Duality of matrix pencils and linearizations

Noferini, Vanni and Poloni, Federico

2013

MIMS EPrint: **2013.17**

Manchester Institute for Mathematical Sciences
School of Mathematics

The University of Manchester

ISSN 1749-9097

Duality of matrix pencils and linearizations

Vanni Noferini* Federico Poloni†

November 19, 2013

In this paper we consider a duality relation on matrix pencils and show that it is a useful tool in the theory of linearizations of matrix polynomials. Exploiting a result that completely characterizes the Kronecker form of dual pencils, we study the behaviour under duality of the spectral structures, including eigenvalues, eigenvectors, Wong chains, and minimal bases. We also present several new applications of this concept, including: constraints on the minimal indices of singular Hamiltonian and symplectic pencils, new sufficient conditions under which pencils in $\mathbb{L}_1$, $\mathbb{L}_2$ linearization spaces are strong linearizations, a new perspective on Fiedler pencils, a link between the Möller-Stetter theorem and some linearizations of matrix polynomials.

Keywords: matrix pencil, Wong chain, linearization, matrix polynomial, singular pencil, Fiedler pencil, pencil duality, Kronecker canonical form

MSC classification: 15A18 15A22

1 Introduction

Consider a pair of matrices of the same size $L_0, L_1 \in \mathbb{C}^{m\times n}$. In matrix theory, a degree-1 polynomial $L_0 + xL_1 \in \mathbb{C}[x]^{m\times n}$ is known as a *matrix pencil* [16]. In this paper we study a concept of duality among matrix pencils, defined as follows. Two pencils $L(x) := L_1x + L_0 \in \mathbb{C}[x]^{m\times n}$ and $R(x) := R_1x + R_0 \in \mathbb{C}[x]^{n\times p}$ are *dual* if the following two conditions hold:

1. $L_1R_0 = L_0R_1$,
2. $\operatorname{rank}\begin{bmatrix} L_1 & L_0 \end{bmatrix} + \operatorname{rank}\begin{bmatrix} R_1 \\ R_0 \end{bmatrix} = 2n.$

*School of Mathematics, University of Manchester, Manchester, England, M13 9PL (vanni.noferini@manchester.ac.uk). Supported by ERC Advanced Grant MATFUN (267526).

†Dipartimento di Informatica, Università di Pisa, Italy (fpoloni@di.unipi.it). Supported in part by the A. von Humboldt Foundation and by INDAM (Istituto Nazionale di Alta Matematica).

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

We emphasize that this paper is not the first study of duality. Dual pencils appear in [22, Section 1.3], where they are given the name of consistent pencils and some of their theoretical properties are stated. Moreover, one can recognize the use of duality (of regular pencils only, which is a less interesting case) in the study of doubling and inverse-free methods [2, 9, 27], as well as in the work [3], which gives an elegant algebraic theory of operations on matrix pencils.

Yet, this technique seems to be underused with respect to its potential and we would like to bring it to the attention of the matrix pencil community. We will argue that it is an elegant tool for the theoretical study of matrix pencils, that allows us to obtain new results and revisit old ones, greatly simplifying the treatment of singular cases.

The structure of the paper is the following: in Section 2, we recall some basic definitions and classical results on matrix pencils and matrix polynomials. In Sections 3 and 4, we state some theoretical results describing how Kronecker canonical form, eigenvectors and minimal bases change under duality. We then show how duality can be used for several tasks in different applications:

- describing the possible Kronecker forms of singular symplectic and Hamiltonian pencils (Section 5);
- revisiting and simplifying proofs about the spectral properties of square (possibly singular) Fiedler pencils (Section 7);
- developing a connection between duality and the vector spaces of linearizations $\mathbb{L}_1$ and $\mathbb{L}_2$ introduced in [25], obtaining new insight for the singular case (Section 8);
- illustrating a connection between the Möller-Stetter theorem and some specific linearization of a matrix polynomial (Section 9).

We conclude the paper by describing two different methods that can be used to compute duals, and showing how they can be combined with the theory presented here to derive old and new linearizations (Section 10).

Most of the theory developed in this paper is applicable to any field $\mathbb{F}$. If the field is not closed, eigenvalues are sought in its algebraic closure. For simplicity, however, our exposition is for $\mathbb{F} = \mathbb{C}$.

2 Preliminaries on matrix pencils and polynomials

In this section, we recall some classical definitions and results on matrix pencils and polynomials. Throughout the paper, the ring of scalar polynomials with coefficients in $\mathbb{C}$ is denoted by $\mathbb{C}[x]$, and the set of those with degree not larger than d by $\mathbb{C}[x]_d$. The notation We denote by $R^{m\times n}$ the set of $m \times n$ matrices with coefficients in R. The dimensions m and n are allowed to be zero [7, page 90].

We denote by $\operatorname{diag}(M_1, M_2, \ldots, M_m)$ the block diagonal matrix formed by concatenating diagonally the (not necessarily square) blocks $M_1, M_2, \ldots, M_m$. We introduce the notation

$$K_{k,k+1}(x) := \begin{bmatrix} 0_{k\times 1} & I_k \end{bmatrix} x - \begin{bmatrix} I_k & 0_{k\times 1} \end{bmatrix}, \quad K_{k+1,k}(x) := K_{k,k+1}(x)^T,$$

where k is allowed to be zero (giving 0×1 and 1×0 blocks, respectively). Moreover, let $J_k^{(\lambda)}$ denote a Jordan block with size k and eigenvalue λ.

The following result about matrix pencils is classical [16, Chapter 12], and reduces to the Jordan canonical form of a matrix when one considers monic square pencils.

Theorem 2.1 (Kronecker canonical form)**.** *For every matrix pencil $L(x) \in \mathbb{C}[x]_1^{m\times n}$, there exist nonsingular matrices $U \in \mathbb{C}^{m\times m}, V \in \mathbb{C}^{n\times n}$ such that $B(x) = UL(x)V$ has the form $B(x) = \operatorname{diag}(B_1(x), B_2(x), \ldots, B_t(x))$, where each block $B_i(x)$ is one among:*

1. *$xI - J_{k_i}^{(\lambda)}$ (Jordan block of size k_i),*
2. *$xJ_{k_i}^{(0)} - I$ (Jordan block at infinity of size k_i),*
3. *$K_{k_i,k_i+1}(x)$ (right singular block of size $k_i \times (k_i + 1)$),*
4. *$K_{k_i+1,k_i}(x)$ (left singular block of size $(k_i + 1) \times k_i$).*

The pencil $B(x)$ is unique up to a permutation of the diagonal blocks. Therefore, the number of blocks of each kind, size and eigenvalue is an invariant of the pencil $L(x)$.

For the sake of brevity, throughout the paper we will often use the acronym KCF.

It is straightforward to generalize the concept of a matrix pencil to polynomials of higher degree. This leads to the definition of *matrix polynomials* [19], $A(x) := \sum_{i=0}^{d} A_i x^i \in \mathbb{C}[x]^{m\times n}$. A matrix polynomial is called *regular* if it is square and $\det A(x)$ is not the zero polynomial, and *singular* otherwise.

Sometimes, in the theory of matrix polynomials it is convenient to allow for a zero leading coefficient (see, e.g., [26]). For this reason, in our exposition we will not exclude this possibility. When we write about a matrix polynomial $A(x) = \sum_{i=0}^{d} A_i x^i$, we agree that the leading factor could be the zero matrix. The natural number d is therefore an arbitrarily fixed *grade*, equal to or larger than the degree, which is attached artificially to the polynomial [26]. However, in most applications the leading coefficient is nonzero: a reader uncomfortable with the concept of grade may simply think of d as the degree.

A finite eigenvalue of a matrix polynomial $A(x)$ is defined as a complex number λ such that the rank of $A(\lambda)$ as a matrix over the field $\mathbb{C}$ is lower than the rank of $A(x)$ as a matrix over the field of rational functions $\mathbb{C}(x)$. Infinite eigenvalues can be defined as zero eigenvalues of $\operatorname{rev} A(x)$, where the operator rev is defined by

$$\operatorname{rev} \sum_{i=0}^{d} A_i x^i := \sum_{i=0}^{d} A_{d-i} x^i.$$

Furthermore, the Jordan invariants can be extended to the polynomial case, resulting in the concepts of *elementary divisors* and *partial multiplicities*; we refer the reader to the classic books [16, 19] for their definitions, which are not needed in detail here.

If $A(x)$ is an $m \times n$ matrix polynomial, then $\ker_{\mathbb{C}(x)} A(x)$ is a subspace of $\mathbb{C}(x)^n$, and it always has a polynomial basis, i.e., a basis composed by vectors $v^{(k)} \in \mathbb{C}[x]^n$. The *degree* of a vector polynomial $v(x) = [v_i(x)] \in \mathbb{C}[x]^n$ is defined [15, Definition 1] as $\max_{i=1}^n \deg v_i(x)$. A *minimal basis* of $A(x)$ [15, Definition 3] is a basis for the subspace $\ker_{\mathbb{C}(x)} A(x)$ composed entirely of vector polynomials such that the sum of the degrees of its column vectors, known as the *order* [15, Definition 2] of the basis, is minimal among all possible polynomial bases. The degrees of the vectors that form a minimal basis, known as *(right) minimal indices*, are uniquely defined independently of the choice of the basis. It is known that minimal bases transform well under multiplication by invertible constant matrices; we give a simple proof in the next Lemma.

Lemma 2.2. *Let $A(x) \in \mathbb{C}[x]^{m\times n}$, $U \in GL_m(\mathbb{C})$ and $V \in GL_n(\mathbb{C})$. If $M(x)$ is a minimal basis for $UA(x)V$, then $VM(x)$ is a minimal basis for $A(x)$, and has the same minimal indices.*

Proof. We recall that the *high order coefficient matrix* [15] of $M(x) \in \mathbb{C}[x]^{n\times p}$ is denoted by $[M]_h$ and is defined as the matrix whose jth column consists of the coefficient of $x^{\deg M_j}$ in M_j, where M_j is the jth column of $M(x)$.

Evidently, $VM(x)$ is a basis of $\ker_{\mathbb{C}(x)} A(x)$, so it suffices to prove minimality and preservation of minimal indices. By [15, Main Theorem], $M(x)$ is minimal if and only if (a) $M(x) \mod p(x)$ has full column rank for all irreducible $p(x) \in \mathbb{C}[x]$ and (b) $[M]_h$ has full column rank. Since $M(x)$ is minimal, $M(x) \mod (x - x_0) = M(x_0)$ has full column rank for all $x_0 \in \mathbb{C}$. Hence $VM(x_0) = VM(x) \mod (x - x_0)$ has full column rank as well. Furthermore, denoting by M_j (resp., VM_j) the jth column of $M(x)$ (resp., $VM(x)$), the two relations $\deg VM_j = \deg M_j$ and $[VM_j]_h = V[M_j]_h$ (hence $[VM]_h = V[M]_h$) can be verified directly by expanding the polynomial vectors M_j and VM_j in powers of x.

