Sapienza Università di Roma

Dipartimento di Scienze Statistiche
Sezione di Statistica Economica ed Econometria

Massimo Franchi **Paolo Paruolo**

Normal forms of regular matrix polynomials via local rank factorization

DSS Empirical Economics and Econometrics Working Papers Series

DSS-E3 WP 2011/1

DSS Empirical Economics and Econometrics
Working Papers Series

Dipartimento di Scienze Statistiche
Sezione di Statistica Economica ed Econometria
“Sapienza” Università di Roma
P.le A. Moro 5 – 00185 Roma - Italia

http://www.dss.uniroma1.it

NORMAL FORMS OF REGULAR MATRIX POLYNOMIALS VIA LOCAL RANK FACTORIZATION

MASSIMO FRANCHI* AND PAOLO PARUOLO†

Abstract. The 'local rank factorization' (LRF) of a regular matrix polynomial at an eigenvalue consists of a sequence of matrix rank factorizations of a certain function of its coefficients; the LRF delivers the local Smith form and extended canonical systems of root functions that correspond to the eigenvalue. In this paper it is shown that by performing the LRF at each finite eigenvalue and at infinity one can contruct the Smith form, Jordan triples and decomposable pairs of the matrix polynomial. When $A(\lambda) = A - \lambda B$, where $A, B \in \mathbb{C}^{p\times p}$, the analysis delivers the Kronecker form of $A(\lambda)$ and strict similarity transformations; for $B = I$ one finds the Jordan form of A and Jordan bases.

Key words. Matrix polynomials; spectral theory; canonical forms; Smith form; Kronecker form; Jordan form; Jordan chains; Jordan pairs; Jordan triples

AMS subject classifications. 15A09, 15A21, 47A46, 47A56

1. Introduction.

Normal forms of matrix polynomials play a central role in linear algebra, with important applications in matrix theory [6, 15, 17, 20], in the study of systems of differential and difference equations [9, 10, 16], in system theory [2, 12, 14], as well as in times-series econometrics [3, 4, 11, 13]. The same tools are also employed in various numerical algorithms, such as the ones in [1, 21, 22] for calculating the coefficients of the principal part of the inverse.

A standard reference for the analytic theory of elementary divisors is Gantmacher [6, ch. VI], who reduces a square matrix to its Jordan normal form via the Smith normal form of a matrix polynomial. A second classical reference is Gohberg, Lancaster and Rodman [9], who collect the spectral properties of monic and non-monic matrix polynomials of arbitrary degree into Jordan triples and decomposable pairs and discuss their relations with the Smith and Kronecker normal forms (and other representations).

In this paper we present a theory of reduction of regular matrix polynomials to normal form based on a 'local rank factorization' (LRF). This procedure consists of a sequence of matrix rank factorizations of a certain function of the coefficients of the matrix polynomial and it delivers the local Smith form and extended canonical systems of left and right root functions that correspond to an eigenvalue. The LRF was introduced in Franchi and Paruolo [5] to study the inversion of a regular analytic matrix function; they also showed how these tools can be used to construct canonical sets of left- and right-Jordan chains and local Jordan triples.

In this paper we show that by performing the LRF of regular matrix pencils $A(\lambda) = A - \lambda B$, where $A, B \in \mathbb{C}^{p\times p}$, at each finite eigenvalue and at infinity one can construct the Kronecker form of $A(\lambda)$ and strict similarity transformations; for $B = I$ one finds the Jordan form of A and Jordan bases. This also gives rise to an enlarged set of necessary and sufficient conditions for a matrix to be diagonalizable, some of which are stated in terms of the subspaces associated with the LRF.

The rest of the paper is organized as follows: Sections 2 and 3 introduce notation

*Corresponding author. Università di Roma "La Sapienza", P.le Aldo Moro 5, 00185 Rome, Italy; e-mail: massimo.franchi@uniroma1.it.

†Università dell'Insubria, Via Monte Generoso 71, 21100 Varese, Italy; e-mail: paolo.paruolo@uninsubria.it.

and review the local analysis in [5] (i.e. the LRF, local Smith form and local Jordan triples). Sections 4 and 5 contain the results of the paper. Specifically, Section 4 presents the global analysis (Smith form, Jordan triples and decomposable pairs). Section 5 considers the particular cases $A(\lambda) = A - \lambda B$, with $A, B \in \mathbb{C}^{p\times p}$, or $A \in \mathbb{C}^{p\times p}$, $B = I$. Section 6 contains an example and Section 7 concludes. Proofs are given within the main text, except for the ones taken from [5], which are collected in Appendix A.

1.1. Notation. The following notation is used throughout the paper: by $a := b$ and $b =: a$ we indicate that a is defined by b; any empty sum is defined equal to 0. For any matrix $\varphi \in \mathbb{C}^{p\times r}$, φ' denotes its conjugate transpose. We indicate by $\operatorname{col}\varphi := \{\varphi v, v \in \mathbb{C}^r\}$ the column space of φ and by $\operatorname{col}\varphi'$ the row space of φ; this is in line with current use, see [19, p. 170]. By $\varphi_\perp$ we indicate a basis of $\operatorname{col}^\perp \varphi$, the orthogonal complement of $\operatorname{col}\varphi$ in $\mathbb{C}^p$, where orthogonality is with respect to the standard inner product in $\mathbb{C}^p$, $\langle x, y\rangle := y'x$. The matrix rank factorization of φ is written as $\varphi = -\alpha\beta'$, where α and β are bases of $\operatorname{col}\varphi$ and $\operatorname{col}\varphi'$, see Theorem 1 in [18], and the negative sign is chosen for convenience in the calculations. When φ has full column rank, we set $\bar\varphi := \varphi(\varphi'\varphi)^{-1}$ and $\bar\varphi' := (\bar\varphi)' = (\varphi'\varphi)^{-1}\varphi'$. With this notation the orthogonal projector matrix onto $\operatorname{col}\varphi$ can be written as $P_\varphi := \bar\varphi\varphi' = \varphi\bar\varphi'$, and we denote by $M_\varphi := I - P_\varphi$ the orthogonal projector matrix onto $\operatorname{col}^\perp \varphi$. Finally $\delta_{n,m}$ is Kronecker's delta, i.e. $\delta_{n,m} = 0$ for $n \neq m$ and $\delta_{m,m} = 1$.

2. Definitions. This section reports the definition of 'local rank factorization' (LRF) and states its link with the order of the pole of the inverse function (Theorem 2.3). The stated results are a particular case of the ones proved in [5] for regular analytic matrix functions. Proofs of this section are contained in the Appendix.

2.1. Preliminaries. Consider a regular $p \times p$ matrix polynomial $A(\lambda)$ and let $\sigma(A) := \{\lambda \in \mathbb{C} : \operatorname{rank} A(\lambda) < p\}$ be the set of its eigenvalues. Assume $\sigma(A)$ is non-empty, i.e. $A(\lambda)$ is not unimodular, and observe that for some $\lambda_u \in \sigma(A)$ one may have $A(\lambda_u) = 0$; in this case all the elements of $A(\lambda)$ contain the common factor $(\lambda - \lambda_u)^k, k > 0$, and one can write $A(\lambda) = (\lambda - \lambda_u)^k \widetilde{A}(\lambda)$ with $\widetilde{A}(\lambda_u) \neq 0$. For ease of exposition, we assume henceforth that this common factor simplification has already been performed on $A(\lambda)$, so that $A(\lambda) \neq 0$ for all $\lambda \in \mathbb{C}$. Let $\operatorname{adj} A(\lambda), |A(\lambda)|$ respectively be the adjoint and the determinant of $A(\lambda)$ and let $d(\lambda)$ be the greatest common divisor of all the minors of order $p - 1$ of $A(\lambda)$; then, see Gantmacher [6, p. 90], one has

$$\operatorname{adj} A(\lambda) = d(\lambda)A^\circ(\lambda), \qquad |A(\lambda)| = d(\lambda)a(\lambda),$$

where $A^\circ(\lambda), a(\lambda)$ are respectively called the reduced adjoint and the minimal polynomial of $A(\lambda)$. Since $d(\lambda)$ contains the common factors between the elements of $\operatorname{adj} A(\lambda)$, one has $A^\circ(\lambda) \neq 0$ for all $\lambda \in \mathbb{C}$. Because λ_u is an eigenvalue of $A(\lambda)$ if and only if it is a root of its minimal polynomial, for $\lambda_u \in \sigma(A)$ one finds

$$a(\lambda) = (\lambda - \lambda_u)^m a_u(\lambda), \qquad a_u(\lambda_u) \neq 0, \qquad m > 0, \tag{2.1}$$

where we call m 'the multiplicity of λ_u' (as a root of the minimal polynomial).

