\alpha_k \neq 0$. For instance, the Chebyshev polynomial recurrence is usually written $T_{n+1}(z) = 2zT_n(z) - T_{n-1}(z)$ but is easily rewritten in the above form by isolating $zT_n(z)$, and all Chebyshev $\alpha_k = 1/2$ for $k > 1$. We give a selection in Table 1, and refer the reader to Section 18.9 of the Digital Library of Mathematical Functions (dlmf.nist.gov) for more. See also [18].

For all such bases, we have the linearization⁶

$$(3.27) \quad z\mathbf{C}_1 - \mathbf{C}_0 = z \left[\begin{array}{c|cccc} \frac{p_5}{\alpha_4} & & & & \\ \hline & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 1 \end{array} \right] - \left[\begin{array}{c|cccc} -p_4 + \frac{\beta_4}{\alpha_4}p_5 & -p_3 + \frac{\gamma_4}{\alpha_4}p_5 & -p_2 & -p_1 & -p_0 \\ \hline & \alpha_3 & \beta_3 & \gamma_3 & \\ & & \alpha_2 & \beta_2 & \gamma_2 \\ & & & \alpha_1 & \beta_1 & \gamma_1 \\ & & & & \alpha_0 & \beta_0 \end{array} \right],$$

(remember that $\alpha_k \neq 0$) and

$$(3.28) \quad \mathbf{X} = [0 \ 0 \ 0 \ 0 \ 1],$$

$$(3.29) \quad \mathbf{Y} = [1 \ 0 \ 0 \ 0 \ 0]^T.$$

As an example showing how to extend this companion matrix to a linearization of a matrix polynomial, a Newton interpolational basis on the nodes $\tau_0, \tau_1, \dots, \tau_5$ has the linearization for matrix polynomials of this grade,

$$(3.30) \quad z \begin{bmatrix} \mathbf{P}_5 & & & & \\ & \mathbf{I}_n & & & \\ & & \mathbf{I}_n & & \\ & & & \mathbf{I}_n & \\ & & & & \mathbf{I}_n \end{bmatrix} - \begin{bmatrix} -\mathbf{P}_4 + \tau_4\mathbf{P}_5 & -\mathbf{P}_3 & -\mathbf{P}_2 & -\mathbf{P}_1 & -\mathbf{P}_0 \\ & \mathbf{I}_n & \tau_3\mathbf{I}_n & & \\ & & \mathbf{I}_n & \tau_2\mathbf{I}_n & \\ & & & \mathbf{I}_n & \tau_1\mathbf{I}_n \\ & & & & \mathbf{I}_n & \tau_0\mathbf{I}_n \end{bmatrix}.$$

3.2. The Bernstein basis. The set of polynomials $\{B_k^\ell(z)\}_{k=0}^\ell = \binom{\ell}{k} z^k(1 - z)^{\ell-k}$ is a set of $\ell + 1$ polynomials each of exact degree ℓ that together forms a basis for polynomials of grade ℓ . Bernstein

⁶For exposition, we follow Peter Lancaster's dictum, namely that the 5×5 case almost always gives the idea.

polynominals have many applications, for example, in Computer-Aided Geometric Design (CAGD), and many important properties including that of optimal condition number over all bases positive on $[0, 1]$. They do not satisfy a simple three term recurrence relation of the form discussed in Section 3.1, although they satisfy an interesting and useful ‘degree-elevation’ recurrence, namely

$$(3.31) \quad (j + 1)B_{j+1}^\ell(z) + (\ell - j)B_j^\ell(z) = \ell B_j^{\ell-1}(z),$$

which specifically demonstrates that a sum of Bernstein polynomials of degree ℓ might have degree strictly less than ℓ . See [14], [15], and [16] for more details of Bernstein bases.