We conclude that $VM(x)$ is minimal and has the same minimal indices of $M(x)$. $\square$

A simple consequence of Lemma 2.2 is that one can determine the minimal basis of a pencil from its KCF; indeed, it is easy to verify that a minimal basis for the pencil $B(x)$ described in Theorem 2.1 is $M(x) = \mathrm{diag}(M_1(x), M_2(x), \ldots, M_t(x))$, where $M_i(x)$ is equal to

$$\begin{bmatrix} x^{k_i} & x^{k_i-1} & \cdots & x & 1 \end{bmatrix}^T \tag{1}$$

if the block $M_i(x)$ if of the form $K_{k_i,k_i+1}(x)$, and the empty vector in $\mathbb{C}^{k\times 0}$ otherwise. Hence, $VM(x)$ is a minimal basis for $L(x)$ and its right minimal indices coincide with the row sizes k_i of the right singular blocks in its KCF.

Similarly, one can define left minimal indices as degrees of a minimal polynomial basis for the left kernel of $A(x)$, and for a pencil they coincide with the column sizes k_i of the $K_{k_i+1,k_i}(x)$ Kronecker blocks.

Let us define the *block transpose* $A^{\mathcal{B}}$ of a matrix A partitioned in blocks A_{ij} as the block matrix whose blocks are A_{ji}. Clearly, this definition depends on the choice of the

block sizes, which should be clear from the context — in this paper, this means most often $m \times n$ blocks. Moreover, given the matrix polynomial $A(x) = \sum_{i=0}^{d} A_i x^i$, we set

$$\operatorname{row}(A) = \begin{bmatrix} A_d & A_{d-1} & \cdots & A_0 \end{bmatrix}; \quad \operatorname{col}(A) = (\operatorname{row}(A))^{\mathcal{B}} = \begin{bmatrix} A_d \\ A_{d-1} \\ \vdots \\ A_0 \end{bmatrix}.$$

If $\operatorname{row}(A)$ has full row rank (or, equivalently, there is no nonzero $w \in \mathbb{C}^m$ such that $w^T A(x) = 0$), we say that $A(x)$ is *row-minimal*. If $\operatorname{col}(A)$ has full column rank (or, equivalently, there is no nonzero $v \in \mathbb{C}^n$ such that $A(x)v = 0$), we say that $A(x)$ is *column-minimal*.

Finally, we define the special matrix

$$J_n := \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}.$$

3 Dual pencils and Kronecker forms

In this section, we derive some basic results concerning the Kronecker form of dual pencils. Such results will be central in the rest of the paper. Most of these results appear in the existing work on dual pencils by Kublanovskaya, Simonova and collaborators: see [21, 22, 32] and the references therein. Nevertheless we have decided to include a self-contained exposition of these results, which seem to have been almost forgotten in the linear algebra community.

With the definitions stated in the previous section, the two conditions that define duality can be rewritten as

1. $\operatorname{row}(L) J_n \operatorname{col}(R) = 0$,
2. $\operatorname{rank} \operatorname{row}(L) + \operatorname{rank} \operatorname{col}(R) = 2n$.

This formulation highlights the special role played by the two matrices $\operatorname{col}(R)$ and $\operatorname{row}(L)$.

The rows of the matrix $\operatorname{row}(L)$ span the left nullspace of $J_n \operatorname{col}(R)$. Thus, given a pencil $R(x)$, we can construct explicitly one of its left duals by taking a basis for the left null space of $J_n \operatorname{col}(R)$ and using it as the rows of $\operatorname{row}(L)$. The dual constructed in this way is row-minimal. Clearly, given any row-minimal left dual $L(x)$, any other left dual of $R(x)$ can be constructed as $ML(x)$, where $M \in \mathbb{C}^{k \times m}$ has full column rank.

Similarly, we can construct a column-minimal right dual of a given pencil $L(x)$ by taking the right nullspace of $\operatorname{row}(L) J_n$. Any other right dual of $L(x)$ can be obtained as $R(x)N$ for a full-row-rank N.

We start from a lemma giving the duals of Kronecker blocks.

Lemma 3.1. *1. Let $B(x)$ be any nonsingular Kronecker block ($\lambda \in \mathbb{C}$ or $\lambda = \infty$). Then, $B(x)$ is a left and right dual of itself.*

2. *A right dual of $K_{k,k+1}(x)$ is $K_{k+1,k+2}(x)$. A left dual of $K_{k,k+1}(x)$ is $K_{k-1,k}(x)$ if $k > 0$, and the empty matrix if $k = 0$.*

3. *A left dual of $K_{k+1,k}(x)$ is $K_{k+2,k+1}(x)$. A right dual of $K_{k+1,k}(x)$ is $K_{k,k-1}(x)$ if $k > 0$, and the empty matrix if $k = 0$.*

Proof. 1. It is easy to check that $B_1B_0 = B_0B_1$, where $B(x) = B_1x + B_0$, since one of the coefficients B_1 or B_0 is $\pm I$. For the same reason, $\operatorname{row}(B)$ and $\operatorname{col}(B)$ have full rank, since they contain an identity block.

2. Note that it is enough to prove the condition on the right dual. We have

$$\begin{bmatrix} 0 & I_k \end{bmatrix} \begin{bmatrix} -I_{k+1} & 0 \end{bmatrix} = \begin{bmatrix} 0 & -I_k & 0 \end{bmatrix} = \begin{bmatrix} -I_k & 0 \end{bmatrix} \begin{bmatrix} 0 & I_{k+1} \end{bmatrix}. \tag{2}$$

Moreover, $\operatorname{rank}\operatorname{row}(K_{k,k+1}(x)) = k$ and $\operatorname{rank}\operatorname{col}(K_{k+1,k+2}(x)) = k+2$ are checked easily, so the rank condition holds.

3. Transpose everything in (2).

□

The following results characterize completely the KCF of dual pencils; it had (implicitly) appeared in [22, Section 1.3.2]: here we give a direct proof.

Theorem 3.2. *Suppose that a pencil $A(x)$ has Kronecker canonical form $UAV = B(x) = \operatorname{diag}(B_1(x), B_2(x), \dots, B_t(x))$. Then,*

1. *a row-minimal left dual of $A(x)$ is $S_l(x)U$, where*

$$S_l(x) = \operatorname{diag}(L_1(x), L_2(x), \dots, L_t(x)),$$

and

$$L_i(x) = \begin{cases} B_i(x) & \text{if } B_i(x) \text{ is any nonsingular Kronecker block,} \\ K_{k-1,k}(x) & \text{if } B_i(x) = K_{k,k+1}(x) \text{ with } k > 0, \\ K_{k+2,k+1}(x) & \text{if } B_i(x) = K_{k+1,k}(x), \\ \text{the } 0 \times 0 \text{ empty matrix} & \text{if } B_i(x) = K_{0,1}(x). \end{cases} \tag{3}$$

2. *All left duals of $A(x)$ have KCF $\operatorname{diag}(S_l(x), K_{1,0}(x), K_{1,0}(x), \dots, K_{1,0}(x))$, where we allow an arbitrary number of $K_{1,0}$ blocks.*

3. *a column-minimal right dual of $A(x)$ is $R(x) = VS_r(x)$, where*

$$S_r(x) = \operatorname{diag}(R_1(x), R_2(x), \dots, R_t(x)),$$

and

$$R_i(x) = \begin{cases} B_i(x) & \text{if } B_i(x) \text{ is any nonsingular Kronecker block,} \\ K_{k+1,k+2}(x) & \text{if } B_i(x) = K_{k,k+1}(x), \\ K_{k,k-1}(x) & \text{if } B_i(x) = K_{k+1,k}(x) \text{ with } k > 0, \\ \text{the } 0 \times 0 \text{ empty matrix} & \text{if } B_i(x) = K_{1,0}(x). \end{cases}$$

4. *All right duals of $A(x)$ have KCF* $\operatorname{diag}(S_r(x), K_{0,1}(x), K_{0,1}(x), \ldots, K_{0,1}(x))$*, where we allow an arbitrary number of $K_{0,1}$ blocks.*

Proof. 1 We have $S_l(x)B(x) = 0$, since it is a product between two conformally partitioned block diagonal matrices and the products between each pair of diagonal blocks vanishes because of Lemma 3.1. Hence $S_l(x)UA(x)V = 0$, but since V is invertible this implies $S_l(x)UA(x) = 0$.

It remains to verify the rank condition. Let $m_i \times n_i$ be the dimension of $L_i(x)$. As argued above in the proof of Lemma 3.1, $\operatorname{rank}\operatorname{row}(L_i) = m_i$ for the first three kinds of blocks in (3), and the same obviously holds for the fourth as well. Again by Lemma 3.1, $\operatorname{rank}\operatorname{col}(B_i) = 2n_i - m_i$ (since it is in the definition of duality) in the first three cases, and the same holds trivially for the fourth case. Now we have $n = \sum_i n_i$, $\operatorname{rank}\operatorname{row}(S_l(x)U) = \operatorname{rank}\operatorname{row}(S_l(x)) = \sum \operatorname{rank}\operatorname{row}(L_i) = \sum m_i$, and $\operatorname{rank}\operatorname{col}(B) = \sum \operatorname{rank}\operatorname{col}(B_i) = \sum(2n_i - m_i) = 2n - m$.

The dual $S_l(x)U$ is row-minimal since each $\operatorname{row}(L_i)$ has full row rank, and hence so does $\operatorname{row}(S_l)$.

2 Let $D(x)$ be a left dual of $A(x)$. All left duals can be written as $D(x) = ML(x)$, where $L(x)$ is a row-minimal dual and M is a full-column-rank matrix. If we complete M to a square invertible matrix as $\begin{bmatrix} M & M' \end{bmatrix}$, then we have

$$D(x) = \begin{bmatrix} M & M' \end{bmatrix} \begin{bmatrix} L(x) \\ 0 \end{bmatrix}.$$

Taking $L(x) = S_l(x)U$, which we know to be minimal, we have

$$D(x) = \begin{bmatrix} M & M' \end{bmatrix} \begin{bmatrix} S_l(x) \\ 0 \end{bmatrix} U,$$

which is a Kronecker canonical form for $D(x)$ with the required form.

3,4 are analogous to 1,2 and we omit the details.

□

In other words, when taking a left dual, the regular part is unchanged, the right minimal indices decrease by 1, and the left minimal indices increase by 1; the converse holds for right duals.

Corollary 3.3. *(Simultaneous Kronecker canonical form) For any dual pair $L(x), R(x)$, there are nonsingular U, V, W such that $UL(x)V$ and $V^{-1}R(x)W$ are both in Kronecker canonical form.*

In the theory of singular pencils, minimal bases and minimal indices play an important role. Therefore, we wish now to investigate how minimal bases change under duality. We note that a similar result is stated in [22, equation (1.3.4)].

Theorem 3.4. *Let $R(x) = R_1x + R_0$ and $L(x) = L_1x + L_0$ be dual. Suppose that the columns of $M(x)$ are a right minimal basis for $R(x)$. For any $(\gamma, \delta) \in \mathbb{C}^2 \setminus \{(0,0)\}$, the columns of $\frac{\gamma R_1+\delta R_0}{\gamma-\delta x}M(x)$ are a right minimal basis of $L(x)$.*

Proof. Notice first that it is sufficient to prove the result for $L(x)$ and $R(x)$ in KCF. Indeed, for generic $L(x), R(x)$, we can always reduce them to simultaneous KCF $UL(x)V$ and $V^{-1}R(x)W$. Then, $W^{-1}M(x)$ is a minimal basis of $V^{-1}R(x)W$; if the result holds for pencils in KCF, then $V^{-1}\frac{\gamma R_1+\delta R_0}{\gamma-\delta x}M(x)$ is a minimal basis for $UL(x)V$. Hence, the same holds for $\frac{\gamma R_1+\delta R_0}{\gamma-\delta x}M(x)$ and $L(x)$.

