Root polynomials and their role in the theory of matrix polynomials

Froilán M. Dopico* Vanni Noferini†

September 21, 2018

Abstract

We give a coherent theory of root polynomials, an algebraic tool useful for the analysis of matrix polynomials. In particular, we first survey results previously appeared in the literature, giving a formal proof for those that lacked one. We next extend some of these results providing some new concepts and related theorems, thus simplifying and expanding the theory. Then, we give some applications of root polynomials, such as the recovery of Jordan chains from linearizations of matrix polynomials, or the behaviour of Jordan chains under rational reparametrization. We also briefly discuss how root polynomials can be used to define eigenvectors and eigenspaces for singular matrix polynomials.

Keywords: matrix polynomials , polynomial matrices , root polynomials , μ-independent set , complete set , maximal set , minimaximal set , eigenvalues , minimal bases , singular , eigenvectors of singular matrix polynomials
MSC: 15A03 , 15A18 , 15A21 , 15A54

1 Introduction

Root polynomials of *regular* matrix polynomials (or, equivalently, polynomial matrices) are introduced in the first chapter of the classical book by Gohberg, Lancaster, and Rodman [11, Section 1.5] as an analytical concept that simplifies the study of Jordan chains [11, Section 1.6]. Although root polynomials constitute a powerful tool for establishing spectral properties of matrix polynomials, they do

*Departamento de Matemáticas, Universidad Carlos III de Madrid, Avda. Universidad 30, 28911 Leganés, Spain. (dopico@math.uc3m.es). Supported by "Ministerio de Economía, Industria y Competitividad of Spain" and "Fondo Europeo de Desarrollo Regional (FEDER) of EU" through grants MTM2015-65798-P and MTM2017-90682-REDT.

†Department of Mathematical Sciences, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, UK. (vnofer@essex.ac.uk).

not seem to be widely used in the literature. In fact, in the mentioned landmark reference [11], root polynomials are only used once after Chapter 1. More precisely, they are used in [11, Corollary 7.11, p. 203] to characterize divisibility in terms of spectral data. Recently, root polynomials have been extended for the first time to *singular* matrix polynomials in [17, Section 8] as an instrumental technical tool for proving the main results in that reference.

Both in [11] and [17], root polynomials are considered more an auxiliary than a central concept and, as a consequence, are treated in a very concise way, which leads in some occasions to imprecise statements and in general to a theory that seems more difficult than it actually is. In this context, this paper has three main goals. First, to establish rigorously and with detail the definition and the most important properties of root polynomials in the general setting of singular matrix polynomials. Second, to show that root polynomials interact naturally with a number of problems that have attracted the attention of the research community in the last years as, for instance, rational transformations of matrix polynomials, with particular emphasis on Möbius transformations, [15, 17, 18], linearizations of matrix polynomials and related recovery properties [1, 2, 4, 5, 14, 16, 19],and dual pencils [20]. We hope that the third goal of this paper will be obtained as a result of the two previous ones, since we expect that our manuscript will encourage the research community to familiarize with, and to use more often, root polynomials, which should be, in our opinion, one of the fundamental tools of any researcher on the theory of matrix polynomials.

We emphasize that, in the case of regular matrix polynomials, root polynomials are very closely related to Jordan chains, as it is clearly explained in [11, Sections 1.5 and 1.6]. The relation of a root polynomial to a Jordan chain is the same of that of a generating function [21] to a sequence; for a survey of good reasons to use generating functions as a tool to manipulate and analyze sequences, see [21]. Therefore, all the results in this paper admit, when specialized to regular matrix polynomials, an immediate translation into the language of Jordan chains. This solves a number of open problems in the literature. For example, how rational transformations of polnoymial matrices change the Jordan chains of the polynomial (the answer to this question is briefly sketched in [17, Remark 8.3] and is considered an open problem in [15, Remark 6.12]), or how Jordan chains of a matrix polynomials can be recovered from the Jordan chains of some of the most relevant linearizations studied in the literature.

On the other hand, one has singular matrix polynomials. Currently, there is not even agreement in the literature on how to define consistently an eigenvector of a *singular* matrix polynomial corresponding to anyone of its eigenvalues, and the first vectors of traditional Jordan chains are eigenvectors [11]. As we show in this paper, root polynomials naturally extend to the singular case. Clearly, we may still view them as generating functions, which is a natural way to extend Jordan chains to the singular case. We feel that this task is beyond the goal of the present paper. However, we stress that the extension of the definition of root polynomials from regular to singular matrix polynomials is performed in this paper in a fully consistent way leading to a concept that is easy to handle. Hence, the theory presented in this paper can be used to define in the future Jordan chains of singular matrix polynomials in a meaningful way, which is of interest in certain applications [3]. For the present, we will restrict here to sketch in Subsection 2.3 how eigenvectors and eigenspaces of singular matrix polynomials can be defined by following the ideas introduced in [17] and further developed in this paper.

The paper is organized as follows. Section 2 revises in its first part some basic results of the theory of matrix polynomials and introduces in its second part the definition of root polynomials and some related concepts. The existence of maximal sets of root polynomials at any finite eigenvalue of a matrix polynomial is established in Section 3, as well as their relationship with partial multiplicities and their behavior under some polynomial equivalences. Section 4 proves some extremality properties of maximal sets of root polynomials that are used in Section 5 to generate many maximal sets of root polynomials from a given maximal set, in particular maximal sets of root polynomials with minimal grade. Root polynomials at infinity are introduced in Section 6. Sections 7, 8, and 9 establish, respectively, how root polynomials change under rational transformations, how root polynomials can be recovered from those of linearizations, and how root polynomials of dual pencils are related each other. Finally, some conclusions are discussed in Section 10.

2 Preliminaries

2.1 Basic results

Although a theory of root polynomials over any field can be developed, it is complicated by the fact that the finite eigenvalues of a matrix polynomial over a generic field $\mathbb{K}$ may lie in the algebraic closure of $\mathbb{K}$. In this paper, we will neglect this complication for simplicity of exposition, and we consider polynomials with coefficients in an algebraically closed field $\mathbb{F}$. Moreover, we will present a theory of right root polynomials and related concept. Indeed, left root polynomials (as well as left eigenvectors of matrix polynomials, left minimal indices, etc.) of a matrix polynomial $P(x)$ can simply be defined as right root polynomials (eigenvectors, minimal indices) of $P(x)^T$. It is therefore sufficient to consider right root polynomials, and, since there is no ambiguity in this paper, in the following we will omit the adjective "right".

Throughout, we shall consider matrix polynomials, possibly rectangular, with elements in the principal ideal domain $\mathbb{F}[x]$; the field of fractions of $\mathbb{F}[x]$ is denoted by $\mathbb{F}(x)$. We denote the set of $m \times n$ such matrix polynomials by $\mathbb{F}[x]^{m\times n}$. We first recall some basic definitions in the theory of matrix polynomials.

Definition 2.1 (Normal rank)**.** Let $P(x) \in \mathbb{F}[x]^{m\times n}$. Then the rank of $P(x)$ over the field $\mathbb{F}(x)$ is called the normal rank of $P(x)$.

A square matrix polynomial $P(x) \in \mathbb{F}[x]^{n\times n}$ such that its normal rank is n is called *regular*. Any matrix polynomial which is not regular is said to be *singular*.

Definition 2.2 (Finite eigenvalues)**.** Let $P(x) \in \mathbb{F}[x]^{m\times n}$ have normal rank r. Then $\mu \in \mathbb{F}$ is called a finite eigenvalue of $P(x)$ if the rank of $P(\mu)$ over $\mathbb{F}$ is strictly less than r.

Recall that in ring theory a *unit* is an invertible element of the ring, i.e.[1], $u \in R$ is a unit if $\exists\ v \in R$ such that $uv = vu = 1_R$. When $R = \mathbb{F}[x]^{n\times n}$ is a square matrix polynomial ring, its units are sometimes called *unimodular* matrix polynomials. It is straightforward to show that the units of $\mathbb{F}[x]^{n\times n}$ are precisely those matrix polynomials whose determinant is a nonzero constant in $\mathbb{F}$. We now expose two fundamental theorems on matrix polynomials and their Smith and local Smith canonical forms [11].

Theorem 2.3 (Smith form)**.** *Let $P(x) \in \mathbb{F}[x]^{m\times n}$. Then there exists two unimodular matrix polynomials $U(x) \in \mathbb{F}[x]^{m\times m}$, $V(x) \in \mathbb{F}[x]^{n\times n}$, such that*

$$S(x) = U(x)P(x)V(x) = \begin{bmatrix} d_1(x) & 0 & \dots & 0 & \dots & 0 \\ 0 & d_2(x) & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \dots & d_r(x) & \dots & 0 \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 \end{bmatrix},$$

where $d_1(x), \dots, d_r(x) \in \mathbb{F}[x]$ are called the invariant polynomials *of $P(x)$ and they are monic polynomials such that $d_k(x)|d_{k+1}(x)$ for all $k = 1, 2, \dots, r-1$. The matrix polynomial $S(x)$ is uniquely determined by $P(x)$ and is called the* Smith canonical form *of $P(x)$. Moreover, factorizing*

$$d_i(x) = \prod_{j\in J} (x - x_j)^{\kappa_{i,(j)}},$$

which is possible for some finite set of indices J as we have assumed that $\mathbb{F}$ is algebraically closed, the factors $(x - x_j)^{\kappa_{i,(j)}}$ such that $\kappa_{i,(j)} > 0$ are called the elementary divisors of $P(x)$ corresponding to the eigenvalue x_j. The nonnegative integers $\kappa_{i,(j)}$ satisfy $\kappa_{i_1,(j)} \le \kappa_{i_2,(j)} \Leftrightarrow i_1 \le i_2$ and are called the partial multiplicities of the eigenvalue x_j. The algebraic multiplicity of an eigenvalue is the sum of its partial multiplicities; the geometric multiplicity of an eigenvalue is the number

[1]It is possible in certain rings to construct u, v such that $uv = 1_R \neq vu$, so $uv = 1_R$ alone would not suffice to consider u, v units; this subtlety does not concern us here.

of nonzero partial multiplicities. If an eigenvalue x_j has geometric multiplicity s, we denote the nonincreasing list of its partial multiplicities by $m_1, \dots, m_s$ with $m_1 = \kappa_{r,(j)} \geq \cdots \geq m_s = \kappa_{r+1-s,(j)}$.

The matrix polynomial $S(x)$ is called the Smith canonical form of $P(x)$, and it is uniquely determined by $P(x)$ (unlike $U(x)$ and $V(x)$).

Theorem 2.4 (Local Smith form)**.** *Suppose that the partial multiplicities of $\mu \in \mathbb{F}$ for a certain matrix polynomial $P(x) \in \mathbb{F}[x]^{m\times n}$ are $\kappa_1, \dots, \kappa_r$ (possibly allowing some of them to be zero). Then, there exist two regular matrix polynomials $A(x) \in \mathbb{F}[x]^{m\times m}$, $B(x) \in \mathbb{F}[x]^{n\times n}$ such that* $\det A(\mu) \det B(\mu) \neq 0$ *and*

$$D(x) = A^{-1}(x)P(x)B^{-1}(x) = \begin{bmatrix} (x-\mu)^{\kappa_1} & 0 & \dots & 0 & \dots & 0 \\ 0 & (x-\mu)^{\kappa_2} & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \dots & (x-\mu)^{\kappa_r} & \dots & 0 \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 \end{bmatrix}.$$

The matrix polynomial $D(x)$ is uniquely determined by $P(x)$ and is called the local Smith form *of $P(x)$ at μ.*

For ease of notation and terminology, given a matrix A with elements in a (not necessarily algebraically closed) field $\mathbb{K}$ we define $\operatorname{span} A$ as the subspace of all the vectors that can be formed as linear combinations over $\mathbb{K}$ of the columns of A, and $\ker A$ as the set of all vectors whose image under the linear map represented by A is the zero vector. Similarly, if A has full column rank, we say that A is a basis of $\operatorname{span} A$.

Next, we recall the notions of minimal bases and minimal indices [10]

Definition 2.5 (Minimal bases and minimal indices)**.** A matrix polynomial $M(x) \in \mathbb{F}[x]^{n\times p}$ of normal rank p is called a minimal basis if the sum of the degrees of its columns (sometimes called its *order*) is minimal among all polynomial bases of $\operatorname{span} M(x) \subseteq \mathbb{F}(x)^p$.

It is proved in [10] that the degrees of the columns of a minimal basis depend only on the subspace $\operatorname{span} M(x)$. A minimal basis such that the degrees of its columns are non-decreasing is called an ordered minimal basis [20]. Moreover,

Definition 2.6. If $M(x)$ is a minimal basis and $\operatorname{span} M(x) = \ker P(x)$ for some $P(x) \in \mathbb{F}[x]^{m\times n}$, we say that $M(x)$ is a minimal basis for $P(x)$. The degrees of the columns of any minimal basis of $P(x)$ are called the (right) minimal indices of $P(x)$.

Minimal bases have the important property that the equation $M(x)v(x) = b(x)$, for any polynomial vector right hand side $b(x)$, always admits a *polynomial* solution

$v(x)$ [10, Main Theorem]. Iin other words, a minimal basis $M(x)$, as a matrix, is left invertible *over the ring* $\mathbb{F}[x]$; in the following we will often use this result without further justification.

2.2 Root polynomials

We now recall some preliminary definitions and basic results, first discussed in [17], which are useful for the theory of root polynomials. In Definition 2.7 we follow the conventon that $\operatorname{span} M = \{0\} \subset \mathbb{F}^n$ for any empty matrix $M \in \mathbb{F}^{n\times 0}$.

Definition 2.7. Let $M(x) \in \mathbb{F}[x]^{n\times p}$ be a minimal basis for $P(x) \in \mathbb{F}[x]^{m\times n}$, and $\mu \in \mathbb{F}$. Then, $\ker_\mu P(x) := \operatorname{span} M(\mu)$.

