Root polynomials and their role in the theory of matrix polynomials

Froilán M. Dopico[a,1], Vanni Noferini[b,*]

[a]*Departamento de Matemáticas, Universidad Carlos III de Madrid, Avda. Universidad 30, 28911 Leganés, Spain.*

[b]*Aalto University, Department of Mathematics and Systems Analysis, P.O. Box 11100, FI-00076, Aalto, Finland.*

Abstract

We develop a complete and rigorous theory of root polynomials of arbitrary matrix polynomials, i.e., either regular or singular, and study how these vector polynomials are related to the spectral properties of matrix polynomials. We pay particular attention to the so-called maximal sets of root polynomials and prove that they carry complete information about the eigenvalues (finite or infinite) of matrix polynomials and that they are related to the matrices that transform any matrix polynomial into its Smith form. In addition, we describe clearly, for the first time in the literature, the extremality properties of such maximal sets and identify some of them whose vectors have minimal grade. Once the main theoretical properties of root polynomials have been established, the interaction of root polynomials with three problems that have attracted considerable attention in the literature is analyzed. More precisely, we study the change of root polynomials under rational transformations, or reparametrizations, of matrix polynomials, the recovery of the root polynomials of a matrix polynomial from those of its most important linearizations, and the relationship between the root polynomials of two dual pencils. We emphasize that for the case of regular matrix polynomials all the results in this paper can be translated into the language of Jordan chains, as a consequence of the well known relationship between root polynomials and Jordan chains. Therefore, a number of open problems are also solved for Jordan chains of regular matrix polynomials. 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, Jordan chains
2000 MSC: 15A03, 15A18, 15A21, 15A54

---

*Corresponding author

[1]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.

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 [13, Section 1.5] as an analytical concept that simplifies the study of Jordan chains [13, Section 1.6]. Although root polynomials constitute a powerful tool for establishing spectral properties of matrix polynomials, they do not seem to be widely used in the literature. In fact, in the mentioned landmark reference [13], root polynomials are only used once after Chapter 1. More precisely, they are used in [13, 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 [19, Section 8] as an instrumental technical tool for proving the main results in that reference.

Both in [13] and [19], 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, [17, 19, 20], linearizations of matrix polynomials and related recovery properties [1, 2, 5, 6, 16, 18, 21], and dual pencils [22]. 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 is clearly explained in [13, Sections 1.5 and 1.6]. The relation of a root polynomial to a Jordan chain is the same as that of a generating function [23] to a sequence; for a survey of good reasons to use generating functions as a tool to manipulate and analyze sequences, see [23]. 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 polynomial matrices change the Jordan chains of the polynomial (the answer to this question is briefly sketched in [19, Remark 8.3] and is considered an open problem in [17, Remark 6.12]), or how Jordan chains of a matrix polynomial 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 any one of its eigenvalues, and the first vectors of traditional Jordan chains are eigenvectors [13, Section 1.4]. 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 [4]. 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 [19] and further developed in this paper.

Since one of our main goals is to develop a consistent, complete and rigorous theory of root polynomials, this paper is relatively long and contains a wealth of results that may puzzle the reader in a quick first reading. In order to guide the reader, we summarize in this and in the next paragraphs, the main definitions and results. The central object of this paper, i.e., a root polynomial of an arbitrary matrix polynomial at an eigenvalue, is introduced in Definition 2.11 together with the definition of its order. We advance that, as we will see throughout the paper, root polynomials are really significative when they form maximal sets of root polynomials of a matrix polynomial at an eigenvalue, a concept that is presented in Definition 2.17 after some other auxiliary notions have been described. The properties that make maximal sets of root polynomials interesting are analyzed in detail in Sections 3 and 4, where they are explicitly constructed in Theorem 3.5, as well as related to the local Smith forms of the considered matrix polynomial. Moreover, Theorem 4.2 establishes the fundamental result that all the maximal sets of root polynomials at the same eigenvalue of a matrix polynomial have the same orders and that such orders are the nonzero partial multiplicities of the eigenvalue. The extremality properties that motivate the use of the name *maximal* sets of root polynomials are described in Theorem 4.1. The property that root polynomials are defined only up to certain additive terms is clearly presented in Proposition 5.1 and Theorem 5.3 in Section 5 and allows us to introduce in Definition 5.4 root polynomials and maximal sets of root polynomials with minimal grade. Since eigenvalues at infinity play an important role in the spectral theory of matrix polynomials, Section 6 defines and establishes the properties of the root polynomials at infinity. Section 6 closes what can be considered the first part of the paper, where the theory and basic properties of root polynomials are developed. In contrast, the remaining sections describe how root polynomials interact with other important ideas from the theory and applications of matrix polynomials and their contents are described in the next paragraph.

We remark that the sections in the second part of the paper, i.e., Sections 7, 8 and 9, are not self contained, in contrast to the previous ones. This is due to the fact that the results developed in Sections 7, 8 and 9 are based on many other results and concepts whose detailed description would make this paper very long. The first problem considered in the second part of the paper concerns rational reparametrizations of matrix polynomials. The effect of rational reparametrizations, in particular of Möbius transformations, on the properties of matrix polynomials has received considerable attention in the last years (see, for instance, [17] and [19] as well as the references therein). In this scenario, Theorems 7.1 and 7.2 in Section 7 establish how root polynomials and maximal sets of root polynomials change under rational reparametrizations. Another hot research topic during the last decade in the area of matrix polynomials has been the development of different classes of linearizations and how to recover features of a matrix polynomial from the corresponding ones of their linearizations. Currently, there are so many classes of linearizations available in the literature that to study how the root polynomials of a matrix polynomial are recovered from all of them is completely out of the scope of this (and of any) paper. Therefore, Section 8 solves the recovery problem for maximal sets of root polynomials only for three of the most important classes of linearizations: for Fiedler linearizations [6] in Theorem 8.5, for linearizations in the spaces $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ [16] in Theorem 8.10, and for block Kronecker linearizations [10] in Theorem 8.11. The last results of this paper are included in Section 9 and deal with the relationship between the root polynomials of two dual pencils, which is carefully described in Theorems 9.4 and 9.6 and Corollary 9.9.

From the point of view of the organization, it only remains to add that the reader can find in Section 2 some basic results of the theory of matrix polynomials that are needed in the rest of the paper and that some conclusions are discussed in Section 10.