Let us now prove the result assuming that $L(x), R(x)$ are in KCF. For each block B_i of type K_{k_i,k_i+1}, with polynomial kernel $M_i(x)$ as in (1), we can verify directly that

$$\widetilde{M_i}(x) := \frac{\gamma R_1 + \delta R_0}{\gamma - \delta x} M_i(x) = \begin{bmatrix} x^{k_i-1} & x^{k_i-2} & \cdots & x & 1 \end{bmatrix}^T.$$

This vector is indeed a minimal basis for the kernel of the K_{k_i-1,k_i} block which is the left dual of B_i. For blocks B_i of types other than $K_{i,i+1}$, neither B_i nor its dual have a polynomial kernel, so we define M_i and $\widetilde{M_i}$ to be empty vectors of the suitable dimensions (both $k_i \times 0$ if B_i is a regular block, or $k_{i+1} \times 0$ and $k_i \times 0$ if B_i is of type K_{k_{i+1},k_i}).

Therefore, the minimal basis $M(x) = \operatorname{diag}(M_1(x), \dots, M_t(x))$ is transformed via $\frac{\gamma R_1+\delta R_0}{\gamma-\delta x}$ into $\widetilde{M}(x) = \operatorname{diag}(\widetilde{M_1}(x), \dots, \widetilde{M_t}(x))$, which we know to be a minimal basis for the pencil in KCF $L(x)$. □

In a similar way, it can be shown how a left minimal basis of $R(x)$ can be constructed starting from a left minimal basis of $L(x)$. We omit the details as they are analogous to Theorem 3.4.

4 Dual pencils and Wong chains

In this section, we agree to the convention that $\frac{1}{0} =: \infty$, useful to analyze infinite eigenvalues. Moreover, we will need the following lemma.

Lemma 4.1. *Let $L(x)$ and $R(x)$ be a pair of dual pencils. Then the following identity holds for all $\alpha, \beta, \gamma, \delta \in \mathbb{C}$:*

$$(\alpha L_1 + \beta L_0)(\gamma R_1 + \delta R_0) = (\gamma L_1 + \delta L_0)(\alpha R_1 + \beta R_0). \tag{4}$$

Proof. If we expand the products, the identity reduces to $(\alpha\delta-\beta\gamma)L_1R_0 = (\alpha\delta-\beta\gamma)L_0R_1$, which holds because $L_1R_0 = L_0R_1$. □

If $L(x)$ and $R(x)$ are square regular pencils, and v is an eigenvector of $R(x)$ with eigenvalue $x = \frac{\alpha}{\beta} \in \mathbb{C} \cup \{\infty\}$, then for each γ, δ such that $\frac{\gamma}{\delta} \neq \frac{\alpha}{\beta}$ Lemma 4.1 implies

$$0 = (\gamma L_1 + \delta L_0)(\alpha R_1 + \beta R_0)v = (\alpha L_1 + \beta L_0)(\gamma R_1 + \delta R_0)v,$$

and thus $w = (\gamma R_1 + \delta R_0)v$ is an eigenvector of $L(x)$ with the same eigenvalue $\frac{\alpha}{\beta}$.

It looks natural to try to generalize this relation to singular pencils. However, an additional difficulty appears in defining the needed quantities. In the regular case, eigenvectors corresponding to eigenvalues of geometric multiplicity 1 are uniquely defined up to a scalar nonzero constant, but this is not the case if $\dim \ker_{\mathbb{C}(x)} R(x) > 0$. Indeed, if $R(\lambda)w = 0$, then for any vector $z(x) \in \ker_{\mathbb{C}(x)} R(x)$ we also have $R(\lambda)(w + z(\lambda)) = 0$, thus $w + z(\lambda)$ is as good a choice as w for an eigenvector.

One could argue that there are similar problems also with the regular case and the Jordan canonical form. For instance, the matrix $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ has a well-defined (up to scalar multiples) eigenvector $\begin{bmatrix} 1 & 0 \end{bmatrix}^T$, but the second vector of the Jordan chain can be freely chosen as $\begin{bmatrix} \alpha & 1 \end{bmatrix}^T$ for any $\alpha \in \mathbb{C}$, so it is *not* uniquely defined up to scalar multiples. In presence of Kronecker blocks with dimensions larger than 1, further vectors appearing in the Kronecker chains become also ill-defined, and it becomes more complicated to quantify exactly how much freedom there is in these choices.

In other words, while the KCF $K(x)$ of a pencil is unique, the transformation matrices U and V are not, and therefore cannot be used to introduce uniquely defined quantities.

Wong chains are an underused tool that can be employed to avoid this problem. They were introduced by K.-T. Wong [35] and recently brought back to the attention of the linear algebra community by T. Berger, A. Ilchmann and S. Trenn [4, 5, 6].

We define here a generalized version of the original concept.

Definition 4.2. Let $R(x) = R_1 x + R_0 \in \mathbb{C}[x]_1^{m\times n}$ be a pencil, and $\lambda = \frac{\alpha}{\beta} \in \mathbb{C} \cup \{\infty\}$. Then for any (γ, δ) such that $\alpha\delta \neq \beta\gamma$, the *Wong chain of* $R(x)$ *attached to* λ is the sequence of vector subspaces $(W_k)_k \subseteq \mathbb{C}^n$ defined by the following recurrence:

$$\begin{aligned} W_0^{(\lambda)} &= \{0\}, \\ W_{k+1}^{(\lambda)} &= (\alpha R_1 + \beta R_0)^{-1}(\gamma R_1 + \delta R_0) W_k^{(\lambda)}. \end{aligned}$$

In the last formula, we have used the following notations to denote how a matrix $M \in \mathbb{C}^{m\times n}$ acts on a vector subspace V: $MV := \{Mv \in \mathbb{C}^m \mid v \in V\}$ (image of $V \subseteq \mathbb{C}^n$ via M), and $M^{-1}V := \{w \in \mathbb{C}^n \mid Mw \in V\}$ (preimage of $V \subseteq C^m$). It is easy to prove that the definition above does not depend on the particular choice of γ, δ.

We point out that our Definition 4.2 is different from the original one: it is a projective generalization in which the chain for a projective point $\frac{\alpha}{\beta}$ can be constructed using another point $\frac{\gamma}{\delta} \neq \frac{\alpha}{\beta}$. In [35, 4, 5] only the special cases $\frac{\alpha}{\beta} = \infty, \frac{\gamma}{\delta} = 0$ and $\frac{\alpha}{\beta} = 0, \frac{\gamma}{\delta} = \infty$ appear, while in [6, 31] the authors allow $\frac{\alpha}{\beta} \in \mathbb{C}, \frac{\gamma}{\delta} = \infty$. As far as we know, the observation that it is possible to change the second base point $\frac{\gamma}{\delta}$ in the definition (without altering the corresponding subspace chain) also appears here for the first time.

The following property is already known for the existing versions of Wong chains [35, 5], but we reprove it using our extended definition.

Lemma 4.3. $W_k^{(\lambda)} \subseteq W_{k+1}^{(\lambda)}$ *for all* k.

Proof. The relation is obvious for $k = 0$. Suppose that $w \in W_k^{(\lambda)}$, and assume $W_{k-1}^{(\lambda)} \subseteq W_k^{(\lambda)}$. Then there exists $w' \in W_{k-1}^{(\lambda)}$ such that $(\alpha R_1 + \beta R_0)w = (\gamma R_1 + \delta R_0)w'$; but since $w' \in W_k^{(\lambda)}$, $w \in W_{k+1}^{(\lambda)}$, which proves the lemma by induction. □

It is also easy to see (and already known for the "standard" Wong chains) that if $W_{k_0}^{(\lambda)} = W_{k_0-1}^{(\lambda)}$ then $\bigcup_k W_k^{(\lambda)} = W_{k_0}^{(\lambda)}$.

Notice that many well-defined quantities associated to a pencil can be expressed in terms of Wong chains, thereby fixing the ill-definition problems that we pointed out in the beginning of this section. We give some examples.

- As we argued before, eigenvectors of a singular pencil $R(x)$ are ill-defined, since given an eigenvector w it can be replaced with $w + z(\lambda)$ for any $z(x) \in \ker_{\mathbb{C}(x)} R(x)$. Nevertheless, the subspace $\operatorname{span} w + \operatorname{span}\{z(\lambda) \mid z(x) \in \ker_{\mathbb{C}(x)} R(x)\}$ is unique. This space is $W_1^{(\lambda)}$; so we see that it is a well-defined and meaningful generalization of the concept of eigenvector to singular pencils.

- Consider a Jordan chain associated to an eigenvalue λ of a regular pencil R, i.e., a sequence of vectors such that $(R_1\lambda + R_0)w_1 = 0$ and $(R_1\lambda + R_0)w_{k+1} = w_k$ for $k = 1, 2, \dots r$. For $k > 1$, the vector w_k can be replaced by any linear combination of itself and other vectors w_j with $j < k$. In other words, the only well-defined quantities are the subspaces $W_k^{(\lambda)} = \operatorname{span}(w_1, w_2, \dots, w_k)$.

- The right minimal indices, which are well defined, can be determined from the fact that, for any λ which is not an eigenvalue of $R(x)$, $\dim W_{k+1}^{(\lambda)} - \dim W_k^{(\lambda)}$ is the number of right Kronecker blocks of size k or greater [6].

One can define *left Wong chains* as the Wong chains of $R(x)^T$, and, using them, reconstruct left minimal indices and generalize left eigenvectors.

This description of Wong chains illustrates their key role in the theory of matrix pencils, in particular when singular pencils are studied. Therefore, it is natural to investigate how they change under duality.

Theorem 4.4. *Let $L(x)$ and $R(x)$ be a pair of dual pencils. Let $(V_k^{(\lambda)})_k$ and $(W_k^{(\lambda)})_k$ be the Wong chains of $L(x)$ and $R(x)$, respectively, attached to $\lambda = \frac{\alpha}{\beta} \in \mathbb{C} \cup \{\infty\}$. The relation*

$$V_k = (\gamma R_1 + \delta R_0)W_k$$

holds for all k, γ, δ provided that $\alpha\delta \neq \beta\gamma$.

Proof. We give a proof by induction on k. For $k = 0$ the relation is obvious since $\{0\} = (\gamma R_1 + \delta R_0)\{0\}$.