Lemma 2.8. *The definition of* $\ker_\mu P(x)$ *is independent of the particular choice of a minimal basis* $M(x)$*.*

Proof. Let $N(x)$ be any other minimal basis of $P(x)$, and write $N(x) = M(x)T(x)$. Hence, $N(\mu) = M(\mu)T(\mu)$. By [20, Lemma 3.6] $T(\mu)$ is manifestly invertible for any μ, and hence, $N(\mu)$ has full column rank. Thus, its columns span the same subspace as those of $M(\mu)$. □

Lemma 2.9. $v \in \ker_\mu P(x) \subseteq \mathbb{F}^n \Leftrightarrow \exists w(x) \in \mathbb{F}[x]^n : P(x)w(x) = 0$ *and* $w(\mu) = v$*.*

Proof. Let $M(x)$ be a minimal basis of $\ker P(x)$. Then one implication is obvious because there exists a polynomial vector $c(x)$ such that $w(x) = M(x)c(x)$, and hence $v = M(\mu)c(\mu)$.

Suppose now $v = M(\mu)c$ for some constant vector c, and define $w(x) = M(x)c$ to conclude the proof. □

Lemma 2.10. $\ker_\mu P(x) \subseteq \ker P(\mu)$*, and equality holds if and only if* μ *is not a finite eigenvalue of* $P(x)$*.*

Proof. $P(x)M(x) = 0 \Rightarrow P(\mu)M(\mu) = 0$ shows the first claim. To prove the second, let $S(x)$ be the Smith form of $P(x)$. Observe that $p = \dim \ker_\mu P(x)$ is the number of zero columns of $S(x)$, whereas $p + s = \dim \ker P(\mu)$ is the number of zero columns of $S(\mu)$. Hence, $s = 0$, i.e., the two dimensions coincide, if and only if μ is not a root of any nonzero invariant polynomial of $P(x)$. The latter property is equivalent to being a finite eigenvalue of $P(x)$. □

Definition 2.11 is central to this paper. It was given in [17, Sec. 8] and generalizes the definition of root polynomial given in [11, Ch. 1] for the regular case, i.e., $p = 0$ and $\ker_\mu P(x) = \{0\}$ for all $\mu \in \mathbb{F}$.

Definition 2.11 (Root polynomials)**.** The polynomial vector $r(x) \in \mathbb{F}[x]^n$ is a root polynomial of order $\ell \geq 1$ at $\mu \in \mathbb{F}$ for $P(x) \in \mathbb{F}[x]^{m\times n}$ if the following conditions hold:

1. $r(\mu) \notin \ker_\mu P(x)$;

2. $P(x)r(x) = (x-\mu)^\ell w(x)$ for some $w(x) \in \mathbb{F}[x]^m$ satisfying $w(\mu) \neq 0$.

Proposition 2.12. *Let $P(x) \in \mathbb{F}[x]^{m\times n}$. Then there exists a root polynomial for $P(x)$ at μ if and only if μ is a finite eigenvalue of $P(x)$.*

Proof. Suppose that μ is not an eigenvalue of $P(x)$, and let $r(x) \in \mathbb{F}[x]^n$ satisfy $P(x)r(x) = (x-\mu)^\ell w(x)$ for some $\ell \geq 1$. Then $P(\mu)r(\mu) = 0$ and hence $r(\mu) \in \ker P(\mu) = \ker_\mu P(x)$ by Lemma 2.10. Hence, no polynomial vector can simultaneously satisfy the two conditions in Definition 2.11.

Conversely suppose that μ is an eigenvalue of $P(x)$, and let $r \in \mathbb{F}^n$ be such that $r \in \ker P(\mu)$ but $r \notin \ker_\mu P(x)$. Note that $P(x)r \neq 0$, or otherwise $r \in \operatorname{span} M(x)$, implying $r \in \ker_\mu P(x)$. Hence there is a positive integer ℓ such that $P(x)r = (x-\mu)^\ell w(x)$ for some $w(x) \in \mathbb{F}[x]^m$, $w(\mu) \neq 0$. □

Definition 2.13. Let $M(x) \in \mathbb{F}[x]^{n\times p}$ be a right minimal basis of $P(x) \in \mathbb{F}[x]^{m\times n}$. The vectors $r_1(x), \ldots, r_s(x) \in \mathbb{F}[x]^n$ are a μ-independent set of root polynomials at μ for $P(x)$ if $r_i(x)$ is a root polynomial at μ for $P(x)$ for each $i = 1, \ldots, s$, and the matrix

$$\begin{bmatrix} M(\mu) & r_1(\mu) & \ldots & r_s(\mu) \end{bmatrix}$$

has full column rank.

Note that Definition 2.13 does not depend on the particular choice of a right minimal basis $M(x)$, since given another basis $N(x) = M(x)T(x)$, for some unimodular [20, Lemma 3.6] $T(x) \in \mathbb{F}[x]^{p\times p}$, one has that

$$\begin{bmatrix} N(\mu) & r_1(\mu) & \ldots & r_s(\mu) \end{bmatrix} = \begin{bmatrix} M(\mu) & r_1(\mu) & \ldots & r_s(\mu) \end{bmatrix} \begin{bmatrix} T(\mu) & \\ & I_s \end{bmatrix},$$

and the rightmost matrix in the above equality is necessarily nonsingular.

Remark 2.14. We take the chance to correct an imprecise statement in [17], where Definition 2.13 is mistakenly given without including the columns of $M(\mu)$. We note that, in spite of the unfortunate misprint, the correct definition is implicitly used (and, in fact, needed), in the proof of [17, Proposition 8.2].

Definition 2.15. Let $M(x) \in \mathbb{F}[x]^{n\times p}$ be a right minimal basis of a matrix polynomial $P(x) \in \mathbb{F}[x]^{m\times n}$. The vectors $r_1(x), \ldots, r_s(x) \in \mathbb{F}[x]^n$ are a complete set of root polynomials at μ for $P(x)$ if they are μ-independent and there does not exist any set of $s+1$ root polynomials at μ for $P(x)$, say $t_i(x)$, such that the matrix

$$\begin{bmatrix} M(\mu) & t_1(\mu) & \ldots & t_{s+1}(\mu) \end{bmatrix}$$

has full column rank.

Proposition 2.16. *Let $M(x) \in \mathbb{F}[x]^{n\times p}$ be a right minimal basis of a matrix polynomial $P(x) \in \mathbb{F}[x]^{m\times n}$. The polynomial vectors $r_1(x), \ldots, r_s(x) \in \mathbb{F}[x]^n$ are a complete set of root polynomials at μ for $P(x)$ if and only if the columns of the matrix*

$$N = \begin{bmatrix} M(\mu) & r_1(\mu) & \ldots & r_s(\mu) \end{bmatrix}$$

form a basis for $\ker P(\mu)$.

Proof. Observe that, if $\{r_i(x)\}_{i=1}^s$ are a complete set of root polynomials, then $P(\mu)N = 0$, by definition of minimal basis and of root polynomial. By Definition 2.13, the matrix N has full column rank. It remains to argue that its columns form a basis of $\ker P(\mu)$: suppose they do not. Then, we can complete them to a basis, i.e., there exists a matrix X such that $\hat{N} = \begin{bmatrix} N & X \end{bmatrix}$, $\hat{N}$ has full column rank, and $P(\mu)\hat{N} = 0 \Rightarrow P(\mu)X = 0$. Let v be any column of X: then $P(\mu)v = 0$, but $P(x)v \neq 0$ since $v \notin \operatorname{span} M(x)$. Hence, v is a root polynomial for $P(x)$ at μ, and the set $v, r_1(x), \ldots, r_s(x)$ is μ-independent, contradicting Definition 2.15.

Conversely, assume that completeness does not hold. Then, there exists a certain matrix $\hat{N}$, with full column rank and $(\dim\ker P(x) + s + 1)$ columns, such that $P(\mu)\hat{N} = 0$. It follows that $\dim\ker P(\mu) > \dim\ker P(x) + s$, and hence, the columns of N cannot be a basis. □

Definition 2.17. Let $M(x) \in \mathbb{F}[x]^{n\times p}$ be a right minimal basis of a matrix polynomial $P(x) \in \mathbb{F}[x]^{m\times n}$. The vectors $r_1(x), \ldots, r_s(x) \in \mathbb{F}[x]^n$ are a maximal set of root polynomials at μ for $P(x)$ if they are complete and their orders as root polynomials for $P(x)$, say, $\ell_1 \geq \cdots \geq \ell_s > 0$, satisfy the following property: for all $j = 1, \ldots, s$, there is no root polynomial $\hat{r}(x)$ of order $\ell > \ell_j$ such that the matrix

$$\begin{bmatrix} M(\mu) & r_1(\mu) & \ldots & r_{j-1}(\mu) & \hat{r}(\mu) \end{bmatrix}$$

has full column rank.

2.3 Root polynomials, quotient spaces, and a definition of eigenvectors and eigenspaces for singular matrix polynomials

The ideas that lead to Definition 2.11 are intimately related to the concept of a quotient space. This subsection explains how and it has a slightly more abstract-algebraic spirit.

We start by recalling the notion of quotient spaces. For any linear subspace $V \subseteq \mathbb{F}^n$, we consider the equivalence relation on $\mathbb{F}^n$

$$\forall \;\; x, y \in \mathbb{F}^n, \quad x \sim y \Leftrightarrow x - y \in V.$$

Then the quotient space $\mathbb{F}^n/V$ is defined as the set of equivalence classes

$$\forall \;\; x \in \mathbb{F}^n, \quad [x] := \{v \in \mathbb{F}^n : v \sim x\}.$$

A quotient space can be given a vector space structure over $\mathbb{F}$ simply by defining

$$\forall \ \ x \in \mathbb{F}^n, \ \forall \ \ \alpha \in \mathbb{F}, \quad \alpha[x] := [\alpha x], \quad [x]+[y] := [x+y].$$

If V is an invariant subspace of some linear endomorphism of $\mathbb{F}^n$, say, A, then $x \sim y \Rightarrow Ax \sim Ay$. Hence, it is consistent to define and write $A[x] := [Ax]$.

Take now $V = \ker_\mu P(x)$. In this setting, asking that, for a root polynomial $r(x)$, $r(\mu) \notin \ker_\mu P(x)$, is equivalent to imposing $[r(\mu)] \neq [0]$, which is the natural extension to the singular case of the condition $r(\mu) \neq 0$ for a regular $P(x)$ (i.e., for $\ker_\mu P(x) = \{0\}$).

Remark 2.18. Note that by Lemma 2.9

$$[r(\mu)]_{\mathbb{F}^n/\ker_\mu P(x)} \neq [0]_{\mathbb{F}^n/\ker_\mu P(x)} \Rightarrow [r(x)]_{\mathbb{F}(x)^n/\ker P(x)} \neq [0]_{\mathbb{F}(x)^n/\ker P(x)}.$$

By the first isomorphism theorem, $\mathbb{F}(x)^n/\ker P(x)$ is isomorphic to the row space of $P(x)$.

Suppose for now, and for simplicity of exposition, that the eigenvalue μ has geometric multiplicity 1. With this assumption, for a regular $P(x)$, an eigenvector $v \in \mathbb{F}^n (= \mathbb{F}^n/\{0\} = \mathbb{F}^n/\ker_\mu P(x))$ is defined uniquely up to a nonzero scalar. For a singular $P(x)$, we claim that our framework also naturally leads to an eigenvector which is uniquely defined up to a nonzero scalar, except that the "natural" space to define the eigenvectors is equivalence classes in the above defined the quotient space: $[v] \in \mathbb{F}^n/\ker_\mu P(x)$. Note that $\ker_\mu P(x)$ is an invariant subspace for $P(\mu)$.

Definition 2.19. If μ is an eigenvalue of $P(x)$, we say that $[v] \in \mathbb{F}^n/\ker_\mu P(x)$ is an eigenvector associated with the eigenvalue μ if

$$P(\mu)[v] = [0], \qquad [v] \neq [0].$$

It is immediate to check that, if $r(x)$ is a root polynomial at μ for $P(x)$, then $[r(\mu)]$ is an eigenvector in the above sense. Conversely, if $[v]$ is an eigenvector, then there is a root polynomial such that $r(\mu) = v$, e.g., $r(x) = v$. To show consistency of the theory, let $q(x)$ be any other root polynomial at μ for $P(x)$; we must verify that it holds $[q(\mu)] = \alpha[r(\mu)]$ for some $0 \neq \alpha \in \mathbb{F}$. The clue is that, by Proposition 2.16, for any $M(x)$ minimal basis of $P(x)$, the matrices

$$\begin{bmatrix} M(\mu) & r(\mu) \end{bmatrix}, \qquad \begin{bmatrix} M(\mu) & q(\mu) \end{bmatrix}$$

are both a basis for $\ker P(\mu)$. This observation implies that

$$\begin{bmatrix} M(\mu) & q(\mu) \end{bmatrix} = \begin{bmatrix} M(\mu) & r(\mu) \end{bmatrix} \begin{bmatrix} I_p & v \\ 0 & \alpha \end{bmatrix}$$

for some $v \in \mathbb{F}^p, 0 \neq \alpha \in \mathbb{F}$. Hence,

$$q(\mu) = M(\mu)v + \alpha r(\mu) \Leftrightarrow [q(\mu)] = \alpha[r(\mu)].$$

These ideas can be extended to the case of eigenvalues of geometric multiplicity $s > 1$; namely, Definition 2.19 still makes sense (hence why have not specified there that the geometric mulitplicity of μ should be 1). Moreover, one can defined eigenspaces as $\operatorname{span} V$ where V is an $n \times s$ matrix whose columns are linearly independent (over $\mathbb{F}$) equivalence classes in $\mathbb{F}^n / \ker_\mu P(x)$. Similarly as above, one can easily check that any other matrix W whose columns are a basis for the same eigenspace can be written as $W = VA$ for some invertible matrix $A \in \mathbb{F}^{s \times s}$. Moreover, an eigenspace can be constructed starting from any complete set of root polynomials, and conversely a complete set of root polynomials can be constructed starting from the columns of a basis for the eigenspace.

3 Existence of maximal sets of root polynomials, and correspondence with partial multiplicities

We start by showing that any matrix polynomial in local Smith form at μ admits a maximal set of root polynomials at μ.

