Linear Algebra and its Applications 656 (2023) 510–540

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

Root vectors of polynomial and rational matrices: Theory and computation

Vanni Noferini [a,*,1], Paul Van Dooren [b,2]

[a] *Aalto University, Department of Mathematics and Systems Analysis, P.O. Box 11100, FI-00076, Aalto, Finland*
[b] *Université catholique de Louvain, Department of Mathematical Engineering, Av. Lemaitre 4, B-1348 Louvain-la-Neuve, Belgium*

ARTICLE INFO

Article history:
Received 2 May 2022
Accepted 11 October 2022
Available online 13 October 2022
Submitted by C. Mehl

MSC:
15A03
15A18
15A21
15A22
65F15

Keywords:
Rational matrix
Root polynomial
Root vector
Maximal set
Eigenvector
Eigenvalue
Minimal basis
Smith form
Local Smith form
Coalescent pole/zero

ABSTRACT

The notion of root polynomials of a polynomial matrix $P(\lambda)$ was thoroughly studied in Dopico and Noferini (2020) [6]. In this paper, we extend such a systematic approach to general rational matrices $R(\lambda)$, possibly singular and possibly with coalescent pole/zero pairs. We discuss the related theory for rational matrices with coefficients in an arbitrary field. As a byproduct, we obtain sensible definitions of eigenvalues and eigenvectors of a rational matrix $R(\lambda)$, without any need to assume that $R(\lambda)$ has full column rank or that the eigenvalue is not also a pole. Then, we specialize to the complex field and provide a practical algorithm to compute them, based on the construction of a minimal state space realization of the rational matrix $R(\lambda)$ and then using the staircase algorithm on the linearized pencil to compute the null space as well as the root polynomials in a given point λ_0. If λ_0 is also a pole, then it is necessary to apply a preprocessing step that removes the pole while making it possible to recover the root vectors of the original matrix: in this case, we study both the relevant theory (over a general field) and

* Corresponding author.
E-mail address: vanni.noferini@aalto.fi (V. Noferini).
[1] Supported by an Academy of Finland grant (Suomen Akatemian päätös 331240).
[2] Supported by an Aalto Science Institute Visitor Programme.

https://doi.org/10.1016/j.laa.2022.10.013
0024-3795/© 2022 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

an algorithmic implementation (over the complex field), still based on minimal state space realizations.
© 2022 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Let $R(\lambda) \in \mathbb{C}(\lambda)^{m\times n}$ be a rational matrix with coefficients in the field of complex numbers $\mathbb{C}$. In linear system theory and in control theory, an important problem is to find the zeros and poles of $R(\lambda)$. In this setting, $R(\lambda)$ contains the information of the inputs and outputs of a system of differential or difference equations [12,20]. It is shown in [15] that the natural frequencies of the dynamical system are captured by the poles, and the frequencies that are blocked by the system are the zeros. When $R(\lambda)$ is singular, that is, either it is not square or it does not have full rank over the field of rational functions, then the matrix has also left and/or right null spaces [7]. These vectors spaces and their so-called minimal bases provide further information on the initial conditions and on the degrees of freedom associated with the response of the system [7,12,15,20]. Additional motivation to study matrix-valued rational functions stems from the nonlinear eigenvalue problem [10]. In this context, one has a locally meromorphic matrix-valued function $M(\lambda)$, that could either be exactly rational [10,16] or possibly approximated by a rational function $R(\lambda)$ [10,11]. One method to approximate (some of the) zeros and eigenvectors of $M(\lambda)$ is to first compute zeros and eigenvectors of $R(\lambda)$.

A point $\lambda_0 \in \mathbb{C}$ may be a simple or multiple zero, a simple or multiple pole, and even a coalescent pole and zero of the rational matrix $R(\lambda)$. The local (at λ_0) Smith-McMillan form of the rational matrix is then relevant. In linear systems theory, it refines the description of how the dynamical system responds at the frequency λ_0 via the *partial multiplicities* of λ_0. In nonlinear eigenvalue problems, the partial multiplicities are also important. They generalize concepts familiar for the classical eigenvalue problem, like the size of the Jordan blocks of a matrix, or (equivalently) the Segré and Weyr characteristics of a matrix. In particular, information about the partial multiplicities is highly relevant for the perturbation theory of an eigenvalue λ_0.

The spectral description of a rational matrix does not, however, end with the partial multiplicities. In linear systems theory, the response also depends on particular *directions*, which are input vectors that *excite* the system's pole or zero of a multiplicity given by each structural index. These vectors arise naturally when one wants to describe the solution set of particular matrix equations involving rational matrices [12,13] and are also used in tangential interpolation problems of high order [8]. Such vectors can be viewed as a generalization of an eigenvector of a first order system of differential equation modeled by the eigenvalue problem $\lambda x - Ax = 0$. Similarly, in nonlinear eigenvalue problems one may be interested in appropriate generalizations of the eigenvectors of a matrix. If the nonlinear matrix is regular, and the eigenvalue is not also a pole, then eigenvectors

can indeed be always defined as usual. In the case of polynomial eigenvalue problems, eigenvectors were also defined for singular matrix polynomials [6, Definition 2.18].

More generally, in this paper we show how to extend the theory of *root polynomials* [6,9] to the case of rational matrices over any field. For a regular polynomial (or rational) matrix, root polynomials can be viewed as generating functions for the set of Jordan chains at a point. In the singular case, they generalize such a concept, and as mentioned above they for example can be used to define eigenvectors as elements of certain quotient space, see [6, Subsection 2.3]. We will show that root polynomials of rational matrices encode all the information on structural indices [6], and that they make it possible to properly define eigenvectors for rational matrices, even when they are singular or when they have coalescent poles and zeros (or even when they present both issues).

The paper is organized as follows. Sections 2 and 3 are theoretical, and there we work over an algebraically closed field $\mathbb{F}$ (not necessarily $\mathbb{C}$); this is not a restriction because the case of a non-closed field can be dealt with by embedding. In Section 2 we recall the basic definitions and background material for the rest of the paper, including the concept of root polynomials at a point λ_0 for a polynomial matrix. In Section 3 we recall the concept of local discrete valuations and use this to define root vectors of a general rational matrix with a coalescent pole and zero at λ_0. We develop a general theory of root vectors at a finite point showing, amongst other things, that they can be assumed without loss of generality to be root polynomials, that they can be used to define eigenvectors even when the rational matrix has a pole at the point of interest (or even when it is singular), and that *maximal sets* of root vectors always encode the information about partial multiplicities. We also mention how to define root vectors at infinity. The rest of the paper is concerned with the computation of root polynomials of polynomial matrices and root vectors of rational matrices. For this we first recall what is a minimal realization of a general rational matrix $R(\lambda)$ in Section 4. Then, in Section 5, we show how to recover maximal sets of root polynomials of $R(\lambda)$ from those of a minimal system matrix, first assuming that λ_0 is not a pole; the case of λ_0 being also a pole is later dealt with by constructing a minimal realization of a second rational matrix $Y(\lambda)$ that has the property of having no poles at λ_0, but whose maximal sets of root vectors at λ_0 are simply related to those of $R(\lambda)$ and allow us to recover them. We then specialize to the case where the system matrix is a minimal realization, and hence a pencil. In the case of $\mathbb{F} \subseteq \mathbb{C}$, algorithms to compute root polynomials of pencils exist [19], and hence overall this yields an algorithm for computing maximal sets of root polynomials of any rational matrix $R(\lambda)$. Such an algorithm is relatively simple when λ_0 is not also a pole, but also works (albeit with additional complications) in the case of coalescent poles/zeros. Finally, we make some concluding remarks in Section 6.

2. Background and definitions

Let $\mathbb{F}$ be an algebraically closed field. The theory that we will develop can be generalized to any field $\mathbb{K}$ by embedding $\mathbb{K}^{m\times n}$ in $\mathbb{F}^{m\times n}$ where $\mathbb{F}$ is the algebraic closure of

$\mathbb{K}$. The assumption is made for notational convenience as it guarantees that any finite eigenvalue is an element of the base field. Given a point $\lambda_0 \in \mathbb{F}$, the local Smith-McMillan form of the $m \times n$ rational matrix $R(\lambda) \in \mathbb{F}(\lambda)^{m\times n}$ at λ_0 is:

$$M(\lambda) \cdot R(\lambda) \cdot N(\lambda) := \left[\begin{array}{ccc|c} (\lambda-\lambda_0)^{\sigma_r} & & 0 & \\ & \ddots & & \\ 0 & & (\lambda-\lambda_0)^{\sigma_1} & \\ \hline & & & 0_{m-r,n-r} \end{array}\right]. \tag{1}$$

In (1), $M(\lambda)$ and $N(\lambda)$ are rational matrices that do not have poles at λ_0 and are invertible at λ_0, r is the rank of $R(\lambda)$ over $\mathbb{F}(\lambda)$ (sometimes called the *normal rank* of $R(\lambda)$) and the integers σ_i are called *structural indices*, or also *partial multiplicities*,[3] of $R(\lambda)$ at the point λ_0. Moreover, they are ordered non-increasingly : $\sigma_1 \geq \ldots \geq \sigma_r$. Observe that negative indices are associated with poles of $R(\lambda)$, positive indices are associated with zeros of $R(\lambda)$, while zero indices are not associated with either.

Other important sets of indices of a general $m\times n$ rational matrix $R(\lambda)$, called minimal indices, are related to its right null space $\ker R(\lambda)$ and left null space $\ker R(\lambda)^T$, which are rational vector spaces over the field $\mathbb{F}(\lambda)$ of rational functions in λ.

Definition 2.1. A matrix polynomial $N(\lambda) \in \mathbb{F}[\lambda]^{n\times p}$ of normal rank p is called a minimal polynomial basis if the sum of the degrees of its columns, called the order of the basis, is minimal among all bases of span $N(\lambda)$. Its ordered column degrees are called the minimal indices of the basis.

It was shown in [7] that the ordered list of indices is independent of the choice of minimal basis of the space. If we define the right null space $\ker R(\lambda)$ and left null space $\ker R(\lambda)^T$ of a $m \times n$ rational matrix $R(\lambda)$ of normal rank r as the vector spaces of rational vectors $x(\lambda)$ and $y(\lambda)$ annihilated by $R(\lambda)$

$$\ker R(\lambda) := \{x(\lambda) \mid R(\lambda)x(\lambda) = 0\}, \qquad \ker R(\lambda)^T := \{y(\lambda) \mid y^{\mathsf{T}}(\lambda)R(\lambda) = 0\}$$

then the minimal indices of any minimal polynomial basis for these spaces, are called the right and left minimal indices of $R(\lambda)$. The respective dimensions of $\ker R(\lambda)$ and $\ker R(\lambda)^T$ are $n - r$ and $m - r$ and the respective indices are denoted by

$$\{\epsilon_1, \ldots, \epsilon_{n-r}\}, \quad \{\eta_1, \ldots, \eta_{m-r}\}.$$

[3] Variations of this nomenclature can be found in the literature. For example, sometimes partial multiplicities are required to be nonzero integers (thus excluding zero indices), or even positive integers (in this case one distinguishes between the partial multiplicities as a zero, corresponding to the positive σ_i, and the partial multiplicities as a pole, corresponding to $-\sigma_j$ for the negative σ_j). In this paper, it is convenient to define the partial multiplicities as the exponents in the local Smith-McMillan form, thus admitting zero or negative integers.

It was also shown in [7] that for any minimal basis $N(\lambda)$, the constant matrix $N(\lambda_0)$ has full column rank for all $\lambda_0 \in \mathbb{F}$ and the highest column degree matrix of $N(\lambda)$ also has full column rank.

2.1. Root polynomials

In [6,9] a rigorous link was established between partial multiplicities and the so-called *root polynomials* but in those references that concept was analyzed for a polynomial matrix $P(\lambda)$ only. Root polynomials had appeared as technical tools for proving other results [9,17], and in [6] it was argued that they can play an important role in the theory of polynomial matrices. Below we give a definition which is equivalent to that given in [6], but rephrased in a way that is more useful for deriving the results given in this paper. The matrix $P(\lambda) \in \mathbb{F}[\lambda]^{m\times n}$ is assumed to be polynomial in the next definition and theorem, proven in [6].

Definition 2.2. Let $N(\lambda)$ be a right minimal basis of $P(\lambda)$ and $\lambda_0 \in \mathbb{F}$; then

- $x(\lambda) \in \mathbb{F}[\lambda]^n$ is a root polynomial at λ_0 of order $k > 0$ if $P(\lambda)x(\lambda) = (\lambda - \lambda_0)^k v(\lambda)$, $v(\lambda_0) \neq 0$, and $[N(\lambda_0)\ x(\lambda_0)]$ has full column rank;
- $\{x_1(\lambda), ..., x_s(\lambda)\}$ is a set of λ_0-independent root polynomials of orders $\{k_1, ..., k_s\}$ if they are root polynomials at λ_0 of orders $\{k_1, ..., k_s\}$ and $[N(\lambda_0)\ x_1(\lambda_0) \dots x_s(\lambda_0)]$ has full column rank;
- a λ_0-independent set is complete if there does not exist any λ_0-independent set of larger cardinality;
- such a complete set is ordered if $k_1 \geq \dots \geq k_s > 0$;
- such a complete ordered set is maximal if there is no root polynomial $\tilde{x}(\lambda)$ of order $k > k_j$ at λ_0 such that $[N(\lambda_0)\ x_1(\lambda_0) \dots x_{j-1}(\lambda_0)\ \tilde{x}(\lambda_0)]$ has full column rank for any j.