A Bernstein companion pencil for $p_5(z) = \sum_{k=0}^5 p_k B_k^5(z)$ is

$$(3.32) \quad z\mathbf{C}_1 - \mathbf{C}_0 = z \begin{bmatrix} -p_4 + \frac{1}{5}p_5 & -p_3 & -p_2 & -p_1 & -p_0 \\ 1 & \frac{2}{4} & & & \\ & 1 & \frac{3}{3} & & \\ & & 1 & \frac{4}{2} & \\ & & & 1 & \frac{5}{1} \end{bmatrix} - \begin{bmatrix} -p_4 & -p_3 & -p_2 & -p_1 & -p_0 \\ 1 & 0 & & & \\ & 1 & 0 & & \\ & & 1 & 0 & \\ & & & 1 & 0 \end{bmatrix},$$

$$(3.33) \quad \mathbf{X} = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 5 & 5 & 5 & 5 \end{bmatrix},$$

$$(3.34) \quad \mathbf{Y} = [1 \ 0 \ 0 \ 0 \ 0]^T.$$

For a construction of rational $\mathbf{E}(z)$ and $\mathbf{F}(z)$ that show this is a local linearization (unless $z = 1$ is an eigenvalue), see [2]. The paper [30] goes further and gives a general method of constructing all strong linearizations for Bernstein matrix polynomials, including this one (which they show is strictly equivalent to the second companion form for the monomial basis, without any restrictions on z or on the leading coefficient). We have $p^{-1}(z) = \mathbf{X}(z\mathbf{C}_1 - \mathbf{C}_0)^{-1}\mathbf{Y}$ if $p(z) \neq 0$. This pencil was first analyzed in [23] and [24]. One of the present authors independently invented and implemented a version of this companion pencil in Maple (except using $\mathbf{P}^T(z)$, and flipped from the above form) in about 2004. For a proof of numerical stability, see the original thesis [23]. The standard triple is, we believe, new to this paper.

EXAMPLE 3.1 (Singular leading coefficient case).

$$(3.35) \quad \mathbf{P}(z) = \begin{bmatrix} 29/100 & -8/25 \\ 7/10 & -1/100 \end{bmatrix} B_0^3(z) + \begin{bmatrix} -41/50 & 41/100 \\ -7/10 & 91/100 \end{bmatrix} B_1^3(z) \\ + \begin{bmatrix} 9/10 & 19/100 \\ 4/5 & 22/25 \end{bmatrix} B_2^3(z) + \begin{bmatrix} 1 & 1 \\ 9851/1980 & 0 \end{bmatrix} B_3^3(z).$$

Expressing equation (3.35) into the monomial basis, we have

$$\mathbf{P}(z) = \begin{bmatrix} 29/100 & -8/25 \\ 7/10 & -1/100 \end{bmatrix} + \begin{bmatrix} 29/100 & -8/25 \\ 7/10 & -1/100 \end{bmatrix} z + \begin{bmatrix} 849/100 & -57/20 \\ 87/10 & -57/20 \end{bmatrix} z^2 + \begin{bmatrix} -89/20 & 99/50 \\ -89/396 & 1/10 \end{bmatrix} z^3.$$

Taking the determinant of the leading coefficient

$$\det \left(\begin{bmatrix} -89/20 & 99/50 \\ -89/396 & 1/10 \end{bmatrix} \right) = (-89/20)(1/10) - (99/50)(-89/396) = 0,$$

we can observe that leading coefficient is singular, and thus, this matrix polynomial cannot be made monic. The standard triple for equation (3.35) is

$$\begin{aligned}
 C_0 &= \begin{bmatrix} -9/10 & -19/100 & 41/50 & -41/100 & -29/100 & 8/25 \\ -4/5 & -22/25 & 7/10 & -91/100 & -7/10 & 1/100 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \\
 C_1 &= \begin{bmatrix} -17/30 & 43/300 & 41/50 & -41/100 & -29/100 & 8/25 \\ 5099/5940 & -22/25 & 7/10 & -91/100 & -7/10 & 1/100 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 & 0 & 3 \end{bmatrix} \\
 X &= \begin{bmatrix} 1/3 & 0 & 2/3 & 0 & 1 & 0 \\ 0 & 1/3 & 0 & 2/3 & 0 & 1 \end{bmatrix} \quad Y = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix},
 \end{aligned}$$

because

$$X(zC_1 - C_0)^{-1}YP(z) = I_2.$$

3.3. The Lagrange interpolational basis. There are by now several Lagrange basis linearizations. The use of barycentric forms means that Lagrange interpolation is efficient and numerically stable and is increasing in popularity [3]. Here is the definition of the first barycentric form for interpolation of polynomials of grade ℓ on the $\ell + 1$ distinct nodes $\tau_k \in \mathbb{C}$, $0 \leq k \leq \ell$. Take the partial fraction decomposition of the reciprocal of the node polynomial