The multiplicity m is equal to the order of the pole of $A^{-1}(\lambda)$ at λ_u. In fact, from the identity $A(\lambda) \operatorname{adj} A(\lambda) = \operatorname{adj} A(\lambda)A(\lambda) = |A(\lambda)|I$, one finds

$$A(\lambda)A^\circ(\lambda) = A^\circ(\lambda)A(\lambda) = (\lambda - \lambda_u)^m a_u(\lambda)I, \qquad A(\lambda_u), A^\circ(\lambda_u), a_u(\lambda_u) \neq 0; \tag{2.2}$$

this implies

$$A^{-1}(\lambda) = \frac{A^{\circ}(\lambda)}{(\lambda - \lambda_u)^m a_u(\lambda)},$$

so that $0 < \lim_{\lambda\to\lambda_u} ||(\lambda-\lambda_u)^m A^{-1}(\lambda)|| < \infty$ and $\lim_{\lambda\to\lambda_u} ||(\lambda-\lambda_u)^{m-1} A^{-1}(\lambda)|| = \infty$ for any matrix norm $||\cdot||$, see [10, p. 219].

2.2. Local rank factorization. The local rank factorization consists of a sequence of matrix rank factorizations. Recall that for any full column rank matrix φ, we indicate by M_φ the orthogonal projector matrix onto $\mathrm{col}^\perp \varphi$.

Definition 2.1 (Local rank factorization of $A(\lambda)$ at λ_u). *Fix $\lambda_u \in \sigma(A)$, let $A(\lambda) = \sum_{n=0}^{\ell} A_n(\lambda-\lambda_u)^n, A_n \in \mathbb{C}^{p\times p}, A_0 \neq 0$, be the representation of $A(\lambda)$ around λ_u and let j be the counter in the recursions.*

Initialization: *Set $j = 0$, $r_0^{\max} := p$, $\mathcal{J}_{-1} = \emptyset$, $\mathcal{J}_0 = 0$, $A_{0,k} := A_{k-1}, k \geq 1$, and $Q_0 := A_0$; perform the matrix rank factorization*

$$Q_0 = -\alpha_0\beta_0', \tag{2.3}$$

where $a_1 := \alpha_0, b_1 := \beta_0$ are full column rank matrices of dimension $p \times r_0$.

Recursion: *While $r_j < r_j^{\max}$, increment the counter j to $j+1$ and, for the updated value of j, define $r_j^{\max} := p - \sum_{i\in\mathcal{J}_{j-1}} r_i$ and*

$$A_{j,k} := A_{j-1,k+1} + A_{j-1,1} \sum_{i\in\mathcal{J}_{j-2}} \bar{\beta}_i\bar{\alpha}_i' A_{i+1,k}, \qquad k \geq 1. \tag{2.4}$$

If $Q_j := M_{a_j} A_{j,1} M_{b_j} \neq 0$ perform the matrix rank factorization

$$Q_j = -\alpha_j\beta_j', \tag{2.5}$$

where α_j, β_j are full column rank matrices of dimension $p \times r_j$, and set $\mathcal{J}_j := (j : \mathcal{J}_{j-1})$, $a_{j+1} := (\alpha_j : a_j)$ and $b_{j+1} := (\beta_j : b_j)$; else let $\mathcal{J}_j := \mathcal{J}_{j-1}$, $a_{j+1} := a_j$ and $b_{j+1} := b_j$.

End: *Set $\mu := j$, $\mathcal{J} := \mathcal{J}_\mu$, $a := a_{\mu+1}$, and $b := b_{\mu+1}$.*

Remark that Definition 2.1 is a restatement of Definition 2.1 in [5], except that the ordering of $\mathcal{J}$, a, b here is reversed.[1]

2.3. Basic properties of the local rank factorization. We collect here the basic properties of the LRF. The inputs of the LRF are the matrix polynomial $A(\lambda)$ and the eigenvalue λ_u; the outputs are the matrices $\alpha_j, \beta_j \in \mathbb{C}^{p\times r_j}$, $A_{j,k} \in \mathbb{C}^{p\times p}$, and the vector of indices $\mathcal{J} = (i_s : \cdots : i_1)$. Remark that $i_1 = 0$ because $A_0 \neq 0$ and $i_s = \mu$ by construction; in the following μ is called the index of the LRF of $A(\lambda)$ at λ_u. In Theorem 2.3 below, it is shown that μ equals the multiplicity m of λ_u, and hence it is finite. Note that $0 \leq r_j \leq r_j^{\max}$, that the end statement is reached when $r_j = r_j^{\max}$ and that $p = \sum_{j\in\mathcal{J}} r_j$.

Remark 2.2. *Observe that a and b are non-singular $p \times p$ matrices with orthogonal blocks, i.e. $\alpha_j'\alpha_k = \beta_j'\beta_k = 0$ for $j \neq k$, $j, k \in \mathcal{J}$.*

Note also that (2.3), (2.5) define α_j, β_j up to a conformable change of bases of the row- and column-spaces, but this does not affect the definition of $A_{j,k}$ in (2.4).

[1] In [5] the vector of indices $\mathcal{J}$ in the present Definition 2.1 is denoted by $\mathcal{J}_\downarrow$.

Because the orthogonal projectors M_φ are also invariant with respect to the choice of basis, the ranks r_j and the index μ in the LRF are determined independently of the choice of basis in α_j, β_j.

The next theorem states that the index of the LRF is equal to the order of the pole.

THEOREM 2.3 (Order of the pole and index of the LRF). *Let $A(\lambda)$ be a regular matrix polynomial and let $\sigma(A)$ be the set of its eigenvalues; then $A^{-1}(\lambda)$ has a pole of order m at $\lambda_u \in \sigma(A)$ if and only if m is the index of the* LRF *of $A(\lambda)$ at λ_u.*

It follows from this theorem that $A^{-1}(\lambda)$ has a pole of order m at λ_u if and only if $r_j < r_j^{\max}$ for $j = 0, \dots, m-1$ and $r_m = r_m^{\max}$. These conditions characterize the structure of $A(\lambda)$ at λ_u through (2.3), (2.5). Observe that (2.3) is a condition on the rank of A_0, which is reduced, $r_0 < p$. Next let $a_{j\perp}$ and $b_{j\perp}$ be bases of $\operatorname{col} M_{a_j}$ and $\operatorname{col} M_{b_j}$ respectively and observe that (2.5) is a condition on the rank of $a'_{j\perp}A_{j,1}b_{j\perp}$, which is reduced ($r_j < r_j^{\max}$) for $j = 1, \dots, m-1$ and full ($r_m = r_m^{\max}$) for $j = m$. In other terms, $A^{-1}(\lambda)$ has a pole of order m at λ_u if and only if the coefficients of $A(\lambda)$ satisfy

$$|A_0| = 0, \qquad |a'_{j\perp}A_{j,1}b_{j\perp}| = 0, \qquad j = 1, \dots, m-1, \qquad \text{and} \qquad |a'_{m\perp}A_{m,1}b_{m\perp}| \neq 0.$$

For $m = 1, 2$ these conditions were derived in [13] in the context of vector autoregressive processes and are called the $I(1)$ and $I(2)$ conditions. Similar conditions are also found in [12].