The following result [6, Theorem 4.1.3 and Theorem 4.2] explains the importance of these maximal sets.

Theorem 2.3. *Let the nonzero partial multiplicities at λ_0 of $P(\lambda)$ be $0 < \sigma_s \leq \dots \leq \sigma_1$, and suppose that $x_1(\lambda), \dots, x_s(\lambda)$ are a complete set of root polynomials at λ_0 for $P(\lambda)$. Then, the following are equivalent:*

1. *$x_1(\lambda), \dots, x_s(\lambda)$ are a maximal set of root polynomials at λ_0 for $P(\lambda)$;*
2. *the orders of such a set are precisely $\sigma_1, \dots, \sigma_s$;*
3. *the sum of the orders of such a set is precisely $\sum_{i=1}^{s} \sigma_i$.*

In the next section, we show how one can extend the theory of root polynomials to rational matrix-valued functions, even when the point λ_0 is a pole. (If λ_0 is a zero, but

not a pole, the task is trivially achieved by multiplying the rational matrix by its least common denominator, as the latter cannot have a root at λ_0 and, hence, it is easy to see that this operation preserves both partial multiplicities at λ_0 and maximal sets of root polynomials at λ_0.)

3. Root vectors of rational matrices

In this section we build a theory of eigenvalues, (right) root vectors and (right) eigenvectors[4] of any rational matrix $R(\lambda)$ over $\mathbb{F}(\lambda)$, without any need to make any of the assumptions that are sometimes made in the literature, e.g., that $R(\lambda)$ has full column rank or that the point of interest is not a pole. To this goal, we first need to recall what a local discrete valuation is.

Definition 3.1. Given $\lambda_0 \in \mathbb{F}$, the *local discrete valuation* κ associated with the point λ_0 is the function

$$\kappa : \mathbb{F}(\lambda) \to \mathbb{Z} \cup \{\infty\}, \qquad 0 \neq r(\lambda) = (\lambda - \lambda_0)^k \frac{p(\lambda)}{q(\lambda)} \mapsto \kappa[r(\lambda)] = k, \quad 0 \mapsto \kappa[0] = \infty,$$

where, in the case of a nonzero rational function $r(\lambda)$, $p(\lambda)$ and $q(\lambda)$ are polynomials both coprime with $(\lambda - \lambda_0)$, and $k \in \mathbb{Z}$.

Note that, while there is a distinct local discrete valuation for each λ_0, in this paper we fix λ_0 and work locally. Thus, for simplicity of notation, we omit the dependence of the function κ on the point λ_0. It is easy to check that the local discrete valuation at λ_0 satisfies the three properties

(i) $\kappa[a(\lambda)b(\lambda)] = \kappa[a(\lambda)] + \kappa[b(\lambda)]$,
(ii) $\kappa[a(\lambda) + b(\lambda)] \geq \min\{\kappa[a(\lambda)], \kappa[b(\lambda)]\}$, and
(iii) $\kappa[a(\lambda)] = \infty \Leftrightarrow a(\lambda) = 0$.

Local discrete valuations can be extended to matrices by taking the elementwise minimum, and it retains certain (somewhat weaker) properties.

Definition 3.2. If $A(\lambda) \in \mathbb{F}(\lambda)^{m\times n}$ and $\lambda_0 \in \mathbb{F}$, then the local discrete valuation at λ_0 of $A(\lambda)$ is $\kappa[A(\lambda)] = \min_{1\leq i\leq m, 1\leq j\leq n} \kappa[A_{ij}(\lambda)]$.

Proposition 3.3. *For any pair of rational matrices $A(\lambda), B(\lambda)$ and any point $\lambda_0 \in \mathbb{F}$ with associated local discrete valuation κ, the following holds :*

[4] The dual concepts of left root vectors and eigenvectors are, of course, trivially obtainable by working with $R(\lambda)^T$.

1. *if the product* $P(\lambda) = A(\lambda)B(\lambda)$ *is defined, then* $\kappa[P(\lambda)] \geq \kappa[A(\lambda)] + \kappa[B(\lambda)]$*;*
2. *if the sum* $S(\lambda) = A(\lambda) + B(\lambda)$ *is defined, then* $\kappa[S(\lambda)] \geq \min\{\kappa[A(\lambda)], \kappa[B(\lambda)]\}$*;*
3. $\kappa[A(\lambda)] = \infty \Leftrightarrow A(\lambda) = 0$*;*
4. *if* $A(\lambda)$ *is square and* $\kappa[A(\lambda)] = \kappa[\det(A(\lambda))] = 0$*, then* $A(\lambda)$ *is invertible over* $\mathbb{F}(\lambda)$ *and* $\kappa[A(\lambda)^{-1}] = \kappa[\det(A(\lambda)^{-1})] = 0$*;*
5. *if* $A(\lambda)$ *is square and satisfies* $\kappa[A(\lambda)] = \kappa[\det(A(\lambda))] = 0$*, then*
   (a) if the product $P(\lambda) = A(\lambda)B(\lambda)$ *is defined, then* $\kappa[P(\lambda)] = \kappa[B(\lambda)]$
   (b) if the product $Q(\lambda) = B(\lambda)A(\lambda)$ *is defined, then* $\kappa[Q(\lambda)] = \kappa[B(\lambda)]$*;*
6. *if* $S(\lambda)$ *and* $L(\lambda)$ *are, respectively, the Smith-McMillan form of* $A(\lambda)$ *and the local (at* λ_0*) Smith-McMillan form of* $A(\lambda)$*, then* $\kappa[A(\lambda)] = \kappa[S(\lambda)] = \kappa[L(\lambda)]$*;*
7. *if* $A(\lambda)$ *is partitioned into the two block submatrices* $B(\lambda), C(\lambda)$ *then* $\kappa[A(\lambda)] = \min\{\kappa[B(\lambda)], \kappa[C(\lambda)]\}$.

Proof.

1. For some indices i, j,

$$\kappa[P(\lambda)] = \kappa[P_{ij}(\lambda)] = \kappa\left[\sum_k A_{ik}(\lambda)B_{kj}(\lambda)\right]$$
$$\geq \min_k \kappa[A_{ik}(\lambda)] + \min_k \kappa[B_{kj}(\lambda)] \geq \kappa[A(\lambda)] + \kappa[B(\lambda)].$$

2.
$$\kappa[A(\lambda) + B(\lambda)] = \min_{i,j} \kappa[A_{ij}(\lambda) + B_{ij}(\lambda)]$$
$$\geq \min_{i,j} \min\{\kappa[A_{ij}(\lambda)], \kappa[B_{ij}(\lambda)]\} \geq \min\{\kappa[A(\lambda)], \kappa[B(\lambda)]\}.$$

3. Obvious.
4. If $A(\lambda)$ were not invertible, then it would follow that $\kappa[\det(A(\lambda))] = \infty$. Moreover,

$$0 = \kappa[\det(A(\lambda))] + \kappa[\det(A(\lambda)^{-1})] \Rightarrow \kappa[\det(A(\lambda)^{-1})] = 0$$

   and

$$0 = \kappa[I] \geq \kappa[A(\lambda)] + \kappa[A(\lambda)^{-1}] \Rightarrow \kappa[A(\lambda)^{-1}] \leq 0.$$

   Suppose now that $\kappa[A(\lambda)^{-1}] < 0$. Then, since each entry of the inverse of $A(\lambda)$ is the ratio of the corresponding entry of the adjoint of $A(\lambda)$ (which must have nonnegative local discrete valuation being a polynomial in scalars with nonnegative local discrete valuation) and the determinant of $A(\lambda)$, we conclude that $\kappa[\det(A(\lambda))] \geq -\kappa[A(\lambda)^{-1}] > 0$, contradicting $\kappa[\det(A(\lambda))] = 0$.
5. We only prove part (a) as part (b) is similar. Under the assumptions, by the previous item $A(\lambda)$ is invertible and $\kappa[A(\lambda)^{-1}] = 0$. Hence, by item 1,

$$P(\lambda) = A(\lambda)B(\lambda) \Rightarrow \kappa[P(\lambda)] \geq \kappa[B(\lambda)]$$

and

$$A(\lambda)^{-1}P(\lambda) = B(\lambda) \Rightarrow \kappa[B(\lambda)] \geq \kappa[P(\lambda)],$$

yielding $\kappa[P(\lambda)] = \kappa[B(\lambda)]$.

6. It suffices to prove $\kappa[S(\lambda)] = \kappa[A(\lambda)]$, as it is clear that $\kappa[S(\lambda)] = \kappa[L(\lambda)]$. We have that $U(\lambda)A(\lambda) = S(\lambda)V(\lambda)$ for some unimodular polynomial matrices $U(\lambda), V(\lambda)$. Clearly, being unimodular implies that $\kappa[\det(U(\lambda))] = \kappa[\det(V(\lambda))] = \kappa[U(\lambda)] = \kappa[V(\lambda)] = 0$. Hence, by the previous item,

$$\kappa[A(\lambda)] = \kappa[U(\lambda)A(\lambda)] = \kappa[S(\lambda)V(\lambda)] = \kappa[S(\lambda)].$$

7. Immediate by the definition of local discrete valuation of a matrix. □

Let $R(\lambda) \in \mathbb{F}(\lambda)^{m\times n}$ and suppose that $N(\lambda)$ is a minimal basis for its right kernel. Given $\lambda_0 \in \mathbb{F}$, define as usual [6,17] $\ker_{\lambda_0} R(\lambda) = \text{span} N(\lambda_0)$. Then,

Definition 3.4. The rational vector $x(\lambda)$ is a *root vector* at λ_0 of order $k > 0$ for $R(\lambda)$ if:

1. $\kappa[x(\lambda)] = 0$ and $x(\lambda_0) \notin \ker_{\lambda_0} R(\lambda)$;
2. $\kappa[R(\lambda)x(\lambda)] = k > 0$.

Moreover, we say that λ_0 is an *eigenvalue* (or zero) of $R(\lambda)$ if there exists a root vector at λ_0 for $R(\lambda)$.

Definition 3.5. Let $N(\lambda)$ be a right minimal basis for $R(\lambda)$, then

- The set of root vectors $\{x_i(\lambda)\}_{i=1}^s$ at λ_0, of orders $k_1, \ldots, k_s$, is λ_0-independent if the matrix $[N(\lambda_0) \quad x_1(\lambda_0) \quad \ldots \quad x_s(\lambda_0)]$ has full column rank.
- A λ_0-independent set of root vectors $\{x_i(\lambda)\}_{i=1}^s$ is complete if there does not exist any λ_0-independent set $\{x_i(\lambda)\}_{i=1}^\sigma$ such that $\sigma > s$.
- Such a complete set is ordered if $k_1 \geq \cdots \geq k_s > 0$.
- Such a complete ordered set is maximal if, for any j, there is no root vector $\tilde{x}(\lambda)$ of order $k > k_j$ such that

$$[N(\lambda_0) \quad x_1(\lambda_0) \quad \ldots \quad x_{j-1}(\lambda_0) \quad \tilde{x}(\lambda_0)]$$

has full column rank.

At this point, we can state and prove Proposition 3.6 below. Obviously, $\kappa[A(\lambda)] \geq 0$ implies that λ_0 is not a pole of $A(\lambda)$, and according to item 4 of Proposition 3.3,

$\kappa[A(\lambda)] = \kappa[\det(A(\lambda))] = 0$ implies that $A(\lambda)$ is regular and that λ_0 is not an eigenvalue of $A(\lambda)$. Hence, Proposition 3.6 describes the behavior of root vectors when a rational matrix is multiplied (from both left and right) by square rational matrices whose local Smith form at λ_0 is trivial.

Proposition 3.6. *Let $R(\lambda), Q(\lambda) \in \mathbb{F}(\lambda)^{m\times n}$ satisfy $Q(\lambda) = A(\lambda)R(\lambda)B(\lambda)$ for some $A(\lambda) \in \mathbb{F}(\lambda)^{m\times m}$ and $B(\lambda) \in \mathbb{F}(\lambda)^{n\times n}$ such that $\kappa[A(\lambda)] = \kappa[B(\lambda)] = \kappa[\det(A(\lambda))] = \kappa[\det(B(\lambda))] = 0$, where κ is the local discrete valuation associated with $\lambda_0 \in \mathbb{F}$. Then:*

1. *If $r(\lambda)$ is a root vector for $Q(\lambda)$ at λ_0 of order $k > 0$, then $B(\lambda)r(\lambda)$ is a root vector for $R(\lambda)$ at λ_0 of the same order;*
2. *If $q(\lambda)$ is a root vector for $R(\lambda)$ at λ_0 of order $k > 0$, then $B(\lambda)^{-1}q(\lambda)$ is a root vector for $Q(\lambda)$ at λ_0 of the same order;*

Proof. It suffices to prove item 1, since item 2 then follows from

$$R(\lambda) = A(\lambda)^{-1}Q(\lambda)B(\lambda)^{-1}.$$

By item 5 in Proposition 3.3, $\kappa[r(\lambda)] = 0 \Rightarrow \kappa[B(\lambda)r(\lambda)] = 0$. By the same result,

$$k = \kappa[Q(\lambda)r(\lambda)] = \kappa[A(\lambda)R(\lambda)B(\lambda)r(\lambda)] = \kappa[R(\lambda)B(\lambda)r(\lambda)].$$

Now suppose for a contradiction that $B(\lambda_0)r(\lambda_0) \in \ker_{\lambda_0} R(\lambda)$. Then, by [6, Lemma 2.9] there is a polynomial vector $v(\lambda)$ such that $R(\lambda)v(\lambda) = 0$ and $v(\lambda_0) = B(\lambda_0)r(\lambda_0)$. (Although in [6] this is proved in the case where $R(\lambda)$ is a polynomial matrix, its proof is only based on the existence of a minimal basis for $\ker R(\lambda)$ which is still true when $R(\lambda)$ is generally rational.) This implies

$$Q(\lambda)B(\lambda)^{-1}v(\lambda) = 0 \Rightarrow B(\lambda_0)^{-1}v(\lambda_0) = r(\lambda_0) \in \ker_{\lambda_0} Q(\lambda),$$

contradicting the assumption. □

We can then proceed similarly to [6], and show the next theorem whose proof is omitted since it is completely analogous to the proof of [6, Theorem 3.4].

