MANCHESTER
1824

The University
of Manchester

Skew-symmetric matrix polynomials and their Smith forms

Mackey, D. Steven and Mackey, Niloufer and Mehl, Christian and Mehrmann, Volker

2012

MIMS EPrint: **2012.70**

Manchester Institute for Mathematical Sciences
School of Mathematics

The University of Manchester

Reports available from: http://eprints.maths.manchester.ac.uk/
And by contacting: The MIMS Secretary
School of Mathematics
The University of Manchester
Manchester, M13 9PL, UK

ISSN 1749-9097

Skew-Symmetric Matrix Polynomials and their Smith Forms

D. Steven Mackey* Niloufer Mackey*
Christian Mehl† Volker Mehrmann†

June 27, 2012

Abstract

We characterize the Smith form of skew-symmetric matrix polynomials over an arbitrary field $\mathbb{F}$, showing that all elementary divisors occur with even multiplicity. Restricting the class of equivalence transformations to unimodular congruences, a Smith-like skew-symmetric canonical form for skew-symmetric matrix polynomials is also obtained. These results are used to analyze the eigenvalue and elementary divisor structure of matrices expressible as products of two skew-symmetric matrices, as well as the existence of structured linearizations for skew-symmetric matrix polynomials. By contrast with other classes of structured matrix polynomials (e.g., alternating or palindromic polynomials), every regular skew-symmetric matrix polynomial is shown to have a structured strong linearization. While there are singular skew-symmetric polynomials of even degree for which a structured linearization is impossible, for each odd degree we develop a skew-symmetric companion form that uniformly provides a structured linearization for every regular and singular skew-symmetric polynomial of that degree. Finally, the results are applied to the construction of minimal symmetric factorizations of skew-symmetric rational matrices.

Key words. matrix polynomial, matrix pencil, compound matrix, Smith form, elementary divisors, invariant polynomials, Jordan structure, skew-symmetric matrix polynomial, structured linearization, companion form, unimodular congruence, skew-symmetric canonical form, Smith-McMillan form, skew-symmetric rational matrix, minimal symmetric factorization.

AMS subject classification. 65F15, 15A18, 15A21, 15A54, 15A57

1 Introduction

Recent papers have analyzed the Smith forms of several important classes of structured matrix polynomials, in particular, of alternating polynomials [37] and palindromic polynomials [38]. The main motivation for these investigations was to probe for any systematic obstructions to the existence of structured linearizations arising from incompatibilities between the elementary divisor structure of polynomials and pencils within the same structure

*Department of Mathematics, Western Michigan University, Kalamazoo, MI 49008, USA, Emails: steve.mackey@wmich.edu, nil.mackey@wmich.edu. Supported by National Science Foundation grant DMS-1016224. Support from Deutsche Forschungsgemeinschaft through DFG Research Center Matheon during research visits to TU Berlin is gratefully acknowledged.

†Institut für Mathematik, MA 4-5, Technische Universität Berlin, 10623 Berlin, Germany, Emails: mehl@math.tu-berlin.de, mehrmann@math.tu-berlin.de. Supported by Deutsche Forschungsgemeinschaft through DFG Research Center Matheon, 'Mathematics for key technologies' in Berlin.

class. Such incompatibilities were indeed found in [37, 38], and are analogous to restrictions on possible Jordan structures and Schur forms of Hamiltonian and symplectic matrices, see e.g., [31, 34], and their generalizations as even, symplectic, or palindromic pencils, see e.g., [22, 48, 49]. The difficulties stem from constraints on the multiplicities of Jordan structures associated with certain critical eigenvalues: at 0 or ∞ for Hamiltonian matrices and alternating matrix polynomials, at ± 1 for symplectic matrices and palindromic matrix polynomials.

It is also well-known that there are restrictions on the possible Jordan structures of skew-Hamiltonian matrices and their pencil extensions: all eigenvalues have even multiplicity, and every Jordan block appears an even number of times [15, 23, 50]. In contrast to the situation for Hamiltonian and symplectic matrices, though, these restrictions on Jordan structures do not raise any significant difficulties for numerical methods. Indeed, skew-Hamiltonian structure is special among matrices with multiple eigenvalues, since there are stable methods for the computation of eigenvalues, invariant subspaces, and Schur forms that can easily separate the paired structures from each other [42, 49, 51]. Recent results on the perturbation theory for skew-Hamiltonian matrices and their pencil extensions [1, 2, 8, 28] add theoretical support for this observed behavior.

These favorable properties for skew-Hamiltonian matrices immediately lead one to ask to what extent these properties might also be present in their natural matrix polynomial generalization, i.e., the class of skew-symmetric matrix polynomials.

The work presented here[1] initiates this inquiry by developing the structured Smith form for skew-symmetric matrix polynomials over an arbitrary field. This canonical form is then used as an investigative tool for the existence of structured linearizations. After a brief review of the relevant concepts in Section 2, we establish the notion of skew-symmetry over arbitrary commutative rings and fields (including those of characteristic 2) in Section 3, extending several well-known properties of skew-symmetric matrices over $\mathbb{R}$ or $\mathbb{C}$ to a general commutative ring. Canonical forms are the subject of Sections 4 and 5. These are then used in Section 6 to characterize the elementary divisor structure of a product of two skew-symmetric matrices, to resolve the structured linearization question for skew-symmetric polynomials, and to construct symmetric factorizations of skew-symmetric rational matrices that are minimal everywhere except for infinity.

2 Notation and Background Results

Throughout this paper we consider $n \times n$ matrix polynomials $P(\lambda)$ with nonzero leading coefficients, i.e.,

$$P(\lambda) = \lambda^k A_k + \cdots + \lambda A_1 + A_0, \quad A_k \neq 0, \quad A_i \in \mathbb{F}^{n\times n}, \quad 1 \leq i \leq k,$$

where $\mathbb{F}$ is an arbitrary field. It is often useful to view $P(\lambda)$ as a polynomial matrix, i.e., as a single matrix with polynomial entries $P_{ij}(\lambda)$. We will switch freely between these two points of view, using whichever is most appropriate for the task at hand.

We use $\mathbb{F}[\lambda]$ for the ring of polynomials in one variable with coefficients from the field $\mathbb{F}$, and $\mathbb{F}(\lambda)$ to denote the field of rational functions over $\mathbb{F}$. A matrix polynomial $P(\lambda)$ is said to be regular if it is invertible when viewed as matrix over $\mathbb{F}(\lambda)$, equivalently if $\det P(\lambda)$ is

[1] By contrast with our earlier work on Smith forms in [37] and [38], this paper contains no joke.

not the identically zero polynomial; otherwise it is said to be singular. The rank of $P(\lambda)$, sometimes called the normal rank, is the rank of $P(\lambda)$ when viewed as a matrix with entries in the field $\mathbb{F}(\lambda)$.

We now gather together some well-known results and tools from matrix theory that are used in this paper. Details can be found in standard monographs like [17, Ch.VI], [19, Part IV], [32].

2.1 Smith form, elementary divisors, and greatest common divisors

Recall that two $m \times n$ matrix polynomials $P(\lambda), Q(\lambda)$ are said to be *unimodularly equivalent*, denoted by $P \sim Q$, if there exist unimodular matrix polynomials $E(\lambda)$ and $F(\lambda)$ of size $m \times m$ and $n \times n$, respectively, such that

$$Q(\lambda) = E(\lambda)P(\lambda)F(\lambda). \tag{2.1}$$

Here, an $n \times n$ matrix polynomial $E(\lambda)$ is called *unimodular* if $\det E(\lambda)$ is a nonzero constant, independent of λ.

Theorem 2.1 (Smith form [16])**.**
Let $P(\lambda)$ be an $m \times n$ matrix polynomial over an arbitrary field $\mathbb{F}$. Then there exists $r \in \mathbb{N}$, and unimodular matrix polynomials $E(\lambda)$ and $F(\lambda)$ of size $m \times m$ and $n \times n$, respectively, such that

$$E(\lambda)P(\lambda)F(\lambda) = \operatorname{diag}(d_1(\lambda), \ldots, d_{\min\{m,n\}}(\lambda)) =: D(\lambda), \tag{2.2}$$

where $d_1(\lambda), \ldots, d_r(\lambda)$ are monic, $d_{r+1}(\lambda), \ldots, d_{\min\{m,n\}}(\lambda)$ are identically zero, and $d_j(\lambda)$ is a divisor of $d_{j+1}(\lambda)$ for $j = 1, \ldots, r-1$. Moreover, $D(\lambda)$ is unique.

The nonzero diagonal elements $d_j(\lambda)$, $j = 1, \ldots, r$ in the Smith form of $P(\lambda)$ are called the *invariant factors* or *invariant polynomials* of $P(\lambda)$, and have an important interpretation in terms of greatest common divisors of minors of $P(\lambda)$ [17, 19, 32]. Recall that a *minor of order k* of an $m \times n$ matrix A is the determinant of a $k \times k$ submatrix of A, i.e., of a matrix obtained from A by deleting $m-k$ rows and $n-k$ columns. For $d(x) \not\equiv 0$ we write $d(x)|p(x)$ to mean that $d(x)$ is a divisor of $p(x)$. When S is a set of scalar polynomials, we write $d|S$ to mean that $d(x)$ divides each element of S, i.e., $d(x)$ is a *common divisor* of the elements of S. The *greatest common divisor* (or GCD) of a set S containing at least one nonzero polynomial is the unique monic polynomial $g(x)$ such that $g(x)|S$, and if $d(x)|S$ then $d(x)|g(x)$. We denote the GCD of S by $\gcd(S)$.

Theorem 2.2 (Characterization of invariant polynomials)**.**
Let $P(\lambda)$ be an $m \times n$ matrix polynomial over an arbitrary field $\mathbb{F}$ with Smith form as in (2.2). Set $p_0(\lambda) \equiv 1$. For $1 \le j \le \min(m,n)$, let $p_j(\lambda) \equiv 0$ if all minors of $P(\lambda)$ of order j are zero; otherwise, let $p_j(\lambda)$ be the GCD of all minors of $P(\lambda)$ of order j. Then the number r in Theorem 2.1 is the largest integer such that $p_r(\lambda) \not\equiv 0$, i.e., $r = \operatorname{rank} P(\lambda)$. Furthermore, the invariant polynomials $d_1(\lambda), \ldots, d_r(\lambda)$ of $P(\lambda)$ are ratios of GCDs given by

$$d_j(\lambda) = \frac{p_j(\lambda)}{p_{j-1}(\lambda)}, \quad j = 1, \ldots, r,$$

while the remaining diagonal entries of the Smith form of $P(\lambda)$ are given by

$$d_j(\lambda) = p_j(\lambda) \equiv 0, \quad j = r+1, \ldots, \min\{m,n\}.$$

The following simple result on GCDs of sets of matrix polynomials will be needed later.

Lemma 2.3. *Suppose* $S = \{p_1(x), p_2(x), \ldots, p_m(x)\}$ *is a finite set of scalar polynomials over an arbitrary field* $\mathbb{F}$*, and let* $\widetilde{S} := \{p_1^2(x), p_2^2(x), \ldots, p_m^2(x)\}$.

(a) *If* $\gcd(S) = 1$*, then* $\gcd(\widetilde{S}) = 1$.

(b) *If* $\gcd(S) = g(x)$*, then* $\gcd(\widetilde{S}) = g^2(x)$.

Proof. (a): Suppose on the contrary that $\gcd(\widetilde{S}) \neq 1$. Then there exists a non-trivial $\mathbb{F}$-irreducible polynomial $r(x)$ that divides each p_i^2. But then the irreducibility of $r(x)$ implies that $r(x)$ must divide each p_i, contradicting the hypothesis that $\gcd(S) = 1$.
(b): By hypothesis, $p_i(x) = g(x)h_i(x)$ for each i, and $\gcd\{h_1(x), h_2(x), \ldots, h_m(x)\} = 1$. Then $\gcd(\widetilde{S}) = \gcd\{g^2h_1^2, g^2h_2^2, \ldots, g^2h_m^2\} = g^2 \cdot \gcd\{h_1^2, h_2^2, \ldots, h_m^2\} = g^2$, by part (a). $\square$

2.2 Elementary divisors and linearizations

If $P(\lambda)$ is a matrix polynomial over a field $\mathbb{F}$ with rank r, then each of its invariant polynomials $d_i(\lambda)$ for $1 \leq i \leq r$ can be uniquely factored as

$$d_i(\lambda) \;=\; q_{i1}(\lambda)^{\alpha_{i1}} \cdot \ldots \cdot q_{i,\ell_i}(\lambda)^{\alpha_{i,\ell_i}},$$

where $\ell_i \geq 0$, $\alpha_{ij} > 0$, $j = 1, \ldots, \ell_i$, and $q_{ij}(\lambda)$, $j = 1, \ldots, \ell_i$ are distinct $\mathbb{F}$-irreducible monic non-constant polynomials. (If $\ell_i = 0$, then $d_i(\lambda) \equiv 1$ by the definition of the empty product.) Then the *elementary divisors* of P are the collection of factors $q_{ij}(\lambda)^{\alpha_{ij}}$ for $j = 1, \ldots, \ell_i$, $i = 1, \ldots, r$ including repetitions [17]. If $q_{ij}(\lambda)^{\alpha_{ij}} = (\lambda - \lambda_0)^{\alpha_{ij}}$ is a power of a linear factor for some $\lambda_0 \in \mathbb{F}$, then λ_0 is called an *eigenvalue* of P and $(\lambda - \lambda_0)^{\alpha_{ij}}$ is called an elementary divisor associated with λ_0.

Infinite elementary divisors of P are obtained via the *reversal*, $\operatorname{rev} P$, defined by

$$(\operatorname{rev} P)(\lambda) := \lambda^k P(1/\lambda), \quad \text{where} \quad k = \deg P. \tag{2.3}$$

The elementary divisors of P associated with $\lambda_0 = \infty$ are then defined to be the same as those associated with the eigenvalue 0 of $\operatorname{rev} P$.

Recall that an $nk \times nk$ pencil $L(\lambda) = \lambda X + Y$ is called a *linearization* for an $n \times n$ matrix polynomial $P(\lambda)$ of degree k if there exist unimodular $nk \times nk$ matrix polynomials $E(\lambda)$ and $F(\lambda)$ such that

$$E(\lambda)L(\lambda)F(\lambda) = \begin{bmatrix} P(\lambda) & 0 \\ 0 & I_{n(k-1)} \end{bmatrix}.$$

A linearization $L(\lambda)$ is called a *strong linearization* if in addition there exist unimodular $nk \times nk$ matrix polynomials $G(\lambda)$ and $H(\lambda)$ such that

$$G(\lambda)\operatorname{rev} L(\lambda)H(\lambda) = \begin{bmatrix} \operatorname{rev} P(\lambda) & 0 \\ 0 & I_{n(k-1)} \end{bmatrix}.$$

This concept was introduced in [18], and named in [30], see also [29]. It is clear from the definition that the (finite and infinite) elementary divisors of $P(\lambda)$ and $L(\lambda)$ are identical if $L(\lambda)$ is a strong linearization for $P(\lambda)$. For the converse, we need an additional condition on the *nullspace* of $P(\lambda)$.