2. Preliminaries

This section is organized in three parts with three very different levels of novelty. Subsection 2.1 includes known results on the theory of matrix polynomials. Its goal is to emphasize those results that are relevant for this paper and to fix the notation that will be used in the rest of the paper. Subsection 2.2 introduces the definition of root polynomials for any matrix polynomial (possibly singular) and compares such definition with the one for regular matrix polynomials given in [13, Ch. 1]. Moreover, Subsection 2.2 stresses that root polynomials are more interesting, or useful, when they form sets with specific properties as, for instance, maximal sets of root polynomials. The results in Subsection 2.2 are taken from the recent reference [19], though more details than in [19] are provided, and we guess that most readers are not familiar with them. Finally, Subsection 2.3 presents some results which are completely new, to the best of our knowledge, and are of a fundamental nature. They show preliminary ideas on how to define consistently eigenvectors and eigenspaces for singular matrix polynomials and how these concepts are related to root polynomials, which indicates an interesting application of root polynomials and illustrates the importance of this concept. These ideas are only sketched in Subsection 2.3 for brevity, but also because they require an abstract-algebraic approach that it is not needed in the rest of the paper. They will be further developed in the future.

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 concepts. Indeed, left root polynomials (as well as left eigenvectors, 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 related concepts, 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.[2], $u \in R$ 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 [13, Chapter S1]. For brevity, these theorems are merged with some basic definitions and notation to be used in the rest of the paper.

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*

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

[2] 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 because, for the ring of $n \times n$ matrices with elements in any commutative ring, $uv = 1$ implies $vu = 1$.

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 nonegative integers $\kappa_{i,(j)}$ satisfy $\kappa_{i_1,(j)} \leq \kappa_{i_2,(j)} \Leftrightarrow i_1 \leq i_2$ 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 of nonzero partial multiplicities. If an eigenvalue x_j has geometric multiplicity s, we denote the nonincreasing list of its partial multiplicities by $m_1, \ldots, m_s$ with $m_1 = \kappa_{r,(j)} \geq \cdots \geq m_s = \kappa_{r+1-s,(j)}$.*

As mentioned in Theorem 2.3, the Smith canonical form $S(x)$ of $P(x)$ is uniquely determined by $P(x)$. In contrast, the unimodular matrices $U(x)$ and $V(x)$ are not unique. The definition of partial multiplicities can be extended from eigenvalues to any $\mu \in \mathbb{F}$ at the cost of allowing all of them to be zero if μ is not an eigenvalue. The use of this fact provides flexibility and simplicity in the statements of some results.

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 $\kappa_1, \ldots, \kappa_r$ (possibly allowing some, even all, 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 μ 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 [12].

Definition 2.5 (Minimal basis)**.** 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 $\operatorname{span} M(x) \subseteq \mathbb{F}(x)^n$.

It is proved in [12] that the ordered degrees of the columns of a minimal basis depend only on the subspace $\operatorname{span} M(x)$. Such degrees are called the *minimal indices* of 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 [22]. These concepts allow us to introduce the following definition.

Definition 2.6 (Minimal bases and indices of matrix polynomials)**.** 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 $b(x) \in \operatorname{span} M(x)$, always admits a *polynomial* solution $v(x)$ [12, Main Theorem]. In 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 [19], which are useful for the theory of root polynomials. Our first definition introduces a subspace of constant vectors that plays a fundamental role for extending the definition of root polynomials from regular to singular matrix polynomials. Note that in Definition 2.7 we follow the convention that $\operatorname{span} M = \{0\} \subset \mathbb{F}^n$ for any empty matrix $M \in \mathbb{F}^{n\times 0}$. Observe, in addition, that in Definition 2.7 $M(x)$ is an empty matrix whenever $P(x)$ is a regular matrix polynomial.

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, $\ker_\mu P(x) := \operatorname{span} M(\mu)$.

The next lemma proves that $\ker_\mu P(x)$ is uniquely determined by $P(x)$ and μ, as the notation we have chosen indicates.

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 [22, 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 characterizes the subspace $\ker_\mu P(x)$ formed by constant vectors in terms of polynomial vectors of the subspace $\ker P(x)$ of the rational vector space $\mathbb{F}(x)^n$.

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. □

The relationship presented in Lemma 2.10 between two subspaces of $\mathbb{F}^n$ is needed in the proofs of some results of this section.

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 [19, Sec. 8] and generalizes the definition of root polynomial given in [13, 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 $w(\mu) \neq 0$.

Observe that the first condition in Definition 2.11 reduces for regular matrix polynomials to the classical condition $r(\mu) \neq 0$ given in [13, Ch. 1]. The next proposition extends a result already known for regular matrix polynomials [13, Section 1.5] and proves that root polynomials exist only at any finite eigenvalue of the considered matrix polynomial.

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 $w(x) \in \mathbb{F}[x]^m$, $w(\mu) \neq 0$. □

We will see in Section 3, as well as in other sections of this paper, that the concept of root polynomial becomes very powerful when sets of root polynomials with particular properties are considered. Therefore, in the next definitions, we introduce sets of root polynomials with different properties that allow us to reach smoothly the key idea of a *maximal set of root polynomials at* μ in Definition 2.17.

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 minimal basis $N(x)$ of $P(x)$, then $N(x) = M(x)T(x)$ for some unimodular matrix $T(x) \in \mathbb{F}[x]^{p\times p}$ [22, Lemma 3.6]. Thus one has that

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

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

Remark 2.14. We take the chance to correct an imprecise statement in [19], 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 [19, 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 $r_1(x), \dots, r_s(x) \in \mathbb{F}[x]^n$ 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

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

has full column rank.

Proposition 2.16 characterizes when a set of polynomial vectors is a complete set of root polynomials in terms of a constant matrix. It will be used very often in the rest of the paper.

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 $r_1(x), \dots, r_s(x) \in \mathbb{F}[x]^n$ are a complete set of root polynomials at μ for $P(x)$ if and only if the columns of the matrix*

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

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 $v, r_1(x), \dots, r_s(x)$ 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 $(\dim \ker_\mu P(x) + s + 1)$ columns, such that $P(\mu)\hat{N} = 0$. It follows that $\dim \ker P(\mu) > \dim \ker_\mu P(x) + s$, and hence, the columns of N cannot be a basis of $\ker P(\mu)$. □

We are finally in the position of introducing a definition central to this paper.

Definition 2.17 (Maximal set of root polynomials)**.** 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 $r_1(x), \dots, r_s(x) \in \mathbb{F}[x]^n$ 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 $j = 1, \dots, s$, there is no root polynomial $\hat{r}(x)$ of order $\ell > \ell_j$ at μ for $P(x)$ such that the matrix

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

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 takes advantage of this relationship to propose a definition of eigenvector associated to an eigenvalue of a singular matrix polynomial. As far as we know, such definition is not available in the literature and hints an application of root polynomials that will be developed in depth in the future. We warn the reader that this section has a slightly more abstract-algebraic spirit than the rest of the paper.