Theorem 3.7. *Let $R(\lambda), Q(\lambda) \in \mathbb{F}(\lambda)^{m\times n}$ satisfy $Q(\lambda) = A(\lambda)R(\lambda)B(\lambda)$ for some $A(\lambda) \in \mathbb{F}(\lambda)^{m\times m}$ and $B(\lambda) \in \mathbb{F}(\lambda)^{n\times n}$ such that $\kappa[A(\lambda)] = \kappa[B(\lambda)] = \kappa[\det(A(\lambda))] = \kappa[\det(B(\lambda))] = 0$, where κ is the local discrete valuation associated with $\lambda_0 \in \mathbb{F}$. Then:*

1. *If $x_1(\lambda), \dots, x_s(\lambda)$ are a maximal (resp. complete, λ_0-independent) set for $Q(\lambda)$ with orders $k_1 \geq \cdots \geq k_s > 0$, then $B(\lambda)x_1(\lambda), \dots, B(\lambda)x_s(\lambda)$ are a maximal (resp. complete, λ_0-independent) set for $R(\lambda)$ with the same orders;*

2. *If $x_1(\lambda), \dots, x_s(\lambda)$ are a maximal (resp. complete, λ_0-independent) set for $R(\lambda)$ with orders $k_1 \geq \cdots \geq k_s > 0$, then $B(\lambda)^{-1}x_1(\lambda), \dots, B(\lambda)^{-1}x_s(\lambda)$ are a maximal (resp. complete, λ_0-independent) set for $Q(\lambda)$ with the same orders.*

At this point we observe that the set of all matrices $M(\lambda)$ satisfying the assumptions $\kappa[M(\lambda)] = \kappa[\det(M(\lambda))] = 0$ (such as $A(\lambda)$ and $B(\lambda)$ in Theorem 3.7) are precisely the unimodular matrices over the local ring

$$\mathcal{R} := \left\{ \frac{a(\lambda)}{b(\lambda)} \in \mathbb{F}(\lambda) : \gcd(a(\lambda), b(\lambda)) = 1 \text{ and } b(\lambda_0) \neq 0 \right\}. \tag{2}$$

It is clear that $\mathcal{R}$ is a principal ideal domain, and that the local Smith-McMillan form at λ_0 is obtained precisely by unimodular equivalence over $\mathcal{R}$. Hence, one can reduce the study of maximal sets to the case of rational matrices in local Smith-McMillan form. In particular, the proofs of [6, Proposition 2.16 and Theorems 3.5, 4.1 and 4.2] generalize to the rational case, as we show below.

Definition 3.8. Let $R(\lambda) \in \mathbb{F}(\lambda)^{m\times n}$. Then,

$$\ker R(\lambda_0) = \{x_0 \in \mathbb{F}^n : \exists x(\lambda) \in \mathbb{F}(\lambda)^n \ s.t.\ \kappa[R(\lambda)x(\lambda)] > 0 \text{ and } x(\lambda_0) = x_0\}$$

where κ is the local discrete valuation at λ_0.

Note that the condition $x(\lambda_0) = x_0$ in Definition 3.8 implies in particular that λ_0 is not a pole of $x(\lambda)$, or equivalently that $\kappa[x(\lambda)] \geq 0$ (and $= 0$ if and only if $x_0 \neq 0$). It is not hard to verify that $\ker R(\lambda_0)$, as defined above, is a subspace of $\mathbb{F}^n$.

Remark 3.9. In the case where λ_0 is not a pole of $R(\lambda)$, i.e., $R(\lambda_0) \in \mathbb{F}^{m\times n}$, Definition 3.8 is equivalent to

$$\ker R(\lambda_0) = \{x_0 \in \mathbb{F}^n : R(\lambda_0)x_0 = 0\}.$$

This is a more familiar expression but, in general, it does not make sense because $R(\lambda_0)$ may not be well defined. In practice, usually the most relevant case is when $\mathbb{F} = \mathbb{C}$ or some subfield of $\mathbb{C}$: in this case, Definition 3.8 is equivalent to

$$\ker R(\lambda_0) = \{x_0 \in \mathbb{F}^n : \exists x(\lambda) \in \mathbb{F}(\lambda)^n \ s.t.\ \lim_{\lambda\to\lambda_0} R(\lambda)x(\lambda) = 0 \text{ and } x(\lambda_0) = x_0\}$$

since $\kappa[R(\lambda)x(\lambda)] > 0 \Leftrightarrow \lim_{\lambda\to\lambda_0} R(\lambda)x(\lambda) = 0$. Related limit-based definitions have appeared elsewhere, at least under the simplifying assumption that $R(\lambda)$ has full column rank, to characterize eigenvectors of rational matrices with the additional condition $x_0 \neq 0$ (clearly such a nonzero x_0 exists only if $\ker R(\lambda_0)$ is not the trivial subspace): see for instance [5, p. 315]. We will give a more general characterization of eigenvectors in

Definition 3.14. We also point out that Definition 3.8 is more general than the limit-based definitions such as the one in [5, p. 315] and, being purely algebraic, is valid also when $\mathbb{F}$ is not a topological space; this choice is coherent with the literature on polynomial matrices that have often been studied over any field.

Theorem 3.10. *Let $R(\lambda) \in \mathbb{F}(\lambda)^{m\times n}$ with minimal basis $N(\lambda) \in \mathbb{F}[\lambda]^{n\times p}$. Let $\{x_i(\lambda)\}_{i=1}^{s}$ be root vectors at λ_0 for $R(\lambda)$; then, they are a complete set if and only if the columns of the matrix*

$$M := [N(\lambda_0) \quad x_1(\lambda_0) \quad \dots \quad x_s(\lambda_0)]$$

are a basis for $\ker R(\lambda_0)$.

Proof. By definition, $\kappa[R(\lambda)N(\lambda)] = \infty$ and $\kappa[R(\lambda)x_i(\lambda)] > 0$. Hence, the columns of M are in $\ker R(\lambda_0)$. Moreover, if the set $\{x_i(\lambda)\}_{i=1}^{s}$ is complete, M has full column rank as any complete set is λ_0-independent. Suppose the columns of M are not a basis for $\ker R(\lambda_0)$: then they can be completed to a basis, say, by adding some vectors $\{v_i\}_{i=1}^{\ell}$. But then there are rational vectors $w_i(\lambda)$ satisfying both $w_i(\lambda_0) = v_i$ and $\kappa[R(\lambda)w_i(\lambda)] > 0$. This implies that $\{x_i(\lambda)\}_{i=1}^{s} \cup \{w_i(\lambda)\}_{i=1}^{\ell}$ is still a λ_0-independent set. Hence, the original set was not complete. Conversely, if the set was not complete, then the matrix M cannot be a basis for $\ker R(\lambda_0)$ because then by starting from a complete set we can find a full column rank matrix $\hat{M}$, having strictly more columns than M, and such that the columns of $\hat{M}$ belong to $\ker R(\lambda_0)$. □

We observe that Theorem 3.10 in particular implies that the cardinality of the subset of root vectors in a complete set is equal to $s := \dim \ker R(\lambda_0) - \dim \ker {}_{\lambda_0} R(\lambda)$.

Theorem 3.11. *Every complete set of root vectors has the same cardinality. And in every maximal set, the orders of the root vectors are the same.*

Proof. The first statement is obvious by Theorem 3.10. The second statement follows from the proof of [6, Theorem 4.1.3], which shows that root polynomials in a maximal set must have the same orders when $R(\lambda)$ is polynomial; however, it can be applied verbatim to the rational case. (More details are in the proof of Theorem 3.12 below.) □

We can now formulate the major result of this section. It extends [6, Theorem 4.1] to the rational case and has essentially the same proof; given its importance, we include a sketch of proof below.

Theorem 3.12. *Let $R(\lambda) \in \mathbb{F}(\lambda)^{m\times n}$ and suppose that the positive partial multiplicities of $\lambda_0 \in \mathbb{F}$ are $\sigma_1 \geq \cdots \geq \sigma_s > 0$. If $R(\lambda)$ has a complete set of root vectors at λ_0 $x_1(\lambda), \dots, x_s(\lambda)$, having orders $k_1 \geq \cdots \geq k_s$, then $\sigma_i \geq k_i$ for all i, and such a set is maximal if and only if $k_i = \sigma_i$ for all $i = 1, \dots, s$.*

Proof. Following [6], let $\{x_i(\lambda)\}_{i=1}^s$ be a complete set and $\{v_i(\lambda)\}_{i=1}^s$ be a maximal set, and suppose that these sets have orders $k_1 \geq \cdots \geq k_s$ and $\tau_1 \geq \cdots \geq \tau_s$, respectively. Suppose for a contradiction that for some $j \geq 0$ it holds $k_i = \tau_i$ for $i \leq j$ but $k_{j+1} > \tau_{j+1}$. Let $\{r_i(\lambda)\}_{i=1}^p$ be a minimal basis for $R(\lambda)$, then by the definition of maximal set the matrix

$$[r_1(\lambda_0) \quad \dots \quad r_p(\lambda_0) \quad v_1(\lambda_0) \quad \dots \quad v_j(\lambda_0) \quad x_k(\lambda_0)]$$

cannot possibly have full column rank, for all $k = 1, \dots, j+1$. This implies that

$$r_1(\lambda_0), \dots, r_p(\lambda_0), x_1(\lambda_0), \dots, x_{j+1}(\lambda_0) \in \operatorname{span}\{r_1(\lambda_0), \dots, r_p(\lambda_0), v_1(\lambda_0), \dots, v_j(\lambda_0)\};$$

but the latter claim exhibits $p+j+1$ linearly independent (by completeness and hence λ_0-independence of $\{x_i(\lambda)\}_{i=1}^s$) vectors all belonging to the same $(p+j)$-dimensional vector space, which is absurd. This shows that the orders of a complete set are elementwise bounded above by the orders of a maximal set. Moreover, if we further assume that also $\{x_i(\lambda)\}_{i=1}^s$ is maximal, then by a symmetric argument the case where $k_{j+1} < \tau_{j+1}$ can also be excluded analogously: hence, any two maximal sets have the same orders.

To conclude the proof, it suffices to exhibit a particular maximal set and check that its orders are the positive partial multiplicities at λ_0. To this goal, let $S(\lambda) = U(\lambda)R(\lambda)V(\lambda)$ be the local Smith-McMillan form at λ_0 of $R(\lambda)$ for some $U(\lambda), V(\lambda)$ invertible over the local ring (2), and suppose that the integers m, s are such the positive partial powers of $(\lambda - \lambda_0)$ appear precisely in positions $1 \leq m, \dots, m+s-1$. It is a simple exercise to prove that $\{e_{i+1-m}\}_{i=1}^s$ is a maximal set of root vectors for $S(\lambda)$ (see also [6, Theorem 3.1] and note that the proof extends to the rational case) and that their orders are precisely the positive partial multiplicities at λ_0. By Theorem 3.7, $\{V(\lambda)e_{i+1-m}\}_{i=1}^s$ is a maximal set of root vectors for $R(\lambda)$ with the same orders. □

In other words, also in the case of rational matrices, maximal sets of root vectors yield the full information about positive partial multiplicities.

Remark 3.13. We point out that the root vectors have their entries in the ring (2). They can therefore always be scaled with an appropriate polynomial factor to make them polynomial root vectors. These scaling factors do not affect any of the theorems derived in this section.

From the theory of root vectors for rational matrices, one can define eigenvectors by generalizing the approach of [6, Definition 2.18]. We restrict ourselves to right eigenvectors; left eigenvectors can be defined as right eigenvectors of $R(\lambda)^T$. Namely, eigenvectors are nonzero vectors in the quotient space of $\ker R(\lambda_0)/\ker_{\lambda_0} R(\lambda)$ (which has a vector space structure). Observe that, if $R(\lambda)$ has full column rank, then $\ker_{\lambda_0} R(\lambda) = \{0\}$ and hence $\ker R(\lambda_0)/\ker_{\lambda_0} R(\lambda) = \ker R(\lambda_0)$.

Definition 3.14. An equivalence class $[v] \in \ker R(\lambda_0)/\ker_{\lambda_0} R(\lambda)$ is called a (right) eigenvector of $R(\lambda)$ associated with λ_0 if $[v] \neq [0]$.