Definition 2.4 ([10], Definition 2.1)**.** *Let $P(\lambda)$ be an $n \times n$ matrix polynomial over an arbitrary field $\mathbb{F}$, and let $\mathbb{F}(\lambda)$ denote the field of rational functions over $\mathbb{F}$. Then*

$$\mathcal{N}_r(P) := \{x(\lambda) \in \mathbb{F}(\lambda)^n \mid P(\lambda)x(\lambda) \equiv 0\}$$

is called the (right) nullspace *of $P(\lambda)$.*

Lemma 2.5 ([10], Lemma 2.3)**.** *Let $P(\lambda) = \lambda^k A_k + \cdots + \lambda A_1 + A_0$, $A_k \neq 0$ be an $n \times n$ matrix polynomial and let $L(\lambda)$ be an $nk \times nk$ matrix pencil. Then $L(\lambda)$ is a strong linearization for $P(\lambda)$ if and only if $L(\lambda)$ and $P(\lambda)$ have the same (finite and infinite) elementary divisors and $\dim \mathcal{N}_r(P) = \dim \mathcal{N}_r(L)$.*

A classical example of a linearization is the "companion" linearization

$$C_1(\lambda) = \lambda \begin{bmatrix} I_n & & & 0 \\ & \ddots & & \\ & & I_n & \\ 0 & & & A_k \end{bmatrix} + \begin{bmatrix} 0 & -I_n & & 0 \\ & \ddots & \ddots & \\ & & 0 & -I_n \\ A_0 & A_1 & \dots & A_{k-1} \end{bmatrix}. \tag{2.4}$$

It was shown in [18] that $C_1(\lambda)$ is always a strong linearization, for any $n \times n$ matrix polynomial $P(\lambda)$. (Only the case $\mathbb{F} = \mathbb{C}$ was considered in [18], but the proof given there generalizes immediately to polynomials over arbitrary fields.)

Remark 2.6. Let s be the *sum* of the degrees of all the elementary divisors (finite and infinite) of a matrix polynomial $P(\lambda)$ of degree k. Then the Index Sum Theorem for matrix polynomials [12], implies that s can never exceed kr, where $r = \operatorname{rank} P(\lambda)$. For square $P(\lambda)$ of size $n \times n$, the Index Sum Theorem further implies that $P(\lambda)$ is regular if and only if $s = kn$. This fact will be important for us in Section 6.3.2.

2.3 Compound matrices and their properties

Recently, compound matrices have proved very effective in obtaining Smith forms of structured matrix polynomials: for T-even and T-odd polynomials in [37], and for T-palindromic polynomials in [38]. They will once again constitute a central tool in this paper. For references on compound matrices, see [21, Section 0.8], [39, Chapter I.2.7], [45, Section 2 and 28].

We use a variation of the notation in [21] for submatrices of an $m \times n$ matrix A. Let $\eta \subseteq \{1, \dots, m\}$ and $\kappa \subseteq \{1, \dots, n\}$ be arbitrary index sets of cardinality $1 \leq j \leq \min(m, n)$. Then $A_{\eta\kappa}$ denotes the $j \times j$ submatrix of A in rows η and columns κ, and the *$\eta\kappa$-minor of order j* of A is $\det A_{\eta\kappa}$. Note that A has $\binom{m}{j} \cdot \binom{n}{j}$ minors of order j. When $\eta = \kappa$, then $A_{\eta\kappa}$ is referred to as a *principal submatrix* of A, and the corresponding minor $\det A_{\eta\kappa}$ is a *principal minor* of A.

Definition 2.7 (Compound Matrices)**.**
Let A be an $m \times n$ matrix with entries in an arbitrary commutative ring, and let $\ell \leq \min(m, n)$ be a positive integer. Then the ℓth *compound matrix* (or the ℓth *adjugate*) of A, denoted $\mathcal{C}_\ell(A)$, is the $\binom{m}{\ell} \times \binom{n}{\ell}$ matrix whose (η, κ)-entry is the $\ell \times \ell$ minor $\det A_{\eta\kappa}$ of A. Here, the index sets $\eta \subseteq \{1, \dots, m\}$ and $\kappa \subseteq \{1, \dots, n\}$ of cardinality ℓ are ordered lexicographically.

Observe that we always have $\mathcal{C}_1(A) = A$, and, if A is square, then $\mathcal{C}_n(A) = \det A$. Basic properties of $\mathcal{C}_\ell(A)$ that we need are collected in the next theorem.

Theorem 2.8 (Properties of compound matrices)**.**
Let A be an $m \times n$ matrix with entries in a commutative ring, and let $\ell \leq \min(m, n)$ be a positive integer. Then

(a) $\mathcal{C}_\ell(A^T) = \big(\mathcal{C}_\ell(A)\big)^T$;

(b) $\mathcal{C}_\ell(\mu A) = \mu^\ell\, \mathcal{C}_\ell(A)$, *where* $\mu \in \mathbb{F}$;

(c) $\det \mathcal{C}_\ell(A) = (\det A)^\beta$, *where* $\beta = \binom{n-1}{\ell-1}$, *provided that* $m = n$;

(d) $\mathcal{C}_\ell(AB) = \mathcal{C}_\ell(A)\,\mathcal{C}_\ell(B)$, *provided that* $B \in \mathbb{F}^{n\times p}$ *and* $\ell \leq \min(m, n, p)$.

We are especially interested in compounds of matrices with polynomial entries. Note that such a compound can be thought of either as a polynomial with matrix coefficients, or as a matrix with polynomial entries, leading to the natural identification $\mathcal{C}_\ell(P)(\lambda) := \mathcal{C}_\ell(P(\lambda))$. The next theorem, established in [37], shows how the first $\ell + 1$ invariant polynomials of $P(\lambda)$ determine the first *two* invariant polynomials of $C_\ell(P)(\lambda)$.

Theorem 2.9 (First two invariant polynomials of the ℓth compound [37])**.**
Suppose the Smith form of an $n \times n$ matrix polynomial $P(\lambda)$ is

$$D(\lambda) \;=\; \operatorname{diag}\big(d_1(\lambda)\,,\,\ldots\,,\,d_{\ell-1}(\lambda)\,,\,d_\ell(\lambda)\,,\,d_{\ell+1}(\lambda)\,,\,\ldots\,,\,d_n(\lambda)\big)\,,$$

and for $2 \leq \ell < n$ denote the Smith form of the ℓth compound $\mathcal{C}_\ell\big(P(\lambda)\big)$ by

$$S(\lambda) = \operatorname{diag}\big(s_1(\lambda)\,,\,s_2(\lambda)\,,\ldots\,,\,s_{\binom{n}{\ell}}(\lambda)\big).$$

Then the first two diagonal entries of $S(\lambda)$ are given by

$$s_1(\lambda) = \left(\prod_{j=1}^{\ell-1} d_j(\lambda)\right)\cdot d_\ell(\lambda) \quad \textit{and} \quad s_2(\lambda) = \left(\prod_{j=1}^{\ell-1} d_j(\lambda)\right)\cdot d_{\ell+1}(\lambda)\,.$$

3 Skew-symmetry over a commutative ring

Our focus in this work is on matrix polynomials that are skew-symmetric over an arbitrary field $\mathbb{F}$. Let us begin by examining the notion of skew-symmetry for a matrix $A \in \mathbb{F}^{n\times n}$ when $\operatorname{char}\mathbb{F} = 2$. In such a field, the condition $A = -A^T$ is equivalent to saying that A is symmetric, and in particular, the entries on the main diagonal are unconstrained, rather than zero. By explicitly requiring the diagonal entries A_{ii} to vanish, we are led to a definition of skew-symmetry that works in arbitrary fields, and additionally, just as well in the more general setting of a commutative ring.

Definition 3.1. *An $n \times n$ matrix A with entries in a commutative ring R is said to be* skew-symmetric *if $A^T = -A$, and $A_{ii} = 0$ for $i = 1, \ldots, n$.*

Such matrices often go by the name "alternate matrix" or "alternating matrix" [14, 26, 33], but this is not universally the case. We have chosen, instead, Artin's usage in [4], where the condition $A_{ii} = 0$ is explicitly invoked to make the development of skew-symmetry work in a uniform way for all fields. Of course, this condition is redundant when $2 \in R$ is not a zero divisor; otherwise it constitutes an independent constraint on the entries of A. Note that we follow the usage of Jacobson [26], where the element $0 \in R$ is included as one of the zero divisors of the ring R.

A skew-symmetric matrix polynomial $P(\lambda)$ over an arbitrary field $\mathbb{F}$ can now be defined as a skew-symmetric matrix (in the sense of definition 3.1) whose entries are polynomials in $\mathbb{F}[\lambda]$. Equivalently, $P(\lambda)$ is skew symmetric when all its coefficient matrices are skew symmetric (again, in the sense of Definition 3.1). For convenience, a self-contained definition is included below:

Definition 3.2. *An $n \times n$ matrix polynomial $P(\lambda) = \sum_0^k \lambda^i A_i$ over a field $\mathbb{F}$ is said to be* skew-symmetric *if*

(a) $P(\lambda)^T = -P(\lambda)$, *equivalently if* $A_j^T = -A_j$ *for* $j = 0, \ldots, k$, and *if*

(b) $P_{ii}(\lambda) \equiv 0$ *for* $i = 1, \ldots, n$.

We now prove several basic properties of skew-symmetric matrices. Our main interest, of course, is in skew-symmetric matrices with polynomial entries. However, these properties are most naturally developed for matrices with entries in an arbitrary commutative ring.

The first result shows that the determinant of an $n \times n$ skew-symmetric matrix is zero whenever n is odd. This is well known for skew-symmetric matrices over the real or complex numbers. In fact, it is easy to prove when R is any field $\mathbb{F}$ with $\operatorname{char} \mathbb{F} \neq 2$, but is not so straightforward when $\mathbb{F}$ has characteristic two, or when the entries of the matrix are from a commutative ring which is not a field.

Lemma 3.3. *Suppose A is an $n \times n$ skew-symmetric matrix with entries in an arbitrary commutative ring R. Then* $\det A = 0$ *whenever n is odd.*

Proof. First recall the familiar proof, valid whenever the element $2 \in R$ is not a zero divisor. In this case we have

$$\begin{aligned} A^T = -A \quad &\Longrightarrow \quad \det(A^T) = \det(-A) \\ &\Longrightarrow \quad \det A = (-1)^n \det A = -\det A \\ &\Longrightarrow \quad 2 \det A = 0\,, \end{aligned}$$

and so $\det A = 0$, since 2 is not a zero divisor.

To obtain an argument that works for an arbitrary commutative ring R, we take a different approach. For a fixed odd n, consider the generic $n \times n$ skew-symmetric matrix

$$G_n(x_{12}, x_{13}, \ldots, x_{n-1,n}) := \begin{bmatrix} 0 & x_{12} & x_{13} & \cdots & & x_{1,n} \\ -x_{12} & 0 & x_{23} & \cdots & & x_{2,n} \\ -x_{13} & -x_{23} & 0 & & \cdots & x_{3,n} \\ \vdots & \vdots & & \ddots & & \vdots \\ & & & & 0 & x_{n-1,n} \\ -x_{1,n} & -x_{2,n} & & \cdots & -x_{n-1,n} & 0 \end{bmatrix}, \tag{3.1}$$

in which each of the $n(n-1)/2$ entries above the diagonal is a distinct variable. Then $\det G_n$ is a polynomial in the ring $\mathcal{R} := \mathbb{Z}[\,x_{12}, x_{13}, \ldots, x_{n-1,n}\,]$. The determinant $\det A$ of an arbitrary $n \times n$ skew-symmetric matrix A with entries from any commutative ring R can be evaluated simply by plugging in the values $a_{ij} \in R$ of the entries of A into the polynomial $\det G_n$. In other words, if $a_{ij} \in R$ are the entries of A, then $A = G_n(a_{12}, a_{13}, \ldots, a_{n-1,n})$, and

$$\det A \;=\; \det\big[\,G_n(a_{12}, a_{13}, \ldots, a_{n-1,n})\,\big] \;=\; \big(\det G_n\big)(a_{12}, a_{13}, \ldots, a_{n-1,n})\,.$$

Observe, however, that the "familiar" argument described in the first paragraph of this proof applies to $\det G_n$ as an element of the particular ring $\mathcal{R}$, because 2 is not a zero divisor in the ring $\mathcal{R}$. Thus when n is odd, $\det G_n$ is the *identically zero* polynomial in $\mathcal{R}$, so that evaluating $\det G_n$ with arbitrary entries from any commutative ring R will always produce the answer $\det A = 0$ in R. □

Next we prove that skew-symmetry is preserved under congruence transformations. Once again, this is immediate if the entries of the matrix are from a commutative ring R in which 2 is not a zero divisor, but it is not straightforward otherwise.

Lemma 3.4. *Let A, B, F be $n \times n$ matrices over an arbitrary commutative ring R such that $B = F^T AF$. If A is skew-symmetric, then so is B.*

Proof. If $2 \in R$ is not a zero divisor, then the result follows immediately from

$$B^T = \big(F^T AF\big)^T = F^T A^T F = -F^T AF = -B\,.$$

If 2 is a zero divisor in R, then we use the same technique as in the proof of Lemma 3.3. To this end, let G_n be the generic $n \times n$ skew-symmetric matrix as in (3.1), and

$$F_n(y_{11}, y_{12}, \ldots, y_{nn}) := \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{nn} \end{bmatrix}$$

the generic $n \times n$ (unstructured) matrix. Then both G_n and F_n can be viewed as matrices with entries in the polynomial ring $\widehat{\mathcal{R}} = \mathbb{Z}[\,x_{12}, x_{13}, \ldots, x_{n-1,n}, y_{11}, y_{12}, \ldots, y_{nn}\,]$, with $n^2 + n(n-1)/2$ independent commuting variables.

The matrices A and F can now be obtained by substituting the entries from A and F into the corresponding variables of the matrix polynomials G_n and F_n, whereas B can be obtained by substituting those entries exactly into the corresponding variables of the matrix polynomial $F_n^T G_n F_n$. By the argument in the first sentence of this proof, $F_n^T G_n F_n$ is skew-symmetric as a matrix with entries in the ring $\widehat{\mathcal{R}}$, since 2 is not a zero divisor in $\widehat{\mathcal{R}}$. Thus it follows, as in the proof of Lemma 3.3, that B is skew-symmetric as a matrix with entries in the ring R. □

Skew-symmetry is also preserved under inverses, as shown in the next lemma.

