Acta Sci. Math. (Szeged)
8z (201t), 1–28

Equivalences of matrix polynomials

Karl-Heinz Förster and Béla Nagy*

Communicated by L. Kérchy

Abstract. We investigate whether variants of equivalence of (singular) matrix polynomials imply those of the first companion linearizations, and the converse question. We study a method of deciding whether two polynomials are strictly equivalent, and which are all the pairs of matrices effecting this equivalence. The corresponding problems for the (strict) similarity of square matrix polynomials are studied. We investigate the strict equivalence of a square polynomial to a polynomial whose all coefficient matrices are diagonal, and study also the singular case.

1. Introduction

We consider in this paper matrices over $\mathbb{C}$ or $\mathbb{C}[\lambda]$. If G is an $m \times k$ matrix, we shall write $G \in \mathbf{M}(m,k)$, and $\mathbf{M}(m)$ will stand for $\mathbf{M}(m,m)$. $\mathrm{Diag}(G_1,\dots,G_r)$ will denote the block diagonal matrix with diagonal blocks $G_1,\dots,G_r$.

Consider a matrix polynomial of exact degree $n \equiv n_A \equiv \deg(A) \geq 1$:

$$A(\lambda) = \sum_{j=0}^{n} A_j \lambda^j, \qquad A_n \neq 0.$$

The matrices A_j are assumed to be (of size or, equivalently, of type) $m \times k$, hence the polynomials are, in general, *singular*. The *first companion linearization* of $A(\lambda)$

Received February 22, 2012, and in revised form April 5, 2012.
AMS Subject Classifications: 47A56, 15A22.
Key words and phrases: singular matrix polynomial, companion linearization, equivalence, strong and strict equivalence, similarity, strictly diagonizable, (block) diagonally strict equivalence.
*Supported by the Hungarian OTKA Grant No. K77748.

is the linear matrix polynomial (pencil)

$$C_A(\lambda) := \begin{pmatrix} \lambda I_k & -I_k & 0 & 0 & \dots & 0 \\ 0 & \lambda I_k & -I_k & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda I_k & -I_k \\ A_0 & A_1 & A_2 & \dots & A_{n-2} & A_{n-1} + \lambda A_n \end{pmatrix} \in \mathbf{M}((n-1)k + m, nk),$$

where $I_k \in \mathbf{M}(k)$ will denote the identity matrix.

We shall work with pairs of matrix polynomials, and we shall apply the notation n_A for the degree of the polynomial A, and m_A and k_A for the number of rows and columns of (the coefficents of) A, respectively, and similar notation for other matrix polynomials.

There have been in the literature several reasonable and useful *concepts of equivalence of matrix polynomials*, and we shall review some of them in a short form. Our main concern will be a study of the *strict equivalence and similarity*, respectively, of polynomials, and we shall recall the exact definitions at the corresponding places. We shall study whether a variant of equivalence implies the corresponding relation for the first companion linearizations, and also the converse question.

We shall prove that *strict equivalence* of (maybe singular) polynomials implies that of the first companion linearizations (and a bit more), but the converse is not valid (though valid for the stricter concept). A variant for the particular case of monic polynomials will also be proved. We shall then study one method of deciding whether two polynomials are strictly *equivalent*, and what are all the pairs of matrices effecting this equivalence. Note that for the classical variant of this problem [for *linear* polynomials (pencils)] Kronecker obtained a complete set of invariants (the solution of the classification problem), see, e.g., [7, Ch. 12]. On the other hand, the classification problem with respect to strict equivalence for polynomials of degree $n \geq 2$ was shown to be "wild" in the terminology of representation theory (see, e.g., Sergeichuk [17], Belitskii and Sergeichuk [1], Byers, Mehrmann and Xu [3] on this kind of classification problems).

We shall study the corresponding problems for the (strict) *similarity* of square matrix polynomials. Note that a very short but remarkable result by Gelfand and Ponomarev [8] showed that the corresponding classification problem is wild even for the case $n = 1$ (see, e.g., also Nathanson [16], and Belitskii and Sergeichuk [1]). Nevertheless, Friedland in a celebrated paper [6] exhibited a process for obtaining a complete set of invariants for this classification problem. It might be interesting for the reader to compare our method with his results as applied to the low dimensional case $m = k = 2$ and two pairs of matrices in [6, pp. 200–209].

In the final section we shall prove a theorem on the strict equivalence of a square polynomial of degree n to a polynomial whose all coefficient matrices are *diagonal*. Note that this problem seems to be interesting for applications even in the low degree case $n = 2$. We shall also apply a result of Mitra [15] for the more general case of a singular polynomial.

We shall exhibit two relevant examples, the second of which will show how *all the pairs of matrices effecting the similarity* of two second degree monic polynomials in the case $m = k = 2$ can be determined.

We note that the so called *semi-scalar equivalence* of matrix polynomials will not be studied in this paper. We refer the reader to [19], [20] and [5], and to the references there.

The authors are much indebted to a referee calling their attention to the paper [18] dealing with the problem of simultaneous similarity for a pair of 4×4 matrices. We feel that our method of handling the low dimensional case $m = k = 2$ is completely different, and the obtained results are significantly more explicit.

2. Equivalence and linearizations

It is instructive to compare some equivalence properties of (in general, singular) matrix polynomials and the corresponding properties of their first companion linearizations. Let us recall these concepts.

The matrix polynomials $A(\lambda), B(\lambda)$ are called *(simply) equivalent*, if

$$E(\lambda)A(\lambda)F(\lambda) = B(\lambda),$$

where the square matrix polynomials $E(\lambda), F(\lambda)$ are *unimodular*, i.e. the functions $\det[E(\lambda)]$, $\det[F(\lambda)]$ are two nonzero constants. Equivalently, (see [7, Ch 6.3]), $A(\lambda), B(\lambda)$ have the same invariant polynomials. Note that the two degrees n_A, n_B can be different.

We want also to study when the first companion linearizations $C_A(\lambda), C_B(\lambda)$ are equivalent.

It is well known that $C_A(\lambda)$ is equivalent to the block diagonal matrix $\mathrm{Diag}[A(\lambda), I_{k_A(n_A-1)}] \in \mathbf{M}(m_A + k_A(n_A - 1), k_A n_A)$; note that the block entry in the left upper corner of this block matrix is of size $m_A \times k_A$. This implies the following fact that will be used in proof of the next theorem: $C_A(\lambda)$ has as invariant polynomials those of $A(\lambda)$ plus, in addition, the invariant polynomial identically 1 in $k_A(n_A - 1)$ copies.

Theorem 1. *Consider the following 3 statements:*

1) *The matrix polynomials $A(\lambda), B(\lambda)$ are equivalent,*
2) *The first companion linearizations $C_A(\lambda), C_B(\lambda)$ are equivalent,*
3) *The degrees n_A, n_B are equal.*

Any two of these together imply the third, but no one alone implies any of the remaining two.

Proof. Assume 1) and 2). From 1) we have $m_A = m_B =: m, k_A = k_B =: k$. From 2) the sizes of the two linearizations are the same, hence $n_A k = n_B k$. Thus 3) holds.

Assume 1) and 3). With the notation m, k as above and $n := n_A = n_B$, both linearizations have the size $[m + (n-1)k] \times nk$. From 1) the matrix polynomials $A(\lambda), B(\lambda)$ have the same invariant polynomials. As mentioned above $C_A(\lambda)$ and $C_B(\lambda)$ have as invariant polynomials those of $A(\lambda)$ (or, which is in our case the same, those of $B(\lambda)$) plus, in addition, the invariant polynomial identically 1 in $k(n-1)$ copies. Hence they have the same invariant polynomials, i.e. they are equivalent, thus 2) holds.

Assume 2) and 3). Applying our standard notation and [7, Ch. 6.3], it follows that the list of the invariant polynomials of $A(\lambda)$ amplified by $k_A(n_A - 1)$ copies of the identically 1 invariant polynomial is the same as the list of the invariant polynomials of $B(\lambda)$ amplified by $k_B(n_B - 1)$ copies of the identically 1 invariant polynomial. Further, the equivalent linearizations $C_A(\lambda), C_B(\lambda)$ have identical sizes, which gives

$$(n_A - 1)k_A + m_A = (n_B - 1)k_B + m_B, \qquad n_A k_A = n_B k_B.$$

Hence we obtain that the assumption $n_A = n_B$ (equivalently, $k_A = k_B$) implies $m_A = m_B$. Thus the sizes of $A(\lambda), B(\lambda)$ are identical, and they have the same list of invariant polynomials. It follows that $A(\lambda), B(\lambda)$ are equivalent, hence 1) holds.