Definition 3.14 implies that eigenvectors associated with λ_0 exist if and only if λ_0 is an eigenvalue, as by Definition 3.4 and Definition 3.8 this is equivalent to the inclusion $\ker_{\lambda_0} R(\lambda) \subset \ker R(\lambda_0)$ being strict; if, on the contrary, λ_0 is not an eigenvalue then $\ker_{\lambda_0} R(\lambda) = \ker R(\lambda_0)$, and hence the quotient space is the trivial vector space $\{[0]\}$ which does not contain nonzero equivalence classes (and thus it cannot contain eigenvectors). Moreover, by Theorem 3.10, it is clear that $[v]$ is an eigenvector if and only if

$$[0] \neq [v] \in \operatorname{span}([x_1(\lambda_0)], \ldots, [x_s(\lambda_0)])$$

where $\{x_i(\lambda)\}_{i=1}^s$ is any complete set of root vectors at λ_0 for $R(\lambda)$. In other words, $[v]$ is an eigenvector if and only if $v = \sum_{i=1}^s x_i(\lambda_0)c_i + w$ for some $c_i \in \mathbb{F}$ such that $\sum_{i=1}^s x_i(\lambda_0)c_i \neq 0$, and some $w \in \ker_{\lambda_0} R(\lambda)$. When $\mathbb{F}$ is a subfield of $\mathbb{C}$ and λ_0 is simple, one can uniquely (up to a nonzero scalar) fix a representative u of the equivalence class $[v]$ by asking that $u \in \ker_{\lambda_0} R(\lambda)^\perp$; see [14]. Finally, we emphasize again that if $R(\lambda)$ has full column rank then $\ker_{\lambda_0} R(\lambda)$ is trivial, and therefore Definition 3.14 collapses to an eigenvector being equal to $v = x(\lambda_0)$ where $x(\lambda)$ is a root vector at λ_0 for $R(\lambda)$ (note that by definition root vectors do not have poles at λ_0 and are nonzero when evaluated at λ_0). If, furthermore, λ_0 is not a pole, then we recover the classical definition: a nonzero vector v such that $R(\lambda_0)v = 0$.

3.1. Root vectors at infinity

We conclude this section by noting that, as in the polynomial case [6], it is possible to define root vectors at ∞ for rational matrices. There are two possible approaches for this. One, natural for polynomial eigenvalue problems, is via the reversal and was described in [6, Sec. 6]. A second one, more natural in the context of control theory as it reveals the Smith-McMillan positive degrees at infinity, is via the local discrete valuation at infinity of a rational function [7,23]. The latter is defined for scalar rational functions as $\kappa[0] = \infty$ for the zero function, and $\kappa[p(\lambda)/q(\lambda)] = \deg q(\lambda) - \deg p(\lambda)$ if $p(\lambda)/q(\lambda) \neq 0$. It is a simple exercise to verify that this is indeed a local discrete valuation in the sense of satisfying the same three properties as the valuation at a finite point. We can, again, extend the function κ to rational matrices by taking the elementwise minimum of its value on the matrix entries. A root vector $x(\lambda)$ of order k at ∞ for $R(\lambda)$ is then a rational vector such that

1. $\kappa[x(\lambda)] = 0$, that is, $x(\lambda)$ is a proper rational vector with at least one biproper entry;
2. $x(\infty) \notin \ker_\infty R(\lambda)$ where (2a) $x(\infty)$ is defined as the zeroth-order coefficient x_0 in the expansion of $x(\lambda) = \sum_{j=0}^\infty x_j \lambda^{-j}$ (which exists because $x(\lambda)$ is proper) and (2b) $\ker_\infty R(\lambda) := \ker_0 R(\frac{1}{\lambda})$ or, equivalently [17], the vector space spanned by the high order coefficients matrix of a minimal basis for $\ker R(\lambda)$;

3. $\kappa[R(\lambda)x(\lambda)] = k > 0$, that is, $R(\lambda)x(\lambda)$ is a strictly proper rational vector with local discrete valuation at ∞ equal to k.

With this definition, $x(\lambda)$ is a root vector of order k at ∞ for $R(\lambda)$ if and only if $x(\lambda^{-1})$ is a root vector of order k at 0 for $R(\lambda^{-1})$. Moreover, ∞-independent, complete and maximal sets for $R(\lambda)$ can then be defined similarly to the finite case, and correspond to 0-independent, complete and maximal sets for $R(\lambda^{-1})$. Hence, the properties of root vectors that we have established for finite points also carry over to ∞, provided that the partial multiplicities at ∞ are defined [2, Sec. 5] as the ones obtained via the Smith form over the local ring of proper rational functions. For reasons of conciseness, we omit the details. We warn the reader that, as we illustrate below, this definition of partial multiplicities at infinity (common in the context of control theory) differs from the one usually given in the context of polynomial eigenvalue problems.

Example 3.15. It should be noted that the definition of a root vector at infinity proposed above does not yield the one given in [6] in the case of polynomial matrices. The reason is that the partial multiplicities at infinity are defined differently in the two approaches (via the reversal or via the Smith-McMillan approach).

For instance consider the regular polynomial matrix $P(\lambda) = \begin{bmatrix} 1 & \lambda \\ 0 & 1 \end{bmatrix}$. According to the reversal definition (and seeing $P(\lambda)$ as grade 1), the partial multiplicities of $P(\lambda)$ at infinity are $0, 2$; according to the Smith-McMillan definition, however, they are $-1, 1$. This difference is reflected in the two different definitions of root polynomial (or root vector) at infinity given in [6] or here. According to the definition in [6], $r(\lambda) = [-\lambda \quad 1]^T$ is a root polynomial at ∞ of order 2 for $P(\lambda)$ (and a maximal set) because it is the 1-reversal of a root polynomial of order 2 of the 2-reversal of $P(\lambda)$. According to the definition proposed above, instead, $r(\lambda)$ is not even a root vector at ∞ for $P(\lambda)$ (its local discrete valuation at ∞ is $-1 \neq 0$). However, $\lambda^{-1}r(\lambda)$ is a root vector of order 1 at ∞ since (1) $\kappa[\lambda^{-1}r(\lambda)] = 0$ (2) $r(\infty) = [-1 \quad 0]^T \notin \ker_\infty P(\lambda) = \{0\}$ (3) $\kappa[P(\lambda)\lambda^{-1}r(\lambda)] = 1$.

4. The system matrix of a rational matrix

Any rational transfer function $R(\lambda)$ can be realized by a *generalized state-space realization* $\{A, B, C, D, E\}$ (see e.g. [12]) such that

$$R(\lambda) = C(\lambda E - A)^{-1}B + D. \tag{3}$$

If $R(\lambda)$ happens to be proper, i.e. bounded at $\lambda = \infty$, then E can be chosen to be the identity and one refers to the system $\{A, B, C, D\}$ as a *state-space realization* of $R(\lambda)$. One often associates with the above realization (3) the pencil

$$S(\lambda) := \left[\begin{array}{c|c} A - \lambda E & B \\ \hline C & D \end{array}\right] \tag{4}$$

called the *system matrix* of the realization. Note that $R(\lambda)$ is the Schur complement of its system matrix $S(\lambda)$. Techniques for constructing a realization (3) are described in e.g. [21]. It is also shown there how to obtain an *irreducible generalized state space realization* of the transfer function $R(\lambda)$. These are realizations (3) for which

$$\left[\begin{array}{c} A-\lambda E \\ \hline C \end{array}\right] \quad \text{and} \quad \left[\begin{array}{c|c} A-\lambda E & B \end{array}\right] \tag{5}$$

have no finite zeros (see also [24]).

More generally, one can consider a *polynomial system matrix*

$$P(\lambda) = \begin{bmatrix} -A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda) \end{bmatrix} \in \mathbb{F}[\lambda]^{(q+m)\times(q+n)} \tag{6}$$

such that $A(\lambda) \in \mathbb{F}[\lambda]^{q\times q}$ is regular and $R(\lambda) = D(\lambda) + C(\lambda)A(\lambda)^{-1}B(\lambda)$. The polynomial system matrix is said to be *minimal* if the Smith normal form of $[-A(\lambda) \quad B(\lambda)]$ is $[I \quad 0]$ and the Smith normal form of $\begin{bmatrix} -A(\lambda) \\ C(\lambda) \end{bmatrix}$ is $\begin{bmatrix} I \\ 0 \end{bmatrix}$. In the following section, we will investigate how root vectors of a rational matrix and its minimal polynomial system matrix relate.

5. Root vectors of a minimal system matrix and those of its transfer function

As was shown in [20], much of the structure of $R(\lambda)$ can be retrieved from that of $P(\lambda)$ as in (6), when the latter is a minimal system matrix. It is shown there that in that case

- The pole structure of $R(\lambda)$ at its finite poles is identical to the zero structure of $A(\lambda)$ at its finite zeros.
- The zero structure of $R(\lambda)$ at its finite zeros is identical to the zero structure of $P(\lambda)$ at its finite zeros.
- The left and right null-space structures of $R(\lambda)$ and $P(\lambda)$ are the same.

In the above relations, *structure* refers to the index sets of the finite elementary divisors and of the left and right minimal bases. In the rest of this section we show there is also a relation between the *root vectors* as well as the index sets. For simplicity, we focus on finite eigenvalues.

Theorem 5.1. *Let $R(\lambda) = D(\lambda) + C(\lambda)A(\lambda)^{-1}B(\lambda) \in \mathbb{F}(\lambda)^{m\times n}$ be a rational matrix and let $P(\lambda)$ as in (6) be a minimal polynomial system matrix of $R(\lambda)$. If $x(\lambda)$ is a root vector of order k at λ_0 for $R(\lambda)$, then $y(\lambda) = \begin{bmatrix} A(\lambda)^{-1}B(\lambda)x(\lambda) \\ x(\lambda) \end{bmatrix}$ is a root vector of order k at λ_0 for $P(\lambda)$.*

Moreover, if $\{x_i(\lambda)\}_{i=1}^s$ *are a maximal set of root vectors at* λ_0 *for* $R(\lambda)$ *and* $y_i(\lambda) = \begin{bmatrix} A(\lambda)^{-1}B(\lambda)x_i(\lambda) \\ x_i(\lambda) \end{bmatrix}$ *then* $\{y_i(\lambda)\}_{i=1}^s$ *are a maximal set of root vectors at* λ_0 *for* $P(\lambda)$.

Proof. It is readily verified that

$$P(\lambda)y(\lambda) = \begin{bmatrix} 0 \\ R(\lambda)x(\lambda) \end{bmatrix} \Rightarrow \kappa[P(\lambda)y(\lambda)] = k.$$

Since by assumption $\kappa[x(\lambda)] = 0$, to show that $\kappa[y(\lambda)] = 0$ it suffices to verify that $\kappa[A(\lambda)^{-1}B(\lambda)x(\lambda)] \geq 0$; this can be proved as in [18, Proposition 2.5] (the result in [18] is stated for $\mathbb{F} = \mathbb{C}$ and proved using limits, but it is not difficult to extend it to any field and by using valuations). At this point, the proof of [1, Theorem 3.8, part (a)] shows that, if $M(\lambda) \in \mathbb{F}[\lambda]^{n\times p}$ is a minimal basis for $\ker R(\lambda)$, then $N(\lambda) = \begin{bmatrix} A(\lambda)^{-1}B(\lambda)M(\lambda) \\ M(\lambda) \end{bmatrix}$ is a basis for $\ker P(\lambda)$ that has full column rank for all $\lambda \in \mathbb{F}$. As a consequence, $x(\lambda_0) \notin \ker_{\lambda_0} R(\lambda) \Rightarrow y(\lambda_0) \notin \ker_{\lambda_0} P(\lambda)$.

Suppose now that $\{x_i(\lambda)\}_{i=1}^s$ are a maximal set for $R(\lambda)$. Then, $\{y_i(\lambda)\}_{i=1}^s$ are λ_0-independent because

$$\begin{aligned} p + s &\geq \operatorname{rank}\begin{bmatrix} N(\lambda_0) & y_1(\lambda_0) & \dots & y_s(\lambda_0) \end{bmatrix} \\ &\geq \operatorname{rank}\begin{bmatrix} M(\lambda_0) & x_1(\lambda_0) & \dots & x_s(\lambda_0) \end{bmatrix} = p + s. \end{aligned}$$

That $\{y_i(\lambda)\}_{i=1}^s$ are also complete (resp. maximal) follows by Theorem 3.10 (resp. Theorem 3.12) and by Rosenbrock's theorem [4]. □

A natural question is whether the converse of Theorem 5.1 holds, i.e., if the bottom block of a root polynomial for $P(\lambda)$ is always a root vector for $R(\lambda)$. The answer is positive if we consider a point λ_0 which is not a pole of $R(\lambda)$.

Proposition 5.2. *Let* $R(\lambda) = D(\lambda) + C(\lambda)A(\lambda)^{-1}B(\lambda)$ *be a rational matrix and let* $P(\lambda)$ *as in* (6) *be a minimal polynomial system matrix of* $R(\lambda)$. *Suppose, moreover, that* λ_0 *is not a pole of* $R(\lambda)$. *If* $y(\lambda)$ *is a root vector of order* k *for* $P(\lambda)$, *then there exists a rational vector* $v(\lambda)$ *such that* $\kappa[v(\lambda)] \geq 0$, $y(\lambda) = \begin{bmatrix} A(\lambda)^{-1}B(\lambda)x(\lambda) + (\lambda - \lambda_0)^k v(\lambda) \\ x(\lambda) \end{bmatrix}$, *and* $x(\lambda)$ *is a root vector of order at least* k *at* λ_0 *for* $R(\lambda)$.

Moreover, if $\{y_i(\lambda)\}_{i=1}^s$ *are a maximal set of root vectors at* λ_0 *for* $P(\lambda)$ *and for all* i *we denote by* $x_i(\lambda)$ *the bottom block of* $y_i(\lambda)$, *then* $\{x_i(\lambda)\}_{i=1}^s$ *are a maximal set of root vectors at* λ_0 *for* $R(\lambda)$.