Lemma 3.5. *Suppose A is an $n \times n$ matrix with entries in a commutative ring R, and A is invertible. If A is skew-symmetric, then so is A^{-1}.*

Proof. The first condition of skew-symmetry, i.e., that $(A^{-1})^T = -A^{-1}$, comes immediately via the simple computation $(A^{-1})^T = (A^T)^{-1} = (-A)^{-1} = -A^{-1}$. The diagonal-entries condition in Definition 3.1 follows from the classical adjoint characterization of inverse matrices, i.e., $A^{-1} = (1/\det A) \cdot \operatorname{adj} A$. The existence of A^{-1} means that $\det A \neq 0$, so n is even by Lemma 3.3. Now observe that each diagonal entry $(\operatorname{adj} A)_{ii}$ is the determinant of an $(n-1) \times (n-1)$ *principal* submatrix of A. Such a submatrix is skew-symmetric of *odd* size, hence by Lemma 3.3 we have $(\operatorname{adj} A)_{ii} = 0$ for $i = 1, \ldots, n$. □

Finally we consider the various compounds of skew-symmetric matrices over an arbitrary commutative ring, and determine what their structure is. Note that the proof is very similar to that used for Lemma 3.5.

Lemma 3.6. *Suppose A is an $n \times n$ skew-symmetric matrix with entries in an arbitrary commutative ring R. Then the ℓth compound $\mathcal{C}_\ell(A)$ is skew-symmetric when ℓ is odd, but symmetric when ℓ is even.*

Proof. Since A is skew-symmetric we have $A^T = -A$. Taking the ℓth compound of both sides gives $\mathcal{C}_\ell(A^T) = \mathcal{C}_\ell(-A)$, which by Theorem 2.8 simplifies to $\mathcal{C}_\ell(A)^T = (-1)^\ell \mathcal{C}_\ell(A)$. Thus when ℓ is even we immediately see that $\mathcal{C}_\ell(A)$ is symmetric. When ℓ is odd, each diagonal entry of $\mathcal{C}_\ell(A)$ is the determinant of an $\ell \times \ell$ *principal* submatrix of A, which by the argument used in the proof of Lemma 3.5 is always zero. Thus $\mathcal{C}_\ell(A)$ is skew-symmetric when ℓ is odd. □

Of course we are primarily interested in the result of Lemma 3.6 when A is a skew-symmetric matrix polynomial, so we illustrate it with the following example.

Example 3.7. Let $\mathbb{F} = \mathbb{Z}_2$ and consider the skew-symmetric matrix polynomial

$$P(\lambda) = \begin{bmatrix} 0 & 1 & \lambda & \lambda+1 \\ 1 & 0 & \lambda^2+\lambda & 1 \\ \lambda & \lambda^2+\lambda & 0 & \lambda \\ \lambda+1 & 1 & \lambda & 0 \end{bmatrix},$$

which we view as a skew-symmetric matrix with entries in the commutative ring $R = \mathbb{Z}_2[\,\lambda\,]$. Then we obtain

$$\mathcal{C}_2\big(P(\lambda)\big) = \begin{bmatrix} 1 & \lambda & \lambda+1 & \lambda^2+\lambda & 1 & \lambda^3 \\ \lambda & \lambda^2 & \lambda^2+\lambda & \lambda^3+\lambda^2 & \lambda^3 & \lambda^2 \\ \lambda+1 & \lambda^2+\lambda & \lambda^2+1 & 0 & \lambda+1 & \lambda^2+\lambda \\ \lambda^2+\lambda & \lambda^3+\lambda^2 & 0 & \lambda^4+\lambda^2 & \lambda^2+\lambda & \lambda^3+\lambda^2 \\ 1 & \lambda^3 & \lambda+1 & \lambda^2+\lambda & 1 & \lambda \\ \lambda^3 & \lambda^2 & \lambda^2+\lambda & \lambda^3+\lambda^2 & \lambda & \lambda^2 \end{bmatrix}$$

and

$$\mathcal{C}_3\big(P(\lambda)\big) = \begin{bmatrix} 0 & \lambda^3+\lambda & \lambda^4+\lambda^2 & \lambda^5+\lambda^3+\lambda^2 \\ \lambda^3+\lambda & 0 & \lambda^4+\lambda^2+\lambda & \lambda^3+\lambda \\ \lambda^4+\lambda^2 & \lambda^4+\lambda^2+\lambda & 0 & \lambda^4+\lambda^2 \\ \lambda^5+\lambda^3+\lambda^2 & \lambda^3+\lambda & \lambda^4+\lambda^2 & 0 \end{bmatrix}.$$

So indeed we see that $\mathcal{C}_2\big(P(\lambda)\big)$ is symmetric (but not skew-symmetric, since it does not have an all-zeroes diagonal), while $\mathcal{C}_3\big(P(\lambda)\big)$ is skew-symmetric.

Remark 3.8. The proof technique used in Lemmas 3.3 and 3.4 is known informally by various names, such as the "method of the generic example" or the "principle of permanence of identities" [4, p. 456-7]. See also [26, p. 334-5] or [33, p. 372-4] for another use of this technique in developing the basic properties of the Pfaffian.

4 Skew-Smith Form

In this section we develop our main result, the "Skew-Smith Form", starting with the following result about the *first two* diagonal entries in the Smith form of a skew-symmetric matrix polynomial.

Lemma 4.1. *Suppose $P(\lambda)$ is an $n \times n$ skew-symmetric matrix polynomial (regular or singular) over an arbitrary field $\mathbb{F}$ with Smith form*

$$D(\lambda) = \operatorname{diag}\big(d_1(\lambda)\,,\, d_2(\lambda)\,,\ldots\,,\, d_n(\lambda)\big)\,,$$

where $n \geq 2$. Then $d_1(\lambda) = d_2(\lambda)$.

Proof. By Theorem 2.2 we know that $d_1(\lambda)$ is the GCD of all the entries of $P(\lambda)$. Now if $d_1(\lambda) = 0$, then $D(\lambda) = 0$, and we are done. So let us assume that $d_1(\lambda) \not\equiv 0$. Letting $g(\lambda)$ be the GCD of all the 2×2 minors of $P(\lambda)$, then we know from Theorem 2.2 that

$$d_1(\lambda)d_2(\lambda) \;=\; g(\lambda)\,. \tag{4.1}$$

We aim to show that $g(\lambda) = d_1^2(\lambda)$; from (4.1) it will then follow that $d_1(\lambda) = d_2(\lambda)$, as desired. Our strategy is to show that $d_1^2(\lambda) \,|\, g(\lambda)$ and $g(\lambda) \,|\, d_1^2(\lambda)$, from which the equality $g(\lambda) = d_1^2(\lambda)$ then follows because both are monic.

Combining $d_1(\lambda) \,|\, d_2(\lambda)$ from the Smith form together with (4.1), we see immediately that $d_1^2(\lambda) \,|\, g(\lambda)$. To see why $g(\lambda) \,|\, d_1^2(\lambda)$, first observe that the GCD of any set S must always divide the GCD of any subset of S; this follows directly from the definition of GCD. Thus if we can find some *subset* of all the 2×2 minors of $P(\lambda)$ whose GCD is $d_1^2(\lambda)$, then $g(\lambda) \,|\, d_1^2(\lambda)$ and we are done. We claim that the subset $\mathcal{P}$ of all *principal* 2×2 minors of $P(\lambda)$ is one such subset. To see this let $\mathcal{U}$ denote the set $\big\{P_{ij}(\lambda)\big\}_{i<j}$ of all polynomial entries in the strictly upper triangular part of $P(\lambda)$. Then clearly $\gcd(\mathcal{U}) = d_1(\lambda)$. But it is easy to see that the subset $\mathcal{P}$ consists of exactly the squares of all the elements of $\mathcal{U}$, i.e.,

$$\mathcal{P} = \Big\{\big(P_{ij}(\lambda)\big)^2\Big\}_{i<j}\,.$$

Then by Lemma 2.3 we have $\gcd(\mathcal{P}) = d_1^2(\lambda)$, and the proof is complete. □

All the necessary tools are now available to prove our main result, the characterization of all possible Smith forms for skew-symmetric matrix polynomials over an arbitrary field.

Theorem 4.2 (Skew-Smith form)**.** *Suppose that*

$$D(\lambda) = \operatorname{diag}\big(d_1(\lambda),\, d_2(\lambda), \ldots,\, d_\ell(\lambda),\, 0, \ldots, 0\big) \tag{4.2}$$

is an $n \times n$ diagonal matrix polynomial over an arbitrary field $\mathbb{F}$, such that $d_j(\lambda)$ is monic for $j = 1, \ldots, \ell$, and $d_j(\lambda) \,|\, d_{j+1}(\lambda)$ for $j = 1, \ldots, \ell - 1$. Then $D(\lambda)$ is the Smith form of some $n \times n$ skew-symmetric matrix polynomial $P(\lambda)$ over $\mathbb{F}$ if and only if the following conditions hold:

(a) $\ell = 2m$ *is even;*

(b) $d_{2i-1}(\lambda) = d_{2i}(\lambda)$ *for* $i = 1, \ldots, m$ (*pairing of adjacent invariant polynomials*).

Proof. ($\Rightarrow$): First observe that if $d_1(\lambda) = 0$, then $D(\lambda) = 0$, and conditions (a) and (b) hold trivially with $\ell = 0$. So from now on let us assume that $d_1(\lambda) \neq 0$.

Next we show that whenever $d_j \neq 0$ for some *odd* index $j = 2i - 1$, then $n > j$, and $d_{2i}(\lambda) = d_{2i-1}(\lambda)$. To see that $d_{2i-1}(\lambda) \neq 0$ implies $n > 2i - 1$, let us suppose instead that $n = 2i - 1$. Then we would have

$$\det P(\lambda) \;=\; c \det D(\lambda) \;=\; d_1(\lambda) d_2(\lambda) \cdots d_{2i-1}(\lambda) \not\equiv 0\,.$$

But this would contradict the fact that $\det P(\lambda)$, being the determinant of an $n \times n$ skew-symmetric matrix over the commutative ring $R = \mathbb{F}[\,\lambda\,]$ with n *odd*, must be identically zero by Lemma 3.3. Thus $n > 2i - 1$.

With $n > 2i - 1$, we can now apply Theorem 2.9 to the compound matrix $\mathcal{C}_{2i-1}\big(P(\lambda)\big)$ to conclude that the first two invariant polynomials of $\mathcal{C}_{2i-1}\big(P(\lambda)\big)$ are

$$s_1(\lambda) = \left(\prod_{j=1}^{2i-2} d_j(\lambda)\right) \cdot d_{2i-1}(\lambda) \quad \text{and} \quad s_2(\lambda) = \left(\prod_{j=1}^{2i-2} d_j(\lambda)\right) \cdot d_{2i}(\lambda)\,.$$

But $\mathcal{C}_{2i-1}\big(P(\lambda)\big)$ is skew-symmetric by Lemma 3.6, so $s_1(\lambda) = s_2(\lambda)$ by Lemma 4.1, and hence $d_{2i}(\lambda) = d_{2i-1}(\lambda) \not\equiv 0$.

Finally, the fact that $d_{2i-1}(\lambda) \neq 0$ implies $d_{2i}(\lambda) \neq 0$ forces ℓ to be *even*, and the proof of the ($\Rightarrow$) direction is complete.

($\Leftarrow$): Suppose $D(\lambda)$ satisfies conditions (a) and (b). Then there is a simple way to construct a skew-symmetric $P(\lambda)$ with Smith form $D(\lambda)$. First observe that for any scalar polynomial $f(\lambda)$, the 2×2 skew-symmetric polynomial

$$\begin{bmatrix} 0 & f(\lambda) \\ -f(\lambda) & 0 \end{bmatrix} =: \operatorname{antidiag}\big(-f(\lambda), f(\lambda)\big)$$

is unimodularly equivalent to $\operatorname{diag}\big(f(\lambda), f(\lambda)\big)$. Thus if we replace each pair of consecutive invariant polynomials $d_{2j-1}(\lambda) = d_{2j}(\lambda)$ in $D(\lambda)$ by the 2×2 skew-symmetric block $\operatorname{antidiag}\big(-d_{2j}(\lambda), d_{2j}(\lambda)\big)$, then we will have a block-diagonal $P(\lambda)$ that is skew-symmetric with Smith form $D(\lambda)$. □

From Theorem 4.2 and the construction of the special skew-symmetric matrix polynomial in the proof of the sufficiency of the conditions (a) and (b) in Theorem 4.2, we immediately obtain the following corollary, giving a structured "Smith-like" canonical form for skew-symmetric matrix polynomials under unimodular equivalence.

Corollary 4.3 (Skew-symmetric canonical form)**.**
Let $P(\lambda)$ be a skew-symmetric $n \times n$ matrix polynomial over an arbitrary field $\mathbb{F}$. Then there exists $r \in \mathbb{N}$ with $2r \leq n$ and unimodular matrix polynomials $E(\lambda), F(\lambda)$ over $\mathbb{F}$ such that

$$E(\lambda)P(\lambda)F(\lambda) = \begin{bmatrix} 0 & d_1(\lambda) \\ -d_1(\lambda) & 0 \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} 0 & d_r(\lambda) \\ -d_r(\lambda) & 0 \end{bmatrix} \oplus 0_{n-2r} =: K(\lambda), \quad (4.3)$$

where d_j is monic for $j = 1, \ldots, r$ and $d_j \,|\, d_{j+1}$ for $j = 1, \ldots, r-1$. Furthermore, the polynomials $d_j(\lambda)$ for $j = 1, \ldots, r$ are uniquely determined by these conditions.

5 Canonical Form under Unimodular Congruence

In Corollary 4.3, both the original polynomial $P(\lambda)$ and the canonical form $K(\lambda)$ are skew-symmetric, so it is natural to wonder whether $K(\lambda)$ can be obtained from $P(\lambda)$ via *structure-preserving* transformations. Here the appropriate structure-preserving transformations are *unimodular congruences*, introduced next.

Definition 5.1. *Let $P(\lambda), Q(\lambda)$ be two $n \times n$ matrix polynomials. Then we say that P and Q are* unimodularly congruent *if there exists a unimodular matrix polynomial $F(\lambda)$ such that*

$$F(\lambda)^T P(\lambda) F(\lambda) = Q(\lambda).$$

The transformation $P \mapsto F^T P F$ is called a unimodular congruence transformation.

Remark 5.2. Taking $F(\lambda)$ to be an *elementary* unimodular (column) transformation (see [19]), i.e., either the interchange of columns ℓ and k, the addition of the $f(\lambda)$-multiple of the kth column to the ℓth column, or multiplication of the kth column by a nonzero constant $a \in \mathbb{F}$, the corresponding unimodular congruence transformation $P(\lambda) \mapsto F(\lambda)^T P(\lambda) F(\lambda)$ simultaneously performs this action on both the columns and the corresponding rows. Such elementary unimodular congruences will be used extensively in the proof of Theorem 5.4.