We start by recalling the notion of quotient spaces. Let W be a vector space over $\mathbb{F}$, e.g., $W = \mathbb{F}^n$. For any linear subspace $V \subseteq W$, we consider the equivalence relation on W

$$\forall \;\; x, y \in W, \quad x \sim y \Leftrightarrow x - y \in V.$$

Then the quotient space W/V is defined as the set of equivalence classes

$$\forall \;\; x \in W, \quad [x] := \{v \in W : v \sim x\}.$$

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

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

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

Now, let $P(x) \in \mathbb{F}[x]^{m\times n}$, let μ be an eigenvalue of $P(x)$ and take $V = \ker_\mu P(x)$. In this setting, for a root polynomial $r(x)$ at μ for $P(x)$, asking that $r(\mu) \notin \ker_\mu P(x)$ 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 $P(x)$ given in [13, Ch. 1] (i.e., for $\ker_\mu P(x) = \{0\}$). Suppose for now, and for simplicity of exposition, that the eigenvalue μ of $P(x) \in \mathbb{F}[x]^{m\times n}$ has geometric multiplicity 1. With this assumption, for a regular $P(x)$, an eigenvector $v \in \ker P(\mu) (= \ker P(\mu)/\{0\} = \ker P(\mu)/\ker_\mu P(x))$ is defined uniquely up to a nonzero scalar by the well-known equation $P(\mu)v = 0$. For a singular $P(x)$, we claim that the quotient space framework also naturally leads to a definition of eigenvector which is unique up to a nonzero scalar, except that the "natural" space to define the eigenvectors is equivalence classes in the above defined quotient space: $[v] \in \ker P(\mu)/\ker_\mu P(x)$.

Definition 2.18 (Eigenvectors of singular matrix polynomials)**.** If μ is an eigenvalue of $P(x)$, we say that $[v] \in \ker P(\mu)/\ker_\mu P(x)$ is an eigenvector associated with the eigenvalue μ if $[v] \neq [0]$.

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 at μ such that $r(\mu) = v$, e.g., $r(x) = v$. Indeed, $[v] \neq [0]$ implies that $P(x)v \neq 0$, and hence, $v \in \ker P(\mu)$ implies that $P(x)v = (x-\mu)^\ell a(x)$ for some $\ell \in \mathbb{N}$ and $a(\mu) \neq 0$. Note also that Definition 2.18 implies $P(\mu)[v] = [0]$, which is a well defined expression since $\ker_\mu P(x)$ is an invariant subspace for $P(\mu)$, and which resembles the definition in the regular case. 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 $M(x) \in \mathbb{F}[x]^{n\times p}$ 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 $s > 1$ as we briefly sketch in this paragraph. Observe first that Definition 2.18 still makes sense (hence why we have not specified there that the geometric multiplicity of μ should be 1). Moreover, one can define eigenspaces as $\operatorname{span} B$ where B is an $n \times s$ matrix whose columns are linearly independent (over $\mathbb{F}$) equivalence classes in $\ker P(\mu)/\ker_\mu P(x)$. 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 $W = BA$ 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. These and other results will be carefully developed in the future.

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

Proposition 2.12 established the existence of root polynomials at any finite eigenvalue of the considered matrix polynomial. This result immediately implies the existence of μ-independent sets of root polynomials. However, the existence of complete and maximal sets of root polynomials has not been yet proved. The main purpose of this section is to prove in Theorem 3.5 that for any finite eigenvalue μ of any matrix polynomial $P(x)$ there exists a maximal set of root polynomials and to construct such a set in terms of the matrix polynomials regular (or invertible) at μ that appear in the transformation of $P(x)$ into its local Smith form at μ. Moreover, we will prove that the orders of the vector polynomials in the maximal set of root polynomials at μ that we construct are precisely the nonzero partial multiplicities of the eigenvalue μ. This result will be completed by proving that any maximal set of root polynomials at μ for $P(x)$ has the same orders in Theorem 4.2 of Section 4. Thus, we see that maximal sets of root polynomials convey the complete information about the spectral structure of the eigenvalue μ. To obtain these results a number of intermediate results, which are also interesting by themselves, are needed.

We start by showing that any matrix polynomial in local Smith form at μ admits a very simple 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 show 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, $\ell \le \ell_j$, leading to a contradiction. □

The results in Proposition 3.2 appeared in [19, 20] and describe how root polynomials at μ are transformed under multiplication by matrix polynomials that are regular at μ, i.e., they evaluate to an invertible matrix at $x = \mu$. These transformation rules combined with Theorem 3.1 will be key for arguing that a maximal set of root polynomials exists at any finite eigenvalue of any matrix polynomial and for constructing such a set. In the rest of the paper, 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 Proposition 3.2 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 to the contrary 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 $\ker_\mu P(x)$. □

Once we have found out how root polynomials at μ are transformed under multiplication by matrix polynomials regular at μ, the next step is to find out how maximal (resp., complete, μ-independent) sets of root polynomials at μ change under these multiplications. This is the goal of Theorem 3.4. Since the definitions of all these sets (recall Definitions 2.13, 2.15 and 2.17) involve a minimal basis of the considered matrix polynomial, we need to prove first the following lemma.

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 [12, 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 $\ell_1 \geq \dots \geq \ell_s > 0$, 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 $\ell_1 \geq \dots \geq \ell_s > 0$, 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 shown 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, which is the most important result in this section.

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. 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 $P(x)$. □

We stress that the proof of Theorem 3.5 is constructive and that given any factorization $P(x) = A(x)D(x)B(x)$, where $D(x)$ is the local Smith form at μ of $P(x)$ and $A(x)$ and $B(x)$ are matrix polynomials regular at μ, a maximal set of root polynomials can be explicitly constructed. Moreover, note that the local Smith form at μ and the matrices $A(x)$ and $B(x)$ can be easily obtained from the global Smith form of $P(x)$ and the corresponding unimodular matrices.

4. Extremality properties of maximal sets of root polynomials

Theorem 3.5 does not establish that the vector polynomials in *any* maximal set of root polynomials at the same eigenvalue μ of a matrix polynomial $P(x)$ have as orders the nonzero partial multiplicities of μ. In this section, we prove this fact in Theorem 4.2 after showing in Theorem 4.1 that maximal sets of root polynomials are the sets with largest orders (when they are ordered) among all the complete sets of root polynomials at μ. This is the property that motivates the name "maximal" and proves that maximal sets of root polynomials are the only ones revealing the complete eigenstructure of the eigenvalue μ.

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 $P(x)$ at μ 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 $\mu \Leftrightarrow \ell_i = \kappa_i$, $i = 1, \ldots, s$;
   3.3 $\widetilde{r}_1(x), \ldots, \widetilde{r}_s(x)$ is a maximal set of root polynomials for $P(x)$ at $\mu \Leftrightarrow \sum_{i=1}^{s} \ell_i = \sum_{i=1}^{s} \kappa_i$.