The following example is studied (in another context) in [4]: Let

$$A(\lambda) := \begin{pmatrix} \lambda^2 & \lambda \\ \lambda & 1 \end{pmatrix}, \qquad B(\lambda) := \begin{pmatrix} \lambda & 1 \\ 0 & 0 \end{pmatrix}.$$

Both polynomials are square, $m = k = 2,\ n_A = 2, n_B = 1$, and they are singular. It is easily seen that the polynomials are equivalent. It was shown in [4] that more is true: they are *strongly equivalent* in the sense that, in addition, the reversed polynomials defined by

$$A^+(\lambda) := \lambda^2 A(1/\lambda), \qquad B^+(\lambda) := \lambda B(1/\lambda)$$

are also equivalent. Since the polynomial $B(\lambda)$ has degree 1, its first companion linearization $C_B(\lambda)$ is itself. So it has the type 2×2. Since $n_A = 2$, the first companion linearization $C_A(\lambda)$ of $A(\lambda)$ has the type 4×4. Thus $C_A(\lambda), C_B(\lambda)$ are *not equivalent.* For this example 1) holds, but neither do 2) or 3).

Consider now the following example. Let

$$A(\lambda) := \begin{pmatrix} 1 & \lambda^2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda \end{pmatrix}, \qquad B(\lambda) := \mathrm{Diag}(1,1,1,1,1,\lambda).$$

These matrix polynomials are regular, and $n_A = 2, n_B = 1$. The first companion linearization $C_A(\lambda)$ is simply equivalent to the block diagonal matrix polynomial

$$\mathrm{Diag}[A(\lambda), I_3],$$

and the first companion linearization $C_B(\lambda)$ is $B(\lambda)$ itself. It follows that the first companion linearizations are of the same size and equivalent, but the polynomials $A(\lambda), B(\lambda)$ are clearly *not equivalent.* In this example 2) holds, but not 1) or 3).

Finally, it is evident that there are very many examples where even $n_A = n_B = 1$, but the matrix polynomials $A(\lambda), B(\lambda)$ are not equivalent and their first companion linearizations (which coincide with the polynomials since $n_A = n_B = 1$) are not equivalent. Hence 3) holds without 1) or 2). ∎

Now we want to study similar questions for the stronger concept of strong equivalence of singular polynomials. Since there are more (reasonable) variants in the literature, we shall accept and fix the following

Definition 1. The $m \times k$ matrix polynomials $A(\lambda)$, $B(\lambda)$ are called *strongly equivalent*, if they are equivalent, and their reversed polynomials $A^+(\lambda)$, $B^+(\lambda)$ defined by

$$A^+(\lambda) := \lambda^{n_A} A(\lambda^{-1}), \qquad B^+(\lambda) := \lambda^{n_B} B(\lambda^{-1})$$

are also equivalent.

With the help of the above theorem we can prove the following

Theorem 2. *Assume that the $m \times k$ matrix polynomials $A(\lambda)$, $B(\lambda)$ satisfy*

$$n_A = n_B, \qquad n_{A^+} = n_{B^+}.$$

Then $A(\lambda)$, $B(\lambda)$ are strongly equivalent if and only if their first companion linearizations $C_A(\lambda)$, $C_B(\lambda)$ are strongly equivalent.

Proof. The preceding theorem applied to the pairs $A(\lambda)$, $B(\lambda)$ and $A^+(\lambda)$, $B^+(\lambda)$ shows that, under our conditions, $A(\lambda), B(\lambda)$ are strongly equivalent if and only if the pairs $C_A(\lambda), C_B(\lambda)$ and $C_{A^+}(\lambda), C_{B^+}(\lambda)$ are equivalent. We shall show now that the companion linearization $C_{A^+}(\lambda)$ is equivalent to the reversed polynomial $[C_A]^+(\lambda)$.

Indeed, it is known (even for the singular case, cf. [12, Lemma 1]) that the pencil $C_{A^+}(\lambda)$ is equivalent to the block diagonal matrix

$$\mathrm{Diag}[A^+(\lambda), I_{k_{A^+}(n_{A^+}-1)}].$$

On the other hand, [9, p. 780] shows that the latter matrix polynomial is equivalent to the reversed pencil $[C_A]^+(\lambda)$ (the proof there is valid also for the singular case). It is clear that the same argument proves the simple equivalence of the companion linearization $C_{B^+}(\lambda)$ to the reversed polynomial $[C_B]^+(\lambda)$. It follows that $C_A(\lambda), C_B(\lambda)$ are strongly equivalent. ■

Corollary. *Under the conditions of the above theorem assume that the matrix polynomials $A(\lambda)$, $B(\lambda)$ are regular, i.e. $m = k$, and their determinants are not identically* 0*. If $A(\lambda)$, $B(\lambda)$ are strongly equivalent, then their first companion linearizations $C_A(\lambda)$, $C_B(\lambda)$ are strictly equivalent, i.e. there are invertible matrices E, F of appropriate types such that*

$$EC_A(\lambda)F = C_B(\lambda).$$

Proof. We have proved that the first companion linearizations $C_A(\lambda), C_B(\lambda)$ are strongly equivalent. If $A(\lambda), B(\lambda)$ are regular, then the pencils $C_A(\lambda), C_B(\lambda)$ are also regular. It follows that they have the same elementary divisors at each finite point and at infinity. Hence (see [7, Ch. 12.2] or [9, p. 781]) the two pencils are strictly equivalent. ■

Remark 1. We shall see in the example after Theorem 4 that if the pencils $C_A(\lambda), C_B(\lambda)$ are regular and strictly equivalent, it does not follow that the matrix polynomials $A(\lambda), B(\lambda)$ are strictly equivalent.

The *leading* coefficients of C_A, C_B are both *invertible* if and only if the leading coefficients of the square polynomials A, B are. (An important special case is when all are *monic* matrix polynomials.) By Theorem 1, if $n_A = n_B$, the two polynomials A, B are equivalent if and only if C_A, C_B are equivalent. This holds if and only if they have the same (finite) elementary divisors. Assuming invertible leading coefficients, by [7, Ch. 12, Theor. 1], this is the case if and only if the pencils C_A, C_B are *strictly equivalent.* As mentioned above, we shall show that A, B *need not be strictly equivalent.*

Definition 2. We say that two first companion linearizations $C_A(\lambda)$ and $C_B(\lambda)$ in $\mathbf{M}((n-1)k+m, nk)$ of two $m \times k$ matrix polynomials $A(\lambda), B(\lambda)$ of degree n are *(block) diagonally strictly equivalent*, if there are correspondingly partitioned block diagonal matrices $X := \mathrm{Diag}(X_1, X_2, \ldots, X_{n-1}, X_n)$ and $Y := \mathrm{Diag}(Y_1, Y_2, \ldots, Y_{n-1}, Y_n)$ satisfying $XC_A(\lambda)Y = C_B(\lambda)$. The size of the block X_n is m, that of all the other indicated blocks is k, and all the blocks are invertible matrices.

Diagonally strict equivalence of first companion linearizations in $\mathbf{M}(n(k-1)+m, nk)$ clearly implies their strict equivalence, but the converse is false.

Theorem 3. *Assume that two $m \times k$ matrix polynomials $A(\lambda), B(\lambda)$ are strictly equivalent, i.e., there are invertible $m \times m, k \times k$ matrices E, F, such that*

$$B(\lambda) = EA(\lambda)F.$$

Then the linear matrix polynomials (pencils) $C_A(\lambda)$ and $C_B(\lambda)$ are block diagonally strictly equivalent. Conversely, if the linear matrix polynomials (pencils) $C_A(\lambda)$ and $C_B(\lambda)$ are block diagonally strictly equivalent, then $A(\lambda)$ and $B(\lambda)$ are strictly equivalent.

Proof. Assume that two $m\times k$ matrix polynomials $A(\lambda), B(\lambda)$ are strictly equivalent. Then $n_A = n_B = n$, and the inverse matrix F^{-1} exists. It is clear that the product of the block matrices

$$\mathrm{Diag}(F^{-1}, F^{-1}, \ldots, F^{-1}, E)C_A(\lambda)\,\mathrm{Diag}(F, F, \ldots, F, F)$$

is equal to $C_B(\lambda)$.