$$(3.36) \quad w(z) = \prod_{k=0}^{\ell} (z - \tau_k),$$

namely

$$(3.37) \quad \frac{1}{w(z)} = \sum_{k=0}^{\ell} \frac{\beta_k}{z - \tau_k}$$

where the coefficients β_k occurring in the partial fraction decomposition are called the *barycentric weights*. A well-known explicit formula for the β_k is

$$(3.38) \quad \beta_k = \prod_{\substack{j=0 \\ j \neq k}}^{\ell} (\tau_k - \tau_j)^{-1}.$$

The Lagrange basis polynomials are normally written

$$(3.39) \quad \ell_k(z) = \beta_k \prod_{\substack{j=0 \\ j \neq k}}^{\ell} (z - \tau_j).$$

For many sets of nodes (Chebyshev nodes on $[-1, 1]$, or roots of unity on the unit disk), the resulting interpolant is also well-conditioned, and can even be ‘better than optimal’ [11], see also [6]. The pencil we use here is ‘too large’ and has (numerically harmless in our experience) spurious roots at infinity⁷; for alternative formulations, see [35, 31]. Then, the pencil is $z\mathbf{C}_1 - \mathbf{C}_0$ where

$$(3.40) \quad z\mathbf{C}_1 - \mathbf{C}_0 = z \begin{bmatrix} 0 & & & & & \\ & 1 & & & & \\ & & 1 & & & \\ & & & 1 & & \\ & & & & 1 & \\ & & & & & 1 \end{bmatrix} - \begin{bmatrix} 0 & -\rho_0 & -\rho_1 & -\rho_2 & -\rho_3 & -\rho_4 \\ \beta_0 & \tau_0 & & & & \\ \beta_1 & & \tau_1 & & & \\ \beta_2 & & & \tau_2 & & \\ \beta_3 & & & & \tau_3 & \\ \beta_4 & & & & & \tau_4 \end{bmatrix}.$$

Rational matrices $\mathbf{E}(z)$ and $\mathbf{F}(z)$ demonstrating that this is indeed a local linearization, except when the nodes are eigenvalues, can be found in the appendix of [2]. More is true: by a block permutation argument, one can find constant matrices \mathbf{E}_k and \mathbf{F}_k with determinant ± 1 such that $\mathbf{E}_k \mathbf{L}(\tau_k) \mathbf{F}_k = \text{diag}(\mathbf{P}_k, \mathbf{I}_{N-n})$, for $0 \leq k \leq \ell$. Interpolating these matrices polynomially (indeed, we already know the correct barycentric weights) to form polynomial matrices $\mathbf{E}(z)$ and $\mathbf{F}(z)$, we have $\mathbf{E}(z) \mathbf{L}(z) \mathbf{F}(z) = \text{diag}(\mathbf{P}(z), \mathbf{I}_{N-m})$, and moreover because the determinants are ± 1 exactly at the nodes (and therefore nonzero in certain small neighbourhoods of the nodes) we see that $\mathbf{L}(z)$ and $\mathbf{P}(z)$ are equivalent on a set Σ (not necessarily simply connected) that contains the nodes τ_k . Since $\mathbf{L}(z)$ and $\mathbf{P}(z)$ are also equivalent on a set that contains everything *except* the nodes, we see by Proposition 2.1 of [13] that $\mathbf{L}(z)$ is a linearization of $\mathbf{P}(z)$.

The \mathbf{X} and \mathbf{Y} for the standard triple are

$$(3.41) \quad \mathbf{X} = [0 \ 1 \ 1 \ 1 \ 1 \ 1],$$

$$(3.42) \quad \mathbf{Y} = [1 \ 0 \ 0 \ 0 \ 0 \ 0]^T.$$

Notice in this case that for the linearization $N = (\ell + 2)n$ while $\deg p \leq \ell$, and therefore, there are at least $2n$ eigenvalues at infinity. This can be inconvenient if n is at all large.