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 sets 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, $u_1, \ldots, u_p, r_1(\mu), \ldots, r_{j+1}(\mu) \in \text{span}\{u_1, \ldots, u_p, t_1(\mu), \ldots, t_j(\mu)\}$. 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 $u_1, \ldots, u_p, \widetilde{r}_1(\mu), \ldots, \widetilde{r}_j(\mu) \in \text{span}\{u_1, \ldots, u_p, r_1(\mu), \ldots, r_{j-1}(\mu)\}$, 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 $u_1, \ldots, u_p, \widetilde{r}_1(\mu), \widetilde{r}_j(\mu), \hat{r}(\mu) \in \text{span}\{u_1, \ldots, u_p, r_1(\mu), \ldots, r_j(\mu)\}$, and again we get to the contradicting conclusion that $p+j+1$ linearly independent vectors all lie in a subspace of dimension $p+j$.

3.3 Again, one implication is trivial. Now suppose that $\sum_{i=1}^{s} \ell_i = \sum_{i=1}^{s} \kappa_i$. 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.

□

As a corollary of Theorems 3.5 and 4.1, we deduce that the orders of the polynomial vectors of any 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 in root polynomials

It turns out that a root polynomial of order ℓ at an eigenvalue μ is in fact defined up to an additive term of the form $(x-\mu)^\ell w(x)$ where $w(x) \in \mathbb{F}[x]^n$ is arbitrary. More formally, we may state that the natural ring where the entries of 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 $p(x)$. This property is established in Proposition 5.1 and will allow us in Definition 5.4 to introduce root polynomials, and maximal sets of root polynomials, with minimal grades. Such root polynomials and maximal sets of root polynomials are optimal in the sense of not containing terms which do not provide any relevant spectral information.

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 $\hat{v}(x) := v(x) + (x-\mu)^\ell w(x)$, where $w(x) \in \mathbb{F}[x]^n$ is any vector polynomial, is a root polynomial at μ for $P(x)$ of order larger than or equal to ℓ.*

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 situations are possible, as illustrated in the next example.

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 at 0 but $r_2(x) + xw(x)$ is a root polynomial of order $2 > 1$ at 0.

We now turn to illustrating how the quotienting operation of Proposition 5.1 acts on a 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 $\ell_1 \geq \dots \geq \ell_s$, 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 root 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 root 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 root 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 $\{1, \dots, s\}$ 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 $\sum_{i=1}^s \kappa_{\sigma(i)} = \sum_{i=1}^s \kappa_i \leq \sum_{i=1}^s \ell_i \leq \sum_{i=1}^s \kappa_i$. Therefore, $\sum_{i=1}^s \kappa_i = \sum_{i=1}^s \ell_i$, implying by item 3.3 in Theorem 4.1 that $\{q_i(x)\}_{i=1}^s$ are a maximal set. The reverse implication can be proved analogously.

□

The results in this section suggest the forewarned definitions of root polynomials and maximal sets of root polynomials with minimal grade.

Definition 5.4. Let $P(x) \in \mathbb{F}[x]^{m \times n}$ and μ be one of its finite eigenvalues.

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

(a) 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 of root polynomials at μ and they satisfy

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

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 and in their applications. As a consequence it has received considerable attention in modern references on matrix polynomials, see, for instance, [5, 16, 17, 18, 19]. This motivates a definition and an analysis of root polynomials at ∞.

For completeness, we start by recalling the definition of eigenvalue at infinity of a matrix polynomial in terms of the "grade" and the corresponding reversal polynomial [8]. Let $P(x) = \sum_{i=0}^{g} P_i x^i \in \mathbb{F}[x]^{m \times n}$ have grade[3] [17, 19] g, and define the g-reversal polynomial of $P(x)$ as follows

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

Note that the operator rev_g is involutory, i.e., $\operatorname{rev}_g \operatorname{rev}_g P(x) \equiv P(x)$, whenever $\operatorname{rev}_g P(x)$ is considered with the same grade g as $P(x)$.

Definition 6.1 (Eigenvalues and multiplicities at infinity)**.** A matrix polynomial $P(x) \in \mathbb{F}[x]^{m \times n}$ with grade g has an eigenvalue at infinity if zero is an eigenvalue of $\operatorname{rev}_g P(x)$. The algebraic, geometric and partial multiplicities of the infinite eigenvalue of $P(x)$ are defined to be identical to the algebraic, geometric and partial multiplicities of the zero eigenvalue of $\operatorname{rev}_g P(x)$.

In order to deal with root polynomials at infinity for any possible grade g, we need to pay attention to the degrees of certain vector polynomials. The following simple and technical lemma will be useful for this purpose.

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

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

[3] 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.

Next, we define root polynomials at infinity for any possible grade.

Definition 6.3 (Root polynomials at infinity)**.** 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 $\operatorname{rev}_g P(x)$.

The next proposition characterizes root polynomials at infinity. Here, and in the rest of the paper, 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" [12] of $M(x)$. Note that the latter is the same as the "columnwise reversal" of $M(x)$ evaluated at $x = 0$ [22].

Proposition 6.4. *Let $P(x) \in \mathbb{F}[x]^{m\times n}$ have grade g. The polynomial vector $r(x)$ is a root polynomial of order ℓ at infinity for the pair $(P(x), g)$ 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 M_h for any minimal basis $M(x)$ of $P(x)$.

Proof. Suppose that $r(x)$ is a root polynomial of order ℓ at infinity for $(P(x), g)$. 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 [19, Sections 6.1–6.2], $r_{\deg r} \in \ker_\infty P(x) \Leftrightarrow \operatorname{rev}_{\deg r} r(0) \in \ker_0 \operatorname{rev}_g P(x)$. $\square$

Once individual root polynomials at infinity have been defined and characterized, we proceed to define different sets of root polynomials at infinity analogously as we did for finite eigenvalues. Note, however, that the definitions are somewhat different.

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

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

has full column rank.

Definition 6.6. Let $M(x)$ be a right minimal basis of $P(x) \in \mathbb{F}[x]^{m\times n}$ and assume that $P(x)$ has grade g. 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 the pair $(P(x), g)$ if they are ∞-independent and there does not exist any set of $s+1$ root polynomials at ∞ for $(P(x), g)$, 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 $P(x) \in \mathbb{F}[x]^{m\times n}$ and assume that $P(x)$ has grade g. 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 the pair $(P(x), g)$ if they are complete and their orders as root polynomials at infinity for $(P(x), g)$, say, $\ell_1 \geq \cdots \geq \ell_s > 0$, 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 $\hat{\rho}$ for $(P(x), g)$ 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. Let $P(x) \in \mathbb{F}[x]^{m\times n}$ have grade g. The polynomial vectors $r_1(x), \dots, r_s(x) \in \mathbb{F}[x]^n$ are a *minimaximal set of root polynomials at* ∞ for $(P(x), g)$ if they are maximal for $(P(x), g)$ and they satisfy

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

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

When a matrix polynomial has infinite eigenvalues, and root polynomials at infinity are brought into the picture, one can prove results completely analogous to those developed for finite eigenvalues (except that the notation gets more complicated). In particular, one can prove that the orders of a maximal set of root polynomials at ∞ for the pair $(P(x), g)$ are the nonzero partial multiplicities at ∞ of $P(x)$ considered as a matrix polynomial of grade g. We omit further details to keep the paper compact.