3. Local Smith form and local Jordan triples. This section illustrates how the LRF can be used to construct the local Smith form and local Jordan triples of $A(\lambda)$. Theorem 3.1, taken from [5], discusses the local Smith form, while Theorem 3.2, which is in part new, discusses Jordan chains and local Jordan triples. All these quantities refer to a given eigenvalue of the spectrum, hence the label 'local'. The proofs of the parts of this section taken from [5] are reported in the Appendix.

3.1. Local Smith form and extended canonical systems of root functions. In the following we use the notation $(\xi_j)_{j=1}^n := (\xi_1 : \cdots : \xi_n)$; similarly for a vector of indices $\mathcal{H} = (i_1 : \cdots : i_n)$, we let $(\xi_j)_{j\in\mathcal{H}} := (\xi_{i_1} : \cdots : \xi_{i_n})$ and $\operatorname{diag}(\xi_j)_{j\in\mathcal{H}} := \operatorname{diag}(\xi_{i_n}, \dots, \xi_{i_1})$. Recall also that $A^\circ(\lambda)$ denotes the reduced adjoint of $A(\lambda)$.

THEOREM 3.1 (Local Smith form and extended canonical systems of root functions). *Given the outputs of the* LRF *of $A(\lambda)$ at λ_u, define*

$$\widetilde{\pi}_j(\lambda) := \alpha_j - \sum_{k=1}^{m-j-1} A_{j+1,k}\bar{\beta}_j(\lambda-\lambda_u)^k, \quad \widetilde{\gamma}_j(\lambda) := \beta_j - \sum_{k=1}^{m-j-1} A'_{j+1,k}\bar{\alpha}_j(\lambda-\lambda_u)^k,$$
$$\pi_j(\lambda) := \frac{A^\circ(\lambda)\widetilde{\pi}_j(\lambda)}{(\lambda-\lambda_u)^{m-j}}, \quad \gamma_j(\lambda) := \frac{A^\circ(\lambda)'\widetilde{\gamma}_j(\lambda)}{(\lambda-\lambda_u)^{m-j}},$$
$$\widetilde{\Pi}(\lambda) := (\widetilde{\pi}_j(\lambda))_{j\in\mathcal{J}}, \quad \widetilde{\Gamma}(\lambda) := (\widetilde{\gamma}_j(\lambda))_{j\in\mathcal{J}},$$
$$\Pi(\lambda) := (\pi_j(\lambda))_{j\in\mathcal{J}}, \quad \Gamma(\lambda) := (\gamma_j(\lambda))_{j\in\mathcal{J}}.$$

Then

i) the local Smith form of $A(\lambda)$ at λ_u is

$$\Lambda(\lambda) = \operatorname{diag}((\lambda-\lambda_u)^j I_{r_j})_{j\in\mathcal{J}};$$

ii) $\Pi(\lambda)$ *and* $\Gamma(\lambda)$ *are respectively an extended canonical systems of right and left root functions of* $A(\lambda)$ *at* λ_u*, i.e.*

$$A(\lambda)\Pi(\lambda) = a_u(\lambda)\widetilde{\Pi}(\lambda)\Lambda(\lambda), \qquad |\Pi(\lambda_u)|, |\widetilde{\Pi}(\lambda_u)| \neq 0, \tag{3.1}$$

$$\Gamma(\lambda)'A(\lambda) = a_u(\lambda)\Lambda(\lambda)\widetilde{\Gamma}(\lambda)', \qquad |\Gamma(\lambda_u)|, |\widetilde{\Gamma}(\lambda_u)| \neq 0, \tag{3.2}$$

where $a_u(\lambda)$ *is as in* (2.1).

Note that from the form of $\Lambda(\lambda)$ one sees that $j \in \mathcal{J}$ is a partial multiplicity of $A(\lambda)$ at λ_u and that there are exactly r_j partial multiplicities that are equal to j; because $p = \sum_{j\in\mathcal{J}} r_j$, this implies that the local Smith form of $A(\lambda)$ at λ_u is completely determined by the LRF through the pairs $(j, r_j), j \in \mathcal{J}$. Moreover, since the ranks r_j are determined uniquely within the LRF, the local Smith form of $A(\lambda)$ at λ_u is uniquely determined by the LRF.

Moreover, Theorem 3.1 shows that one can construct right and left canonical systems of root functions of $A(\lambda)$ at λ_u using $\alpha_j, \beta_j, A_{j,k}$. Because α_j, β_j are defined up to a change of basis, each choice gives a different canonical systems of root functions $\Pi(\lambda), \Gamma(\lambda)$.

For later use it is convenient to collect the partial multiplicities into the vector

$$\kappa := (\kappa_k)_{k=1}^{p} := (j1'_{r_j})_{j\in\mathcal{J}}, \tag{3.3}$$

where 1_s is the $s \times 1$ vector of ones and κ_k indicates the k-th element in κ.

3.2. Local Jordan triple. In this section we construct a local Jordan triple of $A(\lambda)$ at λ_u; in the following we indicate by J_j a Jordan block of dimension j and eigenvalue λ_u. For ease of exposition, we use the convention that whenever J_0 appears in a matrix the corresponding rows and columns should be discarded, e.g. $\mathrm{diag}(I_s \otimes J_1, I_t \otimes J_0) = I_s \otimes J_1$.

THEOREM 3.2 (Canonical set of Jordan chains and local Jordan triple). *Let* $\Pi(\lambda)$ $(\Gamma(\lambda))$ *in Theorem* 3.1 *refer to* λ_u*, let* $x_k(\lambda)$ $(y_k(\lambda))$ *indicate the* k*-th column in* $\Pi(\lambda)$ $(\Gamma(\lambda))$ *and define* $w_{k,i}$ *from* $w_k(\lambda) = \sum_{i=0}^{n} w_{k,i}(\lambda - \lambda_u)^i, n = \deg w_k(\lambda), w = x, y;$ *finally let*

$$X := (X_k)_{k=1}^{p-r_0}, \qquad X_k := (x_{k,i})_{i=0}^{\kappa_k - 1}, \qquad Y := (Y_k)_{k=1}^{p-r_0}, \qquad Y_k := (y_{k,i})_{i=0}^{\kappa_k - 1},$$

where κ_k *is as in* (3.3)*; then the columns in* X (Y) *form a canonical set of right- (left-) Jordan chains of* $A(\lambda)$ *at* λ_u*.*
Moreover, let J_j *be a Jordan block of dimension* j *and eigenvalue* λ_u *and* $J := \mathrm{diag}(I_{r_j} \otimes J_j)_{j\in\mathcal{J}}$*; then* (X, J, Y) *is a Jordan triple of* $A(\lambda)$ *at* λ_u*.*

Proof. From Theorem 3.1 one has that $\Pi(\lambda)$ $(\Gamma(\lambda))$ is an extended canonical system of right- (left-) root functions of $A(\lambda)$, and that $\Lambda(\lambda)$ is its local Smith form. Hence the statement is a direct consequence of the definition of canonical set of right (left) Jordan chains and local Jordan triples, see [9, p. 32-57]. ∎

REMARK 3.3. *In the following the subscript* u *indicates that a quantity refers to the point* λ_u*; that is,* $\Lambda_u(\lambda), \Pi_u(\lambda), \Gamma_u(\lambda)$*, and* (X_u, J_u, Y_u) *indicate respectively the local Smith form, extended canonical systems of right and left root functions and a local Jordan triple of* $A(\lambda)$ *at* λ_u*.*

4. Smith form, Jordan triples and decomposable pairs. This section reports the first part of the original contribution of the paper. Using the local quantities defined in the previous sections, we construct their global counterparts, i.e. the Smith form, finite and infinite Jordan triples and decomposable pairs of the matrix polynomial. These results show how the LRF can be used to obtain these normal forms for any regular matrix polynomial.