3.4. Hermite interpolational basis. The Lagrange linearization of the previous section has been extended to Hermite interpolational bases, where some of the nodes have ‘flowed together’, collapsing to fewer distinct nodes⁸.

⁷This numerical harmlessness needs some explanation. In brief, Lagrange basis matrix polynomial eigenvalues will be well-conditioned only in a compact region determined by the interpolation nodes and are increasingly ill-conditioned toward infinity; in practice, this means only small changes in the data are needed to perturb large finite ill-conditioned eigenvalues out to infinity. Any eigenvalues produced numerically that are well outside the region determined by the interpolation nodes are likely easily perturbed all the way to infinity and can be safely ignored.

⁸A formal definition can be found in [8], for instance. The essential idea is that given two distinct pieces of data $(\tau_k, p(\tau_k))$ and $(\tau_{k+1}, p(\tau_{k+1}))$, we also know the forward difference $(p_{k+1} - p_k)/(\tau_{k+1} - \tau_k)$. In the limit as one node approaches (flows towards) the other, we still know two pieces of information: $p(\tau_k)$ and $p'(\tau_k)$. Hermite interpolation captures this idea.

We suppose that at each remaining distinct node τ_i , $0 \leq i \leq N - 1$, say, there are now $s_i \geq 1$ consecutive pieces of information known, namely $\mathbf{P}(\tau_i)$, $\mathbf{P}'(\tau_i)/1!$, $\mathbf{P}''(\tau_i)/2!$, and so on up to the last one, the value of the $s_i - 1$ -th derivative at $z = \tau_i$, namely $\mathbf{P}^{(s_i-1)}(\tau_i)/(s_i - 1)!$. The integer s_i is called the *confluency* of the node. The known pieces of information are the local Taylor coefficients of the polynomial fitting the data:

$$(3.43) \quad \rho_{i,j} = \frac{f^{(j)}(\tau_i)}{j!}, \quad 0 \leq j \leq s_i - 1.$$

This gives $1 + \ell = \sum s_i$ pieces of information, determining a polynomial of grade ℓ . The barycentric weights, this time doubly indexed as $\beta_{i,j}$, are again computed from the partial fraction decomposition of the reciprocal of the node polynomial

$$(3.44) \quad \frac{1}{w(z)} = \frac{1}{\prod_{i=0}^{N-1} (z - \tau_i)^{s_i}} = \sum_{i=0}^{N-1} \sum_{j=0}^{s_i-1} \frac{\beta_{i,j}}{(z - \tau_i)^{j+1}}.$$

For evaluation of the interpolating polynomial, one should use the first or second barycentric form; see [8] for details. For theoretical work with the Hermite interpolational bases, however, we can define

$$(3.45) \quad H_{i,j}(z) = \sum_{k=0}^{s_i-1-j} \beta_{i,j+k} w(z) (z - \tau_i)^{-k-1}.$$

These polynomials, each of degree ℓ , form a basis (a Hermite interpolational basis, to distinguish from the Hermite orthogonal polynomials) for polynomials of grade ℓ ; moreover, they generalize the Lagrange property in that only one Taylor coefficient at only one node is 1 and all the rest are zero.

Note that the derivative $\mathbf{P}'(z)$ of a matrix polynomial is a straightforward extension to matrices of the ordinary derivative. It is isomorphic to the matrix with entries that are the ordinary derivatives of the original matrix.

The linearization of the previous section changes to the following elegant form. The matrix \mathbf{C}_1 is unchanged,

$$(3.46) \quad \mathbf{C}_1 = \begin{bmatrix} 0 & & & & & \\ & 1 & & & & \\ & & \ddots & & & \\ & & & 1 & & \\ & & & & 1 & \\ & & & & & 1 \end{bmatrix},$$

being $(\ell + 2)$ by $(\ell + 2)$ as before. The matrix \mathbf{C}_0 changes, picking up transposed Jordan-like blocks for each distinct node. For instance, suppose we have two distinct nodes, τ_0 and τ_1 . Suppose further that τ_0 has confluency $s_0 = 3$ while τ_1 has confluency $s_1 = 2$. This means that we know $f(\tau_0)$, $f'(\tau_0)/1!$, $f''(\tau_0)/2!$, $f(\tau_1)$ and $f'(\tau_1)/1!$. Then,