Suppose now that the thesis is true for V_{k-1}, W_{k-1}. Assume that $w \in W_k$, i.e., $\exists w' \in W_{k-1}$ such that $(\alpha R_1 + \beta R_0)w = (\gamma R_1 + \delta R_0)w'$. Then for $v = (\gamma R_1 + \delta R_0)w$ Lemma 4.1 implies $(\alpha L_1 + \beta L_0)v = (\gamma L_1 + \delta L_0)(\alpha R_1 + \beta R_0)w = (\gamma L_1 + \delta L_0)(\gamma R_1 + \delta R_0)w'$, and $(\gamma R_1 + \delta R_0)w' \in V_{k-1}$ by the inductive hypothesis: thus, $v \in V_k$ and therefore $(\gamma R_1 + \delta R_0)W_k \subseteq V_k$.

Conversely, let $v \in V_k$. Then, $(\alpha L_1 + \beta L_0)v = (\gamma L_1 + \delta L_0)v'$ for some $v' \in V_{k-1}$. This implies $\operatorname{row}(L)J \left[\begin{smallmatrix} \beta v - \delta v' \\ \gamma v' - \alpha v \end{smallmatrix}\right] = 0$. Therefore, by the definition of duality and the discussion in the beginning of Section 3, there exists a vector w such that $\left[\begin{smallmatrix} \beta v - \delta v' \\ \gamma v' - \alpha v \end{smallmatrix}\right] = \operatorname{col}(R)w$, that is, $R_1 w = \beta v - \delta v'$, $R_0 w = \gamma v' - \alpha v$.

Observe that $(\alpha R_1 + \beta R_0)w = (\beta\gamma - \alpha\delta)v' = (\alpha\delta - \beta\gamma)(\gamma R_1 + \delta R_0)w'$ for some $w' \in W_{k-1}$, by the inductive hypothesis. Thus, $w \in W_k$. Moreover, $(\gamma R_1 + \delta R_0)w = (\beta\gamma - \alpha\delta)v$. Therefore, $V_k \subseteq (\gamma R_1 + \delta R_0)W_k$, which concludes the proof. □

One may wonder if the hypothesis $\alpha\delta \neq \beta\gamma$ is necessary; the following example shows that it is indeed the case.

Example 4.5. Let $R(x) = \left[\begin{smallmatrix} x & 1 & 0 \\ 0 & 0 & x-1 \end{smallmatrix}\right]$, $L(x) = \left[\begin{smallmatrix} 0 & x-1 \end{smallmatrix}\right]$, $(\alpha, \beta) = (1, 1)$.

The Wong chain for $R(x)$ at $(1, 1)$ is $W_0 = \{0\}$, $W_1 = \ker(R_1 + R_0)$, which is the column space of the matrix $\left[\begin{smallmatrix} 0 & 1 \\ 0 & -1 \\ 1 & 0 \end{smallmatrix}\right]$, and $\mathbb{C}^3 = W_2 = W_3 = \dots$. On the other side, the Wong chain at $(1, 1)$ for $L(x)$ is $V_0 = \{0\}$, $\mathbb{C}^2 = V_1 = V_2 = \dots$. Notice that $(R_1 + R_0)W_1 = \{0\} \neq V_1$.

5 Structured pencils

The theoretical tool that we have developed can be used to derive new results on the minimal indices of structured pencils. We consider here the following structures. A pencil $S(x) = S_1 x + S_0 \in \mathbb{C}[x]^{2n \times 2n}$ is called *symplectic* if it is row-minimal and $S_0 J_n S_0^* = S_1 J_n S_1^*$. A pencil $H(x) = H_1 x + H_0 \in \mathbb{C}[x]^{2n \times 2n}$ is called *Hamiltonian* if it is row-minimal and $H_0 J_n H_1^* + H_1 J_n H_0^* = 0$.

These definitions generalize those given in [24] for regular pencils; however, the row-minimality condition is not present there, since it is automatically satisfied in the regular case. We discuss in Section 5.4 why it has been added here.

5.1 Symplectic pencils

As a first step, we pick out a particular right dual of a symplectic pencil.

Proposition 5.1. *Let $S(x) = S_1 x + S_0$ be a symplectic pencil. Then $T(x) = T_1 x + T_0 := J_n \operatorname{rev}[S(x)]^* = J_n S_1^* + x J_n S_0^*$ is a column-minimal right dual of $S(x)$.*

Proof. Symplecticity implies that $S_1 T_0 = S_0 T_1$. Column-minimality of $T(x)$ follows from row-minimality of $S(x)$, since

$$\operatorname{rank} \operatorname{col}(T) = \operatorname{rank} \begin{bmatrix} 0 & J_n \\ J_n & 0 \end{bmatrix} (\operatorname{row}(S))^* = \operatorname{rank} (\operatorname{row}(S))^* = 2n.$$

□

We can use this result to determine a relation between the left and right minimal bases of $S(x)$. We first need a definition: the *columnwise reversal* of a matrix polynomial $A(x)$ whose columns are $[A_1(x), \dots, A_n(x)]$ is $\operatorname{cwRev} A(x) := [\operatorname{rev} A_1(x), \dots, \operatorname{rev} A_n(x)]$.

Proposition 5.2. *Let $S(x)$ be a symplectic pencil, and $N(x)$ be a left minimal basis of $S(x)$. Then $J_n S_0^* \operatorname{cwRev} N(x)$ is a right minimal basis for $S(x)$.*

Proof. If $N(x)$ is a left minimal basis for $S(x)$, it is a right minimal basis of $S(x)^*$. Therefore, $\operatorname{cwRev} N(x)$ is a right minimal basis for $\operatorname{rev}[S(x)]^*$ and, thus, for $T(x) = J_n \operatorname{rev}[S(x)]^*$. By Theorem 3.4, this implies that $JS_0^* \operatorname{cwRev} N(x)$ is a right minimal basis for $S(x)$. □

A corresponding result on the minimal indices follows immediately.

Corollary 5.3. *Let $S(x)$ be a symplectic pencil, and let its minimal indices be $\nu_1, \nu_2, \ldots, \nu_s$. Then its right minimal indices are $\nu_1 - 1, \nu_2 - 1, \ldots, \nu_s - 1$.*

5.2 Hamiltonian

We proceed with the same strategy for the Hamiltonian case.

Proposition 5.4. *Let $H(x) = H_1 x + H_0$ be Hamiltonian. Then $G(x) = G_1 x + G_0 := J_n[H(-x)]^* = -xJ_n H_1^* + J_n H_0^*$ is a column-minimal right dual of $H(x)$.*

Proof. Hamiltonianity implies that $H_1 G_0 = H_0 G_1$, and

$$\operatorname{rank} \operatorname{col}(G) = \operatorname{rank} \begin{bmatrix} -J_{2n} & 0 \\ 0 & J_{2n} \end{bmatrix} (\operatorname{row}(H))^* = \operatorname{rank} (\operatorname{row}(H))^* = 2n.$$

□

This implies analogous results on the minimal bases and indices of Hamiltonian pencils.

Proposition 5.5. *Let $H(x) = H_1 x + H_0$ be a Hamiltonian pencil. Suppose that $N(x)$ is a left minimal basis for $H(x)$. Then, $J_n H_1^* N(-x)$ is a right minimal basis for $H(x)$.*

Proof. Notice that $N(-x)$ is a right minimal basis for $[H(-x)]^*$; therefore, $N(x)$ is a right minimal basis also for $G(x) = J_n[H(-x)]^*$. Invoking Theorem 3.4, we conclude that $J_n H_1^* N(-x)$ is a right minimal basis for $H(x)$. □

Corollary 5.6. *Let $H(x)$ be Hamiltonian. Suppose that its left minimal indices are $\nu_1, \nu_2, \ldots, \nu_s$. Then its right minimal indices are $\nu_1 - 1, \nu_2 - 1, \ldots, \nu_s - 1$.*

5.3 Explicit constructions

One may wonder if there are further restrictions on the possible minimal indices of symplectic and Hamiltonian pencils. We prove here that the answer is no: for any sequence of positive integers $\nu_1, \nu_2, \ldots, \nu_s$, there exist a symplectic and an Hamiltonian pencil with exactly the ν_i as left minimal indices (and hence, by Corollaries 5.3 and 5.6, $\nu_1 - 1, \nu_2 - 1, \ldots, \nu_s - 1$ as right minimal indices). First of all, we need the following basic building block.

Let $\mathcal{S}_n$ be the $n \times n$ anticyclic up-shift matrix, i.e., the $n \times n$ matrix such that

$$(\mathcal{S}_n)_{ij} = \begin{cases} 1 & j - i = 1, \\ -1 & i = n, j = 1, \\ 0 & \text{otherwise.} \end{cases}$$

For each $n \geq 1$, the $2n \times 2n$ pencil $H(x) = \operatorname{diag}(-\mathcal{S}_{n+1} K_{n+1,n}(-x), K_{n-1,n}(x))$ is Hamiltonian and has a left minimal index n, a right minimal index $n-1$ and no regular part.

Example 5.7. The smallest such examples are

$$\begin{bmatrix} x & 0 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} x & 1 & 0 & 0 \\ 0 & x & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & x \end{bmatrix}.$$

Then, one needs a method to build direct sums of Hamiltonian pencils.

Lemma 5.8. *Let the pencils*

$$\begin{bmatrix} A_i(x) & B_i(x) \\ C_i(x) & D_i(x) \end{bmatrix}, \quad A_i, B_i, C_i, D_i \in \mathbb{C}[x]^{n_i \times n_i}$$

be Hamiltonian, for $i = 1, 2, \ldots, m$. Let $A(x) = \operatorname{diag}(A_1(x), A_2(x), \ldots, A_m(x))$, and define $B(x)$, $C(x)$, and $D(x)$ analogously. Then,

$$\begin{bmatrix} A(x) & B(x) \\ C(x) & D(x) \end{bmatrix}$$

is Hamiltonian.

By taking direct sums of the blocks in Example **??**, one can achieve all the possible combinations of minimal indices allowed by Corollary 5.6.

As for symplectic pencils, one can show that, given a Hamiltonian example $H(x) = H_1 x + H_0$, the pencil $S(x) = (H_1 - H_0)x + (H_1 + H_0)$ is symplectic and has the same minimal indices as $H(x)$.

5.4 Further remarks on symplectic and Hamiltonian pencils

The assumption that symplectic and Hamiltonian pencils must be row-minimal is not classical. The reason is that existing theory focuses on the regular case only [24], for which it is automatically satisfied.

If this assumption is relaxed, some structure on the minimal indices can be lost. Indeed, let us consider non-row-minimal pencils that satisfy $S_0 J_n S_0^* = S_1 J_n S_1^*$ or $H_0 J H_1^* + H_1 J H_0^* = 0$. It is still true that, for each left minimal index $\nu_j \geq 1$, there is a corresponding right minimal index equal to $\nu_j - 1$. However, there is no constraint concerning the right minimal indices corresponding to zero left minimal indices. To illustrate this, consider the following examples.

Example 5.9. Let

$$L(x) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ x & 0 & x & 0 \end{bmatrix}.$$

This pencil satisfies both the symplectic and Hamiltonian equations. By direct inspection, we see that the left minimal indices of $L(x)$ are 0, 0, 1; the right minimal indices are 0, 0, 0.

Example 5.10. Let

$$L(x) = \begin{bmatrix} 0 & 0 \\ 1 & x \end{bmatrix},$$

satisfying again both Hamiltonian and symplectic equations. The only left minimal index is 0. On the other hand, the right minimal index is 1.