That unimodular congruences are indeed structure-preserving in this context is a consequence of the following special case of Lemma 3.4.

Corollary 5.3. *Suppose $P(\lambda)$ and $Q(\lambda)$ are two $n \times n$ matrix polynomials that are unimodularly congruent. If $P(\lambda)$ is skew-symmetric, then so is $Q(\lambda)$.*

The next result shows that the canonical form (4.3) can still be attained when the unimodular equivalence transformations used in Corollary 4.3 are restricted to unimodular congruence transformations. In particular, this provides an alternative proof for Theorem 4.2. But this result has further ramifications, to be briefly discussed immediately after the proof.

Theorem 5.4 (Skew-symmetric canonical form under unimodular congruence)**.**
Let $P(\lambda)$ be a skew-symmetric $n \times n$ matrix polynomial over an arbitrary field $\mathbb{F}$. Then there exists $r \in \mathbb{N}$ with $2r \leq n$ and a unimodular matrix polynomial $F(\lambda)$ over $\mathbb{F}$ such that

$$F(\lambda)^T P(\lambda) F(\lambda) = \begin{bmatrix} 0 & d_1(\lambda) \\ -d_1(\lambda) & 0 \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} 0 & d_r(\lambda) \\ -d_r(\lambda) & 0 \end{bmatrix} \oplus 0_{n-2r} =: K(\lambda)\,,$$

where d_j is monic for $j = 1, \ldots, r$, and $d_j \mid d_{j+1}$ for $j = 1, \ldots, r-1$. Moreover, the canonical form $K(\lambda)$ is unique.

Proof. The argument given here adapts the usual algorithmic proof (see, e.g., [19]) of the existence of a Smith form for general matrix polynomials, restricting the transformations used to unimodular *congruences* only, so that skew-symmetry is preserved (by Corollary 5.3) throughout the reduction. The strategy is first to show that any $n \times n$ skew-symmetric $P(\lambda)$ with $n \geq 3$ can be transformed by unimodular congruence into the block-diagonal form

$$B(\lambda) \;=\; \begin{bmatrix} 0 & d(\lambda) & \\ -d(\lambda) & 0 & \\ & & Q(\lambda) \end{bmatrix}, \tag{5.1}$$

where $Q(\lambda)$ is $(n-2) \times (n-2)$ skew-symmetric, and such that $B(\lambda)$ satisfies

"Property $\mathcal{D}$": $d(\lambda)$ is monic, and $d(\lambda)$ divides all the entries of $Q(\lambda)$.

A simple induction on n then completes the argument.

We begin by defining several basic operations achievable by combinations of elementary unimodular congruences; these operations will be used repeatedly in the reduction of $P(\lambda)$ to the form (5.1). For several of these operations we will need to refer to the set

$$\mathcal{R}_2(P) := \{p_{ij}(\lambda) \mid i = 1, 2 \text{ and } i < j\}$$

of all entries in the first two rows of the upper triangular part of $P(\lambda)$.

- "*Interchange* of the $(1, 2)$ and (i, j) entries":
  Any entry $p_{ij}(\lambda)$ with $i < j$ can be interchanged with $p_{12}(\lambda)$ via the congruence that interchanges rows 1 and i simultaneously with columns 1 and i, followed by a simultaneous interchange of columns (and rows) 2 and j.

- "*Degree-reduction* of the $(1, 2)$-entry":
  `Input:` any skew-symmetric P such that $p_{12} \neq 0$ and $p_{12} \nmid \mathcal{R}_2(P)$.
  `Output:` a unimodular congruence transforming P into a skew-symmetric $\widetilde{P}$ such that $\widetilde{p}_{12} \neq 0$ and $\deg \widetilde{p}_{12} < \deg p_{12}$.

  Since $p_{12} \nmid \mathcal{R}_2(P)$, we can choose some nonzero p_{ij} with $i = 1$ or $i = 2$ and $3 \leq j \leq n$ such that p_{12} does *not* divide p_{ij}. If $i = 2$ then the congruence that interchanges the first two rows (and columns) will move p_{ij} into the first row, and replace p_{12} by the essentially equivalent $-p_{12}$. Thus without loss of generality we may assume that p_{ij} is in the first row, so from now on $i = 1$. There are two cases to consider, depending on the degrees of p_{12} and p_{1j}. If $\deg p_{12} \leq \deg p_{1j}$, then we may write $p_{1j}(\lambda) = q(\lambda)p_{12}(\lambda) + r_{1j}(\lambda)$, where the remainder $r_{1j}(\lambda)$ is nonzero and $\deg r_{1j} < \deg p_{12}$. Now do the elementary unimodular congruence that adds the $-q(\lambda)$-multiple of column 2 to column j (and similarly for rows). This leaves r_{1j} in the $(1, j)$-position, so an interchange of the $(1, 2)$ and $(1, j)$ entries (as described in the previous bullet) completes the degree reduction operation. If $\deg p_{12} > \deg p_{1j}$, then simply doing an interchange of the $(1, 2)$ and $(1, j)$ entries completes the degree reduction operation.

- "*Sweep* of the first two rows (and columns)":
  `Input:` any skew-symmetric P such that $p_{12} \neq 0$ and $p_{12} \mid \mathcal{R}_2(P)$.
  `Output:` a unimodular congruence transforming P into a skew-symmetric $\widetilde{P}$ in the block-diagonal form (5.1), with $\widetilde{p}_{12} = p_{12} \neq 0$.
  (Note that this $\widetilde{P}$ may not necessarily satisfy property $\mathcal{D}$.)

  Since $p_{12} \mid \mathcal{R}_2(P)$, we can write $p_{ij}(\lambda) = p_{12}(\lambda)q_{ij}(\lambda)$ for each $i = 1, 2$ and $3 \leq j \leq n$. Then the elementary unimodular congruences that add the $-q_{1j}$ multiple of column 2 to the jth column for $j = 3, \ldots, n$ (and similarly for rows) will zero out the $(1, 3)$ through $(1, n)$ entries, while leaving the $(2, 3)$ through $(2, n)$ entries unchanged. To zero out the $(2, 3)$ through $(2, n)$ entries, do the elementary unimodular congruences that add the $+q_{2j}$ multiple of column 1 to the jth column (and similarly for rows) for $j = 3, \ldots, n$. The composition of all these congruences constitutes a *sweep* of the first two rows (and columns).

With these operations in hand, we can now describe the key part of the reduction of P to canonical form K, i.e., the reduction of P to (5.1) satisfying property $\mathcal{D}$. Using an

appropriate *interchange*, we initialize the reduction process by ensuring that the $(1,2)$-entry is nonzero, then proceed as follows.

Step 1 (achieving block-diagonal form (5.1)): Check whether p_{12} divides $\mathcal{R}_2(P)$ or not. As long as p_{12} does not divide all the entries of $\mathcal{R}_2(P)$, continue doing *degree reductions* of the $(1,2)$-entry. In the extreme case the $(1,2)$-entry reduces all the way to a nonzero constant (with degree zero); in this case we would clearly have a $(1,2)$-entry that divides every entry of the first two rows. Thus in general we can be sure that after finitely many such degree reductions, one eventually obtains a skew-symmetric $\widetilde{P}$ such that $\widetilde{p}_{12} \mid \mathcal{R}_2(\widetilde{P})$. Once this divisibility property is attained, then doing a *sweep* to $\widetilde{P}$ yields the block-diagonal form (5.1).

Step 2 (block-diagonal form (5.1) with property $\mathcal{D}$): With a skew-symmetric $\widehat{P}$ in block-diagonal form (5.1), check whether $\widehat{p}_{12}$ divides all the entries of the block Q or not. If so, then Step 2 is completed by a congruence that simultaneously scales the first row and column to make $\widehat{p}_{12}$ monic. If not, then pick any nonzero entry $q_{ij}(\lambda)$ of Q such that $\widehat{p}_{12}$ does not divide q_{ij}, and do the congruence that simultaneously adds the ith row of Q (i.e., the $i+2$nd row of $\widehat{P}$) to the first row of $\widehat{P}$, and the ith column of Q (the $i+2$nd column of $\widehat{P}$) to the first column of $\widehat{P}$. Then return to Step 1 with this no-longer-in-block-diagonal-form skew-symmetric polynomial.

Observe that every return from Step 2 to Step 1 results in at least one further degree reduction. Thus after finitely many times through this loop we will achieve block-diagonal form (5.1) that satisfies property $\mathcal{D}$. (At worst, the $(1,2)$-entry in (5.1) reduces all the way to a nonzero constant, which after scaling clearly results in property $\mathcal{D}$ holding.)

The complete reduction of $P(\lambda)$ to canonical form $K(\lambda)$ can now be formulated as an induction on n. Clearly if $n = 1$ or $n = 2$, then the assertion of the theorem is trivially true. So suppose that $n \geq 3$. By the previous paragraph there is a unimodular congruence that transforms $P(\lambda)$ into the block-diagonal $B(\lambda)$ as in (5.1), satisfying property $\mathcal{D}$. The divisibility property $\mathcal{D}$ now implies that $B(\lambda)$ can be factored as

$$B(\lambda) \;=\; d(\lambda) \begin{bmatrix} 0 & 1 & \\ -1 & 0 & \\ & & \widetilde{Q}(\lambda) \end{bmatrix},$$

where $\widetilde{Q}(\lambda)$ is an $(n-2)\times(n-2)$ skew-symmetric polynomial such that $Q(\lambda) = d(\lambda)\widetilde{Q}(\lambda)$. By the induction hypothesis $\widetilde{Q}(\lambda)$ is unimodularly congruent to a skew-symmetric canonical form $\widetilde{K}(\lambda)$, and thus $P(\lambda)$ is unimodularly congruent to the canonical form

$$K(\lambda) \;=\; \begin{bmatrix} 0 & d(\lambda) & \\ -d(\lambda) & 0 & \\ & & d(\lambda)\widetilde{K}(\lambda) \end{bmatrix}.$$

This completes the proof of the existence part of the theorem.

For the uniqueness part of the theorem, observe that the Smith form of K, and hence also the Smith form of P, is just $\operatorname{diag}\big(d_1(\lambda), d_1(\lambda), d_2(\lambda), d_2(\lambda), \ldots, d_r(\lambda), d_r(\lambda), 0_{n-2r}\big)$. Thus the uniqueness of $K(\lambda)$ follows from the uniqueness of the Smith form of $P(\lambda)$. $\square$

Remark 5.5. It is important to note that the result of Theorem 5.4 is far from being just a trivial addendum to Corollary 4.3. By contrast to the situation for skew-symmetric

matrix polynomials, there are *many* classification problems where restricting the class of equivalence transformations leads to the introduction of additional invariants.

For example, consider the problem of classifying symmetric pencils over the field $\mathbb{R}$. It is easy to see that the pencils

$$L_1(\lambda) = \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad L_2(\lambda) = \lambda \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

are strictly equivalent, since there exist nonsingular matrices $P, Q \in \mathbb{R}^{2\times 2}$, e.g., $P = I_2$ and $Q = \operatorname{diag}(1, -1)$, such that $PL_1(\lambda)Q = L_2(\lambda)$. But $L_1(\lambda)$ and $L_2(\lambda)$ are *not congruent*, i.e., there is no single nonsingular matrix $P \in \mathbb{R}^{2\times 2}$ such that $PL_1(\lambda)P^T = L_2(\lambda)$; this follows immediately from Sylvester's Law of Inertia.

Thus an important consequence of Theorem 5.4 is that, for skew-symmetric matrix polynomials, there are no additional invariants introduced when the set of unimodular equivalence transformations is restricted to the set of unimodular congruence transformations.

Remark 5.6. It is worth highlighting a property of skew-symmetric matrix polynomials that was crucial for the success of our reduction of such polynomials to Smith-like form under unimodular congruence in Theorem 5.4; they have *all zeroes on the diagonal.* Attempting to reduce matrix polynomials with other symmetry structures (like symmetric, T-alternating, Hermitian, or $*$-alternating matrix polynomials) in an analogous fashion fails due to the possible presence of nonzero entries on the diagonal of the matrix polynomial. It is an open problem to derive Smith-like forms under unimodular congruence for matrix polynomials with other types of symmetry structures.

6 Applications of the Skew-Smith Form

The Skew-Smith form theorem has a number of important applications that we explore in this section. In particular, we are easily able to extend some well-known properties of the elementary divisors of skew-symmetric *pencils* [50] to general skew-symmetric polynomials. We also investigate the impact of the Skew-Smith form theorem on the question of the existence of structured linearizations for skew-symmetric matrix polynomials. Using the Smith form of structured classes of polynomials as a means to probe for any obstructions to the existence of structured linearizations has been a continuing theme starting with the papers [37, 38], and has been one of our main motivations for studying the Smith form in the first place. In addition, we consider the properties of matrices expressible as the product of two skew-symmetric matrices, the Smith-McMillan form of skew-symmetric *rational* matrices $R(\lambda)$, and the existence of minimal symmetric factorizations of such $R(\lambda)$.

We begin, though, with a well-known result that is a special case of the Skew-Smith theorem. Applying Theorem 4.2 to a matrix polynomial of degree zero, i.e., to a constant matrix, immediately yields the following result:

Corollary 6.1. *Let $A \in \mathbb{F}^{n\times n}$ be a skew-symmetric matrix over an arbitrary field $\mathbb{F}$. Then the rank of A is even.*

Thus Theorem 4.2 gives a new and independent (albeit somewhat roundabout) proof of this well-known fact.

6.1 Elementary divisors of skew-symmetric matrix polynomials

A fundamental property of the elementary divisors of an arbitrary skew-symmetric matrix polynomial $P(\lambda)$ can be immediately deduced from the Skew-Smith form — every elementary divisor of $P(\lambda)$ has *even multiplicity.* Since $\operatorname{rev} P$ is also skew-symmetric, this holds for every infinite elementary divisor of $P(\lambda)$ as well. Thus we obtain the following result.

Theorem 6.2. *Let $P(\lambda)$ be an $n \times n$ skew-symmetric matrix polynomial, regular or singular, over an arbitrary field $\mathbb{F}$. Then every (finite or infinite) elementary divisor of $P(\lambda)$ occurs with even multiplicity.*

When $\mathbb{F}$ is not an algebraically closed field, then the elementary divisors of $P(\lambda)$ will not necessarily all be powers of linear factors. Note, however, that the uniqueness property of Smith forms implies that the Smith form of $P(\lambda)$ is unchanged by viewing $P(\lambda)$ as a matrix polynomial over any field extension $\widetilde{\mathbb{F}} \supseteq \mathbb{F}$, including the algebraic closure $\overline{\mathbb{F}}$. Thus the invariant polynomials of $P(\lambda)$ are insensitive to change of field, and pairing of elementary divisors will still be present, regardless of whether P is viewed as a matrix polynomial over $\mathbb{F}$ or as a matrix polynomial over any extension field $\widetilde{\mathbb{F}}$.