7. Behaviour of root polynomials under rational reparametrizations

The change of the properties of a matrix polynomial under rational transformations, or more precisely rational reparametrizations of its variable, is a classical topic in the theory of matrix polynomials. Such transformations are used very often to convert a problem into another with more favourable properties. Among these transformations, Möbius transformations are particularly important. Modern approaches to these topics and their applications, as well as a wealth of classical and new results can be found in [17, 19]. The importance of these transformations motivates us to study in this section the relationship between the root polynomials, and different sets of root polynomials, of the pair of matrix polynomials $P(x) \in \mathbb{F}[x]^{m\times n}$ and $Q(y) = [d(y)]^g P(x(y))$ where, as in the previous section, g is the grade of $P(x)$ and

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

for some coprime polynomials $n(y), d(y) \in \mathbb{F}[y]$. For brevity, these definitions for $P(x)$, g, $Q(y)$, $x(y)$, $n(y)$ and $d(y)$ will be used throughout this section without being explicitly referred to. Note that Möbius transformations, studied in [17], correspond to $n(y)$ and $d(y)$ both having grade 1.

Our first result shows how one root polynomial of $P(x)$ is related to another root polynomial of $Q(y)$.

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 $a(x)$, $a(\mu) \neq 0$,

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

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 $u_1(x), \ldots, u_p(x)$. Suppose further that $\deg u_i(x) = \beta_i$. It is shown in [19, 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 independent, complete and maximal 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 $i = 1, \ldots, s$,*

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

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

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

we have that:

1. *$r_1(x), \ldots, r_s(x)$ are a μ-independent set of root polynomials at μ for $P(x)$ if and only if $q_1(y), \ldots, q_s(y)$ are a λ-independent set of root polynomials at λ for $Q(y)$;*
2. *$r_1(x), \ldots, r_s(x)$ are a complete set of root polynomials at μ for $P(x)$ if and only if $q_1(y), \ldots, q_s(y)$ are a complete set of root polynomials at λ for $Q(y)$;*
3. *$r_1(x), \ldots, r_s(x)$ are a maximal set of root polynomials at μ for $P(x)$ if and only if $q_1(y), \ldots, q_s(y)$ are a maximal set of root polynomials at λ for $Q(y)$.*

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

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

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 $\dim\ker Q(\lambda) > s+t$, where t is the number of columns of $M(\mu)$ or, equivalently, of $N(\lambda)$. This is a contradiction, because by definition of $Q(y)$, $\dim\ker P(\mu) = \dim\ker Q(\lambda) = s+t$. The converse statement can be shown similarly.

Finally, item 3 is a consequence of [19, Theorem 4.1] and of Theorem 4.2. □

Analogous results to those in Theorems 7.1 and 7.2 hold for the cases $\mu = \infty$ and λ finite, μ finite and $\lambda = \infty$, or $\mu = \lambda = \infty$. They can be proved using a technique analogous to the strategy employed in [19] to deal with infinite elementary divisors. In order to keep the paper concise, we omit the details of the cases involving infinite eigenvalues.

8. Linearizations and recovery properties

Linearizations are a classical tool in the theory of matrix polynomials [13, Sections 1.1 and 7.2]. Moreover, they have attracted considerable attention in the last fifteen years since they are fundamental in the numerical solution of polynomial eigenvalue problems. Thus, since the reference [16] was published, many other papers on linearizations of matrix polynomials have appeared in the literature. As a very small sample, we mention here [6, 10, 18, 21], where the reader can find many other references on this topic. The definition of linearization of a matrix polynomial immediately implies that the linearization and the matrix polynomial have the same eigenvalues with the same partial multiplicities. In contrast, the definition of linearization does not say how to recover any other interesting information of the polynomial from the linearization, as, for instance, eigenvectors, minimal indices or minimal bases. This has motivated a lot of research on the development of recovery procedures for different magnitudes and for different classes of linearizations. However, recovery procedures for Jordan chains of regular matrix polynomials [13, Chapter 1] have not been developed so far for any family of linearizations of matrix polynomials, with the exception of the first and second Frobenius companion linearizations (for which recovery procedures of Jordan chains can be found implicitly in [13, Chapter 1]).

The goal of this section is to provide very simple rules for recovering the root polynomials of a matrix polynomial from those of some of its linearizations. In addition, we discuss how these recovery rules can be used to recover the Jordan chains of a regular matrix polynomial from those of its linearizations. Since there are many families of linearizations of matrix polynomials, we limit ourselves, for brevity, to study three of the families that have received more attention in the literature: Fiedler linearizations [6], linearizations in the spaces $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ [16], and block Kronecker linearizations [10]. However, we emphasize that the tools and strategies that we develop can be extended to many other families of linearizations, though the precise details will depend on each particular family. The alert reader will note that the recovery rules we will obtain for the families of linearizations mentioned above for root polynomials are exactly the same as those for eigenvectors (in the case of regular matrix polynomials) and minimal bases (in the case of singular matrix polynomials). Finally, we warn the reader that in order to keep the paper concise, this section is not self contained and that the definitions and properties of the considered linearizations are not all included here. They have to be found in the original references.

For completeness, we recall the definition of linearization of a matrix polynomial [13, 8].

Definition 8.1 (Linearization)**.** A matrix polynomial $L(x)$ of degree at most 1 is called a *linearization* for $P(x) \in \mathbb{F}[x]^{m\times n}$ if 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 technical tools that we are going to use throughout this section. The first of these tools relates the minimal bases of $P(x)$ and $\operatorname{diag}(I_k, P(x))$, as well as the subspaces introduced in Definition 2.7 for these two matrix polynomials. Recall that such subspaces are used in the definition of root polynomials.

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 $\mu \in \mathbb{F}$, $\ker_\mu Q(x) = \{\begin{bmatrix} 0 \\ v \end{bmatrix} \; : \; v \in \ker_\mu P(x)\}$.*

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 [12], 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. $\square$

Our second tool relates root polynomials of $P(x)$ and $\operatorname{diag}(I_k, P(x))$. Observe that the relation is not as direct as the one obtained for minimal bases in Lemma 8.2.