The recent work [29] presents a procedure to extract in a stable way a Hamiltonian pencil from the so-called extended pencils appearing in many control theory applications; in that setting, one always obtains row-minimal pencils, a condition that appears naturally in the development. In view of these observations, we consider the row-minimality requirement to be the most natural in the study of singular structured pencils.

In principle, one might extend symplectic and Hamiltonian structure to matrix polynomials of higher grade. This can be done by imposing, respectively, the condition

$$P(x)J[P(x)]^* = \operatorname{rev}(P(x))J[\operatorname{rev}(P(x))]^*$$

or the condition

$$P(x)J[P(x)]^* = P(-x)J[P(-x)]^*.$$

Characterizing these structured matrix polynomials is an open problem that we leave for future research.

As a final note, we point out that all the results stated in this section continue to hold if we replace all the conjugate transposes with transposes.

6 Linearizations of matrix polynomials

In many applications, the study of spectral properties of matrix polynomials is a central topic [8, 18, 19, 20, 28, 33]. A common technique to find the eigenvalues of a matrix polynomial is converting to a linear problem using the following method. Given a matrix polynomial $A(x) \in \mathbb{C}[x]_d^{m\times n}$, a pencil $L(x) \in \mathbb{C}[x]^{(m+p)\times(n+p)}$ is called a *linearization* of $A(x)$ if there are $E(x) \in \mathbb{C}[x]^{(m+p)\times(m+p)}, F(x) \in \mathbb{C}[x]^{(n+p)\times(n+p)}$ such that

$$L(x) = E(x)\operatorname{diag}(A(x), I_p)F(x), \tag{5}$$

where $p \geq 0$ and $\det E(x), \det F(x)$ are nonzero constants. When $m = n$, the most natural (and common) choice is $p = n(d-1)$. Linearizations have the same finite elementary divisors (hence the same eigenvalues) as the starting matrix polynomial [19]. Not every linearization preserves the partial multiplicities of the eigenvalue ∞ [17, 23], however. Linearizations that do are called *strong linearizations*, and they satisfy the additional property that $\operatorname{rev} L(x)$ is a linearization for $\operatorname{rev} A(x)$. The following result characterizes the Kronecker form of all strong linearizations [11].

Theorem 6.1. *Let $A(x)$ be a matrix polynomial. A pencil $L(x)$ is a strong linearization of $A(x)$ if and only if:*

1. *The eigenvalues of $L(x)$ and of $A(x)$ coincide.*
2. *For each eigenvalue $\lambda \in \mathbb{C} \cup \{\infty\}$, the sizes of the regular blocks of eigenvalue λ in the KCF of $L(x)$ coincide with the partial multiplicities associated to λ in $A(x)$.*
3. *The numbers of left and right singular blocks in the KCF of $L(x)$ are equal to $\dim \ker_{\mathbb{C}(x)} A(x)$ and $\dim \ker_{\mathbb{C}(x)} A(x)^*$, respectively.*

Note that there is no constraint on the minimal indices, i.e., the sizes of the singular left and right blocks of $L(x)$. Those may indeed vary for different linearizations.

Several different methods to construct linearizations have been studied in the literature; we recall some of the most common ones.

Companion forms [19, Chapter 1] The well-known *companion matrix* of a scalar polynomial generalizes easily to matrix polynomials. The pencil $C(x) = C_1 x + C_0$, where

$$C_1 = \operatorname{diag}(A_d, -I_{(d-1)n}), \quad C_0 = \begin{bmatrix} A_{d-1} & A_{d-2} & \cdots & A_0 \\ I & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & I & 0 \end{bmatrix} \tag{6}$$

is known as *first companion form*[1], while its block transpose $C^{\mathcal{B}}(x) = C_1 x + C_0^{\mathcal{B}}$ is known as *second companion form.*

Vector spaces of linearizations [25] A large family of linearizations for a matrix polynomial $A(x) \in \mathbb{C}[x]_d^{n\times n}$ is found inside the vector space $\mathbb{L}_1$ of pencils $L_1 x + L_0$ that satisfy

$$L_1 \begin{bmatrix} I_{dn} & 0_{dn\times n} \end{bmatrix} + L_0 \begin{bmatrix} 0_{dn\times n} & I_{dn} \end{bmatrix} = (v \otimes I_n)\operatorname{row}(A) \tag{7}$$

for some $v \in \mathbb{C}^d$, where $\otimes$ denotes the Kronecker product. The operation on the left-hand side is called *column-shifted sum* in [25]. A second vector space $\mathbb{L}_2$ is defined as the set of the block transposes of all pencils $L(x) \in \mathbb{L}_1$, or, equivalently, those which satisfy a similar relation, given by block-transposing everything in (7). Note that the first companion form $C(x)$ is in $\mathbb{L}_1$ and the second companion form $C^{\mathcal{B}}(x)$ is in $\mathbb{L}_2$.

The pencils in the intersection $\mathbb{DL} := \mathbb{L}_1 \cap \mathbb{L}_2$ have many interesting properties; for any regular matrix polynomial, almost all of them are strong linearizations.

Fiedler pencils [1] Define the matrices $F_0 := \operatorname{diag}(I_{(d-1)n}, A_0)$, $F_i := \operatorname{diag}(I_{(d-i-1)n}, G_i, I_{n(i-1)})$, for $i = 1, 2, \ldots, d-1$, where

$$G_i := \begin{bmatrix} A_i & I_n \\ I_n & 0_{n\times n} \end{bmatrix},$$

[1] There is no agreement in the literature on the signs in several special pencils, including 6. Our choice is not standard, but minimizes the number of minus signs that we have to keep track of along the paper.

and $F_d := C_1$.

For each permutation σ of $(0, 1, \dots, d-1)$, the pencil $F_d x + \prod_{i=0}^{d-1} F_{\sigma(i)}$ is a linearization; in particular, the two permutations $(d-1, d-2, \dots, 0)$ and $(0, 1, 2, \dots, d-1)$ yield the two companion forms. Another interesting special case is the block pentadiagonal pencil which corresponds to the permutation $(0, 2, \dots, 1, 3, \dots)$. It is remarkable that, after expanding the products, the constant terms of all Fiedler linearizations can be written explicitly by using only 0, I and $A_0, A_1, \dots A_{d-1}$ as blocks. Several additional generalizations of Fiedler pencils exist [1, 34].

7 Fiedler linearizations as duals

In this section, we use the theory developed above to revisit some of the known results for Fiedler pencils. We first need some results on the first companion form, whose proof is readily obtained by direct inspection (see also [11, Lemma 5.1]).

Lemma 7.1. *1. The first companion form $C(x)$ is a strong linearization for any (regular or singular) $A(x)$.*

2. If $A(x)$ is regular and v is an eigenvector of $A(x)$ with eigenvalue $\lambda = \frac{\alpha}{\beta}$, then

$$\begin{bmatrix} \alpha^{d-1} I & \alpha^{d-2}\beta^1 I & \cdots & \alpha\beta^{d-2} I & \beta^{d-1} I \end{bmatrix}^{\mathcal{B}} v \tag{8}$$

is an eigenvector of $C(x)$ with eigenvalue λ. If $\lambda \neq \infty$, we can use the slightly simpler formula

$$\begin{bmatrix} \lambda^{d-1} I & \lambda^{d-2} I & \cdots & \lambda I & I \end{bmatrix}^{\mathcal{B}} v. \tag{9}$$

3. If $M(x)$ is a minimal basis for $A(x)$, then

$$\begin{bmatrix} x^{d-1} I & x^{d-2} I & \cdots & xI & I \end{bmatrix}^{\mathcal{B}} M(x)$$

is a minimal basis for $C(x)$. In particular, the minimal indices of $C(x)$ are obtained increasing by $d-1$ the minimal indices of $A(x)$.

First of all, we give a new proof of the fact that all square Fiedler pencils are linearizations. Our argument follows closely the original proof of [1] for regular pencils, apart from some minor differences in notation. Indeed, [1, Lemma 2.2] is a very special case of our Theorem 3.2, but the argument there works only for regular pencils. The fact that Fiedler pencils are linearizations even in the singular case was first proved in [12], six years after the regular case and with a completely different technique based on keeping track of a large number of unimodular transformations. Duality allows us to reuse almost verbatim the proof for the regular case, instead[2].

[2]Note that we do not cover here the more involved rectangular case, treated in [13].

Theorem 7.2. *For a (possibly singular) matrix polynomial $A(x) \in \mathbb{C}[x]^{n\times n}$, all Fiedler pencils are strong linearizations.*

Proof. For any $j > i$, we set $F_{j:i} = F_{j-1}F_{j-2}\cdots F_i$ for short. Since F_i and F_j commute for any i, j with $|i-j| > 1$, we can always rearrange the product $\prod_{i=0}^{d-1} F_{\sigma(i)}$ in the form

$$F_{c_1:0}F_{c_2:c_1}F_{c_3:c_2}\cdots F_{c_\Gamma:c_{\Gamma-1}}F_{d:c_\Gamma}, \tag{10}$$

for a suitable sequence $0 < c_1 < \cdots < c_\Gamma < d$, with the only operation of changing the order of pairs of commuting matrices. One can see that $c_1, \ldots, c_\Gamma$ are exactly the indices i such that $\sigma(i-1) < \sigma(i)$ (i.e., i is a *consecution* [12]).

We prove the following result by induction on Γ: all Fiedler pencils with Γ consecutions are strong linearizations for $A(x)$, and each of their right singular indices is greater than or equal to $d-1-\Gamma$.

If $\Gamma = 0$, then we have $F_d x + F_{d-1}F_{d-2}\ldots F_0 = C(x)$, the first companion form, so the result follows from Lemma 7.1.

Now, assuming that we have proved the result for a sequence $c_2, c_3, \ldots, c_\Gamma$, with $\Gamma < d-1$, we prepend an extra element c_1 and prove it for the sequence $c_1, c_2, c_3, \ldots, c_\Gamma$. Let $P = F_{c_1:0}$ and $Q = F_{c_2:c_1}F_{c_3:c_2}F_{c_4:c_3}\cdots F_{d:c_\Gamma}$; the latter is nonsingular since all F_i for $i \notin \{0, d\}$ are nonsingular. Note that P commutes with all terms in Q apart from $F_{c_2:c_1}$. Since $F_{c_2:c_1}F_{c_1:0} = F_{c_2:0}$, the Fiedler pencil $F_d x + QP$ is the one associated to $c_2, c_3, \ldots, c_\Gamma$, which is a strong linearization by the inductive hypothesis. Moreover, also from the inductive hypothesis, it is column-minimal, since all its right minimal indices are greater than 0.

We premultiply this pencil by the nonsingular matrix Q^{-1}, to obtain $R(x) := Q^{-1}F_d x + P$, which is still a column-minimal strong linearization for $A(x)$. Now we claim that $L(x) = F_d x + PQ$ is a left dual of $R(x)$. The first condition is verified since F_d and P commute, so we only need to check that $\operatorname{rank}\operatorname{row}(L) = dn$. Due to the structure of F_d, for this to hold it is enough to prove that PQ has a $n \times n$ identity block somewhere in its first block row. But, due to the structure of the involved matrices, $F_{c_1:0}F_{c_2:c_1}F_{c_3:c_2}\cdots F_{c_\Gamma:c_{\Gamma-1}}$ has $\begin{bmatrix} I & 0 & \ldots & 0 \end{bmatrix}$ as its first block row, while $F_{d:c_\Gamma}$ has an identity in the block in position $(1, d-c_\Gamma+1)$.