It is well known, see [9, p. 181], that $\sigma(A)$ is a finite set, i.e. $\sigma(A) = \{\lambda_u\}_{u=1}^s$ for some integer s; because the invariant polynomials of $A(\lambda)$ are the product of its elementary divisors, the Smith form of $A(\lambda)$ is found at the product of the local Smith forms, i.e

$$\Lambda(\lambda) = \prod_{u=1}^{s} \Lambda_u(\lambda), \tag{4.1}$$

where $\Lambda_u(\lambda)$ is as in Theorem 3.1. Because $\Lambda(\lambda)$ contains information on the behaviour of $A(\lambda)$ for finite values of λ, it does not determine $A(\lambda)$ uniquely; that is, one needs to know both finite and infinite spectral data in order to characterize $A(\lambda)$. By definition, the latter are those of the reverse polynomial $A^{\#}(\lambda) := \lambda^{\ell} A(\lambda^{-1})$ at 0. Hence they can be found by applying the LRF in Definition 2.1, and making use of Theorems 3.1, 3.2 replacing $A(\lambda)$ with $A^{\#}(\lambda)$ and λ_u with $\lambda_\infty := 0$.

THEOREM 4.1 (Finite and infinite Jordan triples and decomposable pairs). *Let $\sigma(A) = \{\lambda_u\}_{u=1}^s$ be the set of eigenvalues of $A(\lambda)$ and let (X_u, J_u, Y_u) be the Jordan triple defined in Theorem 3.2 refer to λ_u; further let*

$$X_F := (X_u)_{u=1}^s, \qquad J_F := \operatorname{diag}(J_u)_{u=1}^s, \qquad Y_F := (Y_u)_{u=1}^s;$$

then (X_F, J_F, Y_F) is a finite Jordan triple of $A(\lambda)$.
Moreover, define $(X_\infty, J_\infty, Y_\infty)$ by applying Definition 2.1 and Theorems 3.1, 3.2 replacing $A(\lambda)$ with $A^{\#}(\lambda)$ and λ_u with $\lambda_\infty := 0$, then $(X_\infty, J_\infty, Y_\infty)$ is an infinite Jordan triple of $A(\lambda)$. Finally, defining

$$X := (X_F : X_\infty), \qquad J := \operatorname{diag}(J_F, J_\infty), \qquad Y := (Y_F : Y_\infty),$$

one finds that (X, J) and (Y, J) are respectively a right- and left- decomposable pair of $A(\lambda)$.

Proof. Given the local Jordan triples at each eigenvalue, one can construct a finite Jordan triple by organizing them into (X_F, J_F, Y_F), see [9, p. 50-57]; because an infinite Jordan triple of $A(\lambda)$ is by definition a Jordan triple of the reverse polynomial $A^{\#}(\lambda)$ at 0, see [9, p. 183-185], one can construct $(X_\infty, J_\infty, Y_\infty)$ by performing the LRF of $A^{\#}(\lambda)$ at 0. The last part of the statement follows from Theorem 7.3 in [9, p. 189]. ∎

5. Kronecker and Jordan forms; similarity to a diagonal matrix. In this section we discuss how the local rank factorization can be used to derive *i*) the Kronecker form and strict similarity transformations for linear matrix polynomials and *ii*) the Jordan form and the associated similarity transformations for matrices. This is achieved considering the regular matrix pencil

$$A(\lambda) = A - \lambda B, \qquad A, B \in \mathbb{C}^{p \times p}.$$

5.1. Kronecker and Jordan forms. First observe how the expressions in the LRF and related quantities simplify for this form of $A(\lambda)$. From Definition 2.1, $A_{1,k} = 0, k > 1$, implies

$$A_{j,k} = 0, \qquad k > 1, \qquad A_{j,1} = A_{j-1,1} \sum_{i \in \mathcal{J}_{j-2}} \bar{\beta}_i \bar{\alpha}_i' A_{i+1,1}, \qquad j = 2, \ldots, m;$$

hence for $j \in \mathcal{J}$, $j \neq m$, $\widetilde{\pi}_j(\lambda), \widetilde{\gamma}_j(\lambda)$ are given by

$$\widetilde{\pi}_j(\lambda) = \alpha_j - A_{j+1,1}\bar{\beta}_j(\lambda - \lambda_u), \qquad \widetilde{\gamma}_j(\lambda) = \beta_j - A_{j+1,1}'\bar{\alpha}_j(\lambda - \lambda_u).$$

THEOREM 5.1 (Kronecker form and strict similarity transformations). *Let* $X := (X_F : X_\infty), J := \operatorname{diag}(J_F, J_\infty), Y := (Y_F : Y_\infty)$ *in Theorem* 4.1 *refer to* $A(\lambda) = A - \lambda B$*, where* $A, B \in \mathbb{C}^{p \times p}$*, and define*

$$U := (BX_F : AX_\infty), \qquad V := (Y_F B : Y_\infty A);$$

then $K(\lambda) := \operatorname{diag}(J_F, I) - \lambda \operatorname{diag}(I, J_\infty)$ *is the Kronecker form of* $A(\lambda)$ *and* X, U *and* Y, V *are strict similarity transformations that connect* $A(\lambda)$ *to* $K(\lambda)$*. That is,*

$$A(\lambda)X = UK(\lambda), \qquad |X|, |U| \neq 0, \qquad YA(\lambda) = K(\lambda)V, \qquad |Y|, |V| \neq 0,$$

so that $A(\lambda) = UK(\lambda)X^{-1} = V^{-1}K(\lambda)Y$*.*

Proof. It is well known, see [9, p. 184-185], that $AX_F = BX_F J_F$, $BX_\infty = AX_\infty J_\infty$. Hence

$$\begin{aligned}
A(\lambda)X &= (AX_F : AX_\infty) - \lambda(BX_F : BX_\infty) \\
&= (BX_F J_F : AX_\infty) - \lambda(BX_F : AX_\infty J_\infty) \\
&= (BX_F : AX_\infty) \operatorname{diag}(J_F, I) - \lambda(BX_F : AX_\infty) \operatorname{diag}(I, J_\infty) \\
&= U(\operatorname{diag}(J_F, I) - \lambda \operatorname{diag}(I, J_\infty)),
\end{aligned}$$

where $U := (BX_F : AX_\infty)$. The fact that X, U are non-singular follows from properties of decomposable pairs, see [9, ch. 7]. Similarly, from $Y_F A = J_F Y_F B$, $Y_\infty B = J_\infty Y_\infty A$, see [9, p. 53], one finds

$$YA(\lambda) = (\operatorname{diag}(J_F, I) - \lambda \operatorname{diag}(I, J_\infty))V, \qquad |Y| \neq 0, \qquad |V| \neq 0,$$

where $V := (Y_F B : Y_\infty A)$. ◻

We next consider the special case $B = I$, i.e. $A(\lambda) = A - I\lambda$, $A \in \mathbb{C}^{p \times p}$; in this case the analysis delivers the Jordan form of A and Jordan bases.

COROLLARY 5.2 (Jordan form and similarity transformations). *Let* (X, J, Y) *in Theorem* 4.1 *refer to* $A(\lambda) = A - \lambda I$*, where* $A \in \mathbb{C}^{p \times p}$*; then* $J = J_F$ *is the Jordan form of* A *and* $X = X_F, Y = Y_F$ *are systems of generalized eigenvectors that connect* A *to* J*. That is,*

$$AX = XJ, \qquad |X| \neq 0, \qquad YA = JY, \qquad |Y| \neq 0,$$

so that $A = XJX^{-1} = Y^{-1}JY$*.*

Proof. Let $B = I$ in Theorem 5.1 and observe that $X_\infty, J_\infty, Y_\infty$ are empty; hence $X = X_F, J = J_F, Y = Y_F$ and $U = X, V = Y$. ◻

5.2. Similarity with respect to a diagonal matrix. The next theorem gives necessary and sufficient conditions for A to be diagonalizable; the conditions are expressed in terms of the outputs of the LRF at each eigenvalue and extend the ones given e.g. in [6, page 152].