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 $\hat{a}(\mu) \neq 0$, 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. Item 1 follows immediately from the last equation above, which also implies $P(x)\widetilde{w}(x) = (x-\mu)^\ell \widetilde{a}(x)$. In addition, Lemma 8.2 and item 1 imply that $\widetilde{w}(\mu) \notin \ker_\mu P(x)$ and, so, $\widetilde{a}(x) \neq 0$, since otherwise $\widetilde{w}(x) \in \ker P(x)$ and $\widetilde{w}(\mu) \in \ker_\mu P(x)$, which is a contradiction. This proves item 2. Finally, either $\hat{a}(\mu) \neq 0$, or $\widetilde{a}(\mu) \neq 0$, or both. This completes the result. □

Our last technical tool relates maximal sets of root polynomials of $P(x)$ and $\operatorname{diag}(I_k, P(x))$.

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. *$\widetilde{r}_1(x), \ldots, \widetilde{r}_s(x)$ are 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 $\widetilde{r}_1(x), \ldots, \widetilde{r}_s(x)$ are a set of root polynomials at μ for $P(x)$ of orders $\widetilde{\ell}_i \geq \ell_i$. They are μ-independent, by applying item 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 item 3.1 in Theorem 4.1, by Theorem 4.2 and by the above, we get
$$\sum_{i=1}^{s} \widetilde{\ell}_i \leq \sum_{i=1}^{s} \ell_i \leq \sum_{i=1}^{s} \widetilde{\ell}_i,$$
and hence, equality holds. By items 3.2–3.3 in Theorem 4.1, $\widetilde{r}_1(x), \ldots, \widetilde{r}_s(x)$ are a maximal set of root polynomials at μ for $P(x)$, and $\widetilde{\ell}_i = \ell_i$ for all $i = 1, \ldots, s$.

□

With the previous technical tools at hand, we study first the important class of linearizations known as "Fiedler pencils", see, e.g., [6] and the references cited therein. This approach to linearizing matrix polynomials has originated in the paper [11], and has been generalized in different directions, such as to rectangular matrix polynomials [7], to a wider class of pencils [1, 2] as well as to matrix polynomials expressed in nonmonomial bases [21]. The recent reference [3] presents a unified approach to many families of Fiedler-like pencils in terms of permutations and the block minimal bases pencils introduced in [10]. To keep the paper compact, here we will focus on the original class of Fiedler pencils for matrix polynomials expressed in the monomial basis and for square matrix polynomials, defined as in [6]. 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 $F_\sigma(x)$ [6, 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 [6] 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 (resp. minimaximal) 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 (resp. minimaximal) set of root polynomials at μ for $P(x)$, having orders $\ell_1 \geq \cdots \geq \ell_s$.*

Proof. By [6, Corollary 4.7], two unimodular matrix polynomials $U(x)$, $V(x)$ such that

$$U(x)F_\sigma(x)V(x) = \begin{bmatrix} I_{(g-1)n} & 0 \\ 0 & P(x) \end{bmatrix} =: Q(x) \tag{2}$$

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 $(g - c_1)$ [6, Remark 5.4].

Assume first that $r_1(x), \ldots, r_s(x)$ is a maximal set of root polynomials at μ for $F_\sigma(x)$. 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 $P(x)$ 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 $b_j(x)$ it holds

$$\widetilde{r}_j(x) = \hat{r}_j(x) + (x - \mu)^{\ell_j} b_j(x), \tag{3}$$

for $j = 1, \ldots, s$, which, by Theorem 5.3, implies that $\widetilde{r}_1(x), \ldots, \widetilde{r}_s(x)$ is a maximal set of root polynomials at μ for $P(x)$ with orders $\ell_1 \geq \cdots \geq \ell_s$. Once the result for maximal sets of root polynomials has been proved, the trivial inequality $\deg \widetilde{r}_j(x) \leq \deg r_j(x)$ for $j = 1, \ldots, s$, together with Definition 5.4 prove the result for minimaximal sets of root polynomials. □

Remark 8.6.

- Theorem 8.5 holds, in particular, for the first Frobenius companion linearization (taking $c_1 = 0$) and for the second Frobenius companion linearization (taking $c_1 = g - 1$) [6].
- The result implies that the orders of $r_j(x)$ and $\widetilde{r}_j(x)$ are equal, which is consistent with the well-known fact that the partial multiplicities at μ of $P(x)$ and $F_\sigma(x)$ are equal and with Theorem 4.2.

Next we discuss the recovery of the Jordan chains of the matrix polynomial from those of their Fiedler linearizations.

Remark 8.7 (Recovery of Jordan Chains)**.** For regular matrix polynomials, the recovery result in Theorem 8.5 for root polynomials yields a recovery result for Jordan chains due to the relationship between the coefficients of root polynomials and Jordan chains presented in [13, Sect. 1.5]. More precisely, if $\widetilde{r}_i(x)$ is developed in powers of $(x-\mu)$ and the vector coefficients of $1,(x-\mu),\ldots,(x-\mu)^{\ell_i-1}$ are taken for $i=1,\ldots,s$, then one obtains a canonical set of Jordan chains of $P(x)$ at μ in the sense of [13, Sect. 1.6]. As far as we know, this is the first result about the recovery of Jordan chains of matrix polynomials from those of their linearizations available in the literature. This remark can be also applied to the rest of the linearizations studied in this section.

A second important class of linearizations is given by the $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ linearization spaces: see [16, Definition 3.1]. Again, for simplicity we focus on the case of matrix polynomials expressed in the monomial bases originally discussed in [16]. We note however that an extension to nonmonomial bases is possible [18], and recovery properties for root polynomials can be derived for other bases as well.

The next proposition follows from, and slightly improves, [5, Theorems 4.1 and 4.6]. It is a technical result that allows us to prove very easily Theorem 8.10, which is the main recovery result of root polynomials from linearizations in the spaces $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$.

Proposition 8.8. *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 exist 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

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

denotes the first Frobenius 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 $w^TK=e_1^T$. Then, there exist 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

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

is the second Frobenius 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 [16, 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 8.9. The property of Z being nonsingular in the statement of Proposition 8.8 is known as $L(x)$ having *full Z-rank* [5].

In Theorem 8.10, we state and prove the announced recovery result of root polynomials from linearizations in the spaces $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$.

Theorem 8.10. *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 maximal (resp. 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* g*th block of* $r_j(x)$*. Then,* $\widetilde{r}_1(x), \dots, \widetilde{r}_s(x)$ *is a maximal (resp. minimaximal) set of root polynomials at* μ *for* $P(x)$*, having orders* $\ell_1 \geq \dots \geq \ell_s$*.*
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 maximal (resp. minimaximal) set of root polynomials at* μ *for* $L(x)$*, having orders* $\ell_1 \geq \dots \geq \ell_s$*. Moreover, define* $W = w^T \otimes I_n$ *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 maximal (resp. minimaximal) set of root polynomials at* μ *for* $P(x)$*, having orders* $\ell_1 \geq \dots \geq \ell_s$*.*

Proof.