In the converse direction assume that

$$\mathrm{Diag}(X_1, X_2, \ldots, X_{n-1}, X_n)C_A(\lambda)\,\mathrm{Diag}(Y_1, Y_2, \ldots, Y_{n-1}, Y_n) = C_B(\lambda).$$

If we compare the coefficients of λ, 1 in this equation, we obtain for the block entries

$$X_jY_j = I_k \qquad (j = 1, 2, \ldots, n-1)$$

and

$$X_jY_{j+1} = I_k \qquad (j = 1, 2, \ldots, n-1).$$

Let $F := Y_1$. F is then invertible, and we obtain successively that

$$X_1 = X_2 = \cdots = X_{n-1} = F^{-1}, \qquad Y_1 = Y_2 = \cdots = Y_n = F.$$

Let $E := X_n$. Then we obtain $EA_jF = B_j$ for $j = 0, 1, \dots, n$, hence

$$EA(\lambda)F = B(\lambda).$$

■

Making use of the above theorem, we want to study when two $k \times k$ *monic* matrix polynomials $A(\lambda), B(\lambda)$ of degree n are *strictly similar*, i.e., when a $k \times k$ invertible matrix F exists such that

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

Applying some ideas of [13], we obtain the following

Theorem 4. *For the $k \times k$ monic matrix polynomials $A(\lambda), B(\lambda)$ the following are equivalent:*

1) *There exists a $k \times k$ invertible matrix F such that*

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

2) *Let $C[A]$ denote the first companion matrix for $A(\lambda)$, i.e.*

$$-C[A] := \begin{pmatrix} 0 & -I_k & 0 & 0 & \dots & 0 \\ 0 & 0 & -I_k & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & -I_k \\ A_0 & A_1 & A_2 & \dots & A_{n-2} & A_{n-1} \end{pmatrix},$$

and similarly for $C[B]$, where the matrices I_k are identity matrices of order k. Then there is an invertible $nk \times nk$ matrix G with correspondingly partitioned first block row

$$\begin{pmatrix} G_1 & 0 & \dots & 0 \end{pmatrix},$$

where the blocks are of order k, such that

$$C[B] = GC[A]G^{-1}.$$

3) *The inverse $B(\lambda)^{-1}$ of the matrix polynomial $B(\lambda)$ can be represented with correspondingly partitioned block matrices as*

$$B(\lambda)^{-1} = \begin{pmatrix} F^{-1} & 0 & \dots & 0 \end{pmatrix} (I\lambda - C[A])^{-1} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ F \end{pmatrix}.$$

Proof. Assume 1). With the notation $G := \mathrm{Diag}(F^{-1}, F^{-1}, \ldots, F^{-1})$ (n times), it follows that $C[B] = GC[A]G^{-1}$, and the first block row of G is $\begin{pmatrix} F^{-1} & 0 & \ldots & 0 \end{pmatrix}$. Hence 2) holds.

Assume 2). The first block row of $C[B]G$ is clearly the second block row of G. The first block row of $GC[A]$ is $\begin{pmatrix} 0 & G_1 & 0 & \ldots & 0 \end{pmatrix}$. From $C[B]G = GC[A]$ we obtain that the second block row of G is $\begin{pmatrix} 0 & G_1 & 0 & \ldots & 0 \end{pmatrix}$. Continuing similarly we see that the matrix G must be equal to $\mathrm{Diag}(G_1, G_1, \ldots, G_1)$. Since G is invertible, so is G_1, and we can define $F := G_1^{-1}$. From

$$C[B] = \mathrm{Diag}(F^{-1}, F^{-1}, \ldots, F^{-1})C[A]\,\mathrm{Diag}(F, F, \ldots, F)$$

we then obtain $B(\lambda) = F^{-1}A(\lambda)F$, hence 1) and 2) are equivalent.

Assume 1). Then for every λ not in the spectrum of the polynomial $A(\lambda)$ we have

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

The resolvent representation of a monic polynomial [10, p. 14] shows that

$$A(\lambda)^{-1} = \begin{pmatrix} I_k & 0 & \ldots & 0 \end{pmatrix} (I\lambda - C[A])^{-1} \begin{pmatrix} 0 & \ldots & 0 & I_k \end{pmatrix}^T,$$

where ^{T} denotes transpose. Multiplying by F^{-1}, F from the appropriate sides, we obtain the statement of 3). Reversing the steps of this paragraph, we see that 3) implies 1), hence they are equivalent. ∎

The following example will show, e.g., that $C_A(\lambda)$ can be strictly equivalent to $C_B(\lambda)$ even if $A(\lambda), B(\lambda)$ are *not* strictly equivalent (cf. Remark after Theorem 2).

Example. Consider the case of two *monic* polynomials with the parameters $n = m = k = 2$. Then

$$C_A(\lambda) = \lambda I_4 + \begin{pmatrix} 0 & -I_2 \\ A_0 & A_1 \end{pmatrix} = \lambda I_4 - C[A],$$

and similarly for $C_B(\lambda)$. Hence $C_A(\lambda)$ is strictly equivalent to $C_B(\lambda)$ exactly when there is an invertible 4×4 matrix X satisfying $C_A(\lambda)X = XC_B(\lambda)$, which is equivalent to

$$\begin{pmatrix} 0 & -I_2 \\ A_0 & A_1 \end{pmatrix} X = X \begin{pmatrix} 0 & -I_2 \\ B_0 & B_1 \end{pmatrix}.$$

Introducing the notation

$$\hat{A} := -C[A], \qquad \hat{B} := -C[B],$$

this becomes $\hat{A}X = X\hat{B}$. We want to analyze *all solutions* X (not only the block diagonal ones) of this equation (cf. [7, Ch. 8]).

Consider fixed Jordan forms $J_{\hat{A}}, J_{\hat{B}}$ of the matrices $\hat{A}, \hat{B}$, respectively. We can and will assume that the Jordan blocks are arranged so that the similarity of $\hat{A}$ and $\hat{B}$ implies $J_{\hat{A}} = J_{\hat{B}}$. Denoting the corresponding transformation matrices by U, V, we obtain

$$\hat{A} = UJ_{\hat{A}}U^{-1}, \qquad \hat{B} = VJ_{\hat{B}}V^{-1} = VJ_{\hat{A}}V^{-1}.$$

The general solution of the equation $\hat{A}X = X\hat{B}$ has the form

$$X = U\tilde{X}V^{-1},$$

where $\tilde{X}$ is the general 4×4 matrix commuting with $J_{\hat{A}}$. In our special case $\tilde{X}$ is, in addition, invertible. It is clear that, under our assumption, *there always exists an invertible solution* X.

However, *it may happen that there is no block diagonal solution (as required in the Theorem).* Let $\hat{A}$ and $\hat{B}$ be defined by

$$\hat{A} := \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 1 & 1 & 2 \\ 1 & 1 & 2 & 1 \end{pmatrix}, \qquad \hat{B} := \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 4 & -2 & 2 & -2 \\ 2 & 1 & -2 & 0 \end{pmatrix}.$$

The spectrum of $\hat{A}$ and of $\hat{B}$ is the set $\{-1, 0, 1, 2\}$, hence the matrices are similar. Assume that X is a block diagonal invertible solution of the equation $\hat{A}X = X\hat{B}$ of the required form, say,

$$X := \begin{pmatrix} v & x & 0 & 0 \\ y & z & 0 & 0 \\ 0 & 0 & v & x \\ 0 & 0 & y & z \end{pmatrix}.$$

We obtain that the matrix

$$\hat{A}X - X\hat{B} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -3v + y - 2x & 2x + z + 2v & -v + 2y + 2x & x + 2z + 2v \\ v - 3y - 2z & x + 2z + 2y & 2v - y + 2z & 2x + z + 2y \end{pmatrix}$$

must be the zero matrix of size 4×4. It follows that $v = x = y = z = 0$, hence X is the zero matrix, a contradiction.

3. Equivalence and commutants

In this section we ask when two families $\{A_j \in \mathbf{M}(m,k) : j \in \mathbb{J}\}$ and $\{B_j \in \mathbf{M}(m,k) : j \in \mathbb{J}\}$ are *simultaneously equivalent*, i.e., there exist invertible $E \in \mathbf{M}(m)$ and $F \in \mathbf{M}(k)$ such that

$$B_j = EA_jF \quad \text{for all} \quad j \in \mathbb{J}.$$

Since $\mathbf{M}(m,k)$ is finite dimensional we can without loss of generality assume that the family is finite. In the following considerations we use and extend some ideas of O. E. Brown [2].