Thus, by Theorem 3.1, the regular parts of the KCFs of $L(x)$ and $R(x)$ coincide, all the left minimal indices in $L(x)$ are larger than those of $R(x)$ by 1 , and all its right minimal indices are smaller by 1. In particular, the number of left and right singular blocks is preserved. Hence, by Theorem 6.1, $L(x)$ is a strong linearization for $A(x)$, too. □

Corollary 7.3 ([12])**.** *Let $r_1, r_2, \ldots, r_h$ be the right minimal indices of a matrix polynomial $A(x) \in \mathbb{C}[x]_d^{n\times n}$, and consider a Fiedler pencil $F(x)$ with Γ consecutions associated with $A(x)$. Then, the right minimal indices of $F(x)$ are $r_1+(d-1-\Gamma), r_2+(d-1-\Gamma), \ldots, r_h+(d-1-\Gamma)$.*

Proof. In the proof of Theorem , to construct $F(x)$ we start from $C(x)$ which has right minimal indices $r_1+d-1, r_2+d-1, \ldots, r_h+d-1$, and obtain $F(x)$ after taking a left dual $\Gamma \le d-1$ times. Each of these times the right minimal indices are shortened by 1, hence by keeping track of their lengths we get the above result. □

Note that $d-1-\Gamma$ is the number of *inversions*, i.e., indices $i \in \{1, 2, \dots, d-1\}$ which are not consecutions.

We can expand on the proof of Theorem to find a minimal basis for each Fiedler pencil explicitly.

Theorem 7.4. *Consider the Fiedler pencil $F(x)$ associated to a given permutation with consecutions $c_1, c_2, \dots, c_\Gamma$, and let $T = F_{d:c_\Gamma}^{-1} F_d F_{d:c_{\Gamma-1}}^{-1} F_d \cdots F_{d:c_1}^{-1} F_d$. If $M_C(x)$ is a right minimal basis for the first companion form $C(x)$, then $M_F(x) = T M_C(x)$ is a right minimal basis for $F(x)$.*

Proof. Our plan is following the proof of Theorem 7 and showing how right minimal bases change along the needed duality operations. The result is obvious if $\Gamma = 0$; now, let us suppose that it holds for a permutation with consecutions $c_2 < c_3 < \cdots < c_\Gamma$ and prove it for the same sequence with an additional consecution $c_1 < c_2 < \cdots < c_\Gamma$. Applying Theorem 3.4 with $(\gamma, \delta) = (1, 0)$, we obtain that a right minimal basis for the new permutation is

$$M_F(x) = (Q^{-1} F_d) F_{d:c_\Gamma}^{-1} F_d F_{d:c_{\Gamma-1}}^{-1} F_d \cdots F_{d:c_2}^{-1} F_d M_C(x),$$

with $Q = F_{c_2:c_1} F_{c_3:c_2} F_{c_4:c_3} \cdots F_{d:c_\Gamma}$ as in Theorem 7. Since each term $F_{c_i:c_{i-1}}^{-1}$ appearing in Q^{-1} commutes with F_d and with all the terms $F_{d:c_j}^{-1}$ for $j > i$, we can reorder the factors to obtain

$$\begin{aligned} M_F(x) &= F_{d:c_\Gamma}^{-1} F_d (F_{c_\Gamma:c_{\Gamma-1}}^{-1} F_{d:c_\Gamma}^{-1}) F_d (F_{c_{\Gamma-1}:c_{\Gamma-2}}^{-1} F_{d:c_{\Gamma-1}}^{-1}) F_d \cdots (F_{c_2:c_1}^{-1} F_{d:c_2}^{-1}) F_d M_C(x) \\ &= F_{d:c_\Gamma}^{-1} F_d F_{d:c_{\Gamma-1}}^{-1} F_d F_{d:c_{\Gamma-2}}^{-1} F_d \cdots F_{d:c_1}^{-1} F_d M_C(x) = T. \end{aligned}$$

□

Corollary 7.5. *If $M_A(x)$ is a minimal basis for $A(x)$, then a right minimal basis for $F(x)$ is $M_F(x) = T(x) M_A(x)$, where*

$$T(x) = T \begin{bmatrix} x^{d-1} I & x^{d-2} I & \cdots & xI & I \end{bmatrix}^{\mathcal{B}}.$$

Using the fact that

$$F_{d:c_i}^{-1} F_d = \operatorname{diag}\left(\begin{bmatrix} 0 & 0 & \cdots & A_d \\ -I & 0 & \cdots & A_{d-1} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & -I & A_{c_i} \end{bmatrix}, I \right),$$

one can compute the product one factor at a time, starting from the right, and verify that this means that the i-th block of $(-1)^\Gamma T(x)$ contains $-x^{\mathcal{I}_i}(A_d x^{i-1} + A_{d-1} x^{i-2} + \cdots + A_{d-i+1})$ if i is a consecution and $x^{\mathcal{I}_i} I$ if i is an inversion, where $\mathcal{I}_i$ is the number of inversions j with $j \leq i$ — see for instance the example 11 in the following. This is in agreement with the equivalent expression found in [12], where the polynomials $P_i(x) = A_d x^{i-1} + A_{d-1} x^{i-2} + \cdots + A_{d-i+1}$ are called *Horner shifts* of $A(x)$.

A similar result holds for Wong chains.

Theorem 7.6. *Let $(W_i^{(\lambda)})_i$ be the Wong chain attached to $\lambda \neq \infty$ for the first companion form $C(x)$, and $(V_i^{(\lambda)})_i$ be the same Wong chain for a Fiedler pencil $F(x)$. Then, $V_i^{(\lambda)} = TW_i^{(\lambda)}$.*

Proof. Once again, we follow the proof of Theorem 7 and show how the Wong chains change along the duality operations. Since we suppose $\lambda \neq \infty$, we can use $(\gamma, \delta) = (1, 0)$ in Theorem 4.4, so the Wong chains are multiplied by $Q^{-1}F_d$ at each duality, exactly like minimal bases. The same argument as in the proof of Theorem 7.4 applies, and yields the same matrix T. □

For a regular pencil, $W_1^{(\lambda)}$ is the subspace of all eigenvectors with eigenvalue λ. Hence, using (9), we get the following result.

Corollary 7.7. *If v is an eigenvector of eigenvalue $\lambda \neq \infty$ for a regular matrix polynomial $A(x)$, then the corrisponding eigenvector of $C(x)$ is $T(\lambda)v$.*

Once again, this agrees with the expressions in [12]. Our results on Wong chains, however, are more general and can be applied to defective eigenvalues and singular pencils as well.

With another choice of (γ, δ), we can obtain results with an excluded eigenvalue at 0 rather than ∞.

Theorem 7.8. *Let $(W_i^{(\lambda)})_i$ be the Wong chain attached to $\lambda \neq 0$ for the first companion form $C(x)$, and $(V_i^{(\lambda)})_i$ be the same Wong chain for a Fiedler pencil $F(x)$. Let $\widetilde{T} := F_{c_1:0}F_{c_2:0}\dots F_{c_\Gamma:0}$. Then, $V_i^{(\lambda)} = \widetilde{T}W_i^{(\lambda)}$.*

Proof. We use $(\gamma, \delta) = (0, 1)$, and proceed as in the previous cases. At each step, we premultiply $F_{c_2:0}\dots F_{c_\Gamma:0}$ by $P = F_{c_1:0}$, so the result is even easier to obtain. □

Example 7.9. Consider the Fiedler linearization $F_6x + F_0F_1F_3F_5F_2F_4$ [11, Examples 3.5 and subsequent ones]. It has three consecutions in $c_1 = 1, c_2 = 2, c_3 = 4$. Therefore, its expression as in (10) is $F_6x + F_0F_1F_3F_2F_5F_4 = F_{1:0}F_{2:1}F_{4:2}F_{6:4}$. By Corollary 7.5, a right minimal basis for it is given by $T(x)M_A(x)$, where $M_A(x)$ is a minimal basis for $A(x)$ and

$$T(x) = \begin{bmatrix} x^2I & xI & -xP_2(x) & I & -P_4(x) & -P_5(x) \end{bmatrix}^{\mathcal{B}}. \tag{11}$$

Similarly, all Wong chains with $\lambda \neq \infty$ are obtained from those of $A(x)$ by left multiplication by $T(\lambda)$. Obtaining an analogous result that includes $\lambda = \infty$ with Theorem 7.8 is not more complicated, but we have to move to the projective version with formula (8). It turns out that all eigenvectors with $\lambda = \frac{\alpha}{\beta} \neq 0$ are recovered by left multiplication by $\widetilde{T}(\alpha, \beta)$, with

$$\widetilde{T}(\alpha, \beta) = \begin{bmatrix} \alpha^5 I & \alpha^4\beta I & \widetilde{P}_2(\alpha, \beta) & \alpha^3\beta^2 I & \widetilde{P}_4(\alpha, \beta) & \widetilde{P}_5(\alpha, \beta) \end{bmatrix}.$$

and $\widetilde{P}_i(\alpha, \beta) = \sum_{j=0}^{d-1-i} A_i\alpha^i\beta^{d-1-i}$.

One can obtain a completely analogous set of results for the left eigenvectors and minimal bases, by starting with the second companion form and performing repeatedly right dualities, one for each inversion in the associated permutation.

8 Duality and $\mathbb{L}_1$, $\mathbb{L}_2$ linearization spaces

Let us consider an $n \times n$ matrix polynomial $A(x)$ of grade d. In this section we study the space $\mathbb{L}_1$ associated to $A(x)$ and its connection with duality. An important pencil belonging to $\mathbb{L}_1$ is the first companion form $C(x)$, defined as in (6). Throughout this section, we denote by B any matrix whose columns form a basis of $\ker \operatorname{row}(A)$, and we define $\mu = \dim \operatorname{coker} \operatorname{row}(A)$. Notice that the rank-nullity theorem implies that $B \in \mathbb{C}^{(dn+n)\times(dn+\mu)}$.

The following result holds (see also [22, Section 1.4.3]).

Proposition 8.1. *Let* $M(x) = \begin{bmatrix} 0_{dn,n} & -I_{dn} \end{bmatrix} x + \begin{bmatrix} I_{dn} & 0_{dn,n} \end{bmatrix}$. *The* $dn \times (dn+\mu)$ *pencil* $D(x) = M(x)B$ *is a minimal right dual of* $C(x)$.

Proof. It is straightforward to check that the structure of the first companion form gives $\operatorname{rank} \operatorname{row}(C) = dn - \mu$. The matrix $\operatorname{col}(M)$ has full column rank, implying $\operatorname{rank} \operatorname{col}(D) = \operatorname{rank} B = dn + \mu$. Thus, it suffices to verify that $\operatorname{row}(C) J_{dn} \operatorname{col}(D) = 0$. To this goal, we notice that $\operatorname{row}(C) J_{dn} \operatorname{col}(M) = \begin{bmatrix} \operatorname{row}(A) \\ 0 \end{bmatrix}$, yielding the desired result since $\operatorname{row}(A)B = 0$. □

We now give a sufficient condition for a pencil belonging to $\mathbb{L}_1$ to be a linearization.