THEOREM 5.3 (Diagonalization). *Let $A \in \mathbb{C}^{p\times p}$, indicate by $\sigma(A) = \{\lambda_u\}_{u=1}^{s}$ the set of its eigenvalues and let $\alpha_{u,j}, \beta_{u,j}$, be the outputs of the* LRF *of $A(\lambda) = A - \lambda I$ at $\lambda_u \in \sigma(A)$; then the following statements are equivalent:*

i) *A is diagonalizable,*
ii) *$m_u = 1$ for $u = 1, \ldots, s$;*
iii) $\operatorname{rank} M_{\alpha_{u,0}} M_{\beta_{u,0}} = p - r_{u,0}$ *for $u = 1, \ldots, s$,*
iv) $\operatorname{rank} \alpha'_{u,1}\beta_{u,1} = p - r_{u,0}$ *for $u = 1, \ldots, s$,*
v) $\operatorname{rank} \alpha'_{u,0}\beta_{u,0} = r_{u,0}$ *for $u = 1, \ldots, s$.*

Proof. *i)* $\Leftrightarrow$ *ii)* By [6, page 152] A is diagonalizable if and only if all its elementary divisors are of first degree, i.e. $m_u = 1$ for $u = 1, \ldots, s$. For ease of notation we drop the subscript u in the remaining parts of the proof. *ii)* $\Leftrightarrow$ *iii)*: see Definition 2.1; *iii)* $\Rightarrow$ *iv)*: *iii)* implies *ii)* and hence $M_{\alpha_0} = P_{\alpha_1}$, $M_{\beta_0} = P_{\beta_1}$, so that $M_{\alpha_0} M_{\beta_0} = \bar{\alpha}_1(\alpha'_1\beta_1)\bar{\beta}'_1$. Because $\bar{\alpha}_1$ and $\bar{\beta}_1$ are of full column rank $p - r_0$, one has that *iii)* implies *iv)*; *iv)* $\Rightarrow$ *iii)*: *iv)* implies $M_{\alpha_0} M_{\beta_0} = \bar{\alpha}_1\alpha'_1\beta_1\bar{\beta}'_1$ has rank $p - r_0$ and hence *iii)* holds; *iv)* $\Rightarrow$ *v)*: first observe that $C := (\beta_0 : \alpha_1)$ and $D := (\alpha_0 : \beta_1)$ are non-singular; in fact

$$\begin{pmatrix} \bar{\beta}'_0 \\ \beta'_1 \end{pmatrix} (\beta_0 : \alpha_1) = \begin{pmatrix} I_r & \bar{\beta}'_0\alpha_1 \\ 0 & \beta'_1\alpha_1 \end{pmatrix}, \qquad \begin{pmatrix} \bar{\alpha}'_0 \\ \alpha'_1 \end{pmatrix} (\alpha_0 : \beta_1) = \begin{pmatrix} I_r & \bar{\alpha}'_0\beta_1 \\ 0 & \alpha'_1\beta_1 \end{pmatrix}$$

are non-singular matrices. Note also that $(\bar{\beta}_0 : \beta_1), (\bar{\alpha}_0 : \alpha_1)$ are non-singular by Remark 2.2. Because

$$C'D = \begin{pmatrix} \beta'_0 \\ \alpha'_1 \end{pmatrix} (\alpha_0 : \beta_1) = \begin{pmatrix} \beta'_0\alpha_0 & 0 \\ 0 & \alpha'_1\beta_1 \end{pmatrix},$$

one has $\operatorname{rank} \beta'_0\alpha_0 = r_0$. *v)* $\Rightarrow$ *iv)*: Replacing subscripts 1, 0 with 0, $0\perp$ in the proof of *iv)* $\Rightarrow$ *v)*, one proves that $\operatorname{rank} \alpha'_{0\perp}\beta_{0\perp} = p - r_0$. Because $M_{\alpha_0} = P_{\alpha_{0\perp}}$ and $M_{\beta_0} = P_{\beta_{0\perp}}$, one finds that $M_{\alpha_0} M_{\beta_0} = \bar{\alpha}_{0\perp}(\alpha'_{0\perp}\beta_{0\perp})\bar{\beta}'_{0\perp}$ has rank equal to $p - r_0$, i.e. *iii)* holds, which implies *iv)*. □

6. Example. In this section we illustrate the main results by computing the Jordan form and generalized eigenvectors of the matrix

$$A = \begin{pmatrix} 6 & 2 & 2 \\ -2 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

see [16, p. 238]; from $A(\lambda) = A - I\lambda$ one finds the reduced adjoint and the minimal polynomial

$$A^{\circ}(\lambda) = \begin{pmatrix} (\lambda-2)^2 & 2(\lambda-2) & 2(\lambda-2) \\ -2(\lambda-2) & (\lambda-6)(\lambda-2) & -4 \\ 0 & 0 & (\lambda-4)^2 \end{pmatrix}, \qquad a(\lambda) = (2-\lambda)(\lambda-4)^2.$$

Hence $\sigma(A) = \{2, 4\}$, $s = 2$ and $a_1(\lambda) = (\lambda - 4)^2$, $a_2(\lambda) = 2 - \lambda$ in (2.1) so that $A^{-1}(\lambda)$ has a pole of order $m_1 = 1$ at $\lambda_1 = 2$ and a pole of order $m_2 = 2$ at $\lambda_2 = 4$.

First consider the eigenvalue $\lambda_1 = 2$; the inizialization of the LRF of

$$A(\lambda) = \underbrace{\begin{pmatrix} 4 & 2 & 2 \\ -2 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{A_0} + \underbrace{\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}}_{A_1}(\lambda - 2)$$

at $\lambda_1 = 2$, see Definition 2.1, delivers

$$A_0 = -\underbrace{\begin{pmatrix} 4 & 2 \\ -2 & 0 \\ 0 & 0 \end{pmatrix}}_{\alpha_0}\underbrace{\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & -1 \end{pmatrix}}_{\beta_0'}, \qquad r_0 = 2 < 3 = r_0^{\max}, \qquad \mathcal{J}_0 = 0,$$

which implies

$$M_{a_1} = I - \bar{\alpha}_0\alpha_0' = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \; M_{b_1} = I - \bar{\beta}_0\beta_0' = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix}.$$

Because $A_{1,1} = A_1 = -I$, recursion $j = 1$ delivers

$$M_{a_1}A_{1,1}M_{b_1} = -\underbrace{\begin{pmatrix} 0 \\ 0 \\ \frac{1}{2} \end{pmatrix}}_{\alpha_1}\underbrace{\begin{pmatrix} 0 & -1 & 1 \end{pmatrix}}_{\beta_1'}, \qquad r_1 = 1 = r_1^{\max}, \qquad \mathcal{J}_1 = (1:0).$$

This terminates the recursion because the end condition is satisfied; hence $\mu_1 = m_1 = 1$. Because $\mathcal{J}_1 = (1:0)$, $r_1 = 1$ and $r_0 = 2$, the local Smith form of $A(\lambda)$ at $\lambda_1 = 2$ is equal to, see Theorem 3.1,

$$\Lambda_1(\lambda) = \begin{pmatrix} \lambda - 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and hence $\kappa = (1:0:0)$. Next compute

$$\widetilde{\pi}_1(\lambda) = \alpha_1 = \begin{pmatrix} 0 \\ 0 \\ \frac{1}{2} \end{pmatrix}, \qquad \pi_1(\lambda) = A^{\circ}(\lambda)\widetilde{\pi}_1(\lambda) = \begin{pmatrix} 0 \\ -2 \\ 2 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ \frac{\lambda-6}{2} \end{pmatrix}(\lambda - 2),$$

$$\widetilde{\pi}_0(\lambda) = \alpha_0 + \bar{\beta}_0(\lambda - 2) = \begin{pmatrix} 6-\lambda & 2 \\ -2 & \frac{2-\lambda}{2} \\ 0 & \frac{2-\lambda}{2} \end{pmatrix},$$

$$\pi_0(\lambda) = \frac{A^{\circ}(\lambda)\widetilde{\pi}_0(\lambda)}{\lambda - 2} = -\frac{(\lambda-4)^2}{2}\begin{pmatrix} 2 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix},$$