Theorem 3.1. *Let $S(x) \in \mathbb{F}[x]^{m \times n}$ be in local Smith form at $\mu \in \mathbb{F}$. Then, denoting by r the normal rank of $S(x)$ and by s the geometric multiplicity of μ as an eigenvalue, the s vectors*

$$e_r, e_{r-1}, \dots, e_{r-s+1},$$

where e_i is the ith vector of the canonical basis of $\mathbb{F}^n$, are a maximal set of root polynomials at μ for $S(x)$. Moreover, their orders are the nonzero partial multiplicities of μ as an eigenvalue of $S(x)$.

Proof. It suffices to prove the statement for the case where μ is an eigenvalue, as otherwise $s = 0$ and there is nothing to prove.

By assumption, $S(x)$ is diagonal and $S(x)_{ii} = 0$ if and only if $i > r$. Hence, a minimal basis for $S(x)$ is $M(x) = M(\mu) = \begin{bmatrix} e_{r+1} & \dots e_n \end{bmatrix}$. Moreover, still by assumption, $S(\mu)_{ii} = 0$ if and only if $i > r - s$. Therefore, the vectors $e_r, e_{r-1}, \dots, e_{r-s+1}$ are all root polynomials at μ for $P(x)$, and by a simple direct computation it can be seen that their orders are the nonzero partial multiplicites of μ as a finite eigenvalue of $S(x)$. Further, they are manifestly μ-independent, as the matrix $N = \begin{bmatrix} e_{r+1} & \dots e_n & e_r & \dots & e_{r-s+1} \end{bmatrix}$ is just a column permutation of the matrix $\begin{bmatrix} 0_{r-s \times n-r+s} \\ I_{n-r+s} \end{bmatrix}$. They are a complete set by Proposition 2.16, since $S(\mu)N = 0$ and by assumption $\dim \ker P(\mu) = n - r + s$.

It remains to argue that they are maximal. Suppose they are not . Then, for some $j \le s$, there exists a certain root polynomial $\hat{r}(x)$ of order $\ell > \ell_j$ such that $e_r, \dots, e_{r-j+2}, \hat{r}(x)$ are μ-independent. We deduce that at least one of the first

$r-j+1$ elements of $\hat{r}(\mu)$ is nonzero. Expanding $\hat{r}(x)$ and $S(x)$ in a power series in $(x-\mu)$, it is easily seen that ℓ is bounded above by the minimal exponent κ_i of $S_{ii}(x) = (x-\mu)^{\kappa_i}$, where the minimum is taken over all the values of i such that $\hat{r}(\mu)_i \neq 0$. Hence, $\ell \leq \ell_j$, leading to a contradiction. □

The following results appeared in [17, 18] and are key for arguing that a maximal set of root polynomials exists for any matrix polynomial. We denote by $\operatorname{adj} P(x)$ the adjugate of a square matrix polynomial $P(x)$. Suppose that $A(x)$ and $B(x)$ are square matrix polynomials, and $Q(x) = A(x)P(x)B(x)$. For the proof of the next result the following equations, whose proof is immediate, will be useful:

$$Q(x) \operatorname{adj} B(x) = A(x)P(x)B(x) \operatorname{adj} B(x) = \det B(x) A(x) P(x);$$
$$\operatorname{adj} A(x) Q(x) = \operatorname{adj} A(x) A(x) P(x) B(x) = \det A(x) P(x) B(x).$$

Proposition 3.2. *Let $P(x), Q(x) \in \mathbb{F}[x]^{m\times n}$ and suppose that $Q(x) = A(x)P(x)B(x)$ for some $A(x) \in \mathbb{F}[x]^{m\times m}$ and $B(x) \in \mathbb{F}[x]^{n\times n}$ such that $\det A(\mu) \det B(\mu) \neq 0$. Then:*

- *if $r(x)$ is a root polynomial for $Q(x)$ at μ of order ℓ, then $B(x)r(x)$ is a root polynomial for $P(x)$ at μ of the same order;*
- *if $q(x)$ is a root polynomial for $P(x)$ at μ of order ℓ, then $\operatorname{adj} B(x)q(x)$ is a root polynomial for $Q(x)$ at μ of the same order;*

Proof. Suppose first that $r(x)$ is a root polynomial at μ for $Q(x)$. Note that $Q(x)r(x) = (x-\mu)^\ell w(x)$, $w(\mu) \neq 0$, implies $P(x)B(x)r(x) = (x-\mu)^\ell A^{-1}(x)w(x)$. Observe that the right hand side must be polynomial, since the left hand side is. Hence, either $A^{-1}(x)w(x)$ is polynomial or it has a pole at $x = \mu$. Yet, the latter case is not possible, since $\det A(\mu) \neq 0$ and $w(x)$ is polynomial. Observe further that $A^{-1}(\mu)w(\mu) \neq 0$. Finally, suppose $B(\mu)r(\mu) \in \ker_\mu P(x)$. Then by Lemma 2.9 $\exists v(x) : v(\mu) = B(\mu)r(\mu)$ and $P(x)v(x) = 0$, implying $Q(x)[\operatorname{adj} B(x)v(x)] = \det B(x)A(x)P(x)v(x) = 0$: a contradiction, because $\operatorname{adj} B(\mu)v(\mu) = \det B(\mu)r(\mu)$, which is a nonzero scalar multiple of $r(\mu)$ and hence cannot belong to $\ker_\mu Q(x)$.

Conversely, let $q(x)$ be a root polynomial at μ for $P(x)$. Then, $P(x)q(x) = (x-\mu)^\ell w(x)$, $w(\mu) \neq 0$ yields $Q(x) \operatorname{adj} B(x)q(x) = (x-\mu)^\ell \det B(x)A(x)w(x)$, and $\det B(\mu)A(\mu)w(\mu) \neq 0$ because $A(\mu)$ and $B(\mu)$ are nonsingular by assumption. To conclude the proof suppose that $\operatorname{adj} B(\mu)q(\mu) \in \ker_\mu Q(x)$. Using Lemma 2.9, $\exists v(x) : v(\mu) = \operatorname{adj} B(\mu)q(\mu)$ and $Q(x)v(x) = 0$. Thus, $P(x)[\det A(x)B(x)v(x)] = \operatorname{adj} A(x)Q(x)v(x) = 0$. The latter equation is absurd, because $\det A(\mu)B(\mu)v(\mu) = \det A(\mu) \det B(\mu)q(\mu)$, which, being a nonzero scalar multiple of $q(\mu)$, cannot belong to $\ker_\mu P(x)$. □

Lemma 3.3. *Let $P(x), Q(x) \in \mathbb{F}[x]^{m\times n}$ and suppose that $Q(x) = A(x)P(x)B(x)$ for some $A(x) \in \mathbb{F}[x]^{m\times m}$ and $B(x) \in \mathbb{F}[x]^{n\times n}$ such that $\det A(\mu) \det B(\mu) \neq 0$. Let $M(x)$ be a minimal basis for $P(x)$ and let $N(x)$ be a minimal basis for $Q(x)$. Then $\ker_\mu P(x) = \operatorname{span} M(\mu) = \operatorname{span} B(\mu)N(\mu)$, and $\ker_\mu Q(x) = \operatorname{span} N(\mu) = \operatorname{span} \operatorname{adj} B(\mu)M(\mu)$.*

Proof. The nonsingularity over the field $\mathbb{F}(x)$ of both $B(x)$ and $A(x)$ implies $\dim\ker P(x) = \dim\ker Q(x) = \dim\ker_\mu P(x) = \dim\ker_\mu Q(x)$. Moreover, manifestly $B(x)N(x)$ is a polynomial basis for $\ker P(x)$ and $\operatorname{adj} B(x)M(x)$ is a polynomial basis for $\ker Q(x)$. However, this does not directly imply the statement because neither of those bases is necessarily minimal.

However, changing bases we have $B(x)N(x) = M(x)C(x)$ and $\operatorname{adj} B(x)M(x) = N(x)D(x)$ for some square, invertible, and polynomial [10, Main Theorem] matrices $C(x), D(x)$. Hence $\operatorname{span} \operatorname{adj} B(\mu)M(\mu) \subseteq \ker_\mu Q(x)$ and $\operatorname{span} B(\mu)N(\mu) \subseteq \ker_\mu P(x)$. But $B(\mu)$ is invertible, and hence $\operatorname{rank} \operatorname{adj} B(\mu)M(\mu) = \operatorname{rank} B(\mu)N(\mu) = \dim\ker_\mu P(x) = \dim\ker_\mu Q(x)$, concluding the proof. □

Theorem 3.4. *Let $P(x), Q(x) \in \mathbb{F}[x]^{m\times n}$ and suppose that $Q(x) = A(x)P(x)B(x)$ for some $A(x) \in \mathbb{F}[x]^{m\times m}$ and $B(x) \in \mathbb{F}[x]^{n\times n}$. Assume further that $\det A(\mu) \det B(\mu) \neq 0$. Then:*

- *if $r_1(x), \dots, r_s(x)$ are a maximal (resp., complete, μ-independent) set of root polynomials at μ for $Q(x)$, with orders $\ell_1 \geq \cdots \geq \ell_s > 0$, then $B(x)r_1(x), \dots, B(x)r_s(x)$ are a maximal (resp., complete, μ-independent) set of root polynomials at μ for $P(x)$, with the same orders;*

- *if $q_1(x), \dots, q_s(x)$ are a maximal (resp., complete, μ-independent) set of root polynomials at μ for $P(x)$, with orders $\ell_1 \geq \cdots \geq \ell_s > 0$, then $\operatorname{adj} B(x)q_1(x), \dots, \operatorname{adj} B(x)q_s(x)$ are a maximal (resp., complete, μ-independent) set of root polynomials at μ for $Q(x)$, with the same orders.*

Proof. The fact that the property of being a root polynomial with a certain order is preserved by left and right multiplication by locally invertible matrix polynomials has already been showed in Proposition 3.2. Now we proceed by steps. Denote by $M(x)$ (resp. $N(x)$) a minimal basis for $P(x)$ (resp. $Q(x)$).

Suppose first $\{r_i(x)\}_{i=1}^s$ are μ-independent, and define

$$Y := \begin{bmatrix} N(\mu) & r_1(\mu) & \dots & r_s(\mu) \end{bmatrix}.$$

Note that by assumption Y has full column rank. Then

$$X = \begin{bmatrix} M(\mu) & B(\mu)r_1(\mu) & \dots & B(\mu)r_s(\mu) \end{bmatrix} = B(\mu)Y(C \oplus I_s)$$

for some invertible matrix C and using Lemma 3.3. Therefore, X has full column rank implying by Definition 2.13 that $\{B(x)r_i(x)\}_{i=1}^s$ are μ-independent.

Suppose further that $\{r_i(x)\}_{i=1}^s$ are complete. This is equivalent to assuming that $\ker Q(\mu) = \operatorname{span} Y$. Observing that X has full colum rank and that $\operatorname{rank} X = \operatorname{rank} Y = \dim\ker Q(\mu) = \dim\ker P(\mu)$, it suffices to argue that $P(\mu)X = P(\mu)B(\mu)Y(C \oplus I_s) = A(\mu)^{-1}Q(\mu)Y(C \oplus I_s) = 0$.

Finally, let us suppose that $\{r_i(x)\}_{i=1}^s$ are maximal, whereas $\{B(x)r_i(x)\}_{i=1}^s$ are not. Since in particular the former are complete, from the argument above the latter are as well. Therefore, it must be the case that for some $j \leq s$ there exists some $\hat{r}(x)$ that is a root polynomial of order $\ell > \ell_j$ at μ for $P(x)$ and such that the matrix

$$\hat{X} = \begin{bmatrix} M(\mu) & B(\mu)r_1(\mu) & \dots & B(\mu)r_{j-1}(\mu) & \hat{r}(\mu) \end{bmatrix}$$

has full column rank. However, by Proposition 3.2, $\operatorname{adj} B(x)\hat{r}(x)$ is a root polynomial for $Q(x)$ at μ of order $\ell > \ell_j$. Using Lemma 3.3,

$$\hat{Y} = \begin{bmatrix} N(\mu) & r_1(\mu) & \dots & r_{j-1}(\mu) & \operatorname{adj} B(\mu)\hat{r}(\mu) \end{bmatrix} = \frac{\operatorname{adj} B(\mu)}{\det B(\mu)} \hat{X}(D \oplus I_{j-1} \oplus \det B(\mu))$$

for some square invertible matrix D. This implies that $\hat{Y}$ has full column rank, contradicting the maximality of $\{r_i(x)\}_{i=1}^s$.

We omit the reverse implications as they can be shown analogously. □

Theorem 3.1 and Theorem 3.4 together yield the following result.

Theorem 3.5. *Let $P(x) \in \mathbb{F}[x]^{m \times n}$. Then, denoting by s the geometric multiplicity of μ as an eigenvalue of $P(x)$, there exists a maximal set of root polynomials at μ for $P(x)$, say $r_1(x), \dots, r_s(x)$, such that their orders are precisely the nonzero partial multiplicities of μ as an eigenvalue of $P(x)$.*

Proof. The proof is constructive: let $D(x)$ be the local Smith form at μ of $P(x)$, so that $P(x) = A(x)D(x)B(x)$ for some matrix polynomials of appropriate size and such that $\det A(\mu) \det B(\mu) \neq 0$. Then, by Theorem 3.1 $e_r, e_{r-1}, \dots, e_{r-s+1}$ are a maximal set of root polynomial at μ for $D(x)$. Hence, by Theorem 3.4, $\operatorname{adj} B(x)e_r, \operatorname{adj} B(x)e_{r-1}, \dots, \operatorname{adj} B(x)e_{r-s+1}$ are a maximal set of root polynomials at μ for $P(x)$, and by Proposition 3.2 and Theorem 3.1 their orders are the partial multiplicities of the eigenvalue μ of $P(x)$. □

4 Extremality properties of maximal sets of root polynomials

Theorem 4.1. *Let $P(x) \in \mathbb{F}[x]^{m \times n}$ and $\mu \in \mathbb{F}$ be one of its finite eigenvalues. Then*

1. *all complete sets of root polynomials of $P(x)$ at μ have the same cardinality: in particular, all maximal sets of root polynomials of $P(x)$ have the same cardinality, which we call s;*

2. *all maximal sets of root polynomials of $P(x)$ at μ have the same ordered list of orders, that we call $\ell_1 \geq \cdots \geq \ell_s$;*

3. *let $\widetilde{r}_1(x), \ldots, \widetilde{r}_s(x)$ be a complete set of root polynomials for $P(x)$ at μ with orders $\kappa_1 \geq \cdots \geq \kappa_s$: then*

   3.1 $\ell_i \geq \kappa_i, \quad i = 1, \ldots s;$

   3.2 *$\widetilde{r}_1(x), \ldots, \widetilde{r}_s(x)$ is a maximal set of root polynomials for $P(x)$ at μ $\Leftrightarrow$ $\ell_i = \kappa_i$, $i = 1, \ldots, s$;*

   3.3 *$\widetilde{r}_1(x), \ldots, \widetilde{r}_s(x)$ is a maximal set of root polynomials for $P(x)$ at μ $\Leftrightarrow$ $\sum_i \ell_i = \sum_i \kappa_i$.*

Proof. 1. It is an immediate corollary of Proposition 2.16.