$$(3.47) \quad \mathbf{C}_0 = \begin{bmatrix} 0 & -f''(\tau_0)/2! & -f'(\tau_0)/1! & -f(\tau_0) & -f'(\tau_1)/1! & -f(\tau_1) \\ \beta_{02} & \tau_0 & & & & \\ \beta_{01} & 1 & \tau_0 & & & \\ \beta_{00} & & 1 & \tau_0 & & \\ \beta_{11} & & & & \tau_1 & \\ \beta_{10} & & & & 1 & \tau_1 \end{bmatrix}.$$

Here, $p_k = p^{(k)}(\tau_0)/k!$ is the ordinary coefficient in the expansion $p(z) = \sum_{k=0}^{\ell} p_k(z - \tau_0)^k$. The numerical stability of these Hermite interpolational linearization has been studied briefly [26] but much remains unknown. We confine ourselves in this paper to the study of the standard triple.

To make a linearization for matrix polynomials out of these scalar linearizations, take the Kronecker tensor product with I_n and insert the appropriate matrix polynomial values and derivative values.

EXAMPLE 3.3 (Matrix polynomial case). Let

$$\tau = [0, 1],$$

and

$$\begin{array}{c|c|c} z & P(z) & P'(z) \\ \hline \tau_0 = 0 & \begin{bmatrix} -1 & 0 \\ -1 & 1 \end{bmatrix} & \\ \hline \tau_1 = 1 & \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} & \begin{bmatrix} 1 & -1 \\ -1 & 0 \end{bmatrix} \end{array}.$$

Then, the standard triple is

$$C_0 = \begin{bmatrix} 0 & 0 & -1 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 1 & 1 & -1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad C_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$X = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix} \quad Y = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

The Hermite interpolating polynomial is

$$P(z) = \begin{bmatrix} z - 1 & -2z^2 + 3z \\ -3z^2 + 5z - 1 & 2z^2 - 4z + 1 \end{bmatrix},$$

and the resolvent form is

$$X(zC_1 - C_0)^{-1}Y = \begin{bmatrix} \frac{-2z^2 + 4z - 1}{6z^4 - 21z^3 + 23z^2 - 8z + 1} & \frac{-2z^2 + 3z}{6z^4 - 21z^3 + 23z^2 - 8z + 1} \\ \frac{-3z^2 + 5z - 1}{6z^4 - 21z^3 + 23z^2 - 8z + 1} & \frac{-z + 1}{6z^4 - 21z^3 + 23z^2 - 8z + 1} \end{bmatrix}.$$

Then,

$$\mathbf{X}(z\mathbf{C}_1 - \mathbf{C}_0)^{-1}\mathbf{Y}\mathbf{P}(z) = \mathbf{I}_2,$$

which indicates that the standard triples is correct.

REMARK 3.4. The modified linearizations of [35] also have standard triples that can be used for algebraic linearization and arguably should be tabled here as well. They have the advantage of including fewer eigenvalues at infinity, or no spurious eigenvalues at infinity, which may lead to better algebraic linearizations. However, they are more involved, and we have less numerical experience with them. In particular, we do not understand their dependence on the ordering of the nodes, and so we leave their analysis to a future study.

4. Concluding remarks. The generalized standard triples (or standard quadruples, if you prefer) that we propose in this paper for convenience in algebraic linearization may have other uses. As pointed out on p. 28 of [20], many of the properties stated in that work for monic polynomials are valid for non-monic polynomials with the appropriate changes made. Some caution with the results of this paper is thus mandated.

We have here *defined* these generalized standard triples simply by the resolvent representation for the matrix polynomial equation (1.7), and only for linearizations, which is all we need for algebraic linearization.

The main theorem of the paper, namely Theorem 2.1, gives a universal way to construct this generalized standard triple for polynomial matrices written in terms of any polynomial basis, given a linearization satisfying equation 2.21. We also gave explicit instructions for this construction using any of several polynomial bases, for convenience, which may give insight for further work in this area.