and define the transformations

$$\widetilde{\Pi}(\lambda) = a_1(\lambda)(\widetilde{\pi}_1(\lambda) : \widetilde{\pi}_0(\lambda)), \qquad \Pi(\lambda) = (\pi_1(\lambda) : \pi_0(\lambda));$$

observe that $\widetilde{\Pi}(2), \Pi(2)$ are non-singular and, as stated in Theorem 3.1, they satisfy

$$A(\lambda)\Pi(\lambda) = \widetilde{\Pi}(\lambda)\Lambda_1(\lambda),$$

i.e. they connect $A(\lambda)$ to its local Smith form $\Lambda_1(\lambda)$. A right Jordan pair of $A(\lambda)$ at $\lambda_1 = 2$ is found applying Theorem 3.2; because $p - r_0 = 1, \kappa_1 = 1$ one has

$$X_{F,1} = \begin{pmatrix} 0 \\ -2 \\ 2 \end{pmatrix}, \qquad J_{F,1} = 2.$$

Next perform the LRF of $A(\lambda)$ at $\lambda_2 = 4$; this delivers $\mathcal{J}_2 = (2:0)$, $r_2 = 1, r_1 = 0, r_0 = 2$ and

$$a = \left(\underbrace{\begin{matrix} -\frac{1}{12} \\ -\frac{1}{12} \\ -\frac{1}{12} \end{matrix}}_{\alpha_2} \quad \underbrace{\begin{matrix} 2 & 2 \\ -2 & 0 \\ 0 & -2 \end{matrix}}_{\alpha_0}\right), \qquad b = \left(\underbrace{\begin{matrix} -1 \\ 1 \\ 0 \end{matrix}}_{\beta_2} \quad \underbrace{\begin{matrix} -1 & 0 \\ -1 & 0 \\ 0 & -1 \end{matrix}}_{\beta_0}\right).$$

Hence the local Smith form of $A(\lambda)$ at $\lambda_2 = 4$ is

$$\Lambda_2(\lambda) = \begin{pmatrix} (\lambda-4)^2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad \kappa = (2:0:0),$$

and the transformations are equal to

$$\widetilde{\Pi}(\lambda) = \underbrace{(2-\lambda)}_{a_1(\lambda)} \left(\underbrace{\begin{matrix} -\frac{1}{12} \\ -\frac{1}{12} \\ -\frac{1}{12} \end{matrix}}_{\widetilde{\pi}_2(\lambda)=\alpha_2} \quad \underbrace{\begin{matrix} \frac{8-\lambda}{2} & 2 \\ -\frac{\lambda}{2} & 0 \\ 0 & 2-\lambda \end{matrix}}_{\widetilde{\pi}_0(\lambda)=\alpha_0+\bar{\beta}_0(\lambda-4)}\right), \ \Pi(\lambda) = \left(\underbrace{\begin{matrix} \frac{4-\lambda^2}{12} \\ \frac{-12+10\lambda-\lambda^2}{12} \\ \frac{-16+8\lambda-\lambda^2}{12} \end{matrix}}_{\pi_2(\lambda)} \quad \underbrace{\begin{matrix} \frac{2-\lambda}{2} & 0 \\ \frac{2-\lambda}{2} & 0 \\ 0 & 2-\lambda \end{matrix}}_{\pi_0(\lambda)}\right).$$

One can check that $|\widetilde{\Pi}(4)|, |\Pi(4)| \neq 0$ and $A(\lambda)\widetilde{\Pi}(\lambda) = \Pi(\lambda)\Lambda_2(\lambda)$, i.e. they connect $A(\lambda)$ to its local Smith form $\Lambda_2(\lambda)$. Because $p - r_0 = 1, \kappa_1 = 2$ and

$$\pi_2(\lambda) = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} -\frac{2}{3} \\ \frac{1}{6} \\ 0 \end{pmatrix}(\lambda-4) - \frac{1}{12}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}(\lambda-4)^2,$$

one finds

$$X_{F,2} = \begin{pmatrix} -1 & -\frac{2}{3} \\ 1 & \frac{1}{6} \\ 0 & 0 \end{pmatrix}, \qquad J_{F,2} = \begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix}.$$

Because $s = 2$ this completes the local analysis; hence the Smith form of $A(\lambda)$ is equal to

$$\Lambda(\lambda) = \begin{pmatrix} (\lambda-2)(\lambda-4)^2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

see (4.1), and a Jordan pair, see Theorem 4.1, is given by

$$X = \left(\begin{array}{c|cc} 0 & -1 & -\frac{2}{3} \\ -2 & 1 & \frac{1}{6} \\ -2 & 0 & 0 \end{array}\right), \qquad J = \operatorname{diag}(J_{F,1}, J_{F,2}) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 1 \\ 0 & 0 & 4 \end{pmatrix}.$$

$$X = \big(\underbrace{X_{F,1}}_{\text{col. 1}}\ \underbrace{X_{F,2}}_{\text{cols. 2–3}} \big)$$

One can check that $|X| \neq 0$, $A = XJX^{-1}$, so that X is a similarity transformation that links A to its Jordan form J, see Theorem 5.2. Similarly one can derive the left counterpart of these quantities.

7. Conclusion. The LRF characterizes the structure of a matrix polynomial at an eigenvalue and delivers its local Smith form, extended canonical systems of left and right root functions, canonical sets of left and right Jordan chains and local Jordan triples. Using the local quantities one then derives the global counterparts, i.e. the Smith form, finite and infinite Jordan triples and decomposable pairs. Applying these tools to matrix polynomials of degree one, one obtains the Kronecker form and strict similarity transformations, as well as the Jordan form and systems of generalized eigenvectors. A enlarged characterization of the conditions for a matrix to be diagonalizable are obtained, some of which are stated in terms of subspaces defined by the LRF.

Appendix A. Proofs taken from [5].

LEMMA A.1. *Fix $\lambda_u \in \sigma(A)$, let $\alpha_j, \beta_j, A_{j,k}, \mathcal{J}$ be the outputs of the* LRF *of $A(\lambda)$ at λ_u and let $A^\circ(\lambda) = \sum_{n=0}^{\ell_\circ} A_n^\circ(\lambda - \lambda_u)^n$ be the reduced adjoint of $A(\lambda)$; then for $0 \leq j \leq n \leq \min(m, \mu)$, one has*

$$\alpha_j \beta_j' A_{n-j}^\circ = M_{a_j} \sum_{k=1}^{n-j} A_{j+1,k} A_{n-j-k}^\circ + c_{n,m} M_{a_j}, \tag{A.1}$$

$$A_{n-j}^\circ \alpha_j \beta_j' = \sum_{k=1}^{n-j} A_{n-j-k}^\circ A_{j+1,k} M_{b_j} + c_{n,m} M_{b_j}, \tag{A.2}$$

where $c_{n,m} := -\delta_{n,m} a_u(\lambda_u)$, see (2.1)*, and $\delta_{n,m}$ is Kronecker's delta.*

Proof. Let

$$A(\lambda) = \sum_{n=0}^{\ell} A_n(\lambda - \lambda_u)^n, \qquad A^\circ(\lambda) = \sum_{n=0}^{\ell_\circ} A_n^\circ(\lambda - \lambda_u)^n$$

and write

$$A(\lambda)A^\circ(\lambda) = \sum_{n=0}^{N} L_n(\lambda - \lambda_u)^n, \qquad A^\circ(\lambda)A(\lambda) = \sum_{n=0}^{N} R_n(\lambda - \lambda_u)^n,$$

where

$$N = m + \deg a_u(\lambda), \qquad L_n = \sum_{k=0}^{n} A_k A_{n-k}^\circ, \qquad R_n = \sum_{k=0}^{n} A_k^\circ A_{n-k};$$

then (2.2) implies

$$L_n = R_n = \delta_{n,m} I, \qquad 0 \leq n \leq m, \tag{A.3}$$

where $\delta_{n,m}$ is Kronecker's delta. We wish to show that for $0 \leq j \leq n \leq \min(m, \mu)$ one has

$$M_{a_j} L_n = -\alpha_j \beta_j' A^\circ_{n-j} + M_{a_j} \sum_{k=1}^{n-j} A_{j+1,k} A^\circ_{n-j-k}, \tag{A.4}$$

which, together with $(A.3)$, implies $(A.1)$.