2. Let $\{r_i(x)\}_{i=1}^s$ and $\{t_i(x)\}_{i=1}^s$ be two maximal set of root polynomials at μ for $P(x)$, with orders, resp., $\ell_1 \geq \cdots \geq \ell_s$ and $\rho_1 \geq \cdots \geq \rho_s$. By Definition 2.17, there is no root polynomials at μ for $P(x)$ of order larger than ℓ_1, nor there is one of order larger than ρ_1. Hence, $\ell_1 = \rho_1$.

   Now we proceed by induction. Assume that $\ell_i = \rho_i$ for $i = 1, \ldots, j < s$, and suppose that $\ell_{j+1} \neq \rho_{j+1}$. Then, without loss of generality, let $\ell_{j+1} > \rho_{j+1}$. Let $M(x)$ be a minimal basis for $P(x)$ and denote the columns of $M(\mu)$ by $u_1, \ldots, u_p$. Again by Definition 2.17, the above implies that $u_1, \ldots, u_p, t_1(\mu), \ldots, t_j(\mu), r_k(\mu)$ are linearly dependent for all $k \leq j+1$, since $\ell_k \geq \ell_{j+1} > \rho_{j+1}$. Therefore, $u_1, \ldots, u_p, r_1(\mu), \ldots, r_{j+1}(\mu) \in \operatorname{span}\{u_1, \ldots, u_p, t_1(\mu), \ldots, t_j(\mu)\}$. Hence, there are $p + j + 1$ linearly independent vectors that all lie in a subspace of dimension $j + p$, which is absurd.

3.1 Let $\{r_i(x)\}_{i=1}^s$ be a maximal set of root polynomials at μ for $P(x)$, listed by nonincreasing order ℓ_i. By Definition 2.17, there is no root polynomials at μ for $P(x)$ of order larger than ℓ_1, implying $\ell_1 \geq \kappa_1$. Now by induction suppose that $\ell_i \geq \kappa_i$ for $i = 1, \ldots, j$, but $\ell_j < \kappa_j$. Then $u_1, \ldots, u_p, \widetilde{r}_1(\mu), \ldots, \widetilde{r}_j(\mu) \in \operatorname{span}\{u_1, \ldots, u_p, r_1(\mu), \ldots, r_{j-1}(\mu)\}$, as if not some $\hat{r}_i(x)$, $i \leq j$, may be picked to contradict Definition 2.17 showing that $r_1(x), \ldots, r_s(x)$ are not maximal. Then again we have $j + p$ linearly independent vectors lying in a subspace of dimension $j + p - 1$: a contradiction.

3.2 Again, let $\{r_i(x)\}_{i=1}^s$ be a maximal set of root polynomials at μ for $P(x)$, listed by nonincreasing order ℓ_i. One implication is immediate by item 2. Suppose now that $\ell_i = \kappa_i$ for all i, but$\{\widetilde{r}_i(x)\}_{i=1}^s$ are not a maximal set. Since they are a complete set, it must happen that there exists a μ-independent set of root polynomials $\widetilde{r}_1(x), \ldots, \widetilde{r}_j(x), \hat{r}(x)$ of orders $\ell_1, \ldots, \ell_j, \ell$ with $\ell > \ell_{j+1}$. But in order not to contradict maximality of $\{r_i(x)\}_{i=1}^{s}$, it must be that $u_1, \ldots, u_p, \widetilde{r}_1(\mu), \widetilde{r}_j(\mu), \hat{r}(\mu) \in \mathrm{span}\{u_1, \ldots, u_p, r_1(\mu), \ldots, r_j(\mu)\}$, and again we get to the contradicting conclusion that $p + j + 1$ linearly independent vectors all lie in a subspace of dimesnion $p + j$.

3.3 Again, one implication is trivial. Now suppose that $\sum_i \ell_i = \sum_i \kappa_i$. There are two cases. If $\ell_i = \kappa_i$ for all i, we can use item 3.2; otherwise, there exists at least one j such that $\kappa_j > \ell_j$. But this is impossible because of item 2, concluding the proof.

□

We deduce that the orders of a maximal set of root polynomials at μ for $P(x)$ are precisely the nonzero partial multiplicites of μ as an eigenvalue of $P(x)$.

Theorem 4.2. *Let $P(x) \in \mathbb{F}[x]^{m \times n}$ have an eigenvalue μ with nonzero partial multiplicities $m_1 \geq \cdots \geq m_s$. Then, any maximal set of root polynomials at μ for $P(x)$ have orders $m_1, \ldots, m_s$.*

Proof. It follows from Theorem 3.5 and Theorem 4.1. □

5 Quotienting the terms of degree ℓ or higher

It turns out that a root polynomial of order ℓ is in fact defined up to an additive term of the form $(x - \mu)^\ell w(x)$ where $w(x) \in \mathbb{F}[x]^n$. More formally, we may state that the natural ring where the entries a root polynomial of order ℓ should be "naturally" defined is $\mathbb{F}[x]/\langle x^\ell \rangle$, where $\langle p(x) \rangle$ is the ideal generated by $p(x)$.

Proposition 5.1. *Let $P(x) \in \mathbb{F}[x]^{m \times n}$. If $v(x)$ is any root polynomial of order ℓ at μ for $P(x)$ then $\hat{v}(x) := v(x) + (x - \mu)^\ell w(x)$ (for any vector polynomial $w(x)$) is a root polynomia at μ for $P(x)l$ of order $\geq \ell$.*

Proof. Observe that $P(x)\hat{v}(x) = (x - \mu)^\ell (a(x) + P(x)w(x))$, with $P(x)v(x) = (x - \mu)^\ell a(x)$. Furthermore, $\hat{v}(\mu) = v(\mu)$, hence the former is in $\ker_\mu P(x)$ if and only if the latter is. Moreover, $a(x) + P(x)w(x) \not\equiv 0$, as otherwise $\hat{v}(x) \in \ker P(x)$, implying $\hat{v}(\mu) = v(\mu) \in \ker_\mu P(x)$. □

Proposition 5.1 says that the operation of adding a term of the form $(x-\mu)^\ell w(x)$ to a root polynomial of order ℓ cannot decrease the order. It does not, however, specify whether the order increases or remains equal. It turns out that both situation are possible, as illustrated below.

Example 5.2. Let

$$P(x) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & x^2 & 0 & 0 \\ 0 & 0 & x^5 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

It is easy to check that $r_1(x) = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}^T$ is a root polynomial of order 5 at 0; note that

$$P(x)r_1(x) + P(x)x^5 \begin{bmatrix} w_1(x) \\ w_2(x) \\ w_3(x) \\ w_4(x) \end{bmatrix} = x^5 \begin{bmatrix} w_1(x) \\ x^2 w_2(x) \\ 1 + x^5 w_3(x) \\ 0 \end{bmatrix},$$

showing that the order of $r_1(x) + x^5 w(x)$ must be equal to 5 for any $w(x) \in \mathbb{F}[x]^4$. On the other hand let $r_2(x) = \begin{bmatrix} x & 1 & 0 & 0 \end{bmatrix}^T$ and $w(x) = \begin{bmatrix} -1 & 0 & 0 & 0 \end{bmatrix}^T$, then $r_2(x)$ is a root polynomial of order 1 but $r_2(x) + xw(x)$ is a root polynomial of order $2 > 1$.

We now turn to illustrating how the quotienting operation of Proposition 5.1 acts on set of root polynomials.

Theorem 5.3. *Let $r_1(x), \dots, r_s(x)$ be root polynomials at μ for $P(x)$ having orders $\ell_1 \geq \dots \ell_s$, and write*

$$r_i(x) = \sum_{j=0}^{d_i} v_{i,j}(x-\mu)^j.$$

Defining

$$q_i(x) = \sum_{j=0}^{\ell_i - 1} v_{i,j}(x-\mu)^j$$

(where $v_{i,j} = 0$ for all $j > d_i$, i.e., $q_i(x) = r_i(x)$, if $d_i \leq \ell_i - 1$), it holds that

1. *$r_1(x), \dots, r_s(x)$ are a μ-independent set of roots polynomials at μ if and only if $q_1(x), \dots, q_s(x)$ are;*

2. *$r_1(x), \dots, r_s(x)$ are a complete set of roots polynomials at μ if and only if $q_1(x), \dots, q_s(x)$ are;*

3. *$r_1(x), \dots, r_s(x)$ are a maximal set of roots polynomials at μ if and only if $q_1(x), \dots, q_s(x)$ are.*

Proof.

1. μ-independence is clearly preserved by adding terms of the form $(x-\mu)^{\ell_i} w_i(x)$ to each root polynomial. Indeed, it is a local property at μ, i.e., it only depends on the 0th order coefficients $r_i(\mu) = q_i(\mu) = v_{i,0}$.

2. It is a corollary of Proposition 2.16.

3. Suppose that $\{r_i(x)\}_{i=1}^s$ are a maximal set, and denote the order of $q_i(x)$ by κ_i. By Proposition 5.1, $\kappa_i \geq \ell_i$. On the other hand, $\{q_i(x)\}_{i=1}^s$ are a complete set, because $\{r_i(x)\}_{i=1}^s$ are. Therefore, by item 3.1 in Theorem 4.1, denoting by σ any permutation of $\{1, \ldots, s\}$ such that the orders of $q_{\sigma(i)}(x)$ are listed in nonincreasing order, $\ell_i \geq \kappa_{\sigma(i)}$ for all i. In particular, we have $\sum_i \kappa_{\sigma(i)} = \sum_i \kappa_i \leq \sum_i \ell_i \leq \sum_i \kappa_i$. Therefore, $\sum_i \kappa_i = \sum_i \ell_i$, implying $\kappa_i = \ell_i$ for all i, and by item 3.2 in Theorem 4.1, $\{q_i(x)\}_{i=1}^s$ are a maximal set.

   The reverse implication can be proved analogously.

□

The results in this section suggest the following definition

Definition 5.4. A root polynomial of order ℓ, say, $r(x) \in \mathbb{F}[x]^n$, is said to be minimal if $\deg r(x) < \ell$.

The polynomial vectors $r_1(x), \ldots, r_s(x) \in \mathbb{F}[x]^n$ are a minimaximal set of root polynomials at μ for $P(x)$ if they are maximal and minimal, i.e., they are a maximal set and they satisfy

$$\deg r_i(x) < \ell_i \quad \forall \quad i$$

where $\ell_1 \geq \cdots \geq \ell_s$ are their orders.

6 Root polynomials at infinity

From now on, we formally set $\infty := \frac{1}{0}$ where $1 \in \mathbb{F}$ and $0 \in \mathbb{F}$ are the zero elements for, respectively, multiplication and addition within $\mathbb{F}$. It turns out that the point at infinity has relevance for the spectral theory of matrix polynomials [4, 14, 15, 16, 17], and this motivates a definition and an analysis of root polynomials at ∞.

Let $P(x) = \sum_{i=0}^g P_i x^i \in \mathbb{F}[x]^{m \times n}$ have grade[2] [15, 17] g, and define

$$\operatorname{rev}_g P(x) = \sum_{i=0}^{g} P_{g-i} x^i = x^g P(1/x).$$

Note that rev_g is involutory, i.e., $\operatorname{rev}_g \operatorname{rev}_g P(x) \equiv P(x)$.

Definition 6.1 (Eigenvalues at infinity)**.**

Lemma 6.2. *Let $P(x) \in \mathbb{F}[x]^{m \times n}$ and $g \geq \deg P(x)$. Then $\operatorname{rev}_g P(x) = x^{g - \deg P} A(x)$, $A(x) \in \mathbb{F}[x]^{m \times n}$, and $A(0) = P_{\deg P} \neq 0$.*

[2]The grade is an integer, greater than or equal to the degree, attached to a polynomial (with scalar, vector, or matrix coefficients); the partial multiplicities at infinity depend on the choice of grade, and for this reason we speak about root polynomials at infinity for a pair (matrix polynomial, grade). Although usually the grade coincides with the degree, in certain applications this may not happen.

Proof. Denote $\delta = g - \deg P \geq 0$. By definition, $\operatorname{rev}_g P(x) = \sum_{i=0}^{g} P_{g-i}x^i = x^\delta \sum_{i=0}^{\deg P} P_{\deg P - i}x^i$. $\square$

Definition 6.3. Let $r(x) \in \mathbb{F}[x]^n$ be a polynomial vector of degree $\deg r(x)$. Moreover, let $P(x) \in \mathbb{F}[x]^{m\times n}$ have grade g. We say that $r(x)$ is a root polynomial of order ℓ at infinity for the pair $(P(x), g)$ if $\operatorname{rev}_{\deg r} r(x)$ is a root polynomial of order ℓ at 0 for $\operatorname{rev}_g P(x)$.