For $j \in \mathbb{J} = \{0,1,\dots,l\}$ let $A_j, B_j \in \mathbf{M}(m,k)$ be equivalent and

$$B_j = E_jA_jF_j, \quad j = 0,1,\dots,l,$$

where $E_j \in \mathbf{M}(m)$ and $F_j \in \mathbf{M}(k)$ are invertible. Then $\operatorname{rank}(A_j) = \operatorname{rank}(B_j) =: r_j \le \min(m,k)$, and there exist invertible L_j, M_j in $\mathbf{M}(m)$ and invertible U_j, V_j in $\mathbf{M}(k)$ such that

$$L_jA_jU_j = D_{r_j} = M_jB_jV_j,$$

where D_{r_j} denotes the diagonal matrix in $\mathbf{M}(m,k)$ whose entry in position (p,p) is 1 for $p = 1,\dots,r_j$ and all other entries are zero. Define the following matrices:

$$N_j := M_jE_jL_j^{-1}, \qquad W_j := V_j^{-1}F_j^{-1}U_j, \quad j = 0,1,\dots,l.$$

Then we have

$$N_jD_{r_j} = D_{r_j}W_j, \qquad j = 0,1,\dots,l. \tag{1}$$

Theorem 5. *Let $A_j, B_j \in \mathbf{M}(m,k)$ be equivalent, $j = 0,1,\dots,l$. In the following we use the notation and the relations described above.*

Assume that $(A_0,\dots,A_l)$ and $(B_0,\dots,B_l)$ are simultaneously equivalent, i.e. the matrices E_j, F_j can be chosen independently of j. Then the product matrices in each row below

$$M_j^{-1}N_jL_j, \qquad j = 0,1,\dots,l,$$

$$U_jW_j^{-1}V_j^{-1}, \qquad j = 0,1,\dots,l,$$

are independent of j.

In the converse direction: Assume that N_j and W_j, $j = 0,1,\dots,l$, are invertible matrices in $\mathbf{M}(m)$ and $\mathbf{M}(k)$, respectively, satisfying the intertwining relations (1), and assume that each of the two families of matrices above is independent of the *index j. Then $(A_0, \dots, A_l)$ and $(B_0, \dots, B_l)$ are simultaneously equivalent, and we have*

$$EA_jF = B_j, \qquad j = 0, 1, \dots, l,$$

where $E = M_j^{-1}N_jL_j$ and $F = U_jW_j^{-1}V_j^{-1}$, $j = 0, 1, \dots, l$.

The proof follows from the considerations before the theorem.

Remark 2. This theorem yields a possibility of deciding whether such a pair (E, F) exists in a concrete situation, though the volume of the needed calculations may be large.

Let us start with one pair (A, B) of $m \times k$ matrices, and seek first the class $\mathbf{H}(A, B)$ of *all pairs (E,F) of invertible matrices* in $\mathbf{M}(m) \times \mathbf{M}(k)$ satisfying

$$EAF = B.$$

A, B must have the same rank $r \leq \min(m, k)$. Applying the notation before the theorem (without the subscripts j), we want to determine all the pairs $(N, W) \in \mathbf{M}(m) \times \mathbf{M}(k)$ of invertible matrices satisfying

$$ND_r = D_rW, \qquad \text{where the block matrix} \qquad D_r = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \in \mathbf{M}(m, k),$$

and I_r is the identity in $\mathbf{M}(r)$.

Let N and W have the corresponding block partitions

$$N = (N_{gh}), \qquad W = (W_{gh}) \quad (g, h = 1, 2),$$

where N and W are of sizes m and k, respectively, and N_{11}, W_{11} are of size r. It follows that (with appropriate zero block matrices 0)

$$N_{11} = W_{11}, \quad N_{21} = 0, \quad W_{12} = 0.$$

Hence they have the block structure

$$N = \begin{pmatrix} N_{11} & N_{12} \\ 0 & N_{22} \end{pmatrix}, \qquad W = \begin{pmatrix} N_{11} & 0 \\ W_{21} & W_{22} \end{pmatrix}.$$

This structure shows that the square submatrices N_{11}, N_{22}, W_{22} must be invertible. In the converse direction: if the pair (N, W) has this block structure, then

$$ND_r = \begin{pmatrix} N_{11} & 0 \\ 0 & 0 \end{pmatrix} = D_rW.$$

Pick now *fixed invertible* matrices L, U, M, V satisfying $LAU = D_r = MBV$, and to any pair (N, W) order the pair $(E, \tilde{F})$ defined by

$$E := M^{-1}NL, \qquad \tilde{F} = VWU^{-1}. \tag{2}$$

Note that the entries of the pair $(E, \tilde{F})$ will be linear combinations of the entries of the pair (N, W). Further, there is a 1-1 correspondence between the classes

$$\{(N, W) \in \mathbf{M}(m) \times \mathbf{M}(k) : ND_r = D_r W\},$$

and

$$\mathbf{H}(A, B) = \{(E, F) \in \mathbf{M}(m) \times \mathbf{M}(k) : E, F \text{ invertible}, EAF = B\}.$$

Here $E = M^{-1}NL$ and $F = \tilde{F}^{-1} = UW^{-1}V^{-1}$.

Thus we have obtained in (2) a description of the sought class $\mathbf{H}(A, B)$ as a function of the entries of the corresponding pair (N, W).

Remark 3. Our basic problem of the simultaneous equivalence of pairs of $(l + 1)$-tuples $(A_0, A_1, \ldots, A_l)$ and $(B_0, B_1, \ldots, B_l)$ clearly has as solutions all the matrix pairs (E, F) satisfying

$$(E, F) \in \bigcap_{j=0}^{l} \mathbf{H}(A_j, B_j).$$

The tuples are *not* simultaneously equivalent if and only if the right-hand side is *void*.

Remark 4. Theorem 5 can clearly be formulated in terms of the strict equivalence

$$EA(\lambda)F = B(\lambda)$$

of the corresponding matrix polynomials.

Remark 5. Theorem 5 gives necessary and sufficient conditions in order the finite sequences of matrices $(A_0, A_1, \ldots, A_l)$ and $(B_0, B_1, \ldots, B_l)$ represent the same finite sequence $(T_0, T_1, \ldots, T_l)$ of linear operators between the finite dimensional spaces $\mathbb{C}^k$ and $\mathbb{C}^m$ *with respect to fixed, but possibly different bases in the two spaces.*

Assume now that $k = m$ and allow only the use of the *same basis in both the domain and the range space*. In other words, we then ask when an invertible matrix F can be found such that

$$F^{-1}A_j F = B_j \qquad (j = 0, 1, \ldots, n), \tag{3}$$

i.e. $(A_0, A_1, \dots, A_l)$ and $(B_0, B_1, \dots, B_l)$ are simultaneously similar.

Note that this question is equivalent to the problem of when the *monic* matrix polynomials of degree $l+1$

$$A(\lambda) = \sum_{j=0}^{l} A_j \lambda^j + I\lambda^{l+1}, \qquad B(\lambda) = \sum_{j=0}^{l} B_j \lambda^j + I\lambda^{l+1}$$

are *strictly equivalent*.

Assume that the $k \times k$ matrices A and B are similar, and T is an invertible matrix satisfing

$$AT = TB.$$

If H is any invertible matrix satisfying $AH = HB$, it can be written as $H = TZ$, where Z is a uniquely determined invertible matrix. It follows that

$$TBZ = ATZ = TZB.$$

Since T is invertible, we obtain $BZ = ZB$. Denote by $\mathbf{T}(A, B)$ the *set of all* invertible matrices H satisfying $AH = HB$, and denote by $\mathbf{K}(M)$ the *commutant* of any matrix M, i.e. the set of all invertible matrices L satisfying $LM = ML$. The preceding paragraph then shows that $\mathbf{T}(A, B) = T\mathbf{K}(B)$ for all $T \in \mathbf{T}(A, B)$. In a completely similar way we obtain that $\mathbf{T}(A, B) = \mathbf{K}(A)T$ for all $T \in \mathbf{T}(A, B)$.

Let J_A denote a fixed Jordan form for A, and T_A denote an invertible matrix satisfying

$$AT_A = T_A J_A.$$

It is known that the commutants of any matrix A and of J_A are connected by the following formula:

$$\mathbf{K}(A)T_A = T_A \mathbf{K}(J_A).$$

A description of the set $\mathbf{T}(A, B)$ is obtained in the following

Theorem 6. *The $k \times k$ invertible matrix F satisfies $AF = FB$ if and only if for some (equivalently, for any) fixed joint Jordan form $J = J_A = J_B$ and fixed invertible transformation matrices T_A, T_B as above we have*

$$F \in T_A \mathbf{K}(J) {T_B}^{-1}. \tag{4}$$

Proof. Assume (4). Then

$${T_A}^{-1} F T_B \in \mathbf{K}(J) = {T_A}^{-1} \mathbf{K}(A) T_A.$$

Hence

$$FT_B{T_A}^{-1} \in \mathbf{K}(A).$$

It follows that

$$FT_B{T_A}^{-1}A = AFT_B{T_A}^{-1}.$$

Multiplying by T_A from the right yields

$$FT_BJ = AFT_B.$$

Multiplying by ${T_B}^{-1}$ from the right gives $FB = AF$. By reversing these steps we obtain the proof of the converse.