Note that in the classic paper [50] it is shown that all (finite and infinite) elementary divisors of any skew-symmetric matrix *pencil* over an algebraically closed field (of characteristic *different* from two) occur with even multiplicity. Theorem 6.2 generalizes and extends this result to skew-symmetric matrix polynomials of arbitrary degree k, over an arbitrary field.

Remark 6.3. For $n \times n$ skew-symmetric matrices A with entries in a commutative ring R, Lemma 3.3 shows that $\det A$ is always zero when n is odd. When n is even, then it is well known that $\det A$ is always a perfect square in R, i.e., $\det A = p^2$ for some $p \in R$; this fact is usually proved using the notion of the *Pfaffian* (see [26] or [33]). The results in Theorems 4.2 and 6.2 can be viewed as a kind of generalization and simultaneous refinement of both the perfect square result and the zero determinant result, at least for the case of skew-symmetric matrices $A = P(\lambda)$ with entries in the specific ring $R = \mathbb{F}[\,\lambda\,]$. Observe, in particular, that even when $P(\lambda)$ is singular, with $\det P(\lambda) = 0$ and $r = \operatorname{rank} P(\lambda) < n$, there is still a kind of hidden "perfect squareness" about $P(\lambda)$; for any even ℓ with $0 < \ell \leq r$, the GCD of all $\ell \times \ell$ minors of $P(\lambda)$ will be a nonzero perfect square in $R = \mathbb{F}[\,\lambda\,]$.

6.2 Products of two skew-symmetric matrices

Skew-Hamiltonian matrices are a much-studied class [7, 13, 15, 23, 43, 44, 51, 52] of even-sized structured matrices, closely related to Hamiltonian and symplectic matrices. Among the various ways to characterize these matrices, the one most relevant to this paper is the following: a $2m \times 2m$ matrix W is skew-Hamiltonian if and only if $W = JK$, where J is the particular $2m \times 2m$ skew-symmetric matrix

$$J = \begin{bmatrix} 0 & I_m \\ -I_m & 0 \end{bmatrix}$$

that defines the standard symplectic bilinear form, and K is an arbitrary $2m \times 2m$ skew-symmetric matrix. Up to this point we have been primarily concerned with investigating the spectral properties of skew-symmetric matrix *polynomials* of all sizes, the natural extension of skew-Hamiltonian matrices to the matrix polynomial world.

In this section we consider instead a generalization of skew-Hamiltonian structure in a somewhat different direction, this time *within* the world of matrices. In particular, we consider the class of all matrices expressible as the product of an *arbitrary pair* of $n \times n$ skew-symmetric matrices, where n is even *or* odd. It is known that all nonzero eigenvalues of any such "product-of-two-skew-symmetric-matrices" have even multiplicity [14]. We easily recover this property as a corollary of the Skew-Smith form.

Corollary 6.4. *Let $A := BC$, where $B, C \in \mathbb{F}^{n\times n}$ are any two skew-symmetric matrices over an arbitrary field $\mathbb{F}$. Then every nonzero eigenvalue of A (in any extension field of $\mathbb{F}$) has even multiplicity, while the multiplicity of the eigenvalue 0 has the same parity as n. In particular, if $r = \min\{\operatorname{rank} B, \operatorname{rank} C\}$, then there exists a scalar polynomial $p(\lambda) \in \mathbb{F}[\,\lambda\,]$ of degree $r/2$ such that the characteristic polynomial of A is $\lambda^{n-r}[\,p(\lambda)\,]^2$.*

Proof. Since the characteristic polynomials of BC and CB are identical [21], we may assume without loss of generality that the first term in the product BC has the minimal rank r. Note that $r = \operatorname{rank} B$ is even, by Corollary 6.1. Letting $Q \in \mathbb{F}^{n\times n}$ be any nonsingular matrix such that $Q^{-1}B$ is in row echelon form, then from Lemma 3.4 it follows that $\widetilde{B} := Q^{-1}BQ^{-T}$ is skew-symmetric and block-diagonal of the form $\widetilde{B} = \operatorname{diag}(\widehat{B}, 0_{n-r})$, where $\widehat{B}$ is a *nonsingular* $r \times r$ skew-symmetric matrix. Now A is similar to

$$\widetilde{A} := Q^{-1}AQ = (Q^{-1}BQ^{-T})\,(Q^TCQ) =: \widetilde{B}\widetilde{C}\,,$$

where $\widetilde{C}$ is also skew-symmetric. Partitioning $\widetilde{C}$ conformally with $\widetilde{B} = \operatorname{diag}(\widehat{B}, 0_{n-r})$ as

$$\widetilde{C} = \begin{bmatrix} \widehat{C} & \widehat{D} \\ \widehat{E} & \widehat{F} \end{bmatrix},$$

where $\widehat{C}$ is $r \times r$ and skew-symmetric, we see that

$$\widetilde{A} = \begin{bmatrix} \widehat{B}\widehat{C} & \widehat{B}\widehat{D} \\ 0 & 0 \end{bmatrix}. \tag{6.1}$$

Thus the characteristic polynomial of $\widetilde{A}$ has the form $\lambda^{n-r}q(\lambda)$, where $q(\lambda)$ is the characteristic polynomial of $\widehat{B}\widehat{C}$. Characteristic polynomials are always monic, so the characteristic polynomial of the product $\widehat{B}\widehat{C}$ must be exactly the same as the determinant of the Smith form of the *regular* pencil $\lambda\widehat{B}^{-1} - \widehat{C}$, which is a skew-symmetric pencil by Lemma 3.5. The desired result now follows immediately from Theorem 4.2. □

Remark 6.5. This result for characteristic polynomials, proved in Corollary 6.4 for matrices $A = BC$ over an arbitrary *field*, also holds when B and C are skew-symmetric matrices with entries from an *arbitrary commutative ring*. This was shown in [14].

The elementary divisor structure of skew-Hamiltonian matrices is also well known (every elementary divisor has even multiplicity), and has been proved in various ways [15, 23]. The next result shows the extent to which this even multiplicity property is retained by general "product-of-two-skew-symmetrics" matrices.

Corollary 6.6. *Let $A := BC$, where $B, C \in \mathbb{F}^{n\times n}$ are any two skew-symmetric matrices over an arbitrary field $\mathbb{F}$. Then every elementary divisor of A has even multiplicity, with the possible exception of elementary divisors associated with the eigenvalue 0. When either B or C is nonsingular, then* all *elementary divisors of A have even multiplicity.*

We will see that Corollary 6.6 is rather easy to prove when either B or C is nonsingular. Indeed, this case has been previously treated by a completely different approach in [23]. However, some extra work is needed to handle the case when both B and C are singular, in the form of two preliminary lemmas (both of independent interest). Recall that unimodular equivalence of matrix polynomials P and Q is denoted by $P \sim Q$.

Lemma 6.7 (Exchange Lemma)**.**
Suppose $f, g, h \in \mathbb{F}[\lambda]$ are scalar polynomials over an arbitrary field $\mathbb{F}$ such that f is relatively prime to both g and h. Then

(a) $\begin{bmatrix} fg & 0 \\ 0 & h \end{bmatrix} \sim \begin{bmatrix} g & 0 \\ 0 & fh \end{bmatrix}$,

(b) *For any $k, \ell \in \mathbb{N}$, we have* $\begin{bmatrix} f^k g & 0 \\ 0 & f^\ell h \end{bmatrix} \sim \begin{bmatrix} f^\ell g & 0 \\ 0 & f^k h \end{bmatrix}$.

Proof. (a) Since f is relatively prime to g and h, there exist scalar polynomials $a, b, c, d \in \mathbb{F}[\lambda]$ such that $af + bg \equiv 1$ and $cf + dh \equiv 1$. Now from these scalar polynomials define unimodular $E(\lambda)$ and $F(\lambda)$ as products of elementary unimodulars:

$$E(\lambda) := \begin{bmatrix} 1 & -a \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -f \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix} = \begin{bmatrix} a - afc + c & 1 - af \\ fc - 1 & f \end{bmatrix},$$

$$F(\lambda) := \begin{bmatrix} 1 & 0 \\ dg & 1 \end{bmatrix} \begin{bmatrix} 1 & -bh \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -bh \\ dg & 1 - bhdg \end{bmatrix}.$$

Using the relations $af + bg \equiv 1$ and $cf + dh \equiv 1$, one can simplify $E(\lambda)$ to $\begin{bmatrix} adh + c & bg \\ -dh & f \end{bmatrix}$, and then verify that

$$E(\lambda) \begin{bmatrix} fg & 0 \\ 0 & h \end{bmatrix} F(\lambda) = \begin{bmatrix} g & 0 \\ 0 & fh \end{bmatrix},$$

thus proving (a).

(b) Begin by observing that f being relatively prime to g and h implies that f^m is also relatively prime to g and h, for any $m \in \mathbb{N}$. Now for exponent pairs $k, \ell \in \mathbb{N}$, there are three cases to consider. If $k = \ell$, then (b) holds trivially. If $k > \ell \geq 0$, then using part (a) we see that

$$\begin{bmatrix} f^k g & 0 \\ 0 & f^\ell h \end{bmatrix} = f^\ell \begin{bmatrix} f^{k-\ell} g & 0 \\ 0 & h \end{bmatrix} \sim f^\ell \begin{bmatrix} g & 0 \\ 0 & f^{k-\ell} h \end{bmatrix} = \begin{bmatrix} f^\ell g & 0 \\ 0 & f^k h \end{bmatrix}. \tag{6.2}$$

The same kind of argument as in (6.2) proves (b) for the case $0 \leq k < \ell$, by factoring out f^k rather than f^ℓ. □

Lemma 6.8 (Cancellation Lemma)**.**
Let $A(\lambda)$ and $B(\lambda)$ be regular *matrix polynomials over a field $\mathbb{F}$, with sizes $m \times m$ and $n \times n$, respectively. Suppose that the finite spectra of $A(\lambda)$ and $B(\lambda)$ are disjoint, i.e., suppose that $A(\lambda)$ and $B(\lambda)$ have no common finite eigenvalues in the algebraic closure $\overline{\mathbb{F}}$. Then*

$$P(\lambda) := \begin{bmatrix} A(\lambda) & C(\lambda) \\ 0 & B(\lambda) \end{bmatrix} \sim \begin{bmatrix} A(\lambda) & 0 \\ 0 & B(\lambda) \end{bmatrix}, \tag{6.3}$$

for any *matrix polynomial* $C(\lambda)$ *over* $\mathbb{F}$ *of size* $m \times n$. *Furthermore the equivalence in* (6.3) *may always be achieved by unimodular transformations over* $\mathbb{F}$. *Thus the elementary divisors of* $P(\lambda)$ *are just the union of the elementary divisors of* $A(\lambda)$ *and* $B(\lambda)$.

Proof. Begin by reducing the diagonal blocks $A(\lambda)$ and $B(\lambda)$ of $P(\lambda)$ to their respective Smith forms $S_A(\lambda)$ and $S_B(\lambda)$ via a block-diagonal unimodular transformation of $P(\lambda)$, so we have

$$P(\lambda) \sim \begin{bmatrix} S_A(\lambda) & \widetilde{C}(\lambda) \\ 0 & S_B(\lambda) \end{bmatrix} =: \widetilde{P}(\lambda)\,.$$

Now observe that $A(\lambda)$ and $B(\lambda)$ having disjoint finite spectrum implies that each diagonal entry of S_A is relatively prime to every diagonal entry of S_B. This relative primeness can now be used to eliminate all entries of $\widetilde{C}(\lambda)$ by a kind of unimodular Gaussian elimination. For example, to eliminate the ij-entry $\widetilde{c}_{ij}(\lambda)$ of $\widetilde{C}(\lambda)$, consider the principal submatrix

$$\begin{bmatrix} \widetilde{a}_{ii}(\lambda) & \widetilde{c}_{ij}(\lambda) \\ 0 & \widetilde{b}_{jj}(\lambda) \end{bmatrix}$$

of $\widetilde{P}(\lambda)$ in rows (and columns) i and $m+j$. Here $\widetilde{a}_{ii}(\lambda)$ denotes the (i,i)-entry of $S_A(\lambda)$, and $\widetilde{b}_{jj}(\lambda)$ the (j,j)-entry of $S_B(\lambda)$. Since $\widetilde{a}_{ii}(\lambda)$ and $\widetilde{b}_{jj}(\lambda)$ are relatively prime, there exist polynomials $q(\lambda)$ and $r(\lambda)$ in $\mathbb{F}[\lambda]$ such that $\widetilde{a}_{ii}q + \widetilde{b}_{jj}r \equiv 1$, and hence such that

$$\widetilde{a}_{ii}\, q\, \widetilde{c}_{ij} \;+\; \widetilde{b}_{jj}\, r\, \widetilde{c}_{ij} \;\equiv\; \widetilde{c}_{ij}\;.$$

Thus we have the following unimodular elimination of $\widetilde{c}_{ij}$:

$$\begin{bmatrix} 1 & -r(\lambda)\widetilde{c}_{ij}(\lambda) \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \widetilde{a}_{ii}(\lambda) & \widetilde{c}_{ij}(\lambda) \\ 0 & \widetilde{b}_{jj}(\lambda) \end{bmatrix} \begin{bmatrix} 1 & -q(\lambda)\widetilde{c}_{ij}(\lambda) \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \widetilde{a}_{ii}(\lambda) & 0 \\ 0 & \widetilde{b}_{jj}(\lambda) \end{bmatrix}.$$

Repeating this for each entry of $\widetilde{C}(\lambda)$ yields a unimodular equivalence of the form

$$\begin{bmatrix} I_m & \widetilde{Y}(\lambda) \\ 0 & I_n \end{bmatrix} \begin{bmatrix} S_A(\lambda) & \widetilde{C}(\lambda) \\ 0 & S_B(\lambda) \end{bmatrix} \begin{bmatrix} I_m & \widetilde{X}(\lambda) \\ 0 & I_n \end{bmatrix} = \begin{bmatrix} S_A(\lambda) & 0 \\ 0 & S_B(\lambda) \end{bmatrix},$$

showing that $P(\lambda) \sim \operatorname{diag}\big(S_A(\lambda), S_B(\lambda)\big)$. Finally, a block-diagonal unimodular transformation can be used to "un-Smith" the blocks $S_A(\lambda)$ and $S_B(\lambda)$, thus getting us to $\operatorname{diag}\big(A(\lambda), B(\lambda)\big)$, and completing the proof. □

Remark 6.9. Note that Lemma 6.8 generalizes a well-known matrix result — if square matrices A and B have disjoint spectra, then

$$\begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$

are similar for any C. Also note that the proof idea of Lemma 6.8 can be adapted to show that, under the given hypotheses on $A(\lambda)$ and $B(\lambda)$, the matrix polynomial Sylvester equation

$$A(\lambda)X(\lambda) + Y(\lambda)B(\lambda) + C(\lambda) = 0$$

always has a *polynomial* solution $\big(X(\lambda), Y(\lambda)\big)$, for any $m \times n$ matrix polynomial $C(\lambda)$.