In Section 1.3, we gave an explicit construction for matrices $\mathbf{E}_H(z)$ and $\mathbf{F}_F(z)$ showing that *algebraic* linearizations are, in fact, local linearizations, with $\mathbf{E}_H(z)(z\mathbf{B}_H - \mathbf{A}_H)\mathbf{F}_H(z) = \text{diag}(\mathbf{P}(z), \mathbf{I}, \dots, \mathbf{I})$.

In [10] we find new constructions for matrices $\mathbf{E}(z)$ and $\mathbf{F}(z)$ which likewise show that the companions for the Lagrange and Hermite interpolational bases are, in fact, linearizations (of matrix polynomials of higher grade). A proof for Lagrange interpolational bases was given already in [2], where indeed the linearization was proved to be strong, but the result for Hermite interpolational bases has not yet, to our knowledge, been published. Indeed while we claim that the standard triple given here works for Hermite interpolational bases, we have not given any details of a proof, and we leave that proof for a future paper.

Acknowledgments. We acknowledge the support of Western University, The National Science and Engineering Research Council of Canada, the Ontario Graduate Scholarship (OGS) program, the University of Alcalá, the Ontario Research Centre of Computer Algebra, and the Rotman Institute of Philosophy. Part of this work was developed while RMC was visiting the University of Alcalá, in the frame of the project Giner de los Rios. We thank Françoise Tisseur for her thorough comments on an earlier version of this paper, and likewise Froilán Dopico for a similarly fruitful discussion. Finally, we thank an anonymous referee for improving (and correcting!) several proofs.

REFERENCES

- [1] M. Al-Ammari and F. Tisseur. Standard triples of structured matrix polynomials. *Linear Algebra Appl.*, 437(3):817–834, 2012.
- [2] A. Amiraslani, R. M. Corless, and P. Lancaster. Linearization of matrix polynomials expressed in polynomial bases. *IMA J. Numer. Anal.*, 29(1):141–157, 2008.