■

Remark 6. In the notation of this theorem we have then

$$\mathbf{T}(A, B) = T_A\mathbf{K}(J){T_B}^{-1}.$$

Remark 7. It follows that any family of ordered pairs $\{(A_j, B_j) : \ j \in \mathbb{J}\}$ of similar $k \times k$ matrices is *simultaneously similar*, i.e. there is an invertible matrix F satisfying $A_jF = FB_j$ for every $j \in \mathbb{J}$ if and only if the set

$$\bigcap_{j\in\mathbb{J}} \mathbf{T}(A_j, B_j) = \bigcap_{j\in\mathbb{J}} T_j\mathbf{K}(B_j) = \bigcap_{j\in\mathbb{J}} T_jT_{B_j}\mathbf{K}(J_{B_j}){T_{B_j}}^{-1}$$

is not empty for all invertible T_j satisfying $A_jT_j = T_jB_j$ and all invertible T_{B_j} satisfying $B_jT_{B_j} = T_{B_j}J_{B_j}$. Since $\mathbf{M}(k)$ is finite dimensional we can without loss of generality assume that the family is finite.

4. Equivalence versus similarity for regular polynomials

In the case of a pair of *regular* matrix polynomials *of degree* n there is a connection between their *strict equivalence* and the *similarity* of a pair of suitably defined *monic* polynomials *of degree* n as follows.

Theorem 7. *Assume that the matrix polynomial $A(\lambda) = \sum_{j=0}^{n} A_j\lambda^j$ of degree n is regular, and the $k \times k$ matrix $A(a)$ is invertible for some $a \in \mathbb{C}$. Define the matrix polynomial*

$$\hat{A}_a(\lambda) := \lambda^n A(a)^{-1} \sum_{j=0}^{n} \left(\frac{1}{\lambda} + a\right)^j A_j.$$

Then $\hat{A}_a(\lambda)$ is a monic polynomial (i.e. its coefficient for λ^n is the identity I). Assume further that there are invertible matrices E, F of order k such that the polynomial

$$B(\lambda) := EA(\lambda)F$$

is strictly equivalent to $A(\lambda)$. Then $B(a)$ is invertible, and we can define the matrix polynomial $\hat{B}_a(\lambda)$ exactly as we did for $\hat{A}_a(\lambda)$ before. Then $\hat{B}_a(\lambda)$ is also monic, and we have

$$\hat{B}_a(\lambda) = F^{-1}\hat{A}_a(\lambda)F.$$

In the converse direction: assume that the equality above (with the same notation) holds. Then the coefficient matrices of the powers λ^j satisfy

$$\hat{B}_a^{(j)} = F^{-1}\hat{A}_a^{(j)}F.$$

Further, for every $\lambda \in \mathbb{C}$

$$B(\lambda) = B(a)F^{-1}A(a)^{-1}A(\lambda)F.$$

Hence with the notation $E := B(a)F^{-1}A(a)^{-1}$ we obtain

$$B(\lambda) = EA(\lambda)F, \qquad B_j = EA_jF \quad (j = 0, 1, \dots, n).$$

Proof. It is straightforward to check that the polynomials $\hat{A}_a(\lambda)$, $\hat{B}_a(\lambda)$ are monic. By assumption, the inverse matrices below exist and satisfy

$$B(a)^{-1} = F^{-1}A(a)^{-1}E^{-1}.$$

Hence

$$\hat{B}_a(\lambda) = \lambda^n F^{-1}A(a)^{-1}E^{-1}\sum_{j=0}^{n}\left(\frac{1}{\lambda}+a\right)^j EA_jF = F^{-1}\hat{A}_a(\lambda)F.$$

In the converse direction: we obtain from the above equality $\hat{B}_a^{(j)} = F^{-1}\hat{A}_a^{(j)}F$ for each $j = 0, 1, \dots, n$. A further calculation yields

$$\begin{aligned} B(\lambda) &= B(a)\sum_{j=0}^{n}\hat{B}_a^{(j)}(\lambda - a)^{n-j} \\ &= B(a)\sum_{j=0}^{n}F^{-1}\hat{A}_a^{(j)}(\lambda - a)^{n-j}F = B(a)F^{-1}A(a)^{-1}A(\lambda)F = EA(\lambda)F. \end{aligned}$$

Comparing the coefficients on both sides, we obtain

$$B_j = EA_jF \quad (j = 0, 1, \dots, n).$$

∎

5. The case of quadratic monic matrix polynomials

We start to study the case of two *monic* polynomials

$$A(\lambda) := I_m\lambda^2 + A_1\lambda + \hat{A}_0, \qquad B(\lambda) := I_m\lambda^2 + B_1\lambda + \hat{B}_0$$

with the size parameters $m = k$. They are strictly equivalent if and only if there are invertible matrices $E, G \in \mathbf{M}(m)$ such that

$$I_m\lambda^2 + B_1\lambda + \hat{B}_0 \equiv B(\lambda) = EA(\lambda)G \equiv E(I_m\lambda^2 + A_1\lambda + \hat{A}_0)G.$$

Comparing the first terms on both sides, we obtain $E = G^{-1}$. Applying the notation from Section 3, if J_1 is a fixed joint Jordan form of A_1 and B_1, and T_{A_1}, T_{B_1} are the corresponding fixed transforming matrices, then

$$\begin{aligned} {T_{A_1}}^{-1}(I_m\lambda^2 + A_1\lambda + \hat{A}_0)T_{A_1} &= I_m\lambda^2 + J_1\lambda + A_0, \\ {T_{B_1}}^{-1}(I_m\lambda^2 + B_1\lambda + \hat{B}_0)T_{B_1} &= I_m\lambda^2 + J_1\lambda + B_0. \end{aligned}$$

Here we have introduced the notation

$$A_0 := {T_{A_1}}^{-1}\hat{A}_0 T_{A_1}, \qquad B_0 := {T_{B_1}}^{-1}\hat{B}_0 T_{B_1}.$$

Hence, our basic equality has the equivalent form

$$\begin{aligned} I_m\lambda^2 + J_1\lambda + B_0 &= {T_{B_1}}^{-1}E(I_m\lambda^2 + A_1\lambda + \hat{A}_0)GT_{B_1} \\ &= {T_{B_1}}^{-1}G^{-1}T_{A_1}(I_m\lambda^2 + J_1\lambda + A_0){T_{A_1}}^{-1}GT_{B_1} \\ &= F^{-1}(I_m\lambda^2 + J_1\lambda + A_0)F, \end{aligned} \tag{5}$$

where

$$F := {T_{A_1}}^{-1}GT_{B_1}.$$

It is clear that an invertible matrix F satisfies (5) if and only if

$$F^{-1}J_1F = J_1, \qquad F^{-1}A_0F = B_0.$$

The latter equality holds if and only if $F \in \mathbf{T}(A_0, B_0)$. Thus for a *fixed* common Jordan form J_0 for A_0 and B_0 and for *fixed* transforming matrices T_{A_0}, T_{B_0} the matrix F satisfies (5) if and only if

$$F \in \mathbf{K}(J_1), \qquad {T_{A_0}}^{-1}FT_{B_0} \in \mathbf{K}(J_0)$$

simultaneously. Note that the existence of such an F is equivalent to

$$\mathbf{K}(J_1)T_{B_0} \cap T_{A_0}\mathbf{K}(J_0) \neq \emptyset. \tag{6}$$

So *the two monic polynomials are strictly equivalent if and only if* (6) *holds.*

The low dimensional case $m = k = 2$

Assume now that the dimension k is equal to 2. It is clear that the Jordan form of any matrix in $\mathbf{M}(2)$ is one of the following types:

$$\text{(I)} \quad \begin{pmatrix} d & 0 \\ 0 & d \end{pmatrix}; \qquad \text{(II)} \quad \begin{pmatrix} d & 1 \\ 0 & d \end{pmatrix}; \qquad \text{(III)} \quad \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix},$$

where $d_1 \neq d_2$. The commutant of any matrix in the class (I) is the set $\mathbf{G}_2$ of all invertible matrices in $\mathbf{M}(2)$. The commutant of a matrix in (II) is the set of matrices of the form $\left(\begin{smallmatrix} m & q \\ 0 & m \end{smallmatrix}\right)$ satisfying $m \neq 0$. The commutant of a matrix in (III) is the set of matrices of the form $\left(\begin{smallmatrix} e & 0 \\ 0 & f \end{smallmatrix}\right)$ such that $ef \neq 0$.