Proof. Since $P(\lambda)$ is minimal and λ_0 is not a pole of $R(\lambda)$, it follows by [4] that λ_0 is not an eigenvalue of $A(\lambda)$. Hence, $A(\lambda)$ is unimodular over the local ring (2) and so is $\begin{bmatrix} I & 0 \\ C(\lambda)A(\lambda)^{-1} & I \end{bmatrix}$. Thus, by Proposition 3.6, $y(\lambda)$ is a root vector at λ_0 of order k for

$$\begin{bmatrix} I & 0 \\ C(\lambda)A(\lambda)^{-1} & I \end{bmatrix} P(\lambda) = \begin{bmatrix} -A(\lambda) & B(\lambda) \\ 0 & R(\lambda) \end{bmatrix}.$$

Decomposing $y(\lambda) = \begin{bmatrix} A(\lambda)^{-1}B(\lambda)x(\lambda) + z(\lambda) \\ x(\lambda) \end{bmatrix}$ it follows that

$$\kappa \left[\begin{bmatrix} -A(\lambda)z(\lambda) \\ R(\lambda)x(\lambda) \end{bmatrix} \right] = k.$$

Hence, $\kappa[R(\lambda)x(\lambda)] \geq k$. Moreover, by item 5 in Proposition 3.3, $\kappa[z(\lambda)] = \kappa[-A(\lambda)z(\lambda)] \geq k$ and therefore we can write $z(\lambda) = (\lambda - \lambda_0)^k v(\lambda)$ as sought. Furthermore, it must be $\kappa[x(\lambda)] = 0$, for if the valuation was positive then we would have that

$$\kappa[y(\lambda)] \geq \min\{\kappa[x(\lambda)], \kappa[B(\lambda)x(\lambda)], \kappa[z(\lambda)]\} \geq \min\{k, \kappa[x(\lambda)]\} > 0.$$

The argument to show $y(\lambda_0) \notin \ker_{\lambda_0} P(\lambda) \Rightarrow x(\lambda_0) \notin \ker_{\lambda_0} R(\lambda_0)$ is similar to its analogue in the proof of Theorem 5.1, except that this time we need to apply the proof of [1, Theorem 3.6, part (a)].

Suppose now that $\{y_i(\lambda)\}_{i=1}^s$ are a maximal set. We claim that $\{x_i(\lambda)\}_{i=1}^s$ are λ_0-independent: this suffices to prove that they are maximal, by Rosenbrock's Theorem [4], Theorem 3.10 and Theorem 3.12. To prove the claim, suppose for a contradiction that it is false. Denote by $M(\lambda)$ a minimal basis of $\ker R(\lambda)$ and observe [1] that $N(\lambda) = \begin{bmatrix} A(\lambda)^{-1}B(\lambda)M(\lambda) \\ M(\lambda) \end{bmatrix}$ is a basis for $\ker P(\lambda)$ that has full rank for $\lambda = \lambda_0$. Then, for some nonzero constant vector c, it holds $0 = [M(\lambda_0) \quad x_1(\lambda_0) \quad \dots x_s(\lambda_0)]\, c$ which implies

$$\begin{bmatrix} A(\lambda_0)^{-1}B(\lambda_0) \\ I \end{bmatrix} [M(\lambda_0) \quad x_1(\lambda_0) \quad \dots x_s(\lambda_0)]\, c = [N(\lambda_0) \quad y_1(\lambda_0) \quad \dots y_s(\lambda_0)]\, c = 0. \qquad \square$$

(A special case of) Proposition 5.2 shows that the problem of finding a maximal set of root vectors of a transfer function at a point which is a zero, but not a pole, can be reduced to the corresponding problem of a pencil

$$S(\lambda) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] - \lambda \left[\begin{array}{c|c} E & 0 \\ \hline 0 & 0 \end{array}\right] \tag{7}$$

derived from an irreducible generalized state-space realization. When $\mathbb{F} \subseteq \mathbb{C}$, an algorithm to compute root vectors (in fact, root polynomials, which are of course a special case of root vectors satisfying the extra constraint of being polynomial) of a pencil was recently studied in [19].

If we remove the assumption that λ_0 is not a pole, then the statement of Proposition 5.2 may no longer be true as the following example illustrates.

Example 5.3.

$$R(\lambda) = \begin{bmatrix} 1 & 0 \\ \frac{1}{\lambda} & 1 \end{bmatrix}, \qquad P(\lambda) = \begin{bmatrix} -\lambda & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

It is easy to verify that $P(\lambda)$ is a minimal polynomial system matrix of $R(\lambda)$. The vector $\begin{bmatrix} -1 & 0 & 1 \end{bmatrix}^T$ is a root polynomial of order 1 at 0 for $P(\lambda)$. However, its bottom block $\begin{bmatrix} 0 & 1 \end{bmatrix}^T$ is not a root vector at 0 for $R(\lambda)$.

To overcome this difficulty, we need a technique to remove the poles while keeping the original positive partial multiplicities, and simultaneously being able to recover maximal sets of root vectors. Let us start from a sufficient condition for root polynomial recovery even in the presence of poles.

Lemma 5.4. *Let $R(\lambda) = D(\lambda) + C(\lambda)A(\lambda)^{-1}B(\lambda)$ be a rational matrix and let $P(\lambda)$ as in (6) be a minimal polynomial system matrix of $R(\lambda)$. Suppose that $y(\lambda) = \begin{bmatrix} u(\lambda) \\ x(\lambda) \end{bmatrix}$ is a root vector of order k at λ_0 for $P(\lambda)$ satisfying the additional condition $A(\lambda)u(\lambda) = B(\lambda)x(\lambda)$. Then, $x(\lambda)$ is a root vector of order k at λ_0 for $R(\lambda)$.*

Moreover, suppose that $\{y_i(\lambda)\}_{i=1}^s$ is a maximal set of root vectors at λ_0 for $P(\lambda)$, all satisfying the additional condition $A(\lambda)u_i(\lambda) = B(\lambda)x_i(\lambda)$ where $y_i(\lambda) = \begin{bmatrix} u_i(\lambda) \\ x_i(\lambda) \end{bmatrix}$. Then, $\{x_i(\lambda)\}_{i=1}^s$ is a maximal set of root vectors at λ_0 for $R(\lambda)$.

Proof. By assumption, $u(\lambda) = A(\lambda)^{-1}B(\lambda)x(\lambda)$ and hence

$$P(\lambda)y(\lambda) = \begin{bmatrix} 0 \\ R(\lambda)x(\lambda) \end{bmatrix} \Rightarrow \kappa[R(\lambda)x(\lambda)] = \kappa[P(\lambda)y(\lambda)] = k.$$

Suppose now for a contradiction that $x(\lambda_0) \in \ker_{\lambda_0} R(\lambda)$; then, if $M(\lambda)$ is a minimal basis for $\ker R(\lambda)$, $x(\lambda_0) = M(\lambda_0)c$ for some constant vector c. On the other hand, by the results in [1],

$$N(\lambda) = \begin{bmatrix} A(\lambda)^{-1}B(\lambda)M(\lambda) \\ M(\lambda) \end{bmatrix}$$

is a polynomial basis for $\ker P(\lambda)$ that has full column rank upon evaluation at $\lambda = \lambda_0$, and therefore we obtain the contradiction $y(\lambda_0) = N(\lambda_0)c \Rightarrow y(\lambda_0) \in \ker_{\lambda_0} P(\lambda)$. Finally, clearly $\kappa[x(\lambda)] \geq \kappa[y(\lambda)] = 0$, but if $\kappa[x(\lambda)] > 0$ then $x(\lambda_0) = 0 \in \ker_{\lambda_0} R(\lambda)$, and thus $\kappa[x(\lambda)] = 0$.

Now suppose that $\{y_i(\lambda)\}_{i=1}^s$ are a maximal set of root vectors at λ_0 for $P(\lambda)$ as in the statement. We claim that $\{x_i(\lambda)\}_{i=1}^s$ are λ_0-independent. Hence, by Rosenbrock's Theorem [4], Theorem 3.10 and Theorem 3.12, we conclude that they are also maximal. It remains to prove the claim. To this goal, denote

$$X(\lambda) := [x_1(\lambda) \quad \dots \quad x_s(\lambda)], \qquad Y(\lambda) := [y_1(\lambda) \quad \dots \quad y_s(\lambda)]$$

and define $F(\lambda) := A(\lambda)^{-1}B(\lambda)M(\lambda)$, $U(\lambda) := A(\lambda)^{-1}B(\lambda)X(\lambda)$; as already observed, neither $F(\lambda)$ nor $U(\lambda)$ have poles at λ_0. Note also that, since $\begin{bmatrix} -A(\lambda) \\ C(\lambda) \end{bmatrix}$ has trivial Smith form, it has some polynomial left inverse $L(\lambda)$. Moreover, there is a matrix $W(\lambda)$ with the property that $W(\lambda_0) = 0$ and such that

$$\begin{bmatrix} 0 & 0 \\ 0 & W(\lambda) \end{bmatrix} = P(\lambda) \begin{bmatrix} N(\lambda) & Y(\lambda) \end{bmatrix} = \begin{bmatrix} -A(\lambda) \\ C(\lambda) \end{bmatrix} \begin{bmatrix} F(\lambda) & U(\lambda) \end{bmatrix} + \begin{bmatrix} B(\lambda) \\ D(\lambda) \end{bmatrix} \begin{bmatrix} M(\lambda) & X(\lambda) \end{bmatrix}$$

which implies

$$\begin{bmatrix} F(\lambda_0) & U(\lambda_0) \end{bmatrix} = -L(\lambda_0) \begin{bmatrix} B(\lambda_0) \\ D(\lambda_0) \end{bmatrix} \begin{bmatrix} M(\lambda_0) & X(\lambda_0) \end{bmatrix} =: H \begin{bmatrix} M(\lambda_0) & X(\lambda_0) \end{bmatrix}$$

where H is some constant matrix. Thus,

$$\begin{bmatrix} N(\lambda_0) & Y(\lambda_0) \end{bmatrix} = \begin{bmatrix} F(\lambda_0) & U(\lambda_0) \\ M(\lambda_0) & X(\lambda_0) \end{bmatrix} = \begin{bmatrix} H \\ I \end{bmatrix} \begin{bmatrix} M(\lambda_0) & X(\lambda_0) \end{bmatrix}.$$

Suppose now for a contradiction that there exists a nonzero constant vector c satisfying $\begin{bmatrix} M(\lambda_0) & X(\lambda_0) \end{bmatrix} c = 0$; then,

$$\begin{bmatrix} N(\lambda_0) & Y(\lambda_0) \end{bmatrix} c = \begin{bmatrix} H \\ I \end{bmatrix} 0 = 0,$$

contradicting the λ_0-independence of $\{y_i(\lambda)\}_{i=1}^s$. □

We are now ready to remove the poles as discussed above.

Theorem 5.5. *Let $R(\lambda) \in \mathbb{F}(\lambda)^{m\times n}$ and $\lambda_0 \in \mathbb{F}$. There exists a square and regular $K(\lambda) \in \mathbb{F}[\lambda]^{n\times n}$ such that*

1. *$Y(\lambda) = R(\lambda)K(\lambda) \in \mathbb{F}[\lambda]^{m\times n}$ does not have poles at λ_0;*
2. *the positive partial multiplicities at λ_0 of $Y(\lambda)$ and $R(\lambda)$ coincide;*
3. *the factorization $R(\lambda) = Y(\lambda)K(\lambda)^{-1}$ is right coprime, i.e., the Smith normal form of $\begin{bmatrix} -K(\lambda) \\ Y(\lambda) \end{bmatrix}$ is $\begin{bmatrix} I \\ 0 \end{bmatrix}$;*
4. *if $r(\lambda)$ is a root vector at λ_0 of order k for $Y(\lambda)$ then $K(\lambda)r(\lambda)$ is a root vector at λ_0 of order k for $R(\lambda)$;*
5. *if $u(\lambda)$ is a root vector at λ_0 of order k for $R(\lambda)$ then $K(\lambda)^{-1}u(\lambda)$ is a root vector at λ_0 of order k for $Y(\lambda)$;*
6. *the vectors $\{K(\lambda)r_i(\lambda)\}$ are a maximal set of root vectors at λ_0 for $R(\lambda)$ if and only if the vectors $\{r_i(\lambda)\}$ are a maximal set of root vectors at λ_0 for $Y(\lambda)$, with the same orders.*

It is easy to prove Theorem 5.5 using the Smith form. However, the computation of the Smith form of $R(\lambda)$ (including the unimodular matrices that lead to it) induces the computation of a maximal set of root vectors (they are a certain subset of the columns of one of the unimodular matrices). Instead, we are interested in using Theorem 5.5 as a preprocessing tool to compute maximal sets of root vectors when λ_0 is a pole (that is, we wish to use it to reduce to the case where λ_0 is not a pole). For this reason, we give below a proof which does not rely on the Smith form of $R(\lambda)$. While this is less direct, it has the advantage of not "solving the problem in order to solve the problem" in the sense discussed above. Indeed, this proof will inspire our algorithmic approach.