1. Let $L(x) \in \mathbb{L}_1(P)$ have full Z-rank. By Theorem 3.4, Proposition 8.8, and Definition 5.4, we can easily see that $\{r_i(x)\}_{i=1}^s$ is a maximal (resp. 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 8.8, 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 maximal (resp. 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 $\begin{bmatrix} I_n & 0 \\ -X & -Z \end{bmatrix} (K^{-1} \otimes I_n) r_j(x)$, $j = 1, \ldots, s$, are a maximal (resp. 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 $w^T K = e_1^T$ that holds by definition of K (see the statement of Proposition 8.8).

□

We conclude this section by analyzing a third important class of linearizations discussed in [10]: block Kronecker linearizations. These linearizations include, among many others, Fiedler linearizations modulo permutations and, under mild assumptions, have excellent properties from the point of view of backward errors when they are used for solving numerically polynomial eigenvalue problems. 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 8.11. *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 [10, Definitions 3.1 and 4.1] and a strong linearization of $P(x)$, where $\eta + \epsilon + 1 = g$. Suppose further that $K_1(x)$ has ϵn rows and $K_2(x)$ has ηn rows, where ϵ and η are defined as in [10, Section 4]. 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 (resp. minimaximal) set of root polynomials at μ for $L(x)$, having orders $\ell_1 \geq \cdots \geq \ell_s$. For all $j = 1, \ldots, s$ denote by $\widetilde{r}_j(x)$ the $(\epsilon + 1)$th block of $r_j(x)$. Then, $\widetilde{r}_1(x), \ldots, \widetilde{r}_s(x)$ is a maximal (resp. minimaximal) set of root polynomials at μ for $P(x)$, having orders $\ell_1 \geq \cdots \geq \ell_s$.

Proof. Using [10, Lemma 2.13] and [10, Remark 4.3], 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

Dual pencils were introduced, under the name consistent pencils, in the pioneering work by Kublanovskaya [14] and they were further studied in the long survey [15] (as well as in some references therein written in Russian). A modern approach to dual pencils can be found in [22], which includes detailed and rigorous proofs of results stated in [15] without proofs, as well as many other new results. Kublanovskaya introduced dual pencils as an additional tool that allows extra flexibility over the standard tool of performing strict equivalences on a linearization in the design of algorithms for solving numerically spectral problems related to matrix polynomials. The key property is that two dual pencils have the same eigenvalues while other magnitudes, as minimal indices for instance, are not equal but can be easily related to each other (see [15, pp. 3097-3098] or [22, Theorem 3.3] for a more precise statement). The goal of this section is to study the relationship between the root polynomials of two dual pencils. As in the previous section, we warn the reader that, for brevity, this section is not self contained and that we rely on frequent references to result in [22].

The following definitions and basic results appear in [22] (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 work in [14, 15].

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 $L(x)$.

Among all right dual pencils of a given pencil, column-minimal right dual pencils have particularly nice properties. Thus, we recall their definition.

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. *[22, Lemma 3.1] The pencil $R(x) = R_1 x + R_0$ is a column-minimal right dual of the pencil $L(x) = L_1 x + L_0$ 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 minimal (according to Definition 5.4) 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 minimal root polynomial of order ℓ at μ for $R(x)$, then $w(x) := (\gamma R_1 + \delta R_0)r(x)$ is a minimal 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 minimal 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 minimal root polynomial of order ℓ at μ for $R(x)$.*

Proof. 1. Suppose that $R(x)r(x) = (x-\mu)^\ell a$, $a = R_1r_{\ell-1} \neq 0$. 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 $(\gamma - \delta\mu)R_1r(\mu) = (\gamma R_1 + \delta R_0)r(\mu) = N(\mu)c$ for some constant vector c, where $N(x)$ is a minimal basis for $L(x)$. It is known [22, 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 $M(\mu)c = (\gamma - \delta\mu)(Q_0 + \mu Q_1)R_1r(\mu) = (\gamma - \delta\mu)(r(\mu) + Q_1R(\mu)r(\mu)) = (\gamma - \delta\mu)r(\mu)$, 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 $L(x)r(x) = (x-\mu)^\ell a$, $a = L_1r_{\ell-1} \neq 0$. Expanding $L(x)$ and $r(x)$ in a power series in $(x-\mu)$, this is equivalent to

$$\begin{aligned} L_1\mu r_0 + L_0r_0 &= 0, \\ L_1r_0 + L_1\mu r_1 + L_0r_1 &= 0, \\ &\vdots \\ L_1r_{\ell-2} + L_1\mu r_{\ell-1} + L_0r_{\ell-1} &= 0. \end{aligned}$$

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

$$\begin{bmatrix} r_0 & r_1 & \dots & r_{\ell-1} \\ -\mu r_0 & -r_0 - \mu r_1 & \dots & -r_{\ell-2} - \mu r_{\ell-1} \end{bmatrix} = \begin{bmatrix} R_1 \\ R_0 \end{bmatrix} \begin{bmatrix} w_0 & w_1 & \dots & w_{\ell-1} \end{bmatrix} \tag{4}$$

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 $r(x) = R_1w(x)$ and $(x-\mu)^\ell r_{\ell-1} - xr(x) = R_0w(x)$. 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 (4) by $\begin{bmatrix} Q_0 & -Q_1 \end{bmatrix}$, we see that

$$\begin{aligned} Q_1\mu r_0 + Q_0r_0 &= w_0, \\ Q_1r_0 + Q_1\mu r_1 + Q_0r_1 &= w_1, \\ &\vdots \\ Q_1r_{\ell-2} + Q_1\mu r_{\ell-1} + Q_0r_{\ell-1} &= w_{\ell-1}, \end{aligned}$$

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 $L(x)$ [22, Theorem 3.8], this violates the assumption that $r(x)$ is a root polynomial.

□

A natural question is whether or not Theorem 9.4 can be extended to root polynomials that are not minimal. The next example shows that the answer is negative.

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 Theorem 9.4. 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 strictly larger than ℓ if $r(x)$ has degree higher than $\ell - 1$.

Our next result shows how minimaximal sets of root polynomials change under the operation of duality for dual pairs $L(x), R(x)$ where the right dual $R(x)$ is column-minimal.