Proposition 6.4. *The polynomial vector $r(x)$ is a root polynomial of order ℓ at infinity for $P(x)$ if and only if*

1. $\deg P(x)r(x) = g + \deg r(x) - \ell$ *and*
2. $\rho \notin \ker_\infty P(x)$,

where ρ is the leading coefficient of $r(x)$ and $\ker_\infty P(x) \subseteq \mathbb{F}^n$ is the subspace spanned by the columns of the "high order coefficient matrix" [10] of any minimal basis $M(x)$ of $P(x)$

Proof. Suppose that $r(x)$ is a root polynomial of order ℓ at infinity for $P(x)$. By Definition 6.3,

$$\operatorname{rev}_g P(x) \operatorname{rev}_{\deg r} r(x) = x^{\deg r + g} P(1/x) r(1/x) = x^\ell a(x), \qquad a(0) \neq 0.$$

By Lemma 6.2, the latter equation implies $\deg P(x)r(x) = g + \deg r(x) - \ell$. Conversely, assume that $\deg P(x)r(x) = g + \deg r(x) - \ell$: again using Lemma 6.2, we conclude that $\operatorname{rev}_g P(x) \operatorname{rev}_{\deg r} r(x) = x^\ell a(x)$ for some $a(x)$ such that $a(0) \neq 0$.

To conclude the proof, note that by [17, Sections 6.1–6.2], $r_{\deg r} \in \ker_\infty P(x) \Leftrightarrow \operatorname{rev}_{\deg r} r(0) \in \ker_0 \operatorname{rev}_g P(x)$. $\square$

For the following definitions, given any matrix polynomial $P(x) \in \mathbb{F}[x]^{m\times n}$ with minimal basis $M(x)$ we denote by M_h the "high order coefficient matrix" [10] of $M(x)$. Note that the latter is the same as the "columnwise reversal" of $M(x)$ evaluated at $x = 0$ [20].

Definition 6.5. Let $M(x)$ be a right minimal basis of $P(x)$. The vectors $r_1(x), \dots, r_s(x) \in \mathbb{F}[x]^n$, having leading coefficients $\rho_1, \dots, \rho_s$, are an ∞-independent set of root polynomials at ∞ for $P(x)$ if $r_i(x)$ is a root polynomial at ∞ for $P(x)$ for each $i = 1, \dots, s$, and the matrix

$$\begin{bmatrix} M_h & \rho_1 & \dots & \rho_s \end{bmatrix}$$

has full column rank.

Definition 6.6. Let $M(x)$ be a right minimal basis of $P(x)$. The vectors $r_1(x), \dots, r_s(x) \in \mathbb{F}[x]^n$, having leading coefficients $\rho_1, \dots, \rho_s$, are a complete set of root polynomials at ∞ for $P(x)$ if they are ∞-independent and there does not exist any set of $s+1$ root polynomials at ∞ for $P(x)$, say$\{t_i(x)\}_{i=1}^{s+1}$, having leading coefficients $\{\tau_i\}_{i=1}^{s+1}$, such that the matrix

$$\begin{bmatrix} M_h & \tau_1 & \dots & \tau_{s+1} \end{bmatrix}$$

has full column rank.

Definition 6.7. Let $M(x)$ be a right minimal basis of $P(x)$. The vectors $r_1(x), \dots, r_s(x) \in \mathbb{F}[x]^n$, having leading coefficients $\rho_1, \dots, \rho_s$, are a maximal set of root polynomials at ∞ for $P(x)$ if they are complete and their orders as root polynomials at infinity for $P(x)$, say, $\ell_1 \geq \dots \ell_s > 0$, satisfy the following property: for all $j = 1, \dots, s$, there is no root polynomial at infinity $\hat{r}(x)$ of order $\ell > \ell_j$ and leading coefficient $\hat{\rho}$ such that the matrix

$$\begin{bmatrix} M_h & \rho_1 & \dots & \rho_{j-1} & \hat{\rho} \end{bmatrix}$$

has full column rank.

Definition 6.8. The polynomial vectors $r_1(x), \dots, r_s(x) \in \mathbb{F}[x]^n$ are a mini-maximal set of root polynomials at ∞ for $P(x)$ if they are maximal and they satisfy

$$\deg r_i(x) < \ell_i \quad \forall \;\; i$$

where $\ell_1 \geq \dots \geq \ell_s$ are their orders.

When a matrix polynomial has infinite eigenvalues, and root polynomials at infinity are brought into the pictures, one can prove results completely analogous to those developed for finite eigenvalues (except that the notation gets more complicated). We omit further details to keep the paper compact.

7 Behaviour under rational reparametrizations

In this section we study the root polynomials for the pair of matrix polynomials $P(x) \in \mathbb{F}[x]^{m \times n}$ and $Q(y) = [d(y)]^g P(x(y))$ where g is the grade [15, 17] of $P(x)$ and

$$x(y) = \frac{n(y)}{d(y)}$$

for some coprime polynomials $n(y), d(y) \in \mathbb{F}[y]$. Note that Möbius transformations, studied in [15], correspond to $n(y)$ and $d(y)$ both having grade 1.

Theorem 7.1. *Let $\mu \in \mathbb{F}$ be an eigenvalue of $P(x)$ and let $\lambda \in \mathbb{F}$ be a solution of multiplicity m of the algebraic equation $\mu d(y) = n(y)$. Then,*

$$r(x) = \sum_{i=0}^{d_r} v_i (x - \mu)^i$$

is a root polynomial at μ *for* $P(x)$ *having order* ℓ *if and only if*

$$q(y) = [d(y)]^{d_r} r(x(y))$$

is a root polynomial at λ *for* $Q(y)$ *having order* $m\ell$*.*

Proof. Observe that $q(y)$ is a polynomial vector. Indeed, we get

$$q(y) = \sum_{i=0}^{d_r} v_i [d(y)]^{d_r - i} (n(y) - \mu d(y))^i.$$

Since by assumption $n(y) - \mu d(y) = (y - \lambda)^m w(y)$ for some scalar polynomial $w(y)$, $w(\lambda) \neq 0$, and using $P(x)r(x) = (x - \mu)^\ell a(x)$ for some polynomial vector $a(x)$, $a(\mu) \neq 0$,

$$Q(y)q(y) = [d(y)]^{d_r+g} P(x(y))r(x(y)) = (y - \lambda)^{m\ell} [w(y)]^\ell [d(y)]^{d_r+g-\ell} a(x(y)). \quad (1)$$

Now, since $\deg P(x)r(x) \leq d_r + g$, it must be $\deg a(x) \leq g + d_r - \ell$, and hence, $w(y)^\ell [d(y)]^{d_r+g-\ell} a(x(y))$ is a polynomial vector. Furthermore, $w(\lambda) \neq 0$ by assumption, $d(\lambda) \neq 0$ as otherwise $n(\lambda) = 0$ contradicting coprimality, and $a(x(\lambda)) = a(\mu) \neq 0$.

It remains to show $q(\lambda) \notin \ker_\lambda Q(y)$. Note first that $q(\lambda) = [d(\lambda)]^{d_r} r(\mu) \neq 0$. Let $M(x)$ be a minimal basis for $P(x)$ and denote its columns by $u_i(x), \ldots, u_p(x)$. Suppose further that $\deg u_i(x) = \beta_i$. It is shown in [17, Section 6.1] that the matrix $N(y)$ whose columns are $[d(y)]^{\beta_i} u_i(x(y))$ is a minimal basis for $Q(y)$. Suppose for a contradiction that $q(\lambda) = N(\lambda)c$ for some $c \in \mathbb{F}^p$. Let

$$d = \begin{bmatrix} d(\lambda)^{\beta_1 - d_r} & & \\ & \ddots & \\ & & d(\lambda)^{\beta_p - d_r} \end{bmatrix} c.$$

Then, $r(\mu) = M(\mu)d$ so that $r(\mu) \in \ker_\mu P(x)$, thus concluding the proof of the first implication.

Conversely, supose that $r(x)$ is not a polynomial at μ for $P(x)$ of order ℓ. If it is a root polynomial at μ for $P(x)$ of order $\ell' \neq \ell$ then $q(y)$ is a root polynomial at λ for $Q(y)$ of order $m\ell' \neq m\ell$ by the first part of the proof. Hence, we may assume that $r(x)$ is not a root polynomial at μ for $P(x)$, that is, either $P(\mu)r(\mu) \neq 0$ or $r(\mu) \in \ker_\mu P(x)$. If the former holds, then by (1) and using $x(\lambda) = \mu$ we deduce that $Q(\lambda)q(\lambda) \neq 0$. If the latter holds, then we have that, for some $c \in \mathbb{F}^p$ and d defined as above,

$$q(\lambda) = r(\mu) = M(\mu)d = N(\lambda)c.$$

□

For a Möbius transformation, the equation $n(y) = d(y)\mu$ is linear. Hence, when Theorem 7.1 is applied in this scenario, $m \equiv 1$ regardless of the value of μ.

We now show how sets of root polynomials behave under a rational reparametrization.

Theorem 7.2. *Let $\mu \in \mathbb{F}$ be an eigenvalue of $P(x)$ and let $\lambda \in \mathbb{F}$ be a solution of multiplicity m of the algebraic equation $\mu d(y) = n(y)$. Suppose that, for $i = 1, \dots, s$,*

$$r_i(x) = \sum_{j=0}^{d_i} v_i (x - \mu)^i$$

are root polynomials at μ for $P(x)$ having order $\ell_1 \geq \dots \geq \ell_s$. Then, defining

$$q_i(y) = [d(y)]^{d_i} r_i(x(y))$$

we have that:

1. *$r_1, \dots r_s$ are μ-independent if and only if $q_1, \dots, q_s$ are λ-independent;*
2. *$r_1, \dots r_s$ are complete if and only if $q_1, \dots, q_s$ are complete;*
3. *$r_1, \dots r_s$ are maximal if and only if $q_1, \dots, q_s$ are maximal.*

Proof. In the same notation as in the proof of Theorem 7.1,

$$\begin{bmatrix} N(\lambda) & q_1(\lambda) & \dots q_s(\lambda) \end{bmatrix} = \begin{bmatrix} M(\mu) & r_1(\mu) & \dots r_s(\mu) \end{bmatrix} \begin{bmatrix} d(\lambda)^{\beta_1} & & & & & \\ & \ddots & & & & \\ & & d(\lambda)^{\beta_p} & & & \\ & & & d(\lambda)^{d_1} & & \\ & & & & \ddots & \\ & & & & & d(\lambda)^{d_s} \end{bmatrix},$$

and because $d(\lambda) \neq 0$ item 1 is proved.

To prove item 2, if the $q_i(\lambda)$ are not a complete set, then by the proof of Proposition 2.16 we see that $\dim \ker Q(\lambda) > s$. This is a contradiction, because by definition of $Q(y)$ $\dim \ker P(\mu) = \dim \ker Q(\lambda) = s$. The converse statement can be shown similarly.

Finally, item 3 is a consequence of [17, Theorem 4.1] and of Theorem 4.2. □

Analogous results hold for the cases $\mu = \infty$, $\lambda = \infty$, or both. They can be proved using a technique analogous to the strategy employed in [17] to deal with infinite elementary divisors. We omit the details.

8 Linearizations and recovery properties

The goal of this section is to illustrate how root polynomials can be used to obtain, in a compact and unified way, the recovery of both minimal bases of singular polynomials and Jordan chains of regular matrix polynomials.

Definition 8.1 (Linearization)**.** A matrix polynomial of degree at most 1 $L(x)$ is called a *linearization* for $P(x) \in \mathbb{F}[x]^{m\times n}$ there exist $k \in \mathbb{N}$ and unimodular matrix polynomials $U(x) \in \mathbb{F}[x]^{(m+k)\times(m+k)}$ and $V(x) \in \mathbb{F}[x]^{(n+k)\times(n+k)}$ such that

$$L(x) = U(x)\begin{bmatrix} I_k & 0 \\ 0 & P(x) \end{bmatrix} V(x).$$

Lemma 8.2, Lemma 8.3, and Proposition 8.4 below are the basic tools that we are going to use throughout this section.

Lemma 8.2. *Let $P(x) \in \mathbb{F}[x]^{m\times n}$ and $Q(x) = \begin{bmatrix} I_k & 0 \\ 0 & P(x) \end{bmatrix}$ for some $k \geq 0$. Then $N(x)$ is a minimal basis of $Q(x)$ if and only if $N(x) = \begin{bmatrix} 0 \\ M(x) \end{bmatrix}$ where $M(x)$ is a minimal basis for $P(x)$. Moreover, for any $\mu \in \mathbb{F}$, $\ker_\mu Q(x) = \operatorname{span}\{\begin{bmatrix} 0 \\ v \end{bmatrix}, \ v \in \ker_\mu P(x)\}$.*

Proof. Note that the second statement follows immediately from the first. To prove the first statement, note that $Q(x)\begin{bmatrix} w(x) \\ v(x) \end{bmatrix} = \begin{bmatrix} w(x) \\ P(x)v(x) \end{bmatrix}$, from which it easily follows that any basis for $\ker Q(x)$ is of the form $\hat{B}(x) = \begin{bmatrix} 0 \\ B(x) \end{bmatrix}$ where $B(x)$ is a basis for $\ker P(x)$. To conclude the proof we can invoke any of the characterizations of minimal bases from [10], e.g., a basis is minimal if and only if $F(\mu)$ is full rank for all $\mu \in \mathbb{F}$ and its higher order coefficient matrix is full rank as well. It is therefore easy to see that $B(x)$ is minimal if and only if $\hat{B}(x)$ is. □

Lemma 8.3. *Let $P(x) \in \mathbb{F}[x]^{m\times n}$ and $Q(x) = \begin{bmatrix} I_k & 0 \\ 0 & P(x) \end{bmatrix}$ for some $k \geq 0$. If $w(x) = \begin{bmatrix} \hat{w}(x) \\ \widetilde{w}(x) \end{bmatrix} \in \mathbb{F}[x]^{k+n}$ is a root polynomial at μ of order ℓ for $Q(x)$, then:*

1. *$\hat{w}(x) = (x-\mu)^\ell \hat{a}(x)$ for some polynomial vector $\hat{a}(x)$;*
2. *$\widetilde{w}(x)$ is a root polynomial at μ of order $\ell' \geq \ell$ for $P(x)$;*
3. *either $a(\mu) \neq 0$, or $\widetilde{w}(x)$ is a root polynomial at μ for $P(x)$ having order exactly ℓ, or both.*

Proof. By definition we have that, for some $a(x)$ with $a(\mu) \neq 0$,

$$(x-\mu)^\ell a(x) = Q(x)w(x) = \begin{bmatrix} \hat{w}(x) \\ P(x)\widetilde{w}(x) \end{bmatrix} = (x-\mu)^\ell \begin{bmatrix} \hat{a}(x) \\ \widetilde{a}(x) \end{bmatrix},$$

where in the last step we have just partitioned $a(x)$ appropriately. Thus, either $\hat{a}(\mu) \neq 0$, or $\widetilde{a}(\mu) \neq 0$, or both. This immediately proves the result (using also Lemma 8.2). □

Proposition 8.4. *Let* $P(x) \in \mathbb{F}[x]^{m\times n}$ *and* $Q(x) = \begin{bmatrix} I_k & 0 \\ 0 & P(x) \end{bmatrix}$ *for some* $k \geq 0$. *Suppose* $r_1(x), \ldots, r_s(x)$ *are a maximal set of root polynomials at* μ *for* $Q(x)$ *of orders* $\ell_1 \geq \cdots \geq \ell_s$. *Then*

1. *for all* $i = 1, \ldots, s$, $r_i(x) = \begin{bmatrix} (x-\mu)^{\ell_i} a_i(x) \\ \widetilde{r}_i(x) \end{bmatrix}$ *for some polynomial vectors* $a_1(x), \ldots, a_s(x)$;
2. $\widetilde{r}_1(x), \ldots, \widetilde{r}_s(x)$ *is a maximal set of root polynomials at* μ *for* $P(x)$ *of orders* $\ell_1 \geq \cdots \geq \ell_s$.