Proof of Theorem 5.5. Without loss of generality, we can represent $R(\lambda) = D(\lambda) + C(\lambda)A(\lambda)^{-1}B(\lambda) \in \mathbb{F}(\lambda)^{m\times n}$ where $A(\lambda), B(\lambda), C(\lambda), D(\lambda)$ are matrices over $\mathbb{F}[\lambda]$ of coherent size such that $A(\lambda)$ is invertible and $P(\lambda)$ as in (6) is minimal, i.e., the matrices $\begin{bmatrix}-A(\lambda) & B(\lambda)\end{bmatrix}$ and $\begin{bmatrix}-A(\lambda)\\ C(\lambda)\end{bmatrix}$ have both trivial Smith form over $\mathbb{F}[\lambda]$. Then, by [3, Theorem 3.3], there exists a unimodular polynomial matrix $U(\lambda)$ such that $\begin{bmatrix}-A(\lambda) & B(\lambda)\end{bmatrix}U(\lambda) = \begin{bmatrix}I & 0\end{bmatrix}$. Hence, there exist polynomial matrices $X(\lambda)$ and $Y(\lambda)$ such that

$$P(\lambda)U(\lambda) = \begin{bmatrix} I & 0 \\ X(\lambda) & Y(\lambda)\end{bmatrix}$$

implying by Rosenbrock's theorem [4] that the polynomial matrix $Y(\lambda)$ has the same positive partial multiplicities at λ_0 as $R(\lambda)$, and (by the fact that it is polynomial) no negative partial multiplicities at λ_0. Similarly, for some rational matrices $W(\lambda), Z(\lambda)$, it holds

$$\begin{bmatrix}-A(\lambda) & B(\lambda)\\ 0 & R(\lambda)\end{bmatrix}U(\lambda) = \begin{bmatrix} I & 0 \\ W(\lambda) & Z(\lambda)\end{bmatrix}.$$

On the other hand, over $\mathbb{F}(\lambda)$,

$$\begin{bmatrix}-A(\lambda) & B(\lambda)\\ 0 & R(\lambda)\end{bmatrix} = \begin{bmatrix} I & 0 \\ C(\lambda)A(\lambda)^{-1} & I\end{bmatrix}P(\lambda).$$

Hence, $Z(\lambda) = Y(\lambda)$, implying that $Y(\lambda) = \begin{bmatrix}0 & R(\lambda)\end{bmatrix}U(\lambda)\begin{bmatrix}0\\ I\end{bmatrix}$. In other words, $Y(\lambda) = R(\lambda)K(\lambda)$ for some square polynomial matrix $K(\lambda) := \begin{bmatrix}0 & I\end{bmatrix}U(\lambda)\begin{bmatrix}0\\ I\end{bmatrix}$. By construction, neither $Y(\lambda)$ nor $K(\lambda)$ have poles at λ_0, and the positive partial multiplicities at λ_0 of $R(\lambda)$ and $Y(\lambda)$ coincide. This proves items 1–2.

Moreover, arguing as in [4, Theorem 9], it is not difficult to prove that $K(\lambda)$ is regular. Finally, a slight modification of the arguments in [4] also shows that

$$Q(\lambda) := \begin{bmatrix}-K(\lambda) & I\\ Y(\lambda) & 0\end{bmatrix}$$

is a minimal system matrix of $R(\lambda)$ (proving item 3).

Now suppose that $r(\lambda)$ is a root vector at λ_0 of order k for $Y(\lambda)$. We claim that $q(\lambda) = \begin{bmatrix} r(\lambda) \\ K(\lambda)r(\lambda) \end{bmatrix}$ is a root vector at λ_0 of order k for $Q(\lambda)$. Indeed, clearly $\kappa[q(\lambda)] = \kappa[r(\lambda)] = 0$ and $\kappa[Q(\lambda)q(\lambda)] = k$. Moreover, if $q(\lambda_0) \in \ker_{\lambda_0} Q(\lambda)$, then by [6, Lemma 2.9] there is a polynomial vector $p(\lambda)$ such that $p(\lambda_0) = q(\lambda_0)$ and $Q(\lambda)p(\lambda) = 0$. Partitioning $p(\lambda) = \begin{bmatrix} p_1(\lambda) \\ p_2(\lambda) \end{bmatrix}$ this in turn implies $p_1(\lambda_0) = r(\lambda_0)$ and $Y(\lambda)p_1(\lambda) = 0$ which, again by [6, Lemma 2.9], contradicts the assumption $r(\lambda_0) \notin \ker_{\lambda_0} Y(\lambda)$. At this point it suffices to observe that $q(\lambda)$ also satisfies the extra condition in Lemma 5.4 to prove item 4.

Conversely, assume now that $u(\lambda)$ is a root vector at λ_0 of order k for $R(\lambda)$. Then, by Theorem 5.1, $\begin{bmatrix} K(\lambda)^{-1}u(\lambda) \\ u(\lambda) \end{bmatrix}$ is a root vector at λ_0 of order k for $Q(\lambda)$. Defining $v(\lambda) := K(\lambda)^{-1}u(\lambda)$, so that $u(\lambda) = K(\lambda)v(\lambda)$, note that the previous observation implies $\kappa[v(\lambda)] \geq 0$. Suppose for a contradiction that $v(\lambda_0) \in \ker_{\lambda_0} Y(\lambda)$. Using [6, Lemma 2.9], this implies the existence of a polynomial vector $z(\lambda)$ such that $z(\lambda_0) = v(\lambda_0)$ and $Y(\lambda)z(\lambda) = 0$. Hence, $R(\lambda)[K(\lambda)z(\lambda)] = 0$ and $K(\lambda_0)z(\lambda_0) = u(\lambda_0)$, yielding $u(\lambda_0) \in \ker_{\lambda_0} R(\lambda)$: a contradiction. In particular, it cannot be $\kappa[v(\lambda)] > 0$ lest $v(\lambda_0) = 0 \in \ker_{\lambda_0} Y(\lambda)$, and this concludes the proof of item 5.

For item 6, let us first start from a maximal set $\{r_i(\lambda)\}_{i=1}^s$ for $Y(\lambda)$. Similarly to what we did before, we can construct from them a maximal set $\{q_i(\lambda)\}_{i=1}^s$ for $Q(\lambda)$ that, moreover, (a) satisfy also the special property of Lemma 5.4 (b) have bottom block $K(\lambda)r_i(\lambda)$. Applying Lemma 5.4, we conclude that $\{K(\lambda)r_i(\lambda)\}_{i=1}^s$ is a maximal set for $R(\lambda)$. For the converse implication, let $M(\lambda)$ be a minimal basis for $\ker R(\lambda)$. From the analysis in [1], the fact that $R(\lambda)$ and $Y(\lambda)$ must have the same normal rank, and the fact that $Q(\lambda)$ is minimal, we see that $N(\lambda) = K(\lambda)^{-1}M(\lambda)$ is a polynomial basis for $\ker Y(\lambda)$; and if $N(\lambda_0)$ did not have full column rank, then $M(\lambda_0) = K(\lambda_0)N(\lambda_0)$ also would not: a contradiction. Hence, given a maximal set for $R(\lambda)$ of the form $\{K(\lambda)r_i(\lambda)\}_{i=1}^s$ (any root vector of $R(\lambda)$ can be written in such form by the previous items),

$$\operatorname{rank}\begin{bmatrix} M(\lambda_0) & K(\lambda_0)r_1(\lambda_0) & \dots & K(\lambda_0)r_s(\lambda_0) \end{bmatrix} \leq \operatorname{rank}\begin{bmatrix} N(\lambda_0) & r_1(\lambda_0) & \dots & r_s(\lambda_0) \end{bmatrix};$$

since the matrix on the left hand side has full column rank by assumption, so must have the matrix on the right hand side. Hence, $\{r_i(\lambda)\}_{i=1}^s$ are λ_0-independent. As usual, we now invoke Theorem 3.10 and Theorem 3.12, together with the previously proved fact that $R(\lambda)$ and $Y(\lambda)$ have the same positive partial multiplicities at λ_0, to also establish completeness and maximality. □

Example 5.6. The construction of Theorem 5.5, applied to the pair $R(\lambda), P(\lambda)$ in Example 5.3, leads to picking (for instance)

$$U(\lambda) = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & \lambda \\ 0 & 1 & 0 \end{bmatrix} \Rightarrow K(\lambda) = \begin{bmatrix} 0 & \lambda \\ 1 & 0 \end{bmatrix} \Rightarrow Y(\lambda) = \begin{bmatrix} 0 & \lambda \\ 1 & 1 \end{bmatrix}.$$

The vector $[-1 \quad 1]^T$ is a root vector at 0 of order 1 for $Y(\lambda)$, and we can recover a root vector at 0 of order 1 for $R(\lambda)$ as

$$K(\lambda)\begin{bmatrix}-1\\1\end{bmatrix} = \begin{bmatrix}\lambda\\-1\end{bmatrix}.$$

Remark 5.7. It is worth noting that, in Example 5.6, left multiplication times $K(\lambda)$ increased the degree of the (order 1) root vector from 0 to 1. This is remarkable: while, given $\lambda_0 \in \mathbb{F}$, a root polynomial at λ_0 of order k for a *polynomial* matrix can always be assumed to have degree at most $k-1$ [6, Section 5], it may be necessary (when the eigenvalue is also a pole) that a root polynomial at λ_0 of order k for a *rational* matrix must have degree higher than $k-1$. This is, indeed, shown by Example 5.6, because it is easy to prove that no polynomial vector of degree 0 can be a root vector at 0 for $R(\lambda)$.

However, let $-m$ be the smallest partial multiplicity at λ_0 of $R(\lambda)$; note that $m > 0$ if we assume that λ_0 is also a pole. Then, $(\lambda-\lambda_0)^m R(\lambda)$ is polynomial, and it is not difficult to show that $r(\lambda)$ is a root polynomial of order k at λ_0 for $R(\lambda)$ if and only if it is a root polynomial of order $k+m$ at λ_0 for $(\lambda-\lambda_0)^m R(\lambda)$. Hence, it follows that we can assume without loss of generality that a root polynomial of order k at λ_0 for a rational matrix $R(\lambda)$, whose smallest partial multiplicity at λ_0 is $-m$, has degree at most $k+m-1$. For instance, in Example 5.6, we had $k=m=1$ and indeed we constructed a root polynomial of degree $k+m-1=1$.

Remark 5.8. The ingredients needed for the proof of Theorem 5.5 can be extended to cover the case where the base ring is any principal ideal domain whose field of fractions is $\mathbb{F}(\lambda)$. In particular, we can replace $\mathbb{F}[\lambda]$ with the ring (2), which by construction is a local domain and thus a p.i.d., and reach analogous conclusions following essentially the same arguments.

In particular, in this case, the starting point of the construction is to express $R(\lambda) = D(\lambda)+C(\lambda)A(\lambda)^{-1}B(\lambda)$. Then, we build a matrix $P(\lambda) = \begin{bmatrix}-A(\lambda) & B(\lambda)\\ C(\lambda) & D(\lambda)\end{bmatrix}$ with entries in (2), with $A(\lambda)$ regular, and minimal at λ_0. Minimality at λ_0 is defined by asking that both $[-A(\lambda) \quad B(\lambda)]$ and $\begin{bmatrix}-A(\lambda)\\ C(\lambda)\end{bmatrix}$ have trivial Smith form over the ring (2). (The Smith form over (2) is often referred to as the local Smith form at λ_0.) Note that this is a weaker assumption than $P(\lambda)$ being minimal and polynomial. Moreover, $U(\lambda)$ is unimodular over (2) (i.e. a rational matrix that satisfies the same assumptions of the matrices $A(\lambda), B(\lambda)$ in Theorem 3.7, or in other words is a rational matrix without poles or zeros at λ_0), and $K(\lambda)$ and $Y(\lambda)$ are both rational matrices without poles at λ_0.

The advantage of working locally is that, instead of removing *all* the poles of $R(\lambda)$, one would remove only the poles *at the point of interest*. This can be more efficient if one is only interested in the root vectors at one point.

An even further generalization can be realized working over the localization

$$\mathcal{L} = \left\{ \frac{a(\lambda)}{b(\lambda)} \in \mathbb{F}(\lambda) : \gcd(a(\lambda), b(\lambda)) = 1 \text{ and } b(\lambda) \in S \right\} \tag{8}$$

where S is any multiplicatively closed subset of $\mathbb{F}[\lambda]$ such that $1 \in S, 0 \notin S$. Indeed, it is not difficult to show that $\mathcal{L}$ is again a principal ideal domain whose field of fractions is $\mathbb{F}(\lambda)$.[5]

We formalize the analysis of Remark 5.8 in the following corollary.

Corollary 5.9. *Let $R(\lambda) \in \mathbb{F}(\lambda)^{m\times n}$ and $\lambda_0 \in \mathbb{F}$. Let S be a multiplicatively closed subset of $\mathbb{F}[\lambda]$ such that $1 \in S, 0 \notin S$, and $p(\lambda) \in S \Rightarrow p(\lambda_0) \neq 0$. Moreover, let $\mathcal{L}$ be the associated ring (8). There exists a square and regular $n \times n$ matrix $K(\lambda)$ over (8) such that*

1. *$Y(\lambda) = R(\lambda)K(\lambda)$ has elements in (8) and does not have poles at λ_0;*
2. *the positive partial multiplicities at λ_0 of $Y(\lambda)$ and $R(\lambda)$ coincide;*
3. *the factorization $R(\lambda) = Y(\lambda)K(\lambda)^{-1}$ is right coprime (over (8)), i.e., the Smith normal form (over (8)) of $\begin{bmatrix} -K(\lambda) \\ Y(\lambda) \end{bmatrix}$ is $\begin{bmatrix} I \\ 0 \end{bmatrix}$;*
4. *if $r(\lambda)$ is a root vector at λ_0 of order k for $Y(\lambda)$ then $K(\lambda)r(\lambda)$ is a root vector at λ_0 of order k for $R(\lambda)$;*
5. *if $u(\lambda)$ is a root vector at λ_0 of order k for $R(\lambda)$ then $K(\lambda)^{-1}u(\lambda)$ is a root vector at λ_0 of order k for $Y(\lambda)$;*
6. *the vectors $\{K(\lambda)r_i(\lambda)\}$ are a maximal set of root vectors at λ_0 for $R(\lambda)$ if and only if the vectors $\{r_i(\lambda)\}$ are a maximal set of root vectors at λ_0 for $Y(\lambda)$, with the same orders.*

Corollary 5.9 suggests a procedure to deal with the case when λ_0 is also a pole. The idea is to modify the state space realization of $R(\lambda)$ to construct a rational matrix $Y(\lambda)$ as in Remark 5.8, that has no poles at λ_0 but has simply related (equal up to premultiplication by $K(\lambda)$) maximal sets of root vectors at λ_0 as $R(\lambda)$, with the same orders.