Assume now that for $j = 0$ or $j = 1$ the matrix J_j is in class (I), hence $\mathbf{K}(J_j) = \mathbf{G}_2$. To consider a fixed case, let $j = 0$, hence $\mathbf{K}(J_0) = \mathbf{G}_2$, and write (6) in the form

$$T_{A_0}^{-1}\mathbf{K}(J_1)T_{B_0} \cap \mathbf{G}_2 \neq \emptyset.$$

By the definition of $\mathbf{G}_2$, this holds for any choice of the left-hand side product. Hence there is a matrix $F \in \mathbf{G}_2$ satisfying (5). The case $j = 1$ is completely similar, so we have to consider the remaining 4 possibilities. In each of these fix the appropriate matrices T_{A_0}, T_{B_0} and the following *notation*:

$$T_{A_0} \equiv a = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \qquad T_{B_0} \equiv b = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}.$$

Case 1: Assume first that *both* matrices J_j are in the class (III), and write (6) in the equivalent form

$$T_{B_0} \in \mathbf{K}(J_1)T_{A_0}\mathbf{K}(J_0).$$

Considering the forms of the matrices in $\mathbf{K}(J_j)$, this is equivalent to the existence of two diagonal matrices with nonzero diagonal entries satisfying

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \operatorname{diag}(e_1, e_2) \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \operatorname{diag}(f_1, f_2).$$

Equivalently, the following system of 4 equations must hold:

$$b_{ik} = e_i a_{ik} f_k \quad (i, k = 1, 2).$$

Clearly, here $a_{ik} = 0$ if and only if $b_{ik} = 0$. Assuming the opposite case for *every* pair (i, k), we may define $c_{ik} := b_{ik}/a_{ik} \neq 0$, and we obtain the system of 4 equations

$$c_{ik} = e_i f_k \quad (i, k = 1, 2). \tag{7}$$

If this system has a solution (e_1, e_2, f_1, f_2), then clearly

$$c_{11}c_{22} = e_1 f_1 e_2 f_2 = c_{12}c_{21}.$$

Conversely, if

$$c_{11}c_{22} = c_{12}c_{21} \neq 0,$$

then the choice

$$f_1 := 1, e_i := c_{i1}, f_2 := c_{i2}/e_i \qquad (i = 1, 2)$$

defines also $f_2 \neq 0$ unambiguously, and yields a solution of the system (7).

If there is at least one pair (i, k) such that $b_{ik} = 0$, for symmetry reasons and for simplicity in notation we may assume that $(i, k) = (1, 1)$, i.e. that

$$b_{11} = 0 = a_{11}.$$

Since a, b are invertible 2×2 matrices, we may have at most $b_{22} = 0 = a_{22}$, but the two other pairs are nonzero, hence define the nonzero c_{12}, c_{21} unambiguously. Define now $f_1 := 1$, then $c_{21} = e_2 f_1$ yields $e_2 := c_{21}$. The equation $0 \neq c_{12} = e_1 f_2$ determines the value of the latter product. The equation

$$b_{22} = a_{22} e_2 f_2$$

determines the value $f_2 \neq 0$ uniquely if $c_{22} \neq 0$ is defined, otherwise it is no condition on f_2. In the former case we obtain unambiguously $e_1 := c_{12}/f_2$, in the latter only the product $e_1 f_2 \neq 0$ is determined. In the converse direction: the so defined (e_1, e_2, f_1, f_2) clearly satisfies the system (7). In any case, we have obtained the following result:

Case 1. *If both matrices J_j are in the class (III), and at least one entry $b_{ik} = 0 = a_{ik}$, then* (6) *is satisfied. However, if each $b_{ik} \neq 0 \neq a_{ik}$* $(i, k = 1, 2)$, *then* (6) *is satisfied if and only if $c_{11}c_{22} = c_{12}c_{21} \neq 0$ holds.*

Consider now the next

Case 2. Assume that both matrices J_j are in the class (II), and write (6) in the equivalent form

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} m_1 & q_1 \\ 0 & m_1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} m & q \\ 0 & m \end{pmatrix}.$$

Here we have used the fact that the inverse of a matrix in $\mathbf{K}(J_1)$ is in the same class. We obtain then

$$b_{21} = m_1 a_{21} m.$$

Assume first, in addition, that $a_{21} \neq 0$. Since $m_1 \neq 0 \neq m$, this is equivalent to $b_{21} \neq 0$. Introduce the notation $c_{21} := b_{21}/a_{21}$. We obtain then

$$m_1 m = c_{21}.$$

Comparing the $(1,1)$ entries we obtain $b_{11} = a_{11}m_1 m + a_{21}q_1 m$, whence

$$q_1 m = (b_{11} - a_{11}c_{21})/a_{21}.$$

Similarly, from the $(2,2)$ entries we obtain

$$m_1 q = (b_{22} - a_{22}c_{21})/a_{21}.$$

Finally, the $(1,2)$ entries yield

$$q_1 q = (b_{12} - a_{12}c_{21} - a_{11}m_1 q - a_{22}q_1 m)/a_{21}.$$

Similarly to Case 1, it follows that

$$c_{21}[(b_{12} - a_{12}c_{21} - a_{11}m_1 q - a_{22}q_1 m)/a_{21}] = m_1 m q_1 q = [(b_{11} - a_{11}c_{21})/a_{21}]\cdot$$
$$\cdot[(b_{22} - a_{22}c_{21})/a_{21}].$$

A short calculation shows this to be equivalent to

$$a_{21}^2 \det(b) = b_{21}^2 \det(a). \tag{8}$$

In the converse direction: assume the above equality, and *define* the products $m_1 m, q_1 m, m_1 q, q_1 q$ as they stand above. A similar reasoning as in Case 1 shows that there is a solution (m_1, m, q_1, q) to the system of 4 equations above. From this the pair (m_1, q_1) determines a left-hand factor, and (m, q) determines a right-hand factor in the equality

$$b = \begin{pmatrix} m_1 & q_1 \\ 0 & m_1 \end{pmatrix} a \begin{pmatrix} m & q \\ 0 & m \end{pmatrix}. \tag{9}$$

Assume now that (in the notation of Case 2) $a_{21} = 0 = b_{21}$. From the equality of the entries $(1,1)$ and $(2,2)$ in a and b we obtain

$$a_{jj} m_1 m = b_{jj} \qquad (j = 1, 2),$$

and the invertibility of the matrices a and b shows that $a_{jj} \neq 0 \neq b_{jj}$. It follows that

$$b_{11}/a_{11} = m_1 m = b_{22}/a_{22}.$$

Hence the equality of the fractions above is a necessary condition. From the equality of the entries $(1,2)$ we obtain the equation

$$a_{11}m_1q + a_{22}q_1m = b_{12} - a_{12}b_{11}/a_{11}.$$

However we fix the values m_1 and m satisfying the necessary condition, this equation always has an infinity of pairs (q_1, q) satisfying it. In the converse direction: it is easy to see that each such quadruple (m_1, m, q_1, q) satisfies (9). We have obtained the following result:

Case 2. *If both matrices J_j are in the class (II), and $a_{21} \neq 0$, then* (6) *is satisfied if and only if the equality* (8) *holds. On the other hand, if $a_{21} = 0$ or, equivalently, $b_{21} = 0$, then* (6) *is satisfied if and only if*

$$b_{11}/a_{11} = b_{22}/a_{22}.$$

Case 3. Assume now that the matrix J_1 is in class (III), and J_0 is in class (II). Write (6) in the equivalent form

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} e_1 & 0 \\ 0 & e_2 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} m & q \\ 0 & m \end{pmatrix}.$$

Comparing the corresponding entries on both sides, we obtain

$$a_{j1}e_jm = b_{j1}, \qquad a_{j2}e_jm + a_{j1}e_jq = b_{j2} \qquad (j = 1, 2).$$

Assume first that for *both* values $j = 1, 2$

$$a_{j1} \neq 0 \quad \text{ or, equivalently } \quad b_{j1} \neq 0.$$

Then the 4 products e_jm, e_jq $(j = 1, 2)$ of two factors (each) are determined uniquely:

$$e_jm = b_{j1}/a_{j1}, \quad e_jq = [b_{j2} - a_{j2}b_{j1}/a_{j1}]/a_{j1} \quad (j = 1, 2).$$