Proof. 1. It follows from item 1 in Lemma 8.3.

2. From item 2 in Lemma 8.3, we know that $\widetilde{r}_1(x), \ldots, \widetilde{r}_s(x)$ is a set of root polynomials at μ for $P(x)$ of orders $\widetilde{\ell}_i \geq \ell_i$. They are μ-independent, by applying Proposition 5.1 and Lemma 8.2. They are complete, by Proposition 2.16 and Lemma 8.2. Finally, since the partial multiplicities of μ as an eigenvalue of $Q(x)$ are the same as the partial multiplicites of μ as an eigenvalue of $P(x)$, by Theorem 4.2 and by the above, we get

$$\sum_i \widetilde{\ell}_i \leq \sum_i \ell_i \leq \sum_i \widetilde{\ell}_i,$$

and hence, equality holds. By items 3.2–3.3 in Theorem 4.1, $\widetilde{r}_1(x), \ldots, \widetilde{r}_s(x)$ is a maximal set, and $\widetilde{\ell}_i = \ell_i$ for all $i = 1, \ldots, s$.

□

An important class of linearizations is given by "Fiedler pencils", see, e.g., [5] and the references cited therein. This approach to linearizing matrix polynomials has originated in the paper [9], and has been generalized in different directions, such as to rectangular matrix polynomials [6], to a wider class of pencils [1, 2] as well as to nonmonomial bases [19]. To keep the paper compact, here we will focus on the original class of Fiedler pencils in the monomial basis and for square matrix polynomials, defined as in [5]. The statement of Theorem 8.5 allows for the recovery of root polynomials of the linearized matrix polynomial from those of a Fiedler pencil. It refers to the definition of the consecution-inversion structure of a fiedler pencil $F_\sigma(x)$ [5, Definition 3.3], which is denoted by CISS(σ). Since this definition requires a rather long and technical tour-de-force, we invite the reader to refer to [5] for the details.

Theorem 8.5. *Let $P(x) \in \mathbb{F}[x]^{n\times n}$ having grade $g \geq 2$, and suppose that $F_\sigma(x)$ is the Fiedler pencil of $P(x)$ associated with a bijection σ having CISS(σ)=$(c_1, i_1, \ldots, c_\ell, i_\ell)$. Also, let us block partition any vector of size $ng \times 1$ with g blocks of size $n \times 1$. Let $r_1(x), \ldots, r_s(x)$ be a maximal set of root polynomials at μ for $F_\sigma(x)$, having orders $\ell_1 \geq \cdots \geq \ell_s$. For all $j = 1, \ldots, s$ denote by $\widetilde{r}_j(x)$ the $(g - c_1)$th block of $r_j(x)$. Then, $\widetilde{r}_1(x), \ldots, \widetilde{r}_s(x)$ is a maximal set of root polynomials at μ for $P(x)$, having orders $\ell_1 \geq \cdots \geq \ell_s$.*

Proof. By [5, Corollary 4.7], two unimodular matrix polynomials $U(x)$, $V(x)$ such that

$$U(x)F_\sigma(x)V(x) = \begin{bmatrix} I_{(g-1)n} & 0 \\ 0 & P(x) \end{bmatrix} =: Q(x)$$

are explicitly known. Moreover, if $V_r(x)$ is the rightmost $ng \times n$ block of $V(x)$, and viewing $V_r(x)$ partitioned as a $g \times 1$ block vector with blocks of size $n \times n$, then $V_r(x)$ has exactly one block equal to I_n, located at the block index $(g - c_1)$ [5, Remark 5.4].

By Theorem 3.4, $V^{-1}(x)r_1(x), \ldots, V^{-1}(x)r_s(x)$ is a maximal set of root polynomials at μ for $Q(x)$ of orders $\ell_1 \geq \cdots \geq \ell_s$. By Proposition 8.4, their bottom blocks (say, $\hat{r}_1(x), \ldots, \hat{r}_s(x)$) form a maximal set of root polynomials at μ for $Q(x)$ of the same orders, whereas their other blocks are of the form $(x - \mu)^{\ell_i} a_i(x)$. From the observation above on the form of $V_r(x)$, we have that for some polynomial vector $b(x)$ it holds

$$\widetilde{r}_j(x) = \hat{r}_j(x) + (x - \mu)^{\ell_j} b(x),$$

which by Theorem 5.3 concludes the proof. □

Remark 8.6.
- Theorem 8.5 holds, in particular, for the first companion linearization (taking $c_1 = 0$) and for the second companion linearization (taking $c_1 = g - 1$).
- The result implies that the orders of $r_j(x)$ and $\widetilde{r}_j(x)$ are equal, because the partial multiplicities at μ of $P(x)$ and $F_\sigma(x)$ are.
- For regular matrix polynomials, the recovery result for root polynomials yields a recovery result for Jordan chains.

A second important class of linearizations is given by the $\mathbb{L}_1$ and $\mathbb{L}_2$ linearization spaces: see [14, Definition 3.1]. Again, for simplicity we focus on the case of the monomial bases originally discussed in [14]. We note however that an extension to nonmonomial bases is possible [16], and recovery properties for root polynomials can be derived for other bases as well.

The next proposition follows from, and slightly improves, [4, Theorems 4.1 and 4.6].

Proposition 8.7. *Let $P(x) = \sum_{i=0}^{g} P_i x^i \in \mathbb{F}[x]^{n\times n}$. Let $L(x) \in \mathbb{L}_1(P)$ have a nonzero right ansatz vector $v \in \mathbb{F}^g$. Also let $M \in GL(g, \mathbb{F})$ satisfy $Mv = e_1$. Then, there exists matrices $Y \in \mathbb{F}^{n\times n(g-1)}$, $Z \in \mathbb{F}^{n(g-1)\times n(g-1)}$ such that*

$$L(x) = (M^{-1} \otimes I_n) \begin{bmatrix} I_n & -Y \\ 0 & -Z \end{bmatrix} C_1(x),$$

where

$$C_1(x) = x \begin{bmatrix} P_k & & & \\ & I_n & & \\ & & \ddots & \\ & & & I_n \end{bmatrix} + \begin{bmatrix} P_{g-1} & \dots & P_1 & P_0 \\ -I_n & & & \\ & \ddots & & \\ & & -I_n & \end{bmatrix}$$

denotes the first companion linearization of $P(x)$. Moreover, if Z is nonsingular, then $L(x)$ is a strong linearization of $P(x)$.

Similarly, let $L(x) \in \mathbb{L}_2(P)$ have a nonzero left ansatz vector $w \in \mathbb{F}^g$. Also let $K \in GL(g, \mathbb{F})$ satisfy $w^T K = e_1^T$. Then, there exists matrices $X \in \mathbb{F}^{n(g-1)\times n}$, $Z \in \mathbb{F}^{n(g-1)\times n(g-1)}$ such that

$$L(x) = C_2(x) \begin{bmatrix} I_n & 0 \\ -X & -Z \end{bmatrix} (K^{-1} \otimes I_n),$$

where

$$C_2(x) = x \begin{bmatrix} P_k & & & \\ & I_n & & \\ & & \ddots & \\ & & & I_n \end{bmatrix} + \begin{bmatrix} P_{k-1} & -I_n & & \\ \vdots & & \ddots & \\ P_1 & & & -I_n \\ P_0 & & & \end{bmatrix}$$

is the second companion linearization of $P(x)$. Moreover, if Z is nonsingular, then $L(x)$ is a strong linearization of $P(x)$.

Proof. We only include the proof of the first statement as the second can be shown analogously. By definition of $\mathbb{L}_1(P)$ and M it is readily seen that $(M \otimes I_n)L(x) \in \mathbb{L}_1(P)$ with right ansatz vector e_1. Hence, by [14, Theorem 3.5], there exist Y, Z, of sizes as in the statement, satisfying

$$(M \otimes I_n)L(x) = x \begin{bmatrix} P_k & -Y \\ 0 & -Z \end{bmatrix} + \begin{bmatrix} Y + \begin{bmatrix} P_{k-1} & \dots & P_1 \end{bmatrix} & P_0 \\ Z & 0 \end{bmatrix} = \begin{bmatrix} I_n & -Y \\ 0 & -Z \end{bmatrix} C_1(x).$$

Hence, $L(x)$ is strictly equivalent to $C_1(x)$ if and only if Z is nonsingular. □

Remark 8.8. The property of Z being nonsingular in the statement of Proposition 8.7 is known as $L(x)$ having *full Z-rank*.

Theorem 8.9. *Let $P(x) \in \mathbb{F}[x]^{n\times n}$ have grade $g \geq 2$. Also, let us block partition any vector of size $ng \times 1$ with g blocks of size $n \times 1$.*

1. *Let $L(x) \in \mathbb{L}_1(P)$ have full Z-rank, and let $r_1(x), \dots, r_s(x)$ be a minimaximal set of root polynomials at μ for $L(x)$, having orders $\ell_1 \geq \cdots \geq \ell_s$. For all $j = 1, \dots, s$ denote by $\widetilde{r}_j(x)$ the gth block of $r_j(x)$. Then, $\widetilde{r}_1(x), \dots, \widetilde{r}_s(x)$ is a minimaximal set of root polynomials at μ for $P(x)$, having orders $\ell_1 \geq \cdots \geq \ell_s$.*

2. *Let $L(x) \in \mathbb{L}_2(P)$ have full Z-rank and left ansatz vector w, and let $r_1(x), \dots, r_s(x)$ be a minimaximal set of root polynomials at μ for $L(x)$, having orders $\ell_1 \geq \cdots \geq \ell_s$. Moreover, define $W = w^T \otimes I$ and, for all $j = 1, \dots, s$, $\widetilde{r}_j(x) := W r_j(x)$. Then, $\widetilde{r}_1(x), \dots, \widetilde{r}_s(x)$ is a minimaximal set of root polynomials at μ for $P(x)$, having orders $\ell_1 \geq \cdots \geq \ell_s$.*

Proof. 1. Let $L(x) \in \mathbb{L}_1(P)$ have full Z-rank. By Theorem 3.4, Proposition 8.7, and Definition 5.4, we can easily see that $\{r_i(x)\}_{i=1}^s$ is a minimaximal set of root polynomials at μ for $C_1(x)$, the first companion linearization of $P(x)$. The statement is therefore a corollary of Theorem 8.5.

2. Let $L(x) \in \mathbb{L}_2(P)$ have full Z-rank and left ansatz vector w. By Theorem 3.4, Proposition 8.7, and Definition 5.4, we can easily see that $\begin{bmatrix} I_n & 0 \\ -X & -Z \end{bmatrix} (K^{-1} \otimes I_n) r_1(x), \dots, \begin{bmatrix} I_n & 0 \\ -X & -Z \end{bmatrix} (K^{-1} \otimes I_n) r_s(x)$ is a minimaximal set of root polynomials at μ for $C_2(x)$, the second companion linearization of $P(x)$. Applying Theorem 8.5 we see that the first blocks of $\begin{bmatrix} I_n & 0 \\ -X & -Z \end{bmatrix} (K^{-1} \otimes I_n) r_j(x)$, $j = 1, \dots, s$, are a minimaximal set of root polynomials of $P(x)$. But these first blocks can be explicitly computed as

$$\begin{bmatrix} I_n & 0 \end{bmatrix} (K^{-1} \otimes I_n) r_j(x) = W r_j(x),$$

where we have used the property $w^T K = e_1$ that holds by definition of K (see the statement of Proposition 8.7).

□

We conclude this section by analyzing a third important class of linearizations discussed in [8]: block Kronecker linearizations. Again, we focus on block Kronecker linearizations of square matrix polynomials for simplicity and to keep the paper compact. A generalization to rectangular $P(x)$ is not particularly difficult (although it somewhat complicates the notation).

Theorem 8.10. *Let $P(x) \in \mathbb{F}[x]^{n \times n}$ having grade $g \geq 2$, and suppose that*

$$L(x) = \begin{bmatrix} M(x) & K_2(x)^T \\ K_1(x) & 0 \end{bmatrix} \in \mathbb{F}[x]^{(\eta+\epsilon+1)n \times (\eta+\epsilon+1)n}$$

is a block Kronecker pencil [8, Definitions 3.1 and 5.1] and a linearization of $P(x)$. Suppose further that $K_1(x)$ has ϵn rows and $K_2(x)$ has ηn rows, where ϵ

and η are defined as in [8, Section 5]. Also, let us block partition any vector of size $ng \times 1$ with g blocks of size $n \times 1$. Let $r_1(x), \dots, r_s(x)$ be a maximal set of root polynomials at μ for $L(x)$, having orders $\ell_1 \geq \cdots \geq \ell_s$. For all $j = 1, \dots, s$ denote by $\widetilde{r}_j(x)$ the $(\epsilon + 1)$th block of $r_j(x)$. Then, $\widetilde{r}_1(x), \dots, \widetilde{r}_s(x)$ is a maximal set of root polynomials at μ for $P(x)$, having orders $\ell_1 \geq \cdots \geq \ell_s$.