Theorem 8.2. *Let* $L(x) \in \mathbb{L}_1(A)$ *be such that* $\operatorname{rank} \operatorname{row}(L) = dn - \mu$. *Then,* $L(x)$ *is a left dual of* $D(x)$ *and a strong linearization.*

Proof. By definition, $L(x)$ must satisfy (7), which using the notation of Proposition 8.1 can be rewritten as $\operatorname{row}(L) J_{dn} \operatorname{col}(M) = (v \otimes I_n) \operatorname{row}(A)$. It follows immediately that $\operatorname{row}(L) J_{dn} \operatorname{col}(D) = \operatorname{row}(L) J_{dn} \operatorname{col}(M) B = 0$. Together with the rank condition in the hypothesis, this implies that $L(x)$ is a left dual of $D(x)$. In particular, $C(x)$ and $L(x)$ have the same (finite and infinite) elementary divisors, and the same right minimal indices. Since $L(x)$ is a square pencil, it has the same amount of left and right singular indices; in particular, this numbers coincide with those of $C(x)$, and thus $L(x)$ is a strong linearization by Theorem 6.1. □

This sufficient condition is weaker than the one of having full Z-rank, found in [11]. Indeed, the following inclusions hold for pencils in $\mathbb{L}_1$:

$$\{\text{pencils with full Z-rank}\} \subset \{\text{left duals of } D(x)\} \subset \{\text{strong linearizations of } A(x)\}. \tag{12}$$

Both inclusions are strict: the pencil in [11, Example 2] is an example of left dual of $D(x)$ which has not full Z-rank, and the following example shows that the second inclusion is strict as well.

Example 8.3. Consider the matrix polynomial $A(x) = \begin{bmatrix} 1 & 0 & 0 \\ x & 0 & 0 \\ x^2 & 0 & 0 \end{bmatrix} =: A_0 + A_1 x + A_2 x^2$. $A(x)$ has no elementary divisors, its left minimal indices are $1, 1$, and its right minimal indices are $0, 0$.

The 6×6 pencil

$$L(x) = \begin{bmatrix} A_2x - A_1 & 2A_1x + A_0 \\ H & -Hx \end{bmatrix}, \quad \text{with } H = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

belongs to $\mathbb{L}_1$. It is a strong linearization of $A(x)$ by Theorem 6.1: indeed, it has no elementary divisors and the dimension of its left and right kernels are 2.

The first companion form of $A(x)$ has left minimal indices $1, 1$; therefore, by Theorem 3.2, $D(x)$ has left minimal indices $0, 0$, and all its left dual have left minimal indices $1, 1$. However, $L(x)$ has left minimal indices $0, 2$, thus it cannot be a left dual of $D(x)$.

In the case of a regular $A(x)$, all the inclusions in (12) become equalities:

Theorem 8.4. *Let $A(x) \in \mathbb{C}[x]_d^{n\times n}$ be a regular matrix polynomial, and $L(x) \in \mathbb{L}_1$ for the space $\mathbb{L}_1$ constructed based on $A(x)$. Then, the following are equivalent:*

1. *$L(x)$ is a strong linearization of $A(x)$*
2. *$L(x)$ has full Z-rank ,*
3. *$L(x)$ is regular,*
4. *$L(x)$ is a left dual of $D(x)$,*
5. *$L(x)$ is row-minimal.*

The first three equivalent conditions of the theorem appear already in [25]; here we add (4) and (5). The last condition (5) seems the simplest one to check in practice.

Proof. As said above, the equivalence between (1), (2) and (3) is from [25]. The arrows (5) $\Longrightarrow$ (4) $\Longrightarrow$ (1) do not require regularity of $A(x)$; the first is the statement of Theorem 8.2, and the second is contained in its proof. It remains to prove (1) $\Longrightarrow$ (5): if $L(x)$ is a linearization of a regular matrix polynomial, by the last point of Theorem 6.1 it has no left or right singular Kronecker blocks, hence in particular it has no $K_{1,0}$ blocks and thus is row-minimal. □

Analogous results concerning the link between $\mathbb{L}_2(A)$ and the right duals of the left duals of the second companion form can be easily obtained by adapting the arguments used in this section.

9 Linearizations and Möller-Stetter theorem

The linearization $D(x)$ described in Section 8 can be introduced in an alternative way as the generalization of a construction that is used in commutative computational algebra to find the solution to (scalar) polynomial systems. While the theory developed so far is self-consistent, it is still interesting to investigate the link between the dual linearization $D(x)$ and the Möller-Stetter theorem in algebraic geometry.

In commutative algebra, a polynomial ideal $I = (f_1, \ldots, f_k) \subseteq \mathbb{C}[x_1, \ldots, x_m]$ is called *zero-dimensional* if the system

$$\begin{cases} f_1(x_1, \ldots, x_m) = 0 \\ \qquad\vdots \\ f_k(x_1, \ldots, x_m) = 0 \end{cases}$$

has only a finite number of solutions, or, equivalently [10, Finiteness Theorem, Section 2.5], if the quotient space $\mathbb{C}[x_1, \ldots, x_m]/I$ is finite-dimensional. The elements of the quotient space are usually denoted by the notation $[p] := \{q \in \mathbb{C}[x_1, \ldots, x_m] \mid q = p + r,\ r \in I\}$.

When this holds, we have the following result, which we present for simplicity in the case of univariate polynomials $m = 1$ (see [10, Section 2.4] for the most general version).

Theorem 9.1 (Möller–Stetter)**.** *Let $f, g \in \mathbb{C}[x]$, and denote by (f) the (principal) ideal generated by f. Consider the linear multiplication map $M_g : \mathbb{C}[x]/(f) \to \mathbb{C}[x]/(f)$ defined as $[p] \mapsto [pg]$. When a basis of $\mathbb{C}[x]/(f)$ is chosen, this map is represented by a matrix. Its eigenvalues (counted with multiplicity) are the values of $g(x_i)$, where x_i are the solutions (counted with multiplicity) x_i of the polynomial equation $f = 0$.*

Even though this stronger condition is usually not of interest in the commutative algebra applications, it is possible to prove that M_x is a linearization of the single polynomial equation $f = 0$ that generates the ideal (f is essentially unique, as $\mathbb{C}[x]$ is a PID). In fact, when f is monic and the monomial basis is chosen, the multiplication operator is the companion matrix of f.

We aim to generalize this result to polynomial matrices in order to produce linearizations. Let us consider the space of all row vector polynomials of grade d

$$W := \mathbb{C}[x]_d^{1\times n} = \left\{ \sum_{i=0}^{k} v_i x^i \mid v_i \in \mathbb{C}^{1\times n} \text{ for all } i = 0, 1, \ldots, d \right\}.$$

This space is isomorphic to $\mathbb{C}^{(d+1)n}$, via $\mathcal{R} : v(x) \mapsto \operatorname{row}(v)$. For any row-minimal grade-d matrix polynomial $A(x) \in \mathbb{C}[x]_d^{n\times n}$, the rows of the matrix $\operatorname{row}(A)$ span an n-dimensional subspace $\mathcal{A}$ of $\mathbb{C}^{(d+1)n}$. It makes sense to consider its image under $\mathcal{R}^{-1}$. This is a subspace of W:

$$\operatorname{rowid}(A) := \mathcal{R}^{-1}(\mathcal{A}) = \{r(x) = r^T A(x) \mid r \in \mathbb{C}^n\} \subseteq W.$$

The notation $\operatorname{rowid}(A)$ suggests that it will play the role of the "ideal generated by $A(x)$" in the Möller-Stetter theorem. More formally, we consider the quotient space $Q := W/\operatorname{rowid}(A)$. The elements of Q are the equivalence classes $[w(x)] := \{a(x) \in W \mid a(x) = w(x) + r(x),\ r(x) \in \operatorname{rowid}(A)\}$. Acting with $\mathcal{R}$ we immediately obtain the corresponding quotient set $\mathcal{R}(Q)$, i.e., the set of the equivalence classes defined as $[\operatorname{row}(w(x))] := \{a \in \mathbb{C}^{(d+1)n} \mid a^T = \operatorname{row}(w(x)) + r^T \operatorname{row}(A),\ r \in \mathbb{C}^n\}$.

We can prove the following result.

Theorem 9.2. *Let $A(x) \in \mathbb{C}[x]_d^{n\times n}$ be a row-minimal matrix polynomial of grade d, and B be a matrix whose columns form a basis for* $\ker \operatorname{row}(A)$. *Let $V := \mathbb{C}[x]_{d-1}^{1\times n}$, $W := \mathbb{C}[x]_d^{1\times n}$, and $Q = W/\operatorname{rowid}(A)$. Let $\overline{M}_x$ and $\overline{M}_1$ be the maps "multiplication by x" and "multiplication by 1" between the spaces V and Q, i.e.,*

$$\begin{aligned} \overline{M}_x &: V \to Q \quad v(x) \mapsto [xv(x)], \\ \overline{M}_1 &: V \to Q \quad v(x) \mapsto [v(x)]. \end{aligned}$$

In suitable bases, it holds that $\overline{M}_x - \overline{M}_1 x = D(x)$, where $D(x) = M(x)B$ is the linearization defined in Proposition 8.1.

Since Theorem 8.1 tells us that $D(x)$ is a strong linearization, Theorem 9.2 shows that $\overline{M}_x - \overline{M}_1 x$ has the same eigenvalues of the matrix polynomial $A(x)$, and thus it can be regarded as a generalization of the Möller-Stetter result to (univariate) matrix polynomials. We proceed with the proof.

Proof. The maps $\overline{M}_x$ and $\overline{M}_1$ can be represented as $\pi \circ M_x$ and $\pi \circ M_1$, where

$$\begin{aligned} M_x &: V \to W \quad v(x) \mapsto xv(x), \\ M_1 &: V \to W \quad v(x) \mapsto v(x), \end{aligned}$$

and $\pi : v(x) \mapsto [v(x)]$ is the projection onto the quotient space $Q = W/\operatorname{rowid}(A)$. We use as bases of V and W the image of the canonical basis of $\mathbb{C}^{dn}$ and $\mathbb{C}^{(d+1)n}$ via the isomorphism $\mathcal{R}$; in these bases, M_x and M_1 are represented by right multiplication of row vectors by $\begin{bmatrix} I_{dn} & 0_{dn\times n} \end{bmatrix}$ and $\begin{bmatrix} 0_{dn\times n} & I_{dn} \end{bmatrix}$, respectively. Now we need to choose a basis for Q and work out the matrix corresponding to π.

Thanks to the definition of B, the map $v \mapsto vB$ (i.e., the action of B on row vectors) has kernel equal to $\mathcal{A}$. Hence, by the first isomorphism theorem, its image is isomorphic to $\mathbb{C}^{n(d+1)}/\mathcal{A}$, which is itself isomorphic via $\mathcal{R}$ to $W/\operatorname{rowid} A = Q$. Therefore, the map $v \mapsto vB$ passes to the quotient and becomes a projection onto Q.