With Lemmas 6.7 and 6.8 in hand, we now return to and complete the proof of Corollary 6.6.

Proof. (of **Corollary 6.6**) Consider first the case when at least one of B or C is nonsingular. Then the pencil $\lambda I - A$ is strictly equivalent to either $\lambda B^{-1} - C$ or $\lambda C^{-1} - B$, so the elementary divisors of the matrix $A = BC$ are the same as those of a regular skew-symmetric pencil, which by Theorem 6.2 all have even multiplicity.

So now let us assume that both B and C are singular. Letting $r := \operatorname{rank} B < n$, then by an argument like the one used in the proof of Corollary 6.4, we see that A is similar to a matrix

$$\widetilde{A} = \begin{bmatrix} \widehat{B}\widehat{C} & \widehat{B}\widehat{D} \\ 0 & 0 \end{bmatrix}$$

as in (6.1), where $\widehat{B}$ and $\widehat{C}$ are $r \times r$ skew-symmetric matrices and $\widehat{B}$ is *nonsingular*. From this point we aim to manipulate the regular pencil $L(\lambda) := \lambda I_n - \widetilde{A}$ so as to reveal the elementary divisors of $\widetilde{A}$, and hence of A.

Letting $S(\lambda)$ be the Smith form of $\lambda I_r - \widehat{B}\widehat{C}$, do unimodular transformations on the first r rows and r columns of $L(\lambda)$ to see that

$$L(\lambda) \sim \begin{bmatrix} S(\lambda) & M(\lambda) \\ 0 & \lambda I_{n-r} \end{bmatrix}.$$

Note that *all* elementary divisors in $S(\lambda)$ have even multiplicity, since $\lambda I_r - \widehat{B}\widehat{C}$ is strictly equivalent to the regular skew-symmetric pencil $\lambda \widehat{B}^{-1} - \widehat{C}$. Now by repeated applications of the Exchange Lemma 6.7 to $S(\lambda)$, we can unimodularly transform $S(\lambda)$ to an $r \times r$ diagonal $D(\lambda)$ of the form $D(\lambda) = \operatorname{diag}\big(D_1(\lambda), D_2(\lambda)\big)$, where D_2 contains all the elementary divisors in $S(\lambda)$ associated with the eigenvalue 0, while D_1 contains every other elementary divisor in $S(\lambda)$, all with even multiplicity. (Note that there is "enough room" in the $r \times r$ $D(\lambda)$ to achieve this separation of the elementary divisors from $S(\lambda)$, because $S(\lambda)$ is the Smith form of an $r \times r$ *pencil*, and so possesses at most r elementary divisors.) Thus we see that

$$L(\lambda) \sim \begin{bmatrix} D_1(\lambda) & 0 & \widetilde{M}_1(\lambda) \\ 0 & D_2(\lambda) & \widetilde{M}_2(\lambda) \\ 0 & 0 & \lambda I_{n-r} \end{bmatrix} =: \widetilde{L}(\lambda)\,.$$

Now re-partition $\widetilde{L}(\lambda)$ into

$$\widetilde{L}(\lambda) = \begin{bmatrix} D_1(\lambda) & \widehat{M}(\lambda) \\ 0 & Z(\lambda) \end{bmatrix},$$

where

$$\widehat{M}(\lambda) = \begin{bmatrix} 0 & \widetilde{M}_1(\lambda) \end{bmatrix} \quad \text{and} \quad Z(\lambda) = \begin{bmatrix} D_2(\lambda) & \widetilde{M}_2(\lambda) \\ 0 & \lambda I_{n-r} \end{bmatrix},$$

and note that $Z(\lambda)$ has *only* elementary divisors associated with the eigenvalue 0. Since $D_1(\lambda)$ and $Z(\lambda)$ are regular polynomials with disjoint finite spectra, we can now apply the Cancellation Lemma 6.8 to conclude that $\widetilde{L}(\lambda) \sim \operatorname{diag}\big(D_1(\lambda), Z(\lambda)\big)$.

Therefore the elementary divisor list of $\widetilde{L}(\lambda)$ is just the concatenation of the elementary divisors of $D_1(\lambda)$, none of which are associated with the eigenvalue 0 and all having even

multiplicity, together with the elementary divisors of $Z(\lambda)$, which are all associated with the eigenvalue 0, but have indeterminate multiplicities because of the presence of the off-diagonal block $\widetilde{M}_2(\lambda)$ in $Z(\lambda)$. Since $\widetilde{L}(\lambda)$ is unimodularly equivalent to $L(\lambda) := \lambda I_n - \widetilde{A}$, and $\widetilde{A}$ is similar to A, the desired conclusion about the elementary divisors of A is proved. $\square$

Corollary 6.6 allows for the possibility that elementary divisors of $A = BC$ associated with the eigenvalue 0 may not necessarily have even multiplicity, at least when both B and C are singular. The next example shows that this possible loss of even multiplicity *can* indeed occur; in fact *all* these multiplicities can be odd, even when $A_{n\times n}$ has even size n.

Example 6.10. Consider the 8×8 skew-symmetric matrices

$$B := \begin{bmatrix} 0_3 & I_3 & \\ -I_3 & 0_3 & \\ & & 0_2 \end{bmatrix} \quad \text{and} \quad C := \begin{bmatrix} 0_3 & -E^T & -F^T \\ E & 0_3 & -G^T \\ F & G & H \end{bmatrix},$$

where 0_k is the $k \times k$ zero matrix, and the blocks E, F, G, H are defined as follows:

$$E = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & a \end{bmatrix} \text{ with } a \neq 0\,, \quad F = \begin{bmatrix} 0 & 0 & b \\ 0 & 0 & c \end{bmatrix} \text{ and } G = \begin{bmatrix} 0 & 1 & d \\ 0 & 0 & e \end{bmatrix}$$

with arbitrary $b, c, d, e \in \mathbb{F}$, and $H = \begin{bmatrix} 0 & h \\ -h & 0 \end{bmatrix}$ is an arbitrary 2×2 skew-symmetric matrix. Then the product

$$A := BC = \begin{bmatrix} E & 0 & -G^T \\ 0 & E^T & F^T \\ 0 & 0 & 0 \end{bmatrix}$$

is easily seen to have eigenvalues $\lambda = 0$ with algebraic multiplicity 6, and $\lambda = a$ with algebraic multiplicity 2. A little further manipulation (mainly just permutation of Jordan blocks) shows the elementary divisor list of A to be $\{\lambda - a,\ \lambda - a,\ \lambda,\ \lambda^2,\ \lambda^3\}$, illustrating both the retention of the even multiplicity property for elementary divisors associated with nonzero eigenvalues, as well as the complete *loss* of even multiplicity for elementary divisors associated with the eigenvalue zero. (Note also that both B and C are indeed singular, as is necessary for any example illustrating this loss of even multiplicity.)

6.3 Structured linearizations of skew-symmetric matrix polynomials

Let us begin by briefly recalling some of the background context of known results on the existence of structured linearizations for two other classes of structured matrix polynomials, alternating and palindromic. In [35] it was shown (constructively) that many regular matrix polynomials that are alternating or palindromic have a strong linearization with the *same structure*. However, it was also shown that there exist some alternating and some palindromic polynomials for which a structured strong linearization is impossible. This motivated the subsequent investigation of the possible Smith forms of alternating matrix polynomials (in [37]) and of palindromic matrix polynomials (in [38]), with the goal of systematically probing for any *elementary divisor obstructions* to the existence of structured

linearizations. It was found that the elementary divisors of alternating or palindromic polynomials satisfy certain necessary conditions, but these conditions are somewhat different for even degree vs. odd degree polynomials. This dichotomy between the behavior of even and odd degree led to several conclusions about structured linearizations in [37, 38]:

(a) Since all odd degree polynomials in a given structure class (including structured pencils) have the same elementary divisor constraints, it was natural to conjecture that every odd degree structured polynomial has a structured linearization. That this is indeed the case was demonstrated by the explicit construction of structured companion forms for every odd degree.

(b) There is a mismatch between the allowed multiplicities of elementary divisors (associated with certain critical eigenvalues) for even degree structured matrix polynomials vs. what is possible for pencils of the same structure type. This mismatch constitutes a fundamental obstruction to the existence of structured linearizations for certain even degree matrix polynomials, and thus precludes the existence of structured companion forms in the even degree case. (See [37, 38] for the technical details of these elementary divisor obstructions.)

Theorems 4.2 and 6.2 on the Smith form and elementary divisor structure of skew-symmetric matrix polynomials now put us in a position to consider the analogous issues for this class of structured matrix polynomials. These theorems reveal *no* apparent elementary divisor obstruction to the existence of structured linearizations for any degree; *all* elementary divisors occur with even multiplicity, no matter whether the degree of the skew-symmetric polynomial is even or odd. Thus it would seem reasonable to conjecture that *every* skew-symmetric matrix polynomial, of any degree, has a skew-symmetric (strong) linearization.

In the next two sections we will show that this conjecture is once again true for odd degree, but false for even degree skew-symmetric polynomials. For this structure class, the obstruction to the universal existence of structured linearizations in the even degree case arises not from any incompatibility of elementary divisors between polynomial and pencil, but rather from considerations related to the structure of minimal indices of singular matrix polynomials, a topic which Theorems 4.2 and 6.2 simply do not address. For the purposes of this paper, the result of Lemma 2.5 will suffice to point out some of the problems that can prevent an even degree skew-symmetric polynomial that is singular from having a skew-symmetric linearization. A more detailed treatment of some of the subtle issues involved in linearizing singular matrix polynomials is discussed in [12]. However, if we put aside the singular case and only consider even degree skew-symmetric polynomials that are regular, we find that structured linearizations now always exist. Since the ideas and techniques involved in addressing the odd and even degree cases are rather different, we consider these two cases in separate sections.

6.3.1 The odd degree case

Structured linearizations for any odd degree skew-symmetric matrix polynomial can be found in a manner analogous to that used in [37], [38], and [11] to handle odd degree alternating or palindromic polynomials — by building structured companion forms. A *companion form* for square matrix polynomials $P(\lambda)$ of degree k is a uniform template for

constructing a matrix pencil $\mathcal{C}_P(\lambda)$ directly from the matrix coefficients of P, such that $\mathcal{C}_P(\lambda)$ is a strong linearization for *every* square polynomial P of degree k, regular or singular, over an arbitrary field. A *structured companion form* for structure class $\mathcal{M}$ is a companion form $\mathcal{C}_P(\lambda)$ with the additional property that $\mathcal{C}_P \in \mathcal{M}$ whenever $P \in \mathcal{M}$. Alternating companion forms were constructed for each odd degree in [37], while palindromic companion forms for each odd degree were constructed in [38] and [11].

We show now how to fashion *skew-symmetric companion forms* for any odd degree. To this end, for a general $n \times n$ polynomial $\lambda^k A_k + \cdots + \lambda A_1 + A_0$ with $A_k \neq 0$ and k odd, consider the block-tridiagonal $nk \times nk$ pencil $\mathcal{S}_P(\lambda) = [S_{ij}(\lambda)]_{i,j=1,\ldots,k}$ with $n \times n$ blocks $S_{ij}(\lambda)$ of the form

$$\begin{aligned} S_{jj}(\lambda) &= \begin{cases} \lambda A_j + A_{j-1} & \text{if } j \text{ is odd,} \\ 0 & \text{if } j \text{ is even,} \end{cases} \\ S_{j,j+1}(\lambda) &= S_{j+1,j}(\lambda) = \begin{cases} \lambda I_n & \text{if } j \text{ is odd,} \\ I_n & \text{if } j \text{ is even,} \end{cases} \end{aligned}$$

and $S_{ij}(\lambda) = 0$ for $|i-j| > 1$. These pencils, also considered in [37], are slightly modified versions of certain companion pencils introduced in [3]. As an illustration, here is $S_P(\lambda)$ for $k = 5$:

$$\mathcal{S}_P(\lambda) = \begin{bmatrix} \lambda A_1 + A_0 & \lambda I & & & \\ \lambda I & 0 & I & & \\ & I & \lambda A_3 + A_2 & \lambda I & \\ & & \lambda I & 0 & I \\ & & & I & \lambda A_5 + A_4 \end{bmatrix}.$$

Observe that $\mathcal{S}_P(\lambda)$ is block-symmetric. If $P(\lambda)$ is skew-symmetric, then it is immediate that the pencil

$$\mathcal{K}_P(\lambda) := \operatorname{diag}\big(I_n, -I_n, I_n, -I_n, \ldots, (-1)^{k-1} I_n \big)\, \mathcal{S}_P(\lambda)$$

is skew-symmetric as well. It was shown in [3] that $\mathcal{S}_P(\lambda)$ is always a strong linearization for $P(\lambda)$ for any regular P of odd degree k over the field $\mathbb{F} = \mathbb{C}$. This result was extended in [37] to include regular and singular (square) matrix polynomials over arbitrary fields. Since $\mathcal{K}_P(\lambda)$ is built from $\mathcal{S}_P(\lambda)$ by multiplication with a nonsingular constant matrix, we immediately obtain the following result.

Lemma 6.11. *Let $P(\lambda)$ be any $n \times n$ matrix polynomial of odd degree (not necessarily regular) over an arbitrary field $\mathbb{F}$. Then $\mathcal{K}_P(\lambda)$ is a structured companion form for the class of skew-symmetric matrix polynomials; in particular, $\mathcal{K}_P$ is always a strong linearization for P, and $\mathcal{K}_P$ is skew-symmetric whenever P is skew-symmetric.*

6.3.2 The even degree case

In contrast to the odd degree case, there exist many skew-symmetric matrix polynomials of even degree that do not even admit a linearization that is skew-symmetric, let alone a strong linearization that is skew-symmetric. In the next example we show how to construct such polynomials for any even degree.