Proof. Using [8, Lemma 2.14] and [8, Remark 5.3], one can explicitly write down unimodular matrix polynomials $U(x)$ and $V(x)$ such that

$$U(x)L(x)V(x) = \begin{bmatrix} I_{(\eta+\epsilon)n} & 0 \\ 0 & P(x) \end{bmatrix}.$$

Moreover, denoting by $V_r(x)$ the rightmost $(\eta + \epsilon + 1)n \times n$ block of $V(x)$, and partitioning $V_r(x)$ as $(\eta + \epsilon + 1)$ block vector with $n \times n$ blocks, then $V_r(x)$ has (at least) one block equal to I_n, located at the block index $(\epsilon + 1)$. The result now follows by an argument analogous to that in the proof of Theorem 8.5. □

9 Dual pencils and root polynomials

The following definitions and basic results appear in [20] (for $\mathbb{F} = \mathbb{C}$, but their extension to a generic algebraically closed field does not cause any issues) and are also related to the pioneering work in [12, 13].

Definition 9.1 (Dual pencils)**.** Two matrix polynomials of degree at most 1, $L(x) = L_1 x + L_0 \in \mathbb{F}[x]^{m \times n}$ and $R(x) = R_1 x + R_0 \in \mathbb{F}[x]^{n \times p}$, are called dual if the following two conditions hold:

1. $L_1 R_0 = L_0 R_1$;
2. $\operatorname{rank} \begin{bmatrix} L_1 & L_0 \end{bmatrix} + \operatorname{rank} \begin{bmatrix} R_1 \\ R_0 \end{bmatrix} = 2n.$

In this case we say that $L(x)$ is a left dual of $R(x)$ and that $R(x)$ is a right dual of $L(x)$.

Definition 9.2 (Column-minimal matrix polynomials)**.** The matrix polynomial $P(x) \in \mathbb{F}[x]^{m \times n}$ is column-minimal if it does not have any zero right minimal index, i.e., $P(x)v \neq 0$ for all $v \in \mathbb{F}^n$.

Proposition 9.3. *[20, Lemma 3.1] $R(x)$ is a column-minimal right dual of $L(x)$ if and only if the columns of the matrix*

$$\begin{bmatrix} R_1 \\ R_0 \end{bmatrix}$$

are a basis for $\ker \begin{bmatrix} L_0 & -L_1 \end{bmatrix}$.

We now proceed to study how root polynomials change under the operation of duality. We first focus on dual pairs $L(x), R(x)$ where the right dual $R(x)$ is column-minimal.

Theorem 9.4. *Let $L(x) = L_1x + L_0$ and let $R(x) = R_1x + R_0$ be a column-minimal right dual of $L(x)$. Moreover, let $\mu \in \mathbb{F}$ be an eigenvalue of theirs. Also let $\gamma, \delta \in \mathbb{F}$ be such that $\delta\mu \neq \gamma$ and let $Q(x) = Q_1x + Q_0$ for some Q_1, Q_0 satisfying $Q_0R_1 - Q_1R_0 = I_p$. Then:*

1. *If $r(x) = \sum_{j=0}^{\ell-1} r_j(x-\mu)^j$ is a root polynomial of order ℓ at μ for $R(x)$ then $w(x) := (\gamma R_1 + \delta R_0)r(x)$ is a root polynomial of order ℓ at μ for $L(x)$.*
2. *If $r(x) = \sum_{j=0}^{\ell-1} r_j(x-\mu)^j$ is a root polynomial of order ℓ at μ for $L(x)$, then $w(x) := Q(x)r(x) - (x-\mu)^\ell Q_1 r_{\ell-1}$ is a root polynomial of order ℓ at μ for $R(x)$.*

Proof. 1. Suppose that $R(x)r(x) = (x-\mu)^\ell a$, $a = R_1 r_{\ell-1} \neq 0$. Then, we have

$$L(x)w(x) = L(x)(\gamma R_1+\delta R_0)r(x) = (\gamma L_1+\delta L_0)R(x)r(x) = (x-\mu)^\ell(\gamma L_1+\delta L_0)a.$$

We claim that $w(\mu) = (\gamma R_1 + \delta R_0)r(\mu) \notin \ker_\mu L(x)$. Indeed, otherwise $(\gamma - \delta\mu)R_1 r(\mu) = (\gamma R_1 + \delta R_0)r(\mu) = N(\mu)c$ for some constant vector c, where $N(x)$ is a minimal basis for $L(x)$. It is known [20, Theorem 3.9] that $Q(x)N(x)$ is a minimal basis for $R(x)$. Hence, $M(\mu) = (Q_0 + \mu Q_1)N(\mu)$, and $M(\mu)c = (\gamma - \delta\mu)(Q_0 + \mu Q_1)R_1 r(\mu) = (\gamma - \delta\mu)(r(\mu) + Q_1 R(\mu)r(\mu)) = (\gamma - \delta\mu)r(\mu)$, contradicting the assumption that $r(x)$ is a root polynomial.

It remains to show that $(\gamma L_1+\delta L_0)a \neq 0$. Indeed, if not, then $L(x)w(x) = 0$, and this leads to the same contradiction as above.

2. Suppose that $L(x)r(x) = (x-\mu)^\ell a$, $a = L_1 r_{\ell-1} \neq 0$. Expanding $L(x)$ and $r(x)$ in a power series in $(x-\mu)$, this is equivalent to

$$\begin{aligned} L_1\mu r_0 + L_0 r_0 &= 0, \\ L_1 r_0 + L_1\mu r_1 + L_0 r_1 &= 0, \\ &\vdots \\ L_1 r_{\ell-2} + L_1\mu r_{\ell-1} + L_0 r_{\ell-1} &= 0. \end{aligned}$$

But since $R(x)$ is a column-minimal right dual of $L(x)$, this implies that

$$\begin{bmatrix} r_0 & r_1 & \dots & r_{\ell-1} \\ -\mu r_0 & -r_0 - \mu r_1 & \dots & -r_{\ell-2} - \mu r_{\ell-1} \end{bmatrix} = \begin{bmatrix} R_1 \\ R_0 \end{bmatrix} \begin{bmatrix} w_0 & w_1 & \dots & w_{\ell-1} \end{bmatrix} \tag{2}$$

for some constant vectors $\{w_i\}_{i=1}^{\ell-1}$. Defining $w(x) = \sum_{j=0}^{\ell-1} w_j(x-\mu)^j$, this yields the equations $r(x) = R_1 w(x)$ and $(x-\mu)^\ell r_{\ell-1} - xr(x) = R_0 w(x)$.

Hence, we see that $R(x)w(x) = (x-\mu)^\ell r_{\ell-1}$. Note that $r_{\ell-1} \neq 0$, due to the linear independence of the columns of the matrices that bring a pencil to its Kronecker canonical form. Moreover, premultiplying (2) by $\begin{bmatrix} Q_0 & -Q_1 \end{bmatrix}$, we see that

$$\begin{aligned} Q_1 \mu r_0 + Q_0 r_0 &= w_0, \\ Q_1 r_0 + Q_1 \mu r_1 + Q_0 r_1 &= w_1, \\ &\vdots \\ Q_1 r_{\ell-2} + Q_1 \mu r_{\ell-1} + Q_0 r_{\ell-1} &= w_{\ell-1}, \end{aligned}$$

that is to say $Q(x)r(x) - Q_1(x-\mu)^\ell r_{\ell-1} = w(x)$.

It remains to check that $w_0 \notin \ker_\mu R(x)$. Suppose it does: then, $w_0 = M(\mu)c$ where $M(x)$ is a minimal basis of $R(x)$. But then $r_0 = R_1 M(\mu)c$, and since $R_1 M(x)$ is a minimal basis for $L(x)$ [20, Theorem 3.8], this violates the assumption that $r(x)$ is a root polynomial.

□

Remark 9.5. If in item 1 of Theorem 9.4 one drops the assumption that $r(x)$ has degree at most $\ell - 1$, then it is not necessarily true that $w(x)$ is a root polynomial of order ℓ of $L(x)$, but only that it is a root polynomial of order *at least* ℓ. For an example where the order increases take $L(x) = R(x) = \begin{bmatrix} x & 1 & 0 \\ 0 & x & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\delta = 0$, $\gamma \neq 0$, and $r(x) = \begin{bmatrix} 1 \\ -x \\ x \end{bmatrix}$.

The very same comment holds for item 2 of the same Theorem. Again, taking the same $L(x)$, $R(x)$, and $r(x)$, define $Q(x) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -x & -x \end{bmatrix}$ to check that $w(x)$ may have order $> \ell$ if $r(x)$ has degree higher than $\ell - 1$.

Theorem 9.6. *Let $L(x) = L_1 x + L_0$ and let $R(x) = R_1 x + R_0$ be a column-minimal right dual of $L(x)$. Moreover, let μ an eigenvalue of theirs. Also let $\gamma, \delta \in \mathbb{F}$ be such that $\delta\mu \neq \gamma$ and let $Q(x) = Q_1 x + Q_0$ for some Q_1, Q_0 satisfying $Q_0 R_1 - Q_1 R_0 = I_p$. Then:*

1. *If $r_1(x), \ldots, r_s(x)$ is a minimaximal set at μ for $R(x)$ then $w_j(x) := (\gamma R_1 + \delta R_0) r_j(x)$, $j = 1, \ldots, s$, is a minimaximal set at μ for $L(x)$.*

2. *If $r_1(x), \ldots, r_s(x)$ is a minimaximal set at μ for $L(x)$, then $w_j(x) := Q(x) r_j(x) - (x-\mu)^\ell Q_1 r_{j,\ell-1}$ is a minimaximal set at μ for $R(x)$.*

Proof. 1. Let $M(x)$ be a minimal basis for $R(x)$. We first prove μ-independence. By assumption, the matrix $X = \begin{bmatrix} M(\mu) & r_1(\mu) & \dots & r_s(\mu) \end{bmatrix}$ has full column rank. On the other hand, by [20, Theorem 3.8], a minimal basis for $L(x)$ is $\frac{\gamma R_1 + \delta R_0}{\gamma - \delta x} M(x)$. Now, suppose that the matrix

$$Y = \begin{bmatrix} \frac{\gamma R_1+\delta R_0}{\gamma-\delta\mu} M(\mu) & w_1(\mu) & \dots & w_s(\mu) \end{bmatrix} = (\frac{1}{\gamma - \delta\mu} \oplus I)(\gamma R_1 + \delta R_0) X$$

does not have full colum rank. Then, a nonzero linear combination of the columns of X, say Xc, is in $\ker(\gamma R_1 + \delta R_0)$. We deduce that $R(x)Xc = \left(\frac{\delta x-\gamma}{\delta\mu-\gamma} R(\mu) + \frac{x-\mu}{\gamma-\delta\mu}(\gamma R_1 + \delta R_0)\right) Xc = 0$, and hence, $R(x)$ has a zero minimal index, contradicting the assumption that it is column-minimal.

We next show completeness. Now suppose that the columns of X span $\ker R(\mu)$, but the columns of Y do not span $\ker L(\mu)$. This contradicts [20, Theorem 3.3], that guarantees that $\dim\ker L(\mu) = \dim\ker R(\mu)$.

To show maximality, it suffices to note that the partial multiplicities at μ of $R(x)$ and $L(x)$ coincide and that the orders of $\{r_i(x)\}_{i=1}^s$ and $\{w_i(x)\}_{i=1}^s$ coincide as well.

Finally, minimaximality follows from the fact that both $r_j(x)$ and $w_j(x)$ have degree $\leq \ell_j - 1$.

2. Denote by $M(x)$ a minimal basis or $R(x)$. We start by showing μ-independence. Suppose that the matrix $X = \begin{bmatrix} M(\mu) & w_1(\mu) & \dots & w_s(\mu) \end{bmatrix}$ does not have full column rank, i.e., $Xc = 0$ for some nonzero constant vector c. By [20, Theorem 3.8], a minimal basis for $L(x)$ is $\frac{\gamma R_1+\delta R_0}{\gamma-\delta x} M(x)$. Let now $Y = \frac{\gamma R_1+\delta R_0}{\gamma-\delta\mu} X$. The proof of Theorem 9.4 showed that $(\gamma R_1 + \delta R_0) w_j(\mu) = (\gamma - \delta\mu) r_j(\mu)$, and hence, $Y = (\gamma R_1 + \delta R_0) Xc = 0$ a contradiction, since Y has full column rank by the assumption that $r_j(x)$ are μ-independent.

We next show completeness. Suppose the columns of X do not span $\ker R(\mu)$. Since the columns of Y span $\ker L(\mu)$, and by [20, Theorem 3.3], this is a contradiction.

To show maximality, it suffices to note that the partial multiplicities at μ of $R(x)$ and $L(x)$ coincide and that the orders of the $r_j(x)$ and $w_j(x)$ coincide as well.

Finally, minimaximality follows from the fact that both $r_j(x)$ and $w_j(x)$ have grade $\ell_j - 1$.

□

To extend Theorem 9.4 and Theorem 9.6 to the case of a right dual which is not column-minimimial, we first need the following auxiliary results.