We proceed by induction on j. Let $j = 1$ in (2.4) and observe that $A_{1,k} = A_k$ with $A_0 = -\alpha_0 \beta_0'$; hence $L_n = -\alpha_0 \beta_0' A^\circ_n + \sum_{k=1}^{n} A_{1,k} A^\circ_{n-k}$, which proves that $(A.4)$ holds for $j = 0$ and $M_{a_0} = I$.

Next assume that $(A.4)$ holds for $0 \leq j \leq \ell \leq n < \min(m, \mu)$; we wish to show that it also holds for $j = \ell + 1 \leq \min(m, \mu)$. Consider $(A.4)$ for $j = \ell$; pre-multiply it by $M_{a_{\ell+1}}$ and re-arrange terms to find

$$M_{a_{\ell+1}} L_n = M_{a_{\ell+1}} A_{\ell+1,1} A^\circ_{n-\ell-1} + M_{a_{\ell+1}} \sum_{k=1}^{n-\ell-1} A_{\ell+1,k+1} A^\circ_{n-\ell-1-k} =: U + V \text{ (say)}. \tag{A.5}$$

Using the projection identity $I = M_{b_{\ell+1}} + P_{b_{\ell+1}}$, one finds

$$U = M_{a_{\ell+1}} A_{\ell+1,1} M_{b_{\ell+1}} A^\circ_{n-\ell-1} + M_{a_{\ell+1}} A_{\ell+1,1} P_{b_{\ell+1}} A^\circ_{n-\ell-1} \tag{A.6}$$

and substituting in the first term from (2.5) and in the second term from $P_{b_{\ell+1}} = \sum_{h=0}^{\ell} \bar{\beta}_h \beta_h'$ one obtains

$$U = -\alpha_{\ell+1} \beta_{\ell+1}' A^\circ_{n-\ell-1} + M_{a_{\ell+1}} A_{\ell+1,1} \sum_{h=0}^{\ell} \bar{\beta}_h \beta_h' A^\circ_{n-\ell-1}. \tag{A.7}$$

We next substitute $\beta_h' A^\circ_{n-\ell-1}$ in the last expression with

$$\beta_h' A^\circ_{n-\ell-1} = \bar{\alpha}_h' \sum_{k=1}^{n-\ell-1} A_{h+1,k} A^\circ_{n-\ell-1-k}, \tag{A.8}$$

which is proved as follows. Observe that $(A.4)$ holds by the induction hypothesis with j and n replaced by h and $n - \ell - 1$, because $h \leq \ell$ and $n - \ell - 1 \leq n - 1 < \min(m, \mu)$. Hence

$$M_{a_h} L_{n-\ell-1} = -\alpha_h \beta_h' A^\circ_{n-\ell-1} + M_{a_h} \sum_{k=1}^{n-\ell-1} A_{h+1,k} A^\circ_{n-\ell-1-k}.$$

Because $L_{n-\ell-1} = 0$, see $(A.3)$, pre-multiplying by $\bar{\alpha}_h'$ and observing that $\bar{\alpha}_h' M_{a_h} = \bar{\alpha}_h'$, one obtains $(A.8)$.

Substituting $(A.8)$ in $(A.7)$, and interchanging the order of summation, one has

$$U = -\alpha_{\ell+1} \beta_{\ell+1}' A^\circ_{n-\ell-1} + M_{a_{\ell+1}} \sum_{k=1}^{n-\ell-1} \left(A_{\ell+1,1} \sum_{h=0}^{\ell} \bar{\beta}_h \bar{\alpha}_h' A_{h+1,k} \right) A^\circ_{n-\ell-1-k}. \tag{A.9}$$

Summing $U + V$ and using (2.4) one finds $(A.4)$ for $j = \ell + 1$. Similarly one proves $(A.2)$ using R_n. □

Proof of Theorem 2.3. We want to show that the index of the LRF of $A(\lambda)$ at λ_u is equal to the order of the pole of $A^{-1}(\lambda)$ at λ_u. Next we show that $\mu < m$ and $\mu > m$ both lead to a contradiction, so that one must have $\mu = m$. First suppose $\mu < m$; for $n = 0, \dots, \mu$ one has $\beta_n' A_0^\circ = 0$ by $(A.1)$, i.e. $b' A_0^\circ = 0$ which is a contradiction because $A_0^\circ \neq 0$ and b is non-singular, see Remark 2.2. Next suppose $\mu > m$; let $j = n = m$ in $(A.1)$ to get $\alpha_m \beta_m' A_0^\circ = M_{a_m}$. Hence, $\operatorname{col} \alpha_m \supseteq \operatorname{col} M_{a_m}$; because $\dim \operatorname{col} \alpha_m = r_m$ and $\dim \operatorname{col} M_{a_m} = p - \sum_{n=0}^{m-1} r_n = r_m^{\max}$, this implies $r_m \geq r_m^{\max}$ which is a contradiction because $m < \mu$ implies $r_m < r_m^{\max}$. Hence $\mu = m$. □

Proof of Theorem 3.1. In order to prove (3.1) it is sufficient to show that $\widetilde{\Pi}(\lambda)$ is an extended system of right root functions of $A^\circ(\lambda)$ at λ_u; together with (2.2) this implies that $\Pi(\lambda)$ is an extended canonical system of right root functions of $A(\lambda)$ at λ_u. In order to do so, first we show that

$$A^\circ(\lambda)\widetilde{\pi}_j(\lambda) = (\lambda - \lambda_u)^{m-j}\pi_j(\lambda), \qquad j \in \mathcal{J}. \tag{A.10}$$

Then we group $(A.10)$ together into

$$A^\circ(\lambda)\widetilde{\Pi}(\lambda) = \Pi(\lambda)\Lambda^\circ(\lambda), \tag{A.11}$$

where $\Lambda^\circ(\lambda) := \operatorname{diag}((\lambda - \lambda_u)^{m-j} I_{r_j})_{j\in\mathcal{J}}$, and show that $\widetilde{\Pi}(\lambda_u)$ and $\Pi(\lambda_u)$ are non-singular; this proves that $\widetilde{\Pi}(\lambda)$ is an extended system of root functions of $A^\circ(\lambda)$ at λ_u by condition (2) in Theorem 1.3 in [7], also reported in the Appendix in [8].

We first prove $(A.10)$. Consider first $0 \leq j \leq n \leq m-1$; for $j = n$, $(A.2)$ implies $A_0^\circ \alpha_j = 0$ and hence

$$A^\circ(\lambda)\alpha_j = \sum_{h=1}^{m-j-1} A_h^\circ \alpha_j (\lambda - \lambda_u)^h + (\lambda - \lambda_u)^{m-j} R_0(\lambda)\alpha_j =: U(\lambda) + V(\lambda) \text{ (say)},$$

where $R_0(\lambda_u) = A_{m-j}^\circ$ and

$$A_h^\circ \alpha_j = \sum_{k=1}^{h} A_{h-k}^\circ A_{j+1,k}\bar{\beta}_j + c_{h+j,m}\bar{\beta}_j, \tag{A.12}$$

follows from $(A.2)$ replacing $n - j$ with h. Observe that for $j = 0, \dots, m-1$ and $h = 1, \dots, m-j-1$ one has $c_{h+j,m} = 0$; hence substituting from $(A.12)$ in $U(\lambda)$ and re-arranging terms, one finds

$$U(\lambda) = \sum_{k=1}^{m-j-1} \left(\sum_{h=k}^{m-j-1} A_{h-k}^\circ (\lambda - \lambda_u)^h \right) A_{j+1,k}\bar{\beta}_j.$$