In the next subsection, we will show how to realize this in practice by modifying a state space realization of $R(\lambda)$ in order to build a state space realization of $Y(\lambda)$.

Remark 5.10. While for many results in the present subsection we have assumed that the polynomial system matrix is (globally) minimal, this assumption can be weakened

[5] Proof: It is a known fact in ring theory that any ideal of $\mathcal{L}$ has the form $S^{-1}J$ where J is an ideal of $\mathbb{F}[\lambda]$. But $\mathbb{F}[\lambda]$ is a p.i.d., and hence $J = \langle g(\lambda) \rangle$ for some polynomial $g(\lambda)$. It follows that $S^{-1}J = \langle \frac{g(\lambda)}{1} \rangle$ is also a principal ideal.

to local minimality at λ_0. We omit a detailed discussion to keep the length of the paper reasonable.

Remark 5.11. Theorem 5.5 and Corollary 5.9 can in fact be further generalized to another interesting case, by taking the localization $\mathcal{L}$ to be the ring of proper rational functions

$$\mathcal{L} = \left\{ \frac{a(\lambda)}{b(\lambda)} \in \mathbb{F}(\lambda) : \gcd(a(\lambda), b(\lambda)) = 1 \text{ and } \deg a(\lambda) \leq \deg b(\lambda) \right\} \tag{9}$$

instead of (8). If $\mathcal{L}$ is the ring (9), we can take $\lambda_0 = \infty$ and construct a rational matrix $Y(\lambda) = K(\lambda)R(\lambda)$ with the following properties: (1) $K(\lambda)$ is square and regular, and the elements of $Y(\lambda)$ and $K(\lambda)$ are all proper rational functions (2) $Y(\lambda)$ has no poles at ∞ and has the same positive partial multiplicities at ∞ as $R(\lambda)$ (3) the factorization $R(\lambda) = Y(\lambda)K(\lambda)^{-1}$ is coprime over (9) (4) root vectors (resp. maximal sets of root vectors) at ∞ for $R(\lambda)$ are $K(\lambda)$ times root vectors (resp. maximal sets of root vectors) at infinity for $Y(\lambda)$, and vice versa root vectors (resp. maximal sets of root vectors) at ∞ for $Y(\lambda)$ are $K(\lambda)^{-1}$ times root vectors (resp. maximal sets of root vectors) at ∞ for $R(\lambda)$. (Here, root vectors at infinity should be understood as in the Smith-McMillan approach described in Subsection 3.1.)

This claim could be proved again in the same way as Theorem 5.5. For the sake of completeness, we give below a shorter proof via the Smith form. Write $R(\lambda) = U(\lambda)S_R(\lambda)V(\lambda) \in \mathbb{F}(\lambda)^{m\times n}$ where $S_R(\lambda)$ is the $m \times n$ Smith form of $R(\lambda)$ over the ring (9) and $U(\lambda), V(\lambda)$ are unimodular over (9) (i.e. they are square biproper rational matrices). Then without loss of generality

$$S_R(\lambda) = \operatorname{diag}(\lambda^{-\sigma_r}, \dots, \lambda^{-\sigma_1}, 0, \dots, 0)$$

where $r = \operatorname{rank} R(\lambda)$ and $\sigma_1 \geq \cdots \geq \sigma_r$ are integers, and the notation $\operatorname{diag}(m_1, \dots, m_\ell)$ denotes a (possibly rectangular) diagonal matrix with diagonal elements $m_1, \dots, m_\ell$. Let p be the integer such that $\sigma_p \geq 0 > \sigma_{p+1}$ $(0 \leq p \leq r)$; in other words $r - p$ is the number of poles at ∞ of $R(\lambda)$. Define the $n \times n$ square matrix

$$S_K(\lambda) = \operatorname{diag}(\lambda^{\sigma_r}, \dots, \lambda^{\sigma_{p+1}}, 1, \dots, 1).$$

Now set $K(\lambda) := V(\lambda)^{-1}S_K(\lambda)$ and $Y(\lambda) := R(\lambda)K(\lambda)$. By construction, $Y(\lambda)$ and $K(\lambda)$ are both proper and moreover $Y(\lambda) = U(\lambda)S_Y(\lambda)$ with

$$S_Y(\lambda) = \operatorname{diag}(1, \dots, 1, \lambda^{-\sigma_p}, \dots, \lambda^{-\sigma_1}, 0, \dots, 0).$$

Hence, all the claims made above (considering root vectors at $\lambda_0 = \infty$) follow easily.

5.1. Coalescent pole/zero: an algorithmic removal of poles that allows us to recover maximal sets of root vectors

In this subsection we specialize to $\mathbb{F} \subseteq \mathbb{C}$ and look at the computation of root vectors of a general rational matrix $R(\lambda) \in \mathbb{C}(\lambda)^{m\times n}$ in the case of a coalescent pole and zero at a finite λ_0. We also assume for simplicity that $R(\lambda)$ has no poles or zeros at ∞, and hence E is non-singular. We first recall a result about a so-called feedback transformation applied to the system matrix of an irreducible generalized state-space realization of a given transfer function of McMillan degree d :

$$S_Y(\lambda) := \begin{bmatrix} A+BF-\lambda E & B \\ C+DF & D \end{bmatrix} = \begin{bmatrix} A-\lambda E & B \\ C & D \end{bmatrix} \begin{bmatrix} I & 0 \\ F & I \end{bmatrix} = S_R(\lambda) \begin{bmatrix} I & 0 \\ F & I \end{bmatrix}. \quad (10)$$

Lemma 5.12 *([25]). Given a pair of complex matrices $(\hat{A}, \hat{B})$ of sizes $(d \times d, d \times m)$ and an arbitrary set of complex numbers $\lambda_1, \ldots, \lambda_n$ then there exists a matrix F such that $\hat{A} + \hat{B}F$ has those eigenvalues if and only if the pair $(\hat{A}, \hat{B})$ is controllable.*

Applying this result to the pair $(\hat{A}, \hat{B}) := (E^{-1}A, E^{-1}B)$ shows that this also holds for generalized state-space systems that are irreducible. The minimality of $S_R(\lambda)$ implies that for all finite λ the following matrices have full row and column rank d, respectively :

$$\operatorname{rank} \begin{bmatrix} A-\lambda E & B \end{bmatrix} = \operatorname{rank} \begin{bmatrix} A-\lambda E \\ C \end{bmatrix} = d. \quad (11)$$

The new system matrix $S_Y(\lambda)$ clearly has the same Smith structure at all finite points, but the poles of the corresponding transfer function $Y(\lambda)$ have changed. This is what we exploit in the next Lemma.

Lemma 5.13. *The transfer functions $R(\lambda)$ and $Y(\lambda)$, of the respective system matrices $S_R(\lambda)$ and $S_Y(\lambda)$ given in (10), are related by an invertible right transformation $K(\lambda)$ with system matrix $S_K(\lambda)$*

$$Y(\lambda) = R(\lambda)K(\lambda), \quad S_K(\lambda) = \begin{bmatrix} A+BF-\lambda E & B \\ F & I \end{bmatrix}. \quad (12)$$

If $S_R(\lambda)$ is minimal, then there always exists a matrix F that assigns the eigenvalues of the pencil $(A + BF - \lambda E)$ to lie in any set $\Lambda \subset \mathbb{C}$. Moreover, if the assigned set Λ is disjoint from the poles and zeros of $R(\lambda)$, then

1. *the inverse transformation $K(\lambda)^{-1}$ is given by the system matrix $S_{K^{-1}}(\lambda) = \begin{bmatrix} A-\lambda E & B \\ -F & I \end{bmatrix}$ and the system matrices $S_K(\lambda)$ and $S_{K^{-1}}(\lambda)$ are both minimal;*

2. *the system matrix* $S_Y(\lambda)$ *is minimal and the Smith-McMillan zeros of* $R(\lambda)$ *and* $Y(\lambda)$ *are the same;*
3. *the factorization* $R(\lambda) = Y(\lambda)K(\lambda)^{-1}$ *is a coprime factorization over the ring of functions*

$$\mathcal{R}_\Lambda := \{r(\lambda) \in \mathbb{C}(\lambda) : |r(\lambda_0)| < \infty \ \ \forall \lambda_0 \notin \Lambda\}, \tag{13}$$

which are rational functions that can only have poles in Λ*;*
4. *the vectors* $\{K(\lambda)r_i(\lambda)\}_{i=1}^s$ *are a maximal set of root vectors at* λ_0 *for* $R(\lambda)$ *if and only if the vectors* $\{r_i(\lambda)\}_{i=1}^s$ *are a maximal set of root vectors at* λ_0 *for* $Y(\lambda)$*, with the same orders.*

Proof. We first prove the relations (12). The following expression is known to be a system matrix for the product of the transfer functions with system matrices $S_K(\lambda)$ and $S_R(\lambda)$ (see [22]) :

$$\left[\begin{array}{cc|c} A-\lambda E & 0 & B \\ 0 & I & 0 \\ \hline C & 0 & D \end{array}\right] \left[\begin{array}{cc|c} I & 0 & 0 \\ 0 & A+BF-\lambda E & B \\ \hline 0 & F & I \end{array}\right] = \left[\begin{array}{cc|c} A-\lambda E & BF & B \\ 0 & A+BF-\lambda E & B \\ \hline C & DF & D \end{array}\right],$$

which is system equivalent to (after applying simple system equivalence transformations)

$$\left[\begin{array}{cc|c} A-\lambda E & 0 & 0 \\ 0 & A+BF-\lambda E & B \\ \hline C & C+DF & D \end{array}\right], \quad \text{and} \quad \left[\begin{array}{c|c} A+BF-\lambda E & B \\ \hline C+DF & D \end{array}\right].$$

To prove item 1, we first verify with the same procedure that the product of the transfer functions with system matrices $S_K(\lambda)$ and $S_{K^{-1}}(\lambda)$ result in the identity matrix. The controllability of $S_{K^{-1}}(\lambda)$ follows from that of $S_R(\lambda)$ and the observability of $S_{K^{-1}}(\lambda)$ follows from the fact that the feedforward term BF of the pair $(A-\lambda E, F)$ moves all eigenvalues of $A-\lambda E$ to new locations. Finally, the minimality of $S_K(\lambda)$ follows from that of $S_{K^{-1}}(\lambda)$, since they both have the same McMillan degree.

The system matrix $S_Y(\lambda)$ may lose minimality only if there is a pole zero cancellation, but the choice of assigned spectrum precludes this, since the poles are the generalized eigenvalues of $(A+BF-\lambda E)$ and hence are disjoint from the zeros of $S_Y(\lambda)$. The rest of item 2 follows trivially from the constant invertible transformation described in (10), between both (minimal) system matrices.

Item 3 follows from the fact that both $K(\lambda)$ and $Y(\lambda)$ have their poles at the spectrum of $(A+BF-\lambda E)$, and hence are bounded outside Λ. Moreover, outside Λ, the compound matrix $\begin{bmatrix} K(\lambda) \\ -Y(\lambda) \end{bmatrix}$ is bounded and equal to the Schur complement of the system matrix

$$\left[\begin{array}{c|c} A+BF-\lambda E & B \\ \hline C+DF & D \\ -F & -I \end{array}\right] = \left[\begin{array}{c|c} A-\lambda E & B \\ \hline C & D \\ 0 & -I \end{array}\right] \begin{bmatrix} I & 0 \\ F & I \end{bmatrix},$$

which has full column rank because the original system was minimal.

Finally, for item 4 we observe that maximal sets for $S_R(\lambda)$ and $S_Y(\lambda)$ are simply related by Theorem 3.7. From a maximal set for $R(\lambda)$, we can construct a maximal set for $S_R(\lambda)$ via Theorem 5.1, whence we construct a maximal set for $S_Y(\lambda)$ which in turn induces a maximal set for $Y(\lambda)$ by Proposition 5.2. From a maximal set for $Y(\lambda)$, we can go via Theorem 5.1 to a maximal set for $S_Y(\lambda)$ that, in addition, also *satisfies the extra condition of Lemma 5.4*. The constant transformation that relates $S_Y(\lambda)$ and $S_R(\lambda)$ is such that this special property is preserved: hence, we can then apply Lemma 5.4 to such a maximal set for $S_R(\lambda)$ and recover a maximal set for $R(\lambda)$. We omit the (straightforward, but tedious) algebraic details necessary to verify that the result of this sequence of transformations eventually leads to the simple relation between maximal sets of $R(\lambda)$ and $Y(\lambda)$ that is described in item 4. □

The analysis of the present subsection can be summarized in the following algorithm:

Input A rational matrix $R(\lambda)$ and a coalescent pole/zero λ_0;
Step 1 Form a minimal system matrix $S_R(\lambda)$ for the transfer function $R(\lambda)$;
Step 2 Find F as in Lemma 5.12 and construct a new system matrix $S_Y(\lambda)$ as in Lemma 5.13;
Step 3 Compute, for example via the algorithm proposed in [19], a maximal set of root polynomials for $S_Y(\lambda)$ and denote them in the block form $\left\{\begin{bmatrix} t_i(\lambda) \\ b_i(\lambda) \end{bmatrix}\right\}_{i=1}^{s}$;
Step 4(a) Obtain a maximal set of root vectors for $R(\lambda)$ as $\{K(\lambda)b_i(\lambda)\}_{i=1}^{s}$ where $K(\lambda)$ is as in Lemma 5.13.