Lemma 9.7. *Let $L(x) = L_1x + L_0$ and $R(x) = R_1x + R_0$ be a right dual of $L(x)$. Suppose that $\begin{bmatrix} \hat{R}_1 \\ \hat{R}_0 \end{bmatrix}$ is a basis for the column space of $\begin{bmatrix} R_1 \\ R_0 \end{bmatrix}$ and let $\hat{R}(x) = \hat{R}_1x + \hat{R}_0$. Moreover, denote by B the full row rank matrix such that*

$$\begin{bmatrix} \hat{R}_1 \\ \hat{R}_0 \end{bmatrix} B = \begin{bmatrix} R_1 \\ R_0 \end{bmatrix},$$

let B^R be any right inverse of B, and let C, K be such that $\begin{bmatrix} K & B^R \end{bmatrix}$ is square and

$$\begin{bmatrix} C \\ B \end{bmatrix} \begin{bmatrix} K & B^R \end{bmatrix} = I.$$

Then:

1. *$\hat{R}(x)$ is a column-minimal right dual of $L(x)$;*
2. *if $M(x)$ is a minimal basis for $\hat{R}(x)$, then $\begin{bmatrix} K & B^R M(x) \end{bmatrix}$ is a minimal basis for $R(x)$;*
3. *if $N(x)$ is a minimal basis for $R(x)$, then the matrix obtained by keeping the nonzero columns of $BN(x)$ is a minimal basis for $\hat{R}(x)$.*

Proof. 1. We have $L_1\hat{R}_0 = L_1R_0B = L_0R_1B = L_0\hat{R}_1$ and

$$\operatorname{rank} \begin{bmatrix} \hat{R}_1 \\ \hat{R}_0 \end{bmatrix} = \operatorname{rank} \begin{bmatrix} R_1 \\ R_0 \end{bmatrix}.$$

2. Note first that $R(x)K = \hat{R}(x)BK = 0$. If $M(x)$ is a minimal basis for $\hat{R}(x)$, then $I \oplus M(x)$ is a minimal basis for $\begin{bmatrix} 0 & \hat{R}(x) \end{bmatrix} = R(x) \begin{bmatrix} K & B^R \end{bmatrix}$ which in turn implies that $\begin{bmatrix} K & B^R \end{bmatrix} (I \oplus M(x)) = \begin{bmatrix} K & B^R M(x) \end{bmatrix}$ is a basis for $\ker R(x)$. It is known [7, 10, 20] that if A is invertible and $\widetilde{M}(x)$ is minimal then $A\widetilde{M}(x)$ is also minimal, and this concludes the proof.

3. By reversing the final argument in the proof of item 2., we see that $\begin{bmatrix} C \\ B \end{bmatrix} N(x)$ is a minimal basis for $\begin{bmatrix} 0 & \hat{R}(x) \end{bmatrix}$. By [20, Lemma 3.6] and by the fact that clearly there exists an ordered minimal basis of the latter of the form $I \oplus \widetilde{N}(x)$, it must be that $\begin{bmatrix} CN(x) \\ BN(x) \end{bmatrix} = \begin{bmatrix} T_0 & \widetilde{T}(x) \\ 0 & \hat{T}(x) \end{bmatrix} \Pi$ for some square invertible matrix T_0 and some permutation matrix Π. Hence, $\begin{bmatrix} CN(x) \\ BN(x) \end{bmatrix} \Pi^T \begin{bmatrix} T_0^{-1} & -T_0^{-1}\widetilde{T}(x) \\ 0 & I \end{bmatrix} = I \oplus \hat{T}(x)$ is a basis for $\ker \begin{bmatrix} 0 & \hat{R}(x) \end{bmatrix}$; and again by [20, Lemma 3.6], it is minimal. It follows that $\hat{T}(x)$ is a minimal

basis for $\hat{R}(x)$. Moreover, from the argument above, we see that $\hat{T}(x)$ is, up to a permutation of its columns, precisely the matrix obtained by keeping the nonzero columns of $BN(x)$.

□

Theorem 9.8. *Under the same assumptions, and with the same notations, of Lemma 9.7:*

1. *If $\hat{r}(x) = \sum_{j=0}^{\ell-1} \hat{r}_j (x-\mu)^j$ is a root polynomial of order ℓ at μ for $\hat{R}(x)$ then $r(x) := B^R \hat{r}(x)$ is a root polynomial of order ℓ at μ for $R(x)$.*

2. *If $r(x) = \sum_{j=0}^{\ell-1} r_j (x-\mu)^j$ is a root polynomial of order ℓ at μ for $R(x)$, then $\hat{r}(x) := Br(x)$ is a root polynomial of order ℓ at μ for $\hat{R}(x)$.*

Proof. 1. By assumption, $\hat{R}(x)\hat{r}(x) = (x-\mu)^\ell a(x)$, $a(\mu) \neq 0$. Hence, $R(x)r(x) = R(x)B^R\hat{r}(x) = \hat{R}(x)\hat{r}(x) = (x-\mu)^\ell a(x)$. Assume now for a contradiction that $r(\mu) \in \ker_\mu R(x)$ and let $M(x)$ be a minimal basis for $\hat{R}(x)$, then by Lemma 9.7 we have that, for some vector $\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$, $\hat{r}(\mu) = BB^R\hat{r}(\mu) = B(Kc_1 + B^R M(\mu)c_2) = M(\mu)c_2$, a contradiction.

2. We know $R(x)r(x) = (x-\mu)^\ell a(x)$, $a(\mu) \neq 0$, which yields $\hat{R}(x)Br(x) = R(x)r(x) = (x-\mu)^\ell a(x)$. Moreover, if $M(x)$ is a minimal basis for $\hat{R}(x)$, and denoting by z the number of zero right minimal indices of $R(x)$ (so that K has precisely z columns),

$$\operatorname{rank} \begin{bmatrix} K & B^R M(\mu) & r(\mu) \end{bmatrix} = \operatorname{rank} \begin{bmatrix} I_z & 0 & Cr(\mu) \\ 0 & M(\mu) & \hat{r}(\mu) \end{bmatrix} = z + \operatorname{rank} \begin{bmatrix} M(\mu) & \hat{r}(\mu) \end{bmatrix},$$

concluding the proof.

□

Corollary 9.9. *Let $L(x) = L_1 x + L_0$ and let $R(x) = R_1 x + R_0$ be a right dual of $L(x)$. Given any column-minimal right dual of $L(x)$, say $\hat{R}(x)$, let the matrices B, B^R, K be defined as in Lemma 9.7, and let $\mu \in \mathbb{F}$ be an eigenvalue of $L(x), R(x)$ and $\hat{R}(x)$. Also let $\gamma, \delta \in \mathbb{F}$ be such that $\delta\mu \neq \gamma$ and let $Q(x) = Q_1 x + Q_0$ for some Q_1, Q_0 satisfying $Q_0 R_1 - Q_1 R_0 = B^R B$. Then:*

1. *If $r(x) = \sum_{j=0}^{\ell-1} r_j (x-\mu)^j$ is a root polynomial of order ℓ at μ for $R(x)$ then $w(x) := (\gamma R_1 + \delta R_0) r(x)$ is a root polynomial of order ℓ at μ for $L(x)$.*

2. *If $r(x) = \sum_{j=0}^{\ell-1} r_j (x-\mu)^j$ is a root polynomial of order ℓ at μ for $L(x)$, then $w(x) := Q(x)r(x) - (x-\mu)^\ell Q_1 r_{\ell-1}$ is a root polynomial of order ℓ at μ for $R(x)$.*

3. *If the columns of K, $r_1(x), \dots, r_s(x)$ is a minimaximal set of root polynomials at μ for $R(x)$ then $w_j(x) := (\gamma R_1 + \delta R_0) r_j(x)$, $j = 1, \dots, s$, is a minimaximal set of root polynomials at μ for $L(x)$.*

4. *If $r_1(x), \dots, r_s(x)$ is a minimaximal set of root polynomials at μ for $L(x)$, then the columns of K and $w_j(x) := B^R Q(x) r_j(x) - (x-\mu)^\ell B^R Q_1 r_{j,\ell-1}$, $j = 1, \dots, s$, is a minimaximal set of root polynomials at μ for $R(x)$.*

Proof. It is a corollary of Theorems 9.4, 9.6 and 9.8, see also [20, Theorem 3.9] and [20, Remark 3.10]. □

The results in this section allow to strengthen the recovery theorem for pencils in $\mathbb{L}_1$.

Theorem 9.10. *Let $P(x) \in \mathbb{F}[x]^{n\times n}$ have degree $k \geq 2$. Also, let us block partition any vector of size $nk \times 1$ with k blocks of size $n \times 1$.*

Let $L(x) \in \mathbb{L}_1(P)$ be a left dual of the pencil $D(x)$ defined in [20, Section 8], and let $r_1(x), \dots, r_s(x)$ be a minimaximal set of root polynomials at μ for $L(x)$, having orders $\ell_1 \geq \dots \geq \ell_s$. For all $j = 1, \dots, s$ denote by $\widetilde{r}_j(x)$ the kth block of $r_j(x)$. Then, $\widetilde{r}_1(x), \dots, \widetilde{r}_s(x)$ is a minimaximal set of root polynomials at μ for $P(x)$, having orders $\ell_1 \geq \dots \geq \ell_s$.

Proof. The pencil $D(x) = D_1 x + D_0$ is a column-minimal right dual of $L(x)$. Hence, by Theorem 9.6, a minimaximal set for $D(x)$ is $\{w_i(x)\}_{i=1}^s$ where, for each $1 \leq j \leq s$, $w_j(x) = Q(x) r_j(x) - Q_1 (x-\mu)^{\ell_j} r_{j,\ell_j - 1}$, $Q(x) = Q_1 x + Q_0$ and $Q_0 D_1 - Q_1 D_0 = I$. In turn, the first companion form $C_1(x)$ is a left dual of $D(x)$. Applying Theorem 9.6 with $\gamma = 1, \delta = 0$, we find that a minimaximal set for $C_1(x)$ is $\{D_1 w_i(x)\}_{i=1}^s$.

Note that by the proof of Theorem 9.4 we have, for all $j = 1, \dots, s$, $D_1 w_j(x) = r_j(x)$, and hence, $\{r_i(x)\}_{i=1}^s$ is a minimaximal set for $C_1(x)$. The statement now follows as a corollary of Theorem 8.5. □

10 Conclusions

We have studied the concept of root polynomials, showing that they are a useful tool in the theory of matrix polynomials. Indeed, several known results have been re-derived in a simplified manner. Moreover, we have also obtained a number of new theorems related to root polynomials and their interaction with other theoretical tools. As a particularly meaningful application, we have shown how eigenvectors and eigenspaces can be consistently defined for singular matrix polynomials, as subspaces of certain quotient spaces.

References

[1] M. I. Bueno, and F. De Terán, *Eigenvectors and minimal bases for some families of Fiedler–like linearizations*, Lin. Multilin. Algebra 61(12), 1605–1628, 2014.

[2] M. I. Bueno, F. De Terán, and F. M. Dopico, *Recovery of eigenvectors and minimal bases of matrix polynomials from generalized Fiedler linearizations*, SIAM J. Matrix Anal. Appl. 32, 463–483, 2011.

[3] R. Byers, V. Mehrmann, and H. Xu, *Trimmed linearizations for structured matrix polynomials*, Linear Algebra Appl. 429, 2373–2400, 2008.

[4] F. De Terán, F. M. Dopico, and D. S. Mackey, *Linearizations of singular matrix polynomials and the recovery of minimal indices*, Electr. J. Lin. Algebra 18, 371–402, 2009.

[5] F. De Terán, F. M. Dopico, and D. S. Mackey, *Fiedler companion linearizations and the recovery of minimal indices*, SIAM J. Matrix Anal. Appl. 31, 2181–2204, 2010.

[6] F. De Terán, F. M. Dopico, and D. S. Mackey, *Fiedler companion linearizations for rectangular matrix polynomials*, Linear Algebra Appl. 437, 957–991, 2012.

[7] F. De Terán, F. M. Dopico, D. S. Mackey, and P. Van Dooren, *Matrix polynomials with completely prescriebed eigenstructure*, SIAM J. Matrix Anal. Appl. 36, 302–328, 2015.

[8] F. M. Dopico, P. W. Lawrence, J. Pérez, and P. Van Dooren, *Block Kronecker linearizations of matrix polynomials and their backward errors*, Numer. Math., 2018, to appear. Extended version available as MIMS EPrint 2016.34, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, 2016.

[9] M. Fiedler, *A note on companion matrices*, Linear Algebra Appl. 372, 325–331, 2003.

[10] G. D. Forney Jr., *Minimal bases of rational vector spaces, with applications to multivariable linear systems*, SIAM J. Control 13, 493–520, 1975.

[11] I. Gohberg, P. Lancaster, and L. Rodman, *Matrix Polynomials*, SIAM, Philadelphia, PA, USA, 2009, (unabridged republication of book first published by Academic Press in 1982).

[12] V. N. Kublanovskaya, *The AB-algorihtm and its modifications for spectral problems of linear pencils of matrices*, Numer. Math. 43, 329–342, 1984.

[13] V. N. Kublanovskaya, *Methods and algorithms of solving spectral problems for polynomial and rational matrices (English translation)*, J. Math. Sci. (N. Y.). 96(3), 3085–3287, 1999.

[14] D. S. Mackey, N. Mackey, C. Mehl, and V. Merhmann, *Vector spaces of linearizations for matrix polynomials*, SIAM J. Matrix Anal. Appl. 28, 971–1004, 2006.

[15] D. S. Mackey, N. Mackey, C. Mehl, and V. Merhmann, *Möbius transformations of matrix polynomials*, Linear Algebra Appl. 470, 120–184, 2015.

[16] Y. Nakatsukasa, V. Noferini, and A. Townsend, *Vector spaces of linearizations for matrix polynomials: a bivariate polynomial approach*, SIAM J. Matrix Anal. Appl. 38(1), 1–29, 2017.

[17] V. Noferini, The behaviour of the complete eigenstructure of a polynomial matrix under a generic rational transformation, Electr. J. Lin. Algebra 22, 607–624, 2012.

[18] V. Noferini, *Polynomial Eigenproblems: a root-finding approach*, Ph.D. thesis, University of Pisa, Italy, 2012.

[19] V. Noferini, and J. Pérez, *Fiedler–comrade and Fiedler–Chebyshev pencils*, SIAM J. Matrix Anal. Appl. 37(4), 1600–1624, 2016.

[20] V. Noferini, and F. Poloni, *Duality of matrix pencils, Wong chains, and linearizations*, Linear Algebra Appl. 471, 730–767, 2015.

[21] H. S. Wilf, *generatingfunctionology*, Academic Press, San Diego, CA, USA, 1990.