Theorem 9.6. *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 $r_1(x), \dots, r_s(x)$ is a minimaximal set of root polynomials at μ for $R(x)$ having orders $\ell_1 \geq \dots \geq \ell_s$, 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 $L(x)$ having orders $\ell_1 \geq \dots \geq \ell_s$.*
2. *If $r_1(x), \dots, r_s(x)$ is a minimaximal set of root polynomials at μ for $L(x)$ having orders $\ell_1 \geq \dots \geq \ell_s$, then $w_j(x) := Q(x)r_j(x) - (x-\mu)^{\ell_j} Q_1 r_{j,\ell_j-1}$, $j = 1, \dots, s$, is a minimaximal set of root polynomials at μ for $R(x)$ having orders $\ell_1 \geq \dots \geq \ell_s$.*

Proof. 1. The equality of the orders follows from Theorem 9.4. Let $M(x)$ be a minimal basis for $R(x)$. We first prove μ-independence of $w_1(x), \dots, w_s(x)$. 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 [22, 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 $R(x)Xc = \Big(\frac{\delta x-\gamma}{\delta\mu-\gamma}R(\mu) + \frac{x-\mu}{\gamma-\delta\mu}(\gamma R_1 + \delta R_0)\Big)Xc = 0$, 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 [22, 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. Again, the equality of the orders follows from Theorem 9.4. Denote by $M(x)$ a minimal basis of $R(x)$. We start by showing μ-independence of $w_1(x), \dots, w_s(x)$. 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 [22, 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, $Y = (\gamma R_1+\delta R_0)Xc = 0$ is 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 $\ker L(\mu)$, by [22, 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-minimal, 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 $L_1\hat{R}_0 = L_1R_0B^R = L_0R_1B^R = L_0\hat{R}_1$ 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 [9, 12, 22] 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 [22, 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 [22, 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 relates minimal root polynomials for the pencils $R(x)$ and $\hat{R}(x)$ appearing in Lemma 9.7.

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 presents the announced extension of Theorems 9.4 and 9.6 to the case of a right dual which is not column-minimal.

Corollary 9.9. *Let $L(x) = L_1x + L_0$ and let $R(x) = R_1x + R_0$ 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 $Q(x) = Q_1x + Q_0$ for some Q_1, Q_0 satisfying $Q_0R_1 - Q_1R_0 = B^RB$. 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 $w(x) := B^RBQ(x)r(x) - (x-\mu)^\ell B^RBQ_1r_{\ell-1}$ is a root polynomial of order ℓ at μ for $R(x)$.*
3. *If $r_1(x), \dots, r_s(x)$ is a minimaximal set of root polynomials at μ for $R(x)$ having orders $\ell_1 \geq \cdots \geq \ell_s$, 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 $L(x)$ having orders $\ell_1 \geq \cdots \geq \ell_s$.*
4. *If $r_1(x), \cdots, r_s(x)$ is a minimaximal set of root polynomials at μ for $L(x)$ having orders $\ell_1 \geq \cdots \geq \ell_s$, then $w_j(x) := B^RBQ(x)r_j(x) - (x-\mu)^{\ell_j}B^RBQ_1r_{j,\ell_j-1}$, $j = 1, \dots, s$, is a minimaximal set of root polynomials at μ for $R(x)$ having orders $\ell_1 \geq \cdots \geq \ell_s$.*

Proof. It is a corollary of Theorems 9.4, 9.6 and 9.8, see also [22, Theorem 3.9] and [22, Remark 3.10]. □

The results in this section allow to strengthen in Theorem 9.10 the recovery result of root polynomials from pencils in the space $\mathbb{L}_1(P)$ presented in Theorem 8.10. Observe that the next result does not require the assumption that $L(x)$ has full Z-rank.

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 [22, 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 $\ell_1 \geq \cdots \geq \ell_s$. 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 $\ell_1 \geq \cdots \geq \ell_s$.

Proof. The pencil $D(x) = D_1x + D_0$ 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 $1 \le j \le s$, $w_j(x) = Q(x)r_j(x) - Q_1(x-\mu)^{\ell_j} r_{j,\ell_j-1}$, $Q(x) = Q_1x + Q_0$ and $Q_0D_1 - Q_1D_0 = I$. 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 $\{D_1w_i(x)\}_{i=1}^s$.

Note that by the proof of Theorem 9.4 we have, for all $j = 1, \dots, s$, $D_1w_j(x) = r_j(x)$, 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

In the first part of this paper, we have studied in depth the concept of root polynomials of a matrix polynomial and we have seen that root polynomials are particularly useful in the study of spectral properties of matrix polynomials when they form sets with special features. Thus, we have defined μ-independent, complete and maximal sets of root polynomials and we have proved a number of interesting properties for maximal sets of root polynomials. For instance that the orders of the root polynomials in any maximal set of root polynomials at an eigenvalue μ are precisely the nonzero partial multiplicities of μ, and that maximal sets of root polynomials at μ can be constructed from the matrices that transform a matrix polynomial into its local Smith form at μ. The extremality properties that motivate the name *maximal set of root polynomials* have been explicitly established for the first time in the literature. Moreover *minimaximal* sets of root polynomials have been defined as those maximal sets whose elements have minimal grade, based on the property that a root polynomial is defined only up to the addition of certain arbitrary terms of degree arbitrarily larger than its order. Root polynomials and sets of root polynomials at infinity have been also defined and their properties studied. We believe that all these results should highlight the importance of *maximal and minimaximal sets of root polynomials* in the theory of matrix polynomials.

Based on the properties developed in the first part of the paper as well as on results previously available in the literature, the second part of the paper studies the interaction of root polynomials with rational transformations of matrix polynomials, with linearizarizations of matrix polynomials and with dual pencils. Thus: (1) the change of root polynomials and sets of root polynomials under rational transformations has been precisely described; (2) the recovery rules for maximal sets of root polynomials of a matrix polynomial from those of its Fiedler, $\mathbb{L}_1(P)$, $\mathbb{L}_2(P)$, and block Kronecker linearizations have been proved to be very simple and equal to those of eigenvectors in the case of regular matrix polynomials; and (3) the relationship between the root polynomials of two dual pencils has been established.

Finally, we mention that as a particularly meaningful application of root polynomials, we have shown how eigenvectors and eigenspaces can be consistently defined for singular matrix polynomials, as subspaces of certain quotient spaces. These results on eigenvectors and eigenspaces are still preliminary and will be further developed in the future.

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] M. I. Bueno, F. M. Dopico, J. Pérez, R. Saavedra, and B. Zykoski, *A simplified approach to Fiedler-like pencils via block minimal bases pencils*, Linear Algebra Appl. 547, 45–104, 2018.

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

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

[6] 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.

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

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

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

[10] F. M. Dopico, P. W. Lawrence, J. Pérez, and P. Van Dooren, *Block Kronecker linearizations of matrix polynomials and their backward errors*, Numer. Math. 140, 373–426, 2018.

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

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

[13] 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).

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

[15] 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.

[16] 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.

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

[18] 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.

[19] V. Noferini, *The behaviour of the complete eigenstructure of a polynomial matrix under a generic rational transformation*, Electron. J. Linear Algebra 22, 607–624, 2012.

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

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

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

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