Step 4(a) requires some additional comment, as $K(\lambda)$ is generally rational and only available through its representation $K(\lambda) = I + F(\lambda E - A - BF)^{-1}B$. Thus, one may wonder how to compute $K(\lambda)b_i(\lambda)$ in practice. One option is to obtain $K(\lambda)$ symbolically, but this may be computationally expensive and anyway, in this paper, we focus on numerical computations. Moreover, depending on the application, one may wish to compute a maximal set root vectors of $R(\lambda)$ in the special case where they are polynomial vectors, which is not guaranteed by the form $K(\lambda)b_i(\lambda)$. To achieve these goals, we can replace Step 4(a) by the following "truncated" algorithm.

Step 4(b) Let $-m$ the smallest partial multiplicity in $R(\lambda)$ as in Remark 5.7; note that m can be computed from the staircase form of $A - \lambda E$ (whose calculation can anyway be useful to reduce the overall complexity: see the discussion after Example 5.14 below). Then, assuming for notational simplicity $\lambda_0 = 0$ (if not, we can perform a linear change of variable), denote $M := A + BF$ and

observe that M is not singular because, by construction, we removed the zero eigenvalues of the pencil $\lambda E - M$. Hence,

$$K(\lambda) = I + F(\lambda E - M)^{-1}B = I - F\left(\sum_{j=0}^{\infty} \lambda^j (M^{-1}E)^j\right) M^{-1}B.$$

Suppose now that $b_i(\lambda)$ is a root polynomial of order k_i at $\lambda_0 = 0$ for $Y(\lambda)$, then using Remark 5.7 we can compute a corresponding root polynomial of order k at $\lambda_0 = 0$ for $R(\lambda)$ observing that

$$K(\lambda)b_i(\lambda) \equiv b_i(\lambda) - F\left(\sum_{j=0}^{k_i+m-1} \lambda^j (M^{-1}E)^j\right) M^{-1}Bb_i(\lambda) \mod \lambda^{k_i+m-1}.$$

This operation can be applied to a maximal set of root polynomials for $Y(\lambda)$ (having possibly distinct orders k_i) and produces a maximal set of root polynomials for $R(\lambda)$.

Example 5.14. We illustrate the above procedure for Step 4(b) in the case of the rational matrix of Example 5.3. Let us first set $S_R(\lambda)$ equal to $P(\lambda)$ in Example 5.3 so that

$$A = 0,\ E = 1,\ D = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},\ B = \begin{bmatrix} 1 & 0 \end{bmatrix},\ C = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

We can then pick

$$F = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \Rightarrow S_Y(\lambda) = \begin{bmatrix} -\lambda - 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow Y(\lambda) = \begin{bmatrix} \frac{\lambda}{\lambda+1} & 0 \\ 0 & 1 \end{bmatrix}.$$

At $\lambda_0 = 0$, $S_Y(\lambda)$ has the maximal set of root polynomials $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, whence we extract the maximal set of root polynomials at 0 for $Y(\lambda)$ consisting of the single vector $b(\lambda) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Its order is $k = 1$, while the smallest partial multiplicity at 0 for $R(\lambda)$ is $-m = -1$. (Note that m is the largest size of a Jordan block with eigenvalue 0 in the pencil $\lambda E - A$.) Hence, we can look for a root polynomial at 0 for $R(\lambda)$ of degree at most $k + m - 1 = 1$. In particular, $M = A + BF = -1$ and therefore

$$K(\lambda)b(\lambda) \equiv \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \begin{bmatrix} -1 \\ -1 \end{bmatrix} \left(\sum_{j=0}^{1} (-\lambda)^j\right) \begin{bmatrix} -1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \lambda \\ \lambda - 1 \end{bmatrix} \mod \lambda.$$

Thus, Step 4(b) constructs $\begin{bmatrix} \lambda \\ \lambda - 1 \end{bmatrix}$ as a maximal set of root polynomials at 0 for $R(\lambda)$. Alternatively, one may apply Step 4(a) and compute exactly

$$K(\lambda) = \begin{bmatrix} \frac{\lambda}{\lambda+1} & 0 \\ -\frac{1}{\lambda+1} & 1 \end{bmatrix} \Rightarrow K(\lambda)b(\lambda) = \begin{bmatrix} \frac{\lambda}{\lambda+1} \\ -\frac{1}{\lambda+1} \end{bmatrix},$$

which is a different maximal set of root vectors for $R(\lambda)$. In this toy example, Step 4(a) is not difficult, but generally it can be problematic because an exact, say symbolic, computation of $K(\lambda)$ may be impractical or demanding. Instead, Step 4(b) can always be implemented in a purely numerical manner.

To conclude the paper, we note that the technique described in this subsection allows us to form a rational matrix $Y(\lambda)$ that has the same zero structure and corresponding root polynomials as $R(\lambda)$, but has no coalescent pole/zeros anymore. But if we are only interested in one single point λ_0 which happens to be also a pole of $R(\lambda)$, then there exists a more economical procedure. We can then follow essentially the same idea, but restricted to that pole/zero only. As pointed out in Section 4, any irreducible generalized state-space realization of $R(\lambda)$ can be updated to have the following block triangular form

$$S(\lambda) := \left[\begin{array}{cc|c} A_1 - \lambda E_1 & A_2 - \lambda E_2 & B_1 \\ 0 & A_0 - \lambda E_0 & B_0 \\ \hline C_1 & C_0 & D \end{array}\right], \tag{14}$$

where the spectrum of the pencil $A_0 - \lambda E_0$ is the single point λ_0 and where the pencil $A_1 - \lambda E_1$ has no eigenvalues at λ_0 and hence contains all the remaining eigenvalues of $A - \lambda E$. The pole assignment matrix F is then applied only to the subsystem with pole λ_0 :

$$S_Y(\lambda) = \left[\begin{array}{cc|c} A_1 - \lambda E_1 & A_2 - \lambda E_2 & B_1 \\ 0 & A_0 - \lambda E_0 & B_0 \\ \hline C_1 & C_0 & D \end{array}\right] \left[\begin{array}{cc|c} I & 0 & 0 \\ 0 & I & 0 \\ \hline 0 & F_0 & I \end{array}\right],$$

which again corresponds to a factorization

$$Y(\lambda) = R(\lambda)K(\lambda), \quad \text{with} \quad K(\lambda) := F_0(\lambda E_0 - A_0 - B_0F_0)^{-1}B_0 + I,$$

and the system matrices

$$S_K(\lambda) = \begin{bmatrix} A_0 + B_0F_0 - \lambda E_0 & B_0 \\ F_0 & I \end{bmatrix} \quad \text{and} \quad S_{K^{-1}}(\lambda) = \begin{bmatrix} A_0 - \lambda E_0 & B_0 \\ -F_0 & I \end{bmatrix}.$$

The minimality of these system matrices follows again from the same arguments, and the factorization $R(\lambda) = Y(\lambda)K(\lambda)^{-1}$ is again coprime for very similar reasons. But since we

can only guarantee that the point λ_0 is not a pole anymore of $Y(\lambda)$, we can only guarantee that the root polynomials of $R(\lambda)$ and $Y(\lambda)$ are related at λ_0. Such a procedure could be repeated for each coalescent pole/zero of $R(\lambda)$. Moreover, the "truncated" Step 4(b) can be applied also in this case.

6. Conclusions

In this paper we revisited the notion of root polynomials of a polynomial matrix $P(\lambda)$ at one of its zeros λ_0. This notion was introduced in [9] for regular polynomial matrices and was thoroughly studied in [6] for possibly singular polynomial matrices. An algorithmic approach for root polynomials of pencils was proposed in [19]. We extended the notion of root polynomials to define root vectors of an arbitrary rational matrix $R(\lambda)$ using valuation theory. This extension makes the definition for coalescent poles and zero consistent with applications. In a second part of the paper, we derive computational methods to construct a maximal set of root vectors for a general rational matrix $R(\lambda)$. The approach is based on computing a (locally) minimal state-space realization, applying the algorithm to compute maximal sets of root polynomials of a pencil [19], and finally recovering a maximal set of root vectors for $R(\lambda)$. If λ_0 is a coalescent pole-zero, a preprocessing step implicitly computes a right coprime factorization $R(\lambda) = Y(\lambda)K(\lambda)^{-1}$ such that the maximal sets of root vectors for $Y(\lambda)$ and $R(\lambda)$ are simply related. This can be done directly at the level of state-space realizations.

Declaration of competing interest

There is no competing interest.

Data availability

No data was used for the research described in the article.

Acknowledgements

VN thanks Froilán Dopico, Manme Quintana and Ion Zaballa for useful discussions on Rosenbrock's theorem; in particular, credit for some of the very initial steps within the proof of Theorem 5.5 should be given to Zaballa who proposed them (in a more general context) while performing research jointly with Dopico and VN. Remark 5.10 is due to Quintana, who in addition shared useful comments on a preliminary version of the manuscript. The authors also thank an anonymous referee for useful suggestions and in particular for asking whether Corollary 5.9 is also true over the ring of proper rational functions, which led to the addition of Remark 5.11.

References

[1] A. Amparan, F. Dopico, S. Marcaida, I. Zaballa, On minimal bases and indices of rational matrices and their linearizations, Linear Algebra Appl. 623 (2021) 14–67.
[2] A. Amparan, S. Marcaida, I. Zaballa, Finite and infinite structures of rational matrices: a local approach, Electron. J. Linear Algebra 30 (2015) 196–226.
[3] A. Amparan, S. Marcaida, I. Zaballa, On coprime rational function matrices, Linear Algebra Appl. 50 (2016) 1–31.
[4] W.A. Coppel, Matrices of rational functions, Bull. Aust. Math. Soc. 11 (1974) 89–113.
[5] M. Dahleh, M.A. Dahleh, G. Verghese, Lectures on dynamic systems and control, Massachusetts Institute of Technology, Open Course lecture notes.
[6] F. Dopico, V. Noferini, Root polynomials and their role in the theory of matrix polynomials, Linear Algebra Appl. 584 (2020) 37–78.
[7] G.D. Forney, Minimal bases of rational vector spaces, with applications to multivariable linear systems, SIAM J. Control 13 (1975) 493–520.
[8] K. Gallivan, A. Vandendorpe, P. Van Dooren, Model reduction of MIMO systems via tangential interpolation, SIAM J. Matrix Anal. Appl. 26 (2) (2004) 328–349.
[9] I. Gohberg, P. Lancaster, L. Rodman, Matrix Polynomials, Academic Press, 1982.
[10] S. Güttel, F. Tisseur, The nonlinear eigenvalue problem, Acta Numer. 26 (2017) 1–94.
[11] S. Güttel, R. Van Beeumen, K. Meerbergen, W. Michiels, NLEIGS: a class of fully rational Krylov methods for nonlinear eigenvalue problems, SIAM J. Sci. Comput. 36 (6) (2014) A2842–A2864.
[12] T. Kailath, Linear Systems, Prentice Hall, New York, 1980.
[13] N. Karcanias, Minimal bases of matrix pencils: algebraic, Toeplitz structure and geometric properties, Linear Algebra Appl. 205–206 (1994) 831–868.
[14] M. Lotz, V. Noferini, Wilkinson's bus: weak condition numbers, with an application to singular polynomial eigenproblems, Found. Comput. Math. 20 (6) (2020) 1439–1473.
[15] A. Mac Farlane, N. Karcanias, Poles and zeros of linear multivariable systems: a survey of the algebraic, geometric and complex-variable theory, Int. J. Control 24 (1) (1976) 33–74.
[16] V. Mehrmann, H. Voss, Nonlinear eigenvalue problems: a challenge for modern eigenvalue methods, GAMM-Mitt. 27 (2004) 121–152.
[17] V. Noferini, The behavior of the complete eigenstructure of a polynomial matrix under a generic rational transformation, Electron. J. Linear Algebra 23 (2012) 607–624.
[18] V. Noferini, L. Nyman, J. Perez, M.C. Quintana, Perturbation theory of transfer function matrices, preprint, 2022.
[19] V. Noferini, P. Van Dooren, On computing root polynomials and minimal bases of matrix pencils, arXiv:2110.15416, 2021.
[20] H.H. Rosenbrock, State-Space and Multivariable Theory, Nelson, London, 1970.
[21] P. Van Dooren, The generalized eigenstructure problem in linear system theory, IEEE Trans. Autom. Control AC-19 (1981) 111–129.
[22] P. Van Dooren, Rational and polynomial matrix factorizations via recursive pole-zero cancellation, Linear Algebra Appl. 137/138 (1990) 663–697.
[23] A.I.G. Vardulakis, Linear Multivariable Control, John Wiley and Sons, New York, 1991.
[24] G. Verghese, P. Van Dooren, T. Kailath, Properties of the system matrix of a generalized state space system, Int. J. Control 30 (1979) 235–243.
[25] W.M. Wonham, Linear Multivariable Control: A Geometric Approach, 2nd ed., Springer, New York, 1979.