- [3] J.-P. Berrut and L. N. Trefethen. Barycentric Lagrange interpolation. *SIAM Rev.*, 46(3):501–517, 2004.
- [4] T. Betcke, N.J. Higham, V. Mehrmann, C. Schröder, and F. Tisseur. NLEVP: A collection of nonlinear eigenvalue problems. *ACM Trans. Math. Softw. (TOMS)*, 39(2):7, 2013.
- [5] M.I.B. Cachadina, J. Perez, A. Akshar, D. Mileeva, and R. Kassem. Linearizations for interpolatory bases – A comparison: New families of linearizations. *Electron. J. Linear Algebra*, 36(36):799–833, 2020.
- [6] J.M. Carnicer, Y. Khifar, and J.M. Peña. Optimal stability of the Lagrange formula and conditioning of the Newton formula. *J. Approx. Theory*, 2017.
- [7] E.Y.S. Chan, R.M. Corless, L. Gonzalez-Vega, J. Rafael Sendra, and J. Sendra. Algebraic linearizations for matrix polynomials. *Linear Algebra Appl.*, 563:373–399, 2019.
- [8] R.M. Corless and N. Fillion. Polynomial and rational interpolation. In: *A Graduate Introduction to Numerical Methods*. Springer, 331–401, 2013.
- [9] R.M. Corless, N. Rezvani, and A. Amiraslani. Pseudospectra of matrix polynomials that are expressed in alternative bases. *Math. Comput. Sci.*, 1(2):353–374, 2007.
- [10] R.M. Corless, L.R. Severyi, and B.D. Saunders. Equivalences for linearizations of matrix polynomials. In: *Proceedings of the 2021 on International Symposium on Symbolic and Algebraic Computation*. ACM, July 2021.
- [11] R.M. Corless and S.M. Watt. Bernstein bases are optimal, but, sometimes, Lagrange bases are better. In: *Proceedings of SYNASC, Timisoara*. MIRTON Press, 141–153, 2004.
- [12] F.M. Dopico, P.W. Lawrence, J. Pérez, and P. Van Dooren. Block Kronecker linearizations of matrix polynomials and their backward errors. *Numer. Math.*, 140(2):373–426, 2018.
- [13] F.M. Dopico, S. Marcaida, M.C. Quintana, and P. Van Dooren. Local linearizations of rational matrices with application to rational approximations of nonlinear eigenvalue problems. *Linear Algebra Appl.*, 604:441–475, 2020.
- [14] R.T. Farouki. The Bernstein polynomial basis: A centennial retrospective. *Comput. Aided Geometr. Des.*, 29(6):379–419, 2012.
- [15] R.T. Farouki and T. Goodman. On the optimal stability of the Bernstein basis. *Math. Computat. Amer. Math. Soc.*, 65(216):1553–1566, 1996.
- [16] R.T. Farouki and V.T. Rajan. On the numerical condition of polynomials in Bernstein form. *Comput. Aided Geometr. Des.*, 4(3):191–216, 1987.
- [17] H. Faßbender and P. Saltenberger. On vector spaces of linearizations for matrix polynomials in orthogonal bases. *Linear Algebra Appl.*, 525:59–83, 2017.
- [18] W. Gautschi. *Orthogonal Polynomials in MATLAB: Exercises and Solutions*, vol. 26. SIAM, 2016.
- [19] I. Gohberg, P. Lancaster, and L. Rodman. *Indefinite Linear Algebra and Applications*. Springer, 2005.
- [20] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials*. SIAM Classics in Applied Mathematics, 2009.
- [21] S. Güttel and F. Tisseur. The nonlinear eigenvalue problem. *Acta Numer.*, 26:1–94, 2017.
- [22] D.J. Jeffrey and R.M. Corless. Linear algebra in Maple. In: L. Hogben (editor), *Handbook of Linear Algebra*, chapter 89. Chapman and Hall/CRC, 2013.
- [23] G.F. Jónsson. *Eigenvalue Methods for Accurate Solution of Polynomial equations*. PhD thesis, Center for Applied Mathematics, Cornell University, Ithaca, NY, 2001.
- [24] G.F. Jónsson and S. Vavasis. Solving polynomials with small leading coefficients. *SIAM J. Mat. Anal. Appl.*, 26(2):400–414, 2004.
- [25] V.N. Kublanovskaya. Methods and algorithms of solving spectral problems for polynomial and rational matrices. *J. Math. Sci.*, 96(3):3085–3287, 1999.
- [26] P.W. Lawrence and R.M. Corless. Numerical stability of barycentric Hermite root-finding. In: *Proceedings of the 2011 International Workshop on Symbolic-Numeric Computation*. ACM, 2012, 147–148.
- [27] J. Liesen and C. Mehl. Matrix polynomials. In: L. Hogben (editor), *Handbook of Linear Algebra*, chapter 18. Chapman and Hall/CRC, 2013.
- [28] D.S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Vector spaces of linearizations for matrix polynomials. *SIAM J. Mat. Anal. Appl.*, 28(4):971–1004, 2006.
- [29] D.S. Mackey, N. Mackey, and F. Tisseur. Polynomial eigenvalue problems: Theory, computation, and structure. In: *Numerical Algebra, Matrix Theory, Differential-Algebraic Equations and Control Theory*. Springer, 319–348, 2015.
- [30] D.S. Mackey and V. Perović. Linearizations of matrix polynomials in Bernstein bases. *Linear Algebra Appl.*, 501:162–197, 2016.
- [31] Y. Nakatsukasa, V. Noferini, and A. Townsend. Vector spaces of linearizations for matrix polynomials: A bivariate polynomial approach. *SIAM J. Mat. Anal. Appl.*, 38(1):1–29, 2017.
- [32] L. Robol, R. Vandebril, and P. Van Dooren. A framework for structured linearizations of matrix polynomials in various bases. *SIAM J. Mat. Anal. Appl.*, 38(1):188–216, 2017.
- [33] A. Storjohann. An $O(n^3)$ algorithm for the Frobenius normal form. In: *Proceedings of the 1998 International Symposium on Symbolic and Algebraic Computation*, ACM, 101–105 1998.



- [34] O. Taussky and H. Zassenhaus. On the similarity transformation between a matrix and its transpose. *Pacific J. Math.*, 9(3):893–896, 1959.
- [35] R. Van Beeumen, W. Michiels, and K. Meerbergen. Linearization of Lagrange and Hermite interpolating matrix polynomials. *IMA J. Numer. Anal.*, 35(2):909–930, 2015.