A necessary condition for the solvability of this system of 4 equations (for the factors) is, similarly to the preceding cases,

$$(e_1m)(e_2q) = (e_2m)(e_1q).$$

This holds if and only if

$$a_{11}b_{11} \det \begin{pmatrix} a_{21} & a_{22} \\ b_{21} & b_{22} \end{pmatrix} + a_{21}b_{21} \det \begin{pmatrix} b_{11} & b_{12} \\ a_{11} & a_{12} \end{pmatrix} = 0. \tag{10}$$

In the converse direction: if these equivalent conditions are satisfied, then we can *define* the products $e_j m, e_j q$ $(j = 1, 2)$ as above. A similar reasoning as in Cases 1 and 2 shows that there is a solution (e_1, e_2, m, q) to the system of 4 equations above. From this the pair (e_1, e_2) determines a left-hand factor, and (m, q) determines a right-hand factor in the equality

$$b = \begin{pmatrix} e_1 & 0 \\ 0 & e_2 \end{pmatrix} a \begin{pmatrix} m & q \\ 0 & m \end{pmatrix}.$$

Assume now that

$$a_{11} = 0,$$

which is equivalent to $b_{11} = 0$. Comparison of the $(1, 2)$ entries shows that then $e_1 m = b_{12}/a_{12}$. Comparing the entries $(2, 1)$ of a and b yields $e_2 m = b_{21}/a_{21}$. Hence we obtain

$$e_2 q = [b_{22} - a_{22} e_2 m]/a_{21} \equiv \det \begin{pmatrix} a_{21} & a_{22} \\ b_{21} & b_{22} \end{pmatrix} /a_{21}^2,$$

where the invertibility of a implies that the last denominator is nonzero. The condition

$$(e_1 m)(e_2 q) = (e_2 m)(e_1 q)$$

is also here necessary, and determines $e_1 q$ uniquely as

$$e_1 q = b_{12} \det \begin{pmatrix} a_{21} & a_{22} \\ b_{21} & b_{22} \end{pmatrix} /[a_{12} a_{21} b_{21}].$$

In the converse direction: if we *define* the products $e_j m, e_j q$ $(j = 1, 2)$ as above, then a similar reasoning as above shows that there is a solution (e_1, e_2, m, q) to the system of 4 equations above. The pair (e_1, e_2) determines a left-hand factor, and (m, q) determines a right-hand factor in the equality

$$b = \begin{pmatrix} e_1 & 0 \\ 0 & e_2 \end{pmatrix} a \begin{pmatrix} m & q \\ 0 & m \end{pmatrix}.$$

Finally, note that the remaining possible case

$$a_{21} = 0 \quad \text{instead of} \quad a_{11} = 0$$

can be handled in a completely similar way. We have obtained the following result:

Case 3. *Let the matrix J_1 be in class (III), and J_0 be in class (II). If $a_{11} a_{21} \neq 0$, then* (6) *is satisfied if and only if* (10) *holds. If $a_{11} a_{21} = 0$, then* (6) *is satisfied.*

Case 4. Assume now that the matrix J_1 is in class (II), and J_0 is in class (III). Write (6) in the equivalent form

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} m & q \\ 0 & m \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} e_1 & 0 \\ 0 & e_2 \end{pmatrix}.$$

Comparing the corresponding entries on both sides, we obtain

$$a_{2j} e_j m = b_{2j}, \qquad a_{1j} e_j m + a_{2j} e_j q = b_{1j} \qquad (j = 1, 2).$$

Assume first that for both values $j = 1, 2$

$$a_{2j} \neq 0 \quad \text{or, equivalently,} \quad b_{2j} \neq 0.$$

Then the 4 products $e_j m, e_j q$ $(j = 1, 2)$ of two factors (each) are determined uniquely:

$$e_j m = b_{2j}/a_{2j}, \quad e_j q = [b_{1j} - a_{1j} b_{2j}/a_{2j}]/a_{2j} \quad (j = 1, 2).$$

A necessary condition for the solvability of this system of 4 equations (for the factors) is, similarly to the preceding cases,

$$(e_1 m)(e_2 q) = (e_2 m)(e_1 q).$$

This holds if and only if

$$a_{21} a_{22} \det(b) = b_{21} b_{22} \det(a). \tag{11}$$

In the converse direction: if these equivalent conditions are satisfied, then we can *define* the products $e_j m, e_j q$ $(j = 1, 2)$ as above. A similar reasoning as in Cases 1,2 and 3 shows that there is a solution (e_1, e_2, m, q) to the system of 4 equations above. The pair (m, q) determines a left-hand factor, and (e_1, e_2) determines a right-hand factor in the equality

$$b = \begin{pmatrix} m & q \\ 0 & m \end{pmatrix} a \begin{pmatrix} e_1 & 0 \\ 0 & e_2 \end{pmatrix}.$$

Assume now that

$$a_{22} = 0,$$

which is equivalent to $b_{22} = 0$. Comparison of the $(2, 1)$ entries shows that then $e_1 m = b_{21}/a_{21}$. Comparing the entries $(1, 2)$ of a and b yields $e_2 m = b_{12}/a_{12}$. Hence we obtain

$$e_1 q = [b_{11} - a_{11} e_1 m]/a_{21} \equiv \det \begin{pmatrix} b_{11} & a_{11} \\ b_{21} & a_{21} \end{pmatrix} / a_{21}^2,$$

where the invertibility of a implies that the last denominator is nonzero. The condition

$$(e_1 m)(e_2 q) = (e_2 m)(e_1 q)$$

is also here necessary, and determines $e_2 q$ uniquely as

$$e_2 q = b_{12} \det \begin{pmatrix} a_{21} & b_{21} \\ a_{11} & b_{11} \end{pmatrix} / [a_{12} a_{21} b_{21}].$$

In the converse direction: if we *define* the products $e_j m, e_j q$ $(j = 1, 2)$ as above, then a similar reasoning as above shows that there is a solution (e_1, e_2, m, q) to the system of 4 equations above. The pair (m, q) determines a left-hand factor, and (e_1, e_2) determines a right-hand factor in the equality

$$b = \begin{pmatrix} m & q \\ 0 & m \end{pmatrix} a \begin{pmatrix} e_1 & 0 \\ 0 & e_2 \end{pmatrix}.$$

Finally, note that the remaining possible case

$$a_{21} = 0 \quad \text{instead of} \quad a_{22} = 0$$

can be handled in a completely similar way. We have obtained the following result:

Case 4. *Let the matrix J_1 be in class (II), and J_0 be in class (III). If $a_{21} a_{22} \neq 0$, then* (6) *is satisfied if and only if* (11) *holds. If $a_{21} a_{22} = 0$, then* (6) *is satisfied.*

6. Strictly diagonizable matrix polynomials

The interest for the application of diagonalizable matrix polynomials seems to be existing and even growing since the 1950's till our days (see, also for a quick review, Lancaster and Zaballa [14] and Zuniga Anaya [21] for the case of quadratic polynomials).

For the strict equivalence of quadratic polynomials we have the following

Theorem 8. *Assume that the leading coefficient A_n of the $m \times m$ matrix polynomial*

$$A(\lambda) = A_n \lambda^n + A_{n-1} \lambda^{n-1} + \cdots + A_1 \lambda + A_0$$

is invertible. Then $A(\lambda)$ is strictly equivalent to a diagonal matrix polynomial (i.e., a polynomial whose all coefficient matrices are diagonal) if and only if the matrices

$$A_n^{-1} A_j \qquad (j = 0, 1, \ldots, n)$$

are diagonalizable and commute pairwise.

Proof. Assume that $A(\lambda)$ is strictly equivalent to the diagonal matrix polynomial

$$D(\lambda) = D_n\lambda^n + D_{n-1}\lambda^{n-1} + \cdots + D_1\lambda + D_0.$$

Then there are invertible matrices C, B of order m such that

$$CA(\lambda)B = D(\lambda).$$

Comparing the leading coefficients on both sides, we obtain that $D_n = CA_nB$ is, by assumption, invertible. Hence

$$\begin{aligned} B^{-1}[\lambda^n + A_n^{-1}A_{n-1}\lambda^{n-1} + \cdots + A_n^{-1}A_1\lambda + A_n^{-1}A_0]B \\ = [B^{-1}A_n^{-1}C^{-1}]CA_n[\lambda^n + A_n^{-1}A_{n-1}\lambda^{n-1} + \cdots + A_n^{-1}A_1\lambda + A_n^{-1}A_0]B \\ = D_n^{-1}CA(\lambda)B = I\lambda^n + D_n^{-1}[D_{n-1}\lambda^{n-1} + \cdots + D_1\lambda + D_0]. \end{aligned}$$

The coefficient matrices $D_n^{-1}D_j$ $(j = 0, 1, \ldots, n-1)$ on the right-hand side are diagonal. From comparing coefficients, so are the coefficient matrices

$$B^{-1}A_n^{-1}A_jB \quad (j = 0, 1, \ldots, n)$$

on the left-hand side. Hence the matrices $A_n^{-1}A_j$ $(j = 0, 1, \ldots, n)$ are simultaneously diagonalizable. [11, Theorem 1.3.19] shows that they are (a fortiori) diagonalizable and commute pairwise.