Example 6.12. Let n be odd and k be even, and consider *any* $n \times n$ skew-symmetric matrix polynomial $P(\lambda)$ of degree k, over an arbitrary field $\mathbb{F}$. Note that by Lemma 3.3

we have $\det P(\lambda) \equiv 0$, so such a $P(\lambda)$ is necessarily singular. Next assume that $P(\lambda)$ has a skew-symmetric linearization; i.e., suppose there exists a $kn \times kn$ skew-symmetric pencil $L(\lambda)$ over $\mathbb{F}$ and unimodular matrix polynomials $E(\lambda)$ and $F(\lambda)$ such that

$$E(\lambda)L(\lambda)F(\lambda) = \begin{bmatrix} P(\lambda) & 0 \\ 0 & I_{(k-1)n} \end{bmatrix}. \tag{6.4}$$

Then clearly (6.4) implies that

$$\operatorname{rank}(L) \;=\; (k-1)n \;+\; \operatorname{rank}(P)\,. \tag{6.5}$$

Note that (6.5) is equivalent to one of the necessary conditions listed in Lemma 2.5, namely that $\dim \mathcal{N}_r(L) = \dim \mathcal{N}_r(P)$. Now if both $P(\lambda)$ and $L(\lambda)$ are skew-symmetric, then by Corollary 6.1 both will have *even* rank. But this would contradict (6.5), since $(k-1)n$ is odd. Thus $P(\lambda)$ cannot have any skew-symmetric linearization. Indeed, these examples show that skew-symmetric companion forms for matrix polynomials of any even degree cannot exist.

On the other hand, if $P(\lambda)$ is a skew-symmetric polynomial of even degree that is *regular*, then there always exists a skew-symmetric strong linearization.

Theorem 6.13. *Let $P(\lambda)$ be a regular skew-symmetric $n \times n$ matrix polynomial of even degree $k = 2s$, over an arbitrary field $\mathbb{F}$. Then $P(\lambda)$ has a skew-symmetric strong linearization.*

Proof. By Corollary 4.3, $P(\lambda)$ is unimodularly equivalent to a matrix polynomial of the form

$$\begin{bmatrix} 0 & d_1(\lambda) \\ -d_1(\lambda) & 0 \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} 0 & d_r(\lambda) \\ -d_r(\lambda) & 0 \end{bmatrix}, \tag{6.6}$$

where d_j is monic with $g_j = \deg d_j \geq 0$ for $j = 1, \ldots, r$, and $d_j \,|\, d_{j+1}$ for $j = 1, \ldots, r-1$. Note that $n = 2r$ must be even, because $P(\lambda)$ is regular.

Let m be the sum of the degrees of the finite elementary divisors of $P(\lambda)$; then $m = 2\ell$ is even, because all elementary divisors occur with even multiplicity by Theorem 6.2, and $\ell = g_1 + g_2 + \cdots + g_r$ by (6.6). Similarly the sum of the degrees of the infinite elementary divisors of $P(\lambda)$ is even, say $\delta = 2\beta$, and because of even multiplicity the list of infinite elementary divisors can be partitioned into two identical copies of a list with (nonzero) degrees $\alpha_1, \alpha_2, \ldots, \alpha_\nu$. Then $\beta = \alpha_1 + \alpha_2 + \cdots + \alpha_\nu$ and $\ell = g_1 + g_2 + \cdots + g_r$, and the sum of the degrees of *all* the elementary divisors of $P(\lambda)$ is $m + \delta = 2\ell + 2\beta = kn = 2sn$ by Remark 2.6, so $\ell + \beta = ns$.

Now define $L(\lambda)$ to be the block-diagonal $ns \times ns$ pencil

$$L(\lambda) = \operatorname{diag}\big(D_1(\lambda), \ldots, D_r(\lambda), N_1(\lambda), \ldots, N_\nu(\lambda)\big)\,,$$

where $D_i(\lambda)$ for $i = 1, \ldots, r$ is empty if $g_i = \deg d_i = 0$, but if $g_i \geq 1$ then $D_i(\lambda)$ is the $g_i \times g_i$ companion linearization (2.4) for the monic 1×1 matrix polynomial $d_i(\lambda)$. The blocks $N_j(\lambda)$ for $j = 1, \ldots, \nu$ are the $\alpha_j \times \alpha_j$ unimodular matrices

$$N_j(\lambda) \;=\; \begin{bmatrix} 1 & \lambda & & \\ & 1 & \ddots & \\ & & \ddots & \lambda \\ & & & 1 \end{bmatrix},$$

each with exactly one elementary divisor, an infinite elementary divisor with degree α_j. Finally consider the $nk \times nk$ pencil

$$\widetilde{L}(\lambda) := \begin{bmatrix} 0 & L(\lambda) \\ -L^T(\lambda) & 0 \end{bmatrix}.$$

Then clearly $\widetilde{L}(\lambda)$ is skew-symmetric, and the list of elementary divisors of $\widetilde{L}(\lambda)$ consists of two identical copies of the list of elementary divisors of $L(\lambda)$. Thus by the construction of $L(\lambda)$ we see that $P(\lambda)$ and $\widetilde{L}(\lambda)$ have exactly the same (finite and infinite) elementary divisors. Since $P(\lambda)$ is regular by hypothesis and $\widetilde{L}(\lambda)$ is regular by Remark 2.6, both have trivial nullspaces, so $\dim \mathcal{N}_r(P) = 0 = \dim \mathcal{N}_r(L)$. Therefore $\widetilde{L}(\lambda)$ is a strong linearization for $P(\lambda)$ by Lemma 2.5. □

Although Theorem 6.13 shows that skew-symmetric strong linearizations always exist for any regular skew-symmetric matrix polynomial $P(\lambda)$ of even degree, it is unclear how such linearizations can be constructed in general without first computing either the Smith form of $P(\lambda)$, or the skew-symmetric canonical form of $P(\lambda)$ as in Corollary 4.3 or Theorem 5.4. However, for skew-symmetric polynomials $P(\lambda)$ over the field $\mathbb{F} = \mathbb{C}$ or $\mathbb{F} = \mathbb{R}$, we can use the pencils in the vector space $\mathbb{DL}(P)$ (introduced in [36]) as a source of easy-to-construct structured linearizations for *regular* $P(\lambda)$ of even (or odd) degree k. It was shown in [20] that for any square polynomial P, every pencil in $\mathbb{DL}(P)$ is always block-symmetric. As a consequence we see that when $P(\lambda)$ is skew-symmetric, then every $L(\lambda) \in \mathbb{DL}(P)$ is also skew-symmetric. All that remains, then, is to find a *regular* pencil $L(\lambda)$ in $\mathbb{DL}(P)$, since regularity of L was shown in [36] to be equivalent to L being a strong linearization for P. This can always be achieved by choosing an $L(\lambda) \in \mathbb{DL}(P)$ that has an associated ansatz vector $v \in \mathbb{F}^k$ such that the roots of the corresponding v-polynomial are disjoint from the spectrum of P. We refer the reader to [36] and [20] for further details.

6.4 Skew-symmetric rational matrices

Our results on matrix polynomials can readily be extended to rational matrices, where the corresponding canonical form is usually referred to as the Smith-McMillan form [9, 24, 40, 41, 47]. Recall that a rational matrix over a field $\mathbb{F}$ is a matrix with entries from the field of rational functions $\mathbb{F}(\lambda)$.

Theorem 6.14 (Smith-McMillan form [24])**.** *Let $R(\lambda)$ be a $n \times n$ rational matrix over an arbitrary field $\mathbb{F}$. Then there exists $r \in \mathbb{N}$ and unimodular matrix polynomials $E(\lambda), F(\lambda)$ over $\mathbb{F}$ such that*

$$E(\lambda)R(\lambda)F(\lambda) = \operatorname{diag}\left(\frac{\nu_1(\lambda)}{\mu_1(\lambda)}, \ldots, \frac{\nu_r(\lambda)}{\mu_r(\lambda)}\right) \oplus 0_{n-r}, \tag{6.7}$$

where ν_j, μ_j are monic and pairwise co-prime polynomials in $\mathbb{F}[\lambda]$ for $j = 1, \ldots, r$, satisfying the divisibility chain conditions $\nu_j \,|\, \nu_{j+1}$ and $\mu_{j+1} \,|\, \mu_j$ for $j = 1, \ldots, r-1$. Moreover, the canonical form (6.7) *is unique.*

6.4.1 Smith-McMillan form for skew-symmetric rational matrices

As a direct consequence of the Skew-Smith form for skew-symmetric matrix polynomials in Theorem 4.2, we immediately obtain the Smith-McMillan form for skew-symmetric rational

matrices, i.e., rational matrices $R(\lambda) = [r_{ij}(\lambda)]$ satisfying $R(\lambda) = -R(\lambda)^T$ and $r_{ii}(\lambda) \equiv 0$ for $i = 1, \ldots, n$. Here $r_{ii}(\lambda) \equiv 0$ means that each diagonal entry $r_{ii}(\lambda)$ is the zero element in the field $\mathbb{F}(\lambda)$, not just that $r_{ii}(\lambda)$ is the identically zero function on $\mathbb{F}$.

Corollary 6.15 (Smith-McMillan form for skew-symmetric rational matrices)**.** *Let $R(\lambda) = [r_{ij}(\lambda)]$ be a skew-symmetric $n \times n$ rational matrix over an arbitrary field $\mathbb{F}$. Then there exists $r \in \mathbb{N}$ with $2r \leq n$ and unimodular matrix polynomials $E(\lambda), F(\lambda)$ over $\mathbb{F}$ such that*

$$E(\lambda)P(\lambda)F(\lambda) = \begin{bmatrix} \frac{\nu_1(\lambda)}{\mu_1(\lambda)} & 0 \\ 0 & \frac{\nu_1(\lambda)}{\mu_1(\lambda)} \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} \frac{\nu_r(\lambda)}{\mu_r(\lambda)} & 0 \\ 0 & \frac{\nu_r(\lambda)}{\mu_r(\lambda)} \end{bmatrix} \oplus 0_{n-2r}, \tag{6.8}$$

where ν_j, μ_j are monic and pairwise co-prime polynomials in $\mathbb{F}[\lambda]$ for $j = 1, \ldots, r$, satisfying the divisibility chain conditions $\nu_j \,|\, \nu_{j+1}$ and $\mu_{j+1} \,|\, \mu_j$ for $j = 1, \ldots, r-1$. Moreover, the polynomials $\nu_j(\lambda), \mu_j(\lambda)$ for $j = 1, \ldots, r$ are uniquely determined by these conditions.

Proof. Let $\psi(\lambda)$ be the least common multiple of all denominator polynomials in the entries $r_{ij}(\lambda)$, $i, j = 1, \ldots, n$ and form the skew-symmetric matrix polynomial $P(\lambda) = \psi(\lambda)R(\lambda)$. Then by Theorem 4.2 there exists $r \in \mathbb{N}$ with $\ell = 2r \leq n$ and unimodular matrix polynomials $E(\lambda), F(\lambda)$ over the field $\mathbb{F}$ such that $E(\lambda)P(\lambda)F(\lambda)$ is in the form (4.2), where d_j is monic for $j = 1, \ldots, \ell$, $d_j \,|\, d_{j+1}$ for $j = 1, \ldots, \ell - 1$, and $d_{2i-1}(\lambda) = d_{2i}(\lambda)$ for $i = 1, \ldots, r$. Cancelling common factors in the rational functions $\frac{d_j(\lambda)}{\psi(\lambda)}$ yields rational functions $\frac{\nu_i(\lambda)}{\mu_i(\lambda)}$ having numerators and denominators with the desired properties. The uniqueness of the polynomials $\nu_i(\lambda)$ and $\mu_i(\lambda)$ follows immediately from the uniqueness of the general Smith-McMillan form (Theorem 6.14). □

As a corollary of Theorem 5.4 we analogously obtain the corresponding Smith-McMillan-like canonical form under unimodular congruence.

Corollary 6.16 (Skew-symmetric Smith-McMillan-like canonical form)**.** *Let $R(\lambda)$ be a skew-symmetric $n \times n$ rational matrix over an arbitrary field $\mathbb{F}$. Then there exists $r \in \mathbb{N}$ with $2r \leq n$ and a unimodular matrix polynomial $F(\lambda)$ over $\mathbb{F}$ such that*

$$F^T(\lambda)P(\lambda)F(\lambda) = \begin{bmatrix} 0 & \frac{\nu_1(\lambda)}{\mu_1(\lambda)} \\ -\frac{\nu_1(\lambda)}{\mu_1(\lambda)} & 0 \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} 0 & \frac{\nu_r(\lambda)}{\mu_r(\lambda)} \\ -\frac{\nu_r(\lambda)}{\mu_r(\lambda)} & 0 \end{bmatrix} \oplus 0_{n-2r} =: \Sigma(\lambda), \tag{6.9}$$

where ν_j, μ_j are monic and pairwise co-prime polynomials in $\mathbb{F}[\lambda]$ for $j = 1, \ldots, r$, satisfying the divisibility chain conditions $\nu_j \,|\, \nu_{j+1}$ and $\mu_{j+1} \,|\, \mu_j$ for $j = 1, \ldots, r-1$. Moreover, the skew-symmetric canonical form $\Sigma(\lambda)$ is unique.

6.4.2 Minimal symmetric factorizations of skew-symmetric rational matrices

The Smith-McMillan form is an important tool in the construction of (minimal) factorizations of complex matrix-valued functions, see [6, 25, 47]. To introduce this concept, let us first consider the notion of *McMillan degree*. Let $R(\lambda) = [r_{ij}(\lambda)]$ be an $n \times n$ rational matrix over $\mathbb{C}$. We will assume that $R(\lambda)$ is regular, i.e., $\det R(\lambda)$ does not vanish identically. Let $\mu(\lambda) = \mu_1(\lambda) \cdots \mu_n(\lambda)$, where $\mu_i(\lambda)$ are the denominator polynomials of the Smith-McMillan form of $R(\lambda)$ as in (6.7), and let $\lambda_0 \in \mathbb{C}$ be a root of μ. Then λ_0 is called a *finite pole* of $R(\lambda)$, and the multiplicity of λ_0 as a root of μ is called the *local pole multiplicity*

$\delta(R,\lambda_0)$ of λ_0 as a pole of $R(\lambda)$. We say that $\lambda_0=\infty$ is a pole of $R(\lambda)$ if $\lambda_0=0$ is a pole of $R(\frac{1}{\lambda})$, and the *local pole multiplicity* $\delta(R,\infty)$ of $\lambda_0=\infty$ as a pole of $R(\lambda)$ is by definition the local pole multiplicity of $\lambda_0=0$ of $R(\frac{1}{\lambda})$. Finally, the *McMillan degree* of R is defined as

$$\delta(R)=\sum_{\lambda_0\in\mathbb{C}\cup\{\infty\}}\delta(R,\lambda_0),$$

where we define $\delta(R,\lambda_0):=0$ whenever λ_0 is not a pole of R. The McMillan degree has an important meaning in the theory of realizations. It is well known [5, 6, 25, 47] that if $R(\lambda)=[r_{ij}(\lambda)]$ is *proper*, i.e., that for each r_{ij} the degree of the numerator does not exceed the degree of the denominator, then $R(\lambda)$ has *realizations* of the form

$$R(\lambda)=D+C(\lambda I_k-A)^{-1}B\,.$$

If the dimension k is minimal among all such realizations of $R(\lambda)$, then k is equal to the McMillan degree $\delta(R)$ and the poles of $R(\lambda)$ coincide with the eigenvalues of A.