To complete the proof, it now suffices to compose the two maps. We conclude that $\begin{bmatrix} I_{dn} & 0_{dn\times n} \end{bmatrix} B$ and $\begin{bmatrix} 0_{dn\times n} & I_{dn} \end{bmatrix} B$ represent $\overline{M}_x$ and $\overline{M}_1$. □

10 Constructing duals

The relation $L_1 R_0 = L_0 R_1$ has been studied estensively in the context of pencil arithmetic and inverse-free matrix iterative algorithms [2, 3, 9, 14, 27]. Two main techniques exist for constructing L_0, L_1 starting from R_0, R_1 (or vice versa).

QR factorization [22, Section 1.5.4.7], [2, 3, 27] Construct the QR factorization

$$\operatorname{col}(R) = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \begin{bmatrix} U \\ 0 \end{bmatrix},$$

and take $L_0 = Q_{12}^*$, $L_1 = -Q_{22}^*$. In practice, a QRP factorization should be used, since $\operatorname{col}(R)$ being close-to-rank-deficient is a concern here.

Enforcing an identity block [9, 30] Suppose that the identity matrix is a submatrix of $\mathrm{col}(R)$, for a pencil $R \in \mathbb{C}^{n\times p}[x]$. Then, we can select a permutation matrix $\Pi \in \mathbb{C}^{2n\times 2n}$ and $X \in \mathbb{C}^{(2n-p)\times p}$ such that

$$\mathrm{col}(R) = \Pi \begin{bmatrix} I_p \\ X \end{bmatrix}.$$

Then, the identity

$$0 = \left(\begin{bmatrix} -X & I_{2n-p} \end{bmatrix} \Pi^{-1} \right) \Pi \begin{bmatrix} I_p \\ X \end{bmatrix}$$

holds, and thus we can choose

$$\begin{bmatrix} L_0 & -L_1 \end{bmatrix} = \begin{bmatrix} -X & I_{2n-p} \end{bmatrix} \Pi^{-1}.$$

Slightly generalizing, if a $p \times p$ submatrix Y of $\mathrm{col}(R)$ is known to be nonsingular, we have

$$\mathrm{col}(R) = \Pi \begin{bmatrix} Y \\ Z \end{bmatrix} = \Pi \begin{bmatrix} I \\ ZY^{-1} \end{bmatrix} Y$$

and thus

$$\begin{bmatrix} L_0 & -L_1 \end{bmatrix} = \begin{bmatrix} -ZY^{-1} & I \end{bmatrix} \Pi^{-1}.$$

Example 10.1. Let us start from the first companion form of a matrix polynomial with grade $d = 3$, for which

$$\mathrm{row}(C) = \begin{bmatrix} A_3 & & \\ & -I & \\ & & -I \\ A_2 & A_1 & A_0 \\ I & & \\ & I & \end{bmatrix}$$

The $3n \times 3n$ matrix formed by the block rows number 2, 3 and 5 is nonsingular. Therefore, we choose a permutation Π that rearranges the block in the new order $(5, 2, 3, 1, 4, 6)$. In this way,

$$X = ZY^{-1} = \begin{bmatrix} A_3 & & \\ A_2 & A_1 & A_0 \\ & I & \end{bmatrix} \begin{bmatrix} I & & \\ & -I & \\ & & -I \end{bmatrix}^{-1},$$

and

$$\begin{bmatrix} L_0 & -L_1 \end{bmatrix} = \begin{bmatrix} -X & I \end{bmatrix} \Pi^{-1} = \begin{bmatrix} I & & & & -A_3 & \\ & A_1 & A_0 & I & -A_2 & \\ & I & & & & I \end{bmatrix},$$

which recovers a pencil belonging to a generalized Fiedler family [1, Example 2.5].

11 Conclusions and acknowledgements

In this paper, we brought some attention on the duality of matrix pencils and on Wong chains, two concepts which have been introduced in the past but whose use is not common in this research area. We have given several examples in the study of matrix pencils where these ideas give a more manageable framework. They have allowed us to derive new results and to simplify proofs of, and shed more light into, known properties. It remains to check if the more recent advances on rectangular Fiedler pencils [13] can be embedded in the same framework; although in principle there are no major obstructions, there are some additional complications due to the varying dimensions of the blocks in their definition.

Moreover, we find the connection to the Möller-Stetter theorem a promising new point of view to look at linearizations. We are currently investigating other extensions of this approach, to see if it can be extended to a more systematic derivation.

We thank Nick Higham and Françoise Tisseur for useful comments and suggestions on an older version of this manuscript. We are grateful to two anonymous referees for their remarks that helped to improve the presentation.

We wish to thank Paul Van Dooren, who pointed out the paper [21], and Froilán Dopico, who referred us to [22]. Indeed, we were initially unaware of the work by Kublanovskaya and collaborators, and we had derived independently many of these theoretical results before realizing they were already known to these authors.

References

[1] E. N. Antoniou and S. Vologiannidis. A new family of companion forms of polynomial matrices. *Electron. J. Linear Algebra*, 11:78–87 (electronic), 2004.

[2] Zhaojun Bai, James Demmel, and Ming Gu. An inverse free parallel spectral divide and conquer algorithm for nonsymmetric eigenproblems. *Numer. Math.*, 76(3):279–308, 1997.

[3] Peter Benner and Ralph Byers. An arithmetic for matrix pencils: theory and new algorithms. *Numer. Math.*, 103(4):539–573, 2006.

[4] Thomas Berger, Achim Ilchmann, and Stephan Trenn. The quasi-Weierstraß form for regular matrix pencils. *Linear Algebra Appl.*, 436(10):4052–4069, 2012.

[5] Thomas Berger and Stephan Trenn. The quasi-Kronecker form for matrix pencils. *SIAM J. Matrix Anal. Appl.*, 33(2):336–368, 2012.

[6] Thomas Berger and Stephan Trenn. Addition to "The quasi-Kronecker form for matrix pencils" [5]. *SIAM J. Matrix Anal. Appl.*, 34(1):94–101, 2013.

[7] Dennis S. Bernstein. *Matrix mathematics*. Princeton University Press, Princeton, NJ, second edition, 2009. Theory, facts, and formulas.

[8] Timo Betcke, Nicholas J. Higham, Volker Mehrmann, Christian Schröder, and Françoise Tisseur. NLEVP: a collection of nonlinear eigenvalue problems. *ACM Trans. Math. Software*, 39(2):Art. 7, 28, 2013.

[9] E. K.-W. Chu, H.-Y. Fan, and W.-W. Lin. A structure-preserving doubling algorithm for continuous-time algebraic Riccati equations. *Linear Algebra Appl.*, 396:55–80, 2005.

[10] David A. Cox, John Little, and Donal O'Shea. *Using algebraic geometry*, volume 185 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2005.

[11] Fernando De Terán, Froilán M. Dopico, and D. Steven Mackey. Linearizations of singular matrix polynomials and the recovery of minimal indices. *Electron. J. Linear Algebra*, 18:371–402, 2009.

[12] Fernando De Terán, Froilán M. Dopico, and D. Steven Mackey. Fiedler companion linearizations and the recovery of minimal indices. *SIAM J. Matrix Anal. Appl.*, 31(4):2181–2204, 2009/10.

[13] Fernando De Terán, Froilán M. Dopico, and D. Steven Mackey. Fiedler companion linearizations for rectangular matrix polynomials. *Linear Algebra Appl.*, 437(3):957–991, 2012.

[14] James Demmel, Ioana Dumitriu, and Olga Holtz. Fast linear algebra is stable. *Numer. Math.*, 108(1):59–91, 2007.

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

[16] F. R. Gantmacher. *The theory of matrices. Vol. 1.* AMS Chelsea Publishing, Providence, RI, 1998. Translated from the Russian by K. A. Hirsch, Reprint of the 1959 translation.

[17] I. Gohberg, M. A. Kaashoek, and P. Lancaster. General theory of regular matrix polynomials and band Toeplitz operators. *Integral Equations Operator Theory*, 11(6):776–882, 1988.

[18] Israel Gohberg, Peter Lancaster, and Leiba Rodman. *Indefinite linear algebra and applications*. Birkhäuser Verlag, Basel, 2005.

[19] Israel Gohberg, Peter Lancaster, and Leiba Rodman. *Matrix polynomials. Reprint of the 1982 original.* Society for Industrial and Applied Mathematics (SIAM), 2009.

[20] Tadeusz Kaczorek. *Polynomial and rational matrices*. Communications and Control Engineering Series. Springer-Verlag London Ltd., London, 2007. Applications in dynamical systems theory.

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

[22] 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.

[23] P. Lancaster and P. Psarrakos. A note on weak and strong linearizations of regular matrix polynomials. Technical Report MIMS-EPrint 2006.72, Manchester Institute for Mathematical Sciences, 2006.

[24] Wen-Wei Lin, Volker Mehrmann, and Hongguo Xu. Canonical forms for Hamiltonian and symplectic matrices and pencils. *Linear Algebra Appl.*, 302/303:469–533, 1999. Special issue dedicated to Hans Schneider (Madison, WI, 1998).

[25] D. Steven Mackey, Niloufer Mackey, Christian Mehl, and Volker Mehrmann. Vector spaces of linearizations for matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 28(4):971–1004 (electronic), 2006.

[26] D. Steven Mackey, Niloufer Mackey, Christian Mehl, and Volker Mehrmann. Smith forms of palindromic matrix polynomials. *Electron. J. Linear Algebra*, 22:53–91, 2011.

[27] A. N. Malyshev. Calculation of invariant subspaces of a regular linear matrix pencil. *Sibirsk. Mat. Zh.*, 30(4):76–86, 217, 1989.

[28] V. Mehrmann and H. Voss. Nonlinear eigenvalue problems: A challenge for modern eigenvalue methods. *Mitt. der Ges. für Angewandte Mathematik und Mechanik*, 27:1029–1051, 2005.

[29] Volker Mehrmann and Federico Poloni. Robust control via the computation of permuted graph bases. Technical Report 19, Technische Universität Berlin, 2012.

[30] Volker Merhmann and Federico Poloni. Doubling algorithms with permuted Lagrangian graph bases. *SIAM J. Matrix Anal. Appl.*, 33(3):780–805, 2012.

[31] Federico Poloni and Timo Reis. A deflation approach for large-scale Lur'e equations. *SIAM J. Matrix Anal. Appl.*, 33(4):1339–1368, 2012.

[32] V. N. Simonova. *Modification of the AB-Algorithms and Their Applications to the Solution of System of Nonlinear Algebraic Equations (in Russian)*. PhD thesis, Leningrad University, Leningrad, 1990.

[33] Françoise Tisseur and Karl Meerbergen. The quadratic eigenvalue problem. *SIAM Rev.*, 43(2):235–286, 2001.

[34] S. Vologiannidis and E. N. Antoniou. A permuted factors approach for the linearization of polynomial matrices. *Math. Control Signals Systems*, 22(4):317–342, 2011.

[35] Kai-Tak Wong. The eigenvalue problem $\lambda Tx + Sx$. *J. Diff. Eqns.*, 16:270–280, 1974.