Assume now that the matrices $A_n^{-1}A_j$; $(j = 0, 1, \ldots, n)$ are diagonalizable and commute pairwise. By [11, Theorem 1.3.19], they are simultaneously diagonalizable. Hence there are matrices B, d_j $(j = 0, 1, \ldots, n)$ of order m such that B is invertible, the d_j are diagonal, and

$$B^{-1}A_n^{-1}A_jB = d_j \quad (j = 0, 1, \ldots, n).$$

Defining $C := B^{-1}A_n^{-1}$, we obtain that C is invertible, and

$$CA(\lambda)B = B^{-1}A_n^{-1}\sum_{j=0}^{n} A_j\lambda^jB = \sum_{j=0}^{n} d_j\lambda^j =: d(\lambda).$$

■

Remark 8. *The commutation relations* $A_n^{-1}A_jA_n^{-1}A_r = A_n^{-1}A_rA_n^{-1}A_j$ *can clearly be written equivalently as*

$$A_jA_n^{-1}A_r = A_rA_n^{-1}A_j \qquad (j, r = 0, 1, \ldots, n).$$

If the matrix polynomial $A(\lambda)$ *is monic, then the above relations simply mean that the coefficients* A_j, A_r *are (pairwise) commuting. Further, it is clear that a completely similar theorem can be proved if we require only that one coefficient* A_q $(q = 0, 1, \ldots, n)$ *be invertible instead of the invertibility of* A_n.

Applying a nice result of S. K. Mitra [15, Theorem 4.1], the following result holds for the case of a *singular polynomial* $A(\lambda)$ *of matrices of sizes* $m \times k$ where we can and will assume $m \le k$ without restricting the generality.

Theorem 9. *Let* $A_0, \dots, A_n$ *be complex matrices in* $\mathbf{M}(m,k)$ *and*

$$A(\lambda) = A_n\lambda^n + A_{n-1}\lambda^{n-1} + \dots + A_1\lambda + A_0.$$

Then the following statements are equivalent:

1) *There are invertible matrices* $E \in \mathbf{M}(m), F \in \mathbf{M}(k)$ *such that the (strictly equivalent) matrix polynomial*

$$D(\lambda) := EA(\lambda)F$$

is diagonal in the sense that for each coefficient matrix D_j *of* $D(\lambda)$ *the entry in position (p,q) is* 0 *for* $p \neq q$*,*

2) *For some* $\lambda \neq 0$ *the matrix* $A(\lambda)$ *satisfies for each* $i = 0, 1, \dots, n,\ j = 0, 1, \dots, n$ *the following relations:*

$$rg(A_i) \subset rg(A(\lambda)), \qquad rg(A_i^T) \subset rg(A(\lambda)^T),$$

and there is one generalized inverse $A(\lambda)^-$ *of* $A(\lambda)$ *for which*

$$A_iA(\lambda)^-A_j = A_jA(\lambda)^-A_i, \qquad A_jA(\lambda)^- \quad \textit{is diagonalizable.}$$

Here rg *denotes the range (or column) space,* ^{T} *denotes the transpose, and* M^- *denotes a generalized inverse of the matrix* M*, i.e. a matrix satisfying* $MM^-M = M$*.*

Proof. 1) $\Rightarrow$ 2): Denote by $d_{j,p}$ the diagonal entries of D_j, $j = 0, 1, \dots, n$, $p = 1, 2, \dots, \min(m,k)$.

Choose a complex number $\lambda \neq 0$ such that for $p = 1, 2, \dots, \min(m,k)$

$$d_{0,p} + \lambda d_{1,p} + \dots + \lambda^n d_{n,p} \neq 0 \quad \text{if} \quad (d_{0,p}, d_{1,p}, \dots, d_{n,p}) \neq (0, 0, \dots, 0),$$

and define the $k \times m$ diagonal matrix D^- by

$$D^-(p,p) = (d_{0,p} + \lambda d_{1,p} + \dots + \lambda^n d_{n,p})^{-1} \quad \text{if} \quad d_{0,p} + \lambda d_{1,p} + \dots + \lambda^n d_{n,p} \neq 0$$

and all other entries equal to 0.

Then

$$rg(D_i) \subset rg(D(\lambda)), \qquad rg(D_i^T) \subset rg(D(\lambda)^T),$$

and D^- is a generalized inverse of $D(\lambda)$ for which

$$D_iD^-D_j = D_jD^-D_i, \qquad D_jD^- \quad \text{is diagonalizable,}$$

for each $i = 0, 1, \dots, n$, $j = 0, 1, \dots, n$.

Now $FD^{-}E$ is a generalized inverse of $A(\lambda)$ and $A_j = E^{-1}D_jF^{-1}$ for $j = 0, 1, \dots, n$. The assertions in 2) follow immediately.

2) $\Rightarrow$ 1): Apply Theorem 4.1 in [15] to the $(n+2)$-tuple $(A(\lambda), A_0, A_1, \dots, A_n)$.

■

References

[1] G. R. Belitskii and V. V. Sergeichuk, Complexity of matrix problems, *Linear Algebra Appl.*, **361** (2003), 203–222.

[2] O. E. Brown, The equivalence of triples of bilinear forms, *Bull. Amer. Math. Soc.*, **37** (1931), 424–426.

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

[4] F. de Teran and F. M. Dopico, Sharp lower bounds for the dimension of linearizations of matrix polynomials, *Electr. J. Linear Algebra*, **17** (2008), 518–531.

[5] J. A. Dias da Silva and T. J. Laffey, On simultaneous similarity of matrices and related questions, *Linear Algebra Appl.*, **291** (1999), 167–184.

[6] S. Friedland, Simultaneous similarity of matrices, *Advances in Math.*, **50** (1983), 189–265.

[7] F. R. Gantmakher, *The Theory of Matrices*, Chelsea, New York, 1977 (2 volumes).

[8] I. M. Gelfand and V. A. Ponomarev, Remarks on the classification of pairs of commuting linear transformations in a finite dimensional space, *Functional Anal. Applications*, **3** (1969), 81–82 (in Russian).

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

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

[11] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge Univ. Press, Cambridge, 1986.

[12] V. B. Khazanov, Spectral properties of lambda-matrices, *LOMI Steklov*, **111** (1981), 180–194 (in Russian).

[13] P. Lancaster and U. Prells, Isospectral families of high-order systems, *Z. Angew. Math. Mech.*, **87** (2007), 219–234.

[14] P. Lancaster and I. Zaballa, Diagonalizable quadratic eigenvalue problems, *Mech. Systems and Signal Processing*, **23** (2009), 1134–1144.

[15] S. K. Mitra, Simultaneous diagonalization of rectangular matrices, *Linear Alg. Appl.*, **47** (1982), 139–150.

[16] M. Nathanson, Classification problems in K-categories, *Fundamenta Math.*, **105** (1980), 187–197.

[17] V. V. Sergeichuk, Classification problems for systems of forms and linear mappings, *Math. USSR-Izv.*, **31** (1988), 481–501.

[18] V. V. Sergeichuk and D. V. Galinskii, Classification of pairs of linear operators in a four dimensional vector space, *Infinite Groups and Related Algebraic Structures*, Inst. Mat. Ukrain. Akad. Nauk, Kiev, 1993, 413–430 (in Russian).

[19] B. Z. Shavarovskii, On some "tame" and "wild" aspects of the problem of semiscalar equivalence of polynomial matrices, *Matem. Zametki*, **76** (2004), 119–132.

[20] B. Z. Shavarovskii, Similarity transformations of decomposable matrix polynomials and related questions, *Computational Mathematics and Mathematical Physics*, **49** (2009), 1469–1482.

[21] J. C. Zuniga Anaya, Diagonalization of quadratic matrix polynomials, *Systems Control Letters*, **59** (2010), 105–113.

K.-H. Förster, Institute of Mathematics, TU Berlin, Germany;
e-mail: foerster@math.tu-berlin.de

B. Nagy, Department of Analysis, Institute of Mathematics, Budapest University of Technology and Economics, Budapest, H-1521, Hungary; *e-mail*: bnagy@math.bme.hu