One can write $\sum_{h=k}^{m-j-1} A_{h-k}^\circ (\lambda - \lambda_u)^h = (\lambda - \lambda_u)^k A^\circ(\lambda) - (\lambda - \lambda_u)^{m-j} R_k(\lambda)$, where $R_k(\lambda_u) = A_{m-j-k}^\circ$; hence $U(\lambda)$ becomes

$$U(\lambda) = A^\circ(\lambda) \sum_{k=1}^{m-j-1} A_{j+1,k}\bar{\beta}_j (\lambda - \lambda_u)^k - (\lambda - \lambda_u)^{m-j} \sum_{k=1}^{m-j-1} R_k(\lambda) A_{j+1,k}\bar{\beta}_j,$$

so that $U(\lambda) + V(\lambda)$ is

$$A^\circ(\lambda)\alpha_j = A^\circ(\lambda) \sum_{k=1}^{m-j-1} A_{j+1,k}\bar{\beta}_j (\lambda - \lambda_u)^k + (\lambda - \lambda_u)^{m-j}\pi_j(\lambda), \tag{A.13}$$

where

$$\pi_j(\lambda) := R_0(\lambda)\alpha_j - \sum_{k=1}^{m-j-1} R_k(\lambda)A_{j+1,k}\bar{\beta}_j. \tag{A.14}$$

By taking all the terms in $(A.13)$ that multiply $A^\circ(\lambda)$ on the l.h.s., one finds $(A.10)$, where $\widetilde{\pi}_j(\lambda) := \alpha_j - \sum_{k=1}^{m-j-1} A_{j+1,k}(\lambda-\lambda_u)^k\bar{\beta}_j$.

Next we have to show that $\widetilde{\Pi}(\lambda_u)$ and $\Pi(\lambda_u)$ are non-singular. Because $\widetilde{\pi}_j(\lambda_u) = \alpha_j$ one has $\widetilde{\Pi}(\lambda_u) = a$ and hence it is non-singular, see Remark 2.2. Next consider the block $\pi_j(\lambda_u)$ in $\Pi(\lambda_u) = (\pi_j(\lambda_u))_{j\in\mathcal{J}}$. Using $(A.14)$ and $R_k(\lambda_u) = A^\circ_{m-j-k}$ for $k = 0, \dots, m-j-1$, one finds

$$\pi_j(\lambda_u) = A^\circ_{m-j}\alpha_j - \sum_{k=1}^{m-j-1} A^\circ_{m-j-k}A_{j+1,k}\bar{\beta}_j,$$

where, substituting from $(A.12)$ for $h = m-j$, one finds

$$\pi_j(\lambda_u) = A^\circ_0 A_{j+1,m-j}\bar{\beta}_j - a_u(\lambda_u)\bar{\beta}_j = -a_u(\lambda_u)(\bar{\beta}_m\bar{\alpha}'_m A_{j+1,m-j}\bar{\beta}_j + \bar{\beta}_j)$$

where we have substituted $A^\circ_0 = -a_u(\lambda_u)\bar{\beta}_m\bar{\alpha}'_m$, which is shown as follows: let $n = j$ in $(A.1)$ and $(A.2)$; for $0 \le j \le m$, one then has

$$\alpha_j\beta'_jA^\circ_0 = c_{j,m}M_{a_j} \text{ and } A^\circ_0\alpha_j\beta'_j = c_{j,m}M_{b_j};$$

these imply $\beta'_jA^\circ_0 = 0$, $A^\circ_0\alpha_j = 0$ for $j = 0, \dots, m-1$ and $\beta'_mA^\circ_0\alpha_m = c_{m,m}I_{r_m}$ and thus $A^\circ_0 = \sum_{j=0}^m P_{\beta_j}A^\circ_0\sum_{j=0}^m P_{\alpha_j} = -a_u(u)\bar{\beta}_m\bar{\alpha}'_m$. This also gives $\pi_m(\lambda_u) = A^\circ_0\alpha_m = -a_u(\lambda_u)\bar{\beta}_m$. Hence one finds

$$\Pi(\lambda_u) = -a_u(\lambda_u)K\begin{pmatrix} I_{r_m} & H \\ 0 & I_{p-r_m}\end{pmatrix},$$

where $a_u(\lambda_u) \neq 0$, $K := (\bar{\beta}_j)_{j\in\mathcal{J}}$ and $H := \bar{\alpha}'_m(A_{m,1}\bar{\beta}_{m-1} : \cdots : A_{1,m}\bar{\beta}_0)$. Because K is non-singular, see Remark 2.2, and the upper triangular matrix is non-singular, one concludes that $\Pi(\lambda_u)$ is non-singular. Similarly one proves (3.2). □

REFERENCES

[1] K. Avrachenkov, M. Haviv, and P. Howlett, *Inversion of analytic matrix functions that are singular at the origin*, SIAM Journal of Matrix Analysis and Applications, 22 (2001), pp. 1175–89.

[2] H. Bart, I. Gohberg, M. Kaashoek, and A. Ran, *Factorization of Matrix and Operator Functions: The State Space Method*, Birkhäuser, Basel-Boston-Berlin, 2008.

[3] M. Franchi, *A representation theory for polynomial cofractionality in vector autoregressive models*, Econometric Theory, 26 (2010), pp. 1201–1217.

[4] M. Franchi and P. Paruolo, *A characterization of vector autoregressive processes with common cyclical features*, Journal of Econometrics, 163 (2011), pp. 105–117.

[5] ———, *Inversion of regular analytic matrix functions: local Smith form and subspace duality*, Linear Algebra and its Applications, doi:10.1016/j.laa.2011.05.005 (2011).

[6] F. Gantmacher, *The Theory of Matrices*, vol. I, AMS Chelsea Publishing, 1959.

[7] I. Gohberg, M. Kaashoek, and F. van Schagen, *On the local theory of regular analytic matrix functions*, Linear Algebra and its Applications, 182 (1993), pp. 9–25.

[8] ———, *Partially Specified Matrices and Operators: Classification, Completion, Applications*, Birkhäuser, Basel-Boston-Berlin, 1995.

[9] I. Gohberg, P. Lancaster, and L. Rodman, *Matrix polynomials*, Academic Press, New York, 1982.
[10] ———, *Invariant subspaces of matrices with applications*, Wiley, New York, 1986.
[11] N. Haldrup and M. Salmon, *Representation of I(2) cointegrated systems using the Smith-McMillan form*, Journal of Econometrics, 44 (1998), pp. 303–325.
[12] P. Howlett, *Input retrieval in finite dimensional linear systems*, Journal of Australian Mathematical Society (Series B), 23 (1982), pp. 357–82.
[13] S. Johansen, *A representation of vector autoregressive processes integrated of order 2*, Econometric Theory, 8 (1992), pp. 188–202.
[14] T. Kailath, *Linear systems*, Prentice Hall, Englewood Cliffs, N.J., 1980.
[15] P. Lancaster, *Linearization of regular matrix polynomials*, Electronic Journal of Linear Algebra, 17 (2008), pp. 21–27.
[16] P. Lancaster and M. Tismenetsky, *The theory of matrices*, Academic Press, 1985.
[17] C. Langenhop, *The Laurent expansion for a nearly singular matrix*, Linear Algebra and its Applications, 4 (1971), pp. 329–40.
[18] G. Marsaglia and G. P. H. Styan, *Equalities and inequalities for ranks of matrices*, Linear and Multilinear Algebra, 2 (1974), pp. 269–292.
[19] C. Meyer, *Matrix Analysis and Applied Linear Algebra*, Society for Industrial and Applied Mathematics (SIAM), 2001.
[20] P. Schweitzer and G. Stewart, *The Laurent expansion of pencils that are singular at the origin*, Linear Algebra and its Applications, 183 (1993), pp. 237–54.
[21] J. Wilkening, *An algorithm for computing Jordan chains and inverting analytic matrix functions*, Linear Algebra and its Applications, 427 (2007), pp. 6–25.
[22] J. Wilkening and J. Yu, *A local construction of the Smith normal form of a matrix polynomial*, Journal of Symbolic Computation, 46 (2011), pp. 1–22.