Using this notion, a factorization

$$R(\lambda)=R_1(\lambda)R_2(\lambda) \tag{6.10}$$

with $n\times n$ rational matrices $R_1(\lambda),R_2(\lambda)$ is called *minimal* if $\delta(R)=\delta(R_1)+\delta(R_2)$, or, equivalently, if

$$\delta(R,\lambda)=\delta(R_1,\lambda)+\delta(R_2,\lambda),\quad \text{for all } \lambda\in\mathbb{C}\cup\{\infty\}.$$

Loosely speaking, this condition means that there is no cancellation of poles and zeros between the factors $R_1(\lambda)$ and $R_2(\lambda)$. It was noted in [27] that the class of factorizations that are minimal on all of $\mathbb{C}\cup\{\infty\}$ is not adequate for every application. Therefore, the slightly weaker concept of *minimality on a set* $\sigma\subseteq\mathbb{C}\cup\{\infty\}$ was considered in [27]. We say that the factorization (6.10) is *minimal at* $\lambda_0\in\mathbb{C}\cup\{\infty\}$ if

$$\delta(R,\lambda_0)=\delta(R_1,\lambda_0)+\delta(R_2,\lambda_0)\,,$$

and it is called *minimal on the set* $\sigma\subseteq\mathbb{C}\cup\{\infty\}$ if it is minimal at all $\lambda_0\in\sigma$. In the special case $\sigma=\mathbb{C}$, we will say that the factorization (6.10) is *minimal everywhere except at* ∞.

For the case of $n\times n$ skew-symmetric rational matrices, it is natural to consider factorizations of the form

$$R(\lambda)=R_1(\lambda)R_2(\lambda)=R_2(\lambda)^TJR_2(\lambda),$$

where $J\in\mathbb{C}^{n\times n}$ is a constant skew-symmetric matrix. Such factorizations are called *symmetric factorizations*. In [46] minimal symmetric factorizations for real skew-symmetric rational matrices were constructed in a rather technical process. If we drop the requirement that the factorization has to be minimal everywhere, then applying Corollary 6.16 we immediately obtain the following factorization result.

Corollary 6.17 (Minimal symmetric factorization of skew-symmetric rational matrices)**.** *Let $R(\lambda)$ be an $n\times n$ regular skew-symmetric rational matrix over $\mathbb{C}$. Then there exists a symmetric factorization*

$$R(\lambda)=R_1(\lambda)R_2(\lambda)=R_2(\lambda)^TJR_2(\lambda)$$

which is minimal everywhere except for ∞, where $R_2(\lambda)$ is an $n\times n$ rational matrix over $\mathbb{C}$ and J is an invertible real constant $n\times n$ skew-symmetric matrix. If $R(\lambda)$ is real, then $R_2(\lambda)$ can also be chosen to be real.

Proof. Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$. Then Corollary 6.16 implies that there exists a unimodular matrix polynomial $F(\lambda)$ over $\mathbb{F}$ such that

$$R(\lambda) = F^T(\lambda)\Sigma(\lambda)F(\lambda)$$

with $\Sigma(\lambda)$ as in (6.9) with $n = 2r$. Setting

$$\begin{aligned} \widetilde{R}_2(\lambda) &= \begin{bmatrix} 0 & \frac{\nu_1(\lambda)}{\mu_1(\lambda)} \\ -1 & 0 \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} 0 & \frac{\nu_r(\lambda)}{\mu_r(\lambda)} \\ -1 & 0 \end{bmatrix}, \\ J &= \underbrace{\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}}_{r \text{ times}}, \end{aligned}$$

and $R_2(\lambda) = \widetilde{R}_2(\lambda)F(\lambda)$, we have a factorization of the desired form. Clearly, $R_2(\lambda)$ is real if $R(\lambda)$ is. The assertion on the minimality follows immediately from the fact that the local pole multiplicities of finite poles are invariant under multiplication with unimodular matrix polynomials, see [5, 25], so that $\delta(R_2, \lambda_0) = \delta(\widetilde{R}_2, \lambda_0)$ for any $\lambda_0 \in \mathbb{C}$. $\square$

7 Conclusions

The elementary divisor structure of skew-symmetric matrix polynomials over an arbitrary field has been completely analyzed in this paper, via a characterization of all possible Smith forms of such polynomials. The use of the properties of compound matrices as a key tool in this investigation constitutes the third success (along with [37, 38]) of this approach to analyzing structured matrix polynomials, and now firmly establishes this as a standard technique for addressing such problems.

We have shown that a Smith-like canonical form that is itself skew-symmetric can be achieved for any skew-symmetric matrix polynomial, even when the unimodular transformations used are restricted to the subclass of unimodular congruences. This implies that no additional invariants are introduced into the classification of skew-symmetric matrix polynomials by the restriction of unimodular equivalences to congruences.

Three significant applications of these results have also been presented: the characterization of the eigenvalue and elementary divisor structure of matrices expressible as the product of two skew-symmetric matrices, the construction of symmetric factorizations of skew-symmetric rational matrices that are minimal everywhere except for infinity, and the investigation of the existence of structured linearizations for skew-symmetric matrix polynomials. Previous results on structured linearizations for alternating [37] and palindromic polynomials [38] exhibit a clear dichotomy between the behavior of the even and odd degree cases, due to incompatibilities between the elementary divisor structures of even and odd degree structured polynomials. Such elementary divisor incompatibilities do *not* occur for skew-symmetric matrix polynomials; all elementary divisors have even multiplicity, regardless of the degree of the polynomial. However, our results on structured linearizations of skew-symmetric polynomials still show a clear difference between the even and odd degree cases. Every odd degree skew-symmetric polynomial has a structured strong linearization; indeed, a skew-symmetric companion form that provides a single uniform template for constructing a structured strong linearization for *every* skew-symmetric matrix polynomial,

regular or singular, regardless of the underlying field, has been constructed for each odd degree. By contrast, although every *regular* skew-symmetric polynomial is shown to have a structured strong linearization, there are large classes of *singular* skew-symmetric matrix polynomials of each even degree for which skew-symmetric linearizations are shown to be impossible. Thus we see that skew-symmetric companion forms for matrix polynomials of any even degree cannot exist.

Acknowledgements. We thank Leiba Rodman for bringing to our attention the application to symmetric factorizations of skew-symmetric rational matrices, and we thank André Ran and Leiba Rodman for helpful discussions.

References

[1] B. Adhikari. *Backward perturbation and sensitivity analysis of structured polynomial eigenvalue problem*. PhD thesis, IIT Guwahati, Dept. of Mathematics, 2008.

[2] B. Adhikari and R. Alam. Structured backward errors and pseudospectra of structured matrix pencils. *SIAM J. Matrix Anal. Appl.*, 31:331–359, 2009.

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

[4] M. Artin. *Algebra*. Prentice-Hall, Englewood Cliffs, New Jersey, 1991.

[5] H. Bart, I. Gohberg, and M.A. Kaashoek. *Minimal factorization of matrix and operator functions*, volume 1 of *Operator Theory: Advances and Applications*. Birkhäuser, Basel, Switzerland, 1979.

[6] H. Bart, I. Gohberg, M.A. Kaashoek, and A.C.M. Ran. *A state space approach to canonical factorization with applications*, volume 178 of *Operator Theory: Advances and Applications*. Birkhäuser, Basel, Switzerland, 2008.

[7] P. Benner, V. Mehrmann, and H. Xu. A numerically stable, structure preserving method for computing the eigenvalues of real Hamiltonian or symplectic pencils. *Numer. Math.*, 78(3):329–358, 1998.

[8] R. Byers and D. Kressner. Structured condition numbers for invariant subspaces. *SIAM J. Matrix Anal. Appl.*, 28:326–347, 2006.

[9] W. A. Coppel. Matrices of rational functions. *Bull. Austral. Math. Soc.*, 11:89–113, 1974.

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

[11] F. De Terán, F. Dopico, and D. S. Mackey. Palindromic companion forms for matrix polynomials of odd degree. *J. Comput. Appl. Math.*, 236(6):1464–1480, 2011.

[12] F. De Terán, F. Dopico, and D. S. Mackey. Private communication. Unpublished manuscript, 2011.

[13] D.Z. Djoković, J. Patera, P. Winternitz, and H. Zassenhaus. Normal forms of elements of classical real and complex Lie and Jordan algebras. *J. Math. Phys.*, 24:1363–1374, 1983.

[14] D. Doković. On the product of two alternating matrices. *Amer. Math. Monthly*, 98(10):935–936, 1991.

[15] H. Faßbender, D. S. Mackey, N. Mackey, and H. Xu. Hamiltonian square roots of skew–Hamiltonian matrices. *Linear Algebra Appl.*, 287:125–159, 1999.

[16] G. Frobenius. Theorie der linearen Formen mit ganzen Coefficienten. *J. Reine Angew. Math. (Crelle)*, 86:146–208, 1878.

[17] F. R. Gantmacher. *Theory of Matrices.* Chelsea, New York, 1959.

[18] 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.

[19] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials.* Academic Press, New York, 1982.

[20] N. J. Higham, D. S. Mackey, N. Mackey, and F. Tisseur. Symmetric linearizations for matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 29(1):143–159, 2006.

[21] R. A. Horn and C. R. Johnson. *Matrix Analysis.* Cambridge University Press, Cambridge, 1985.

[22] R. A. Horn and V. V. Sergeichuk. Congruences of a square matrix and its transpose. *Linear Algebra Appl.*, 389:347–353, 2004.

[23] Kh. D. Ikramov. Hamiltonian square roots of skew-Hamiltonian matrices revisited. *Linear Algebra Appl.*, 325:101–107, 2001.

[24] V. Ionescu, C. Oară, and M. Weiss. General matrix pencil techniques for the solution of algebraic Riccati equations: A unified approach. *IEEE Trans. Automat. Control*, 42(8), 1997.

[25] V. Ionescu, C. Oară, and M. Weiss. *Generalized Riccati Theory and Robust Control: A Popov Function Approach.* John Wiley & Sons, Chichester, 1999.

[26] N. Jacobson. *Basic Algebra I.* W.H. Freeman and Co., San Francisco, 1974.

[27] M.A. Kaashoek, A.C.M. Ran, and L. Rodman. Local minimal factorizations of rational matrix functions in terms of null and pole data: formulas for factors. *Integral Equations Operator Theory*, 16:98–130, 1993.

[28] M. Karow, D. Kressner, and F. Tisseur. Structured eigenvalue condition numbers. *SIAM J. Matrix Anal. Appl.*, 28:1052–1068, 2006.

[29] P. Lancaster. Linearization of regular matrix polynomials. *Electron. J. Linear Algebra*, 17:21–27, 2008.

[30] P. Lancaster and P. Psarrakos. A note on weak and strong linearizations of regular matrix polynomials. MIMS EPrint 2006.72, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, 2005.

[31] P. Lancaster and L. Rodman. *The Algebraic Riccati Equation.* Oxford University Press, Oxford, 1995.

[32] P. Lancaster and M. Tismenetsky. *The Theory of Matrices.* Academic Press, Orlando, 2nd edition, 1985.

[33] S. Lang. *Algebra.* Addison-Wesley, Reading, Mass., 1965.

[34] W.-W. Lin, V. Mehrmann, and H. Xu. Canonical forms for Hamiltonian and symplectic matrices and pencils. *Linear Algebra Appl.*, 301-303:469–533, 1999.

[35] D.S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Structured polynomial eigenvalue problems: Good vibrations from good linearizations. *SIAM J. Matrix Anal. Appl.*, 28(4):1029–1051, 2006.

[36] D.S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Vector spaces of linearizations for matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 28(4):971–1004, 2006.

[37] D.S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Jordan structures of alternating matrix polynomials. *Linear Algebra Appl.*, 432(4):867–891, 2010.

[38] D.S. Mackey, N. Mackey, C. Mehl, and V. Mehrmann. Smith forms for palindromic matrix polynomials. *Electron. J. Linear Algebra*, 22:53–91, 2011.

[39] M. Marcus and H. Minc. *A Survey of Matrix Theory and Matrix Inequalities.* Allyn and Bacon, Boston, 1964.

[40] B. McMillan. Introduction to formal realizability theory – i. *Bell System Tech. J.*, 31:217–279, 1952.

[41] B. McMillan. Introduction to formal realizability theory – ii. *Bell System Tech. J.*, 31:541–600, 1952.

[42] V. Mehrmann, C. Schröder, and V. Simoncini. An implicitly-restarted Krylov subspace method for real symmetric/skew-symmetric eigenproblems. *Linear Algebra Appl.*, 436:4070–4087, 2012.

[43] R. V. Patel. Computation of the stable deflating subspace of a symplectic pencil using structure preserving orthogonal transformations. In *Proceedings 31st Annual Allerton Conf. Comm., Cont. Comp.*, University of Illinois, Urbana-Champaign, IL, Sep. 29 - Oct. 1, 1993.

[44] R. V. Patel. On computing the eigenvalues of a symplectic pencil. *Linear Algebra Appl.*, 188/189:591–611, 1993.

[45] V. V. Prasolov. *Problems and Theorems in Linear Algebra.* American Mathematical Society, Providence, 1994.

[46] A. C. M. Ran, L. Rodman, and J.E. Rubin. Direct complements of invariant Lagrangian subspaces and minimal factorizations of skew-symmetric rational matrix functions. *Linear Algebra Appl.*, 180:61–94, 1993.

[47] H. H. Rosenbrock. *State Space and Multivariable Theory.* Wiley, New York, 1970.

[48] C. Schröder. A canonical form for palindromic pencils and palindromic factorizations. Technical Report 316, DFG Research Center MATHEON, *Mathematics for key technologies* in Berlin, TU Berlin, Str. des 17. Juni 136, D-10623 Berlin, Germany, 2006. `url: http://www.matheon.de/`.

[49] C. Schröder. *Palindromic and Even Eigenvalue Problems - Analysis and Numerical Methods.* PhD thesis, Technical University Berlin, Germany, 2008.

[50] R.C. Thompson. Pencils of complex and real symmetric and skew matrices. *Linear Algebra Appl.*, 147:323–371, 1991.

[51] C. F. Van Loan. A symplectic method for approximating all the eigenvalues of a Hamiltonian matrix. *Linear Algebra Appl.*, 61:233–251, 1984.

[52] D. S. Watkins. On Hamiltonian and symplectic Lanczos processes. *Linear Algebra Appl.*, 385:23–45, 2004.