

**Extending the notions of companion  
and infinite companion to matrix  
polynomials**

*Marc Van Barel*

*Vlastimil Pták*

*Zdeněk Vavřín*

*Report TW 261, July 1997*



Katholieke Universiteit Leuven  
Department of Computer Science  
Celestijnenlaan 200A – B-3001 Heverlee (Belgium)

# Extending the notions of companion and infinite companion to matrix polynomials

*Marc Van Barel*

*Vlastimil Pták*

*Zdeněk Vavřín*

*Report TW 261, July 1997*

Department of Computer Science, K.U.Leuven

## Abstract

An extended infinite companion matrix  $\tilde{C}_\infty(D)$  and an infinite companion matrix  $C_\infty(D)$  for a (nonmonic in general) matrix polynomial  $D$  is introduced and the finite companion matrix  $C(D)$  is generalized to the nonmonic case. These matrices generalize all properties of the infinite and finite companion (Frobenius) matrix corresponding to a scalar polynomial. In particular,  $C_\infty(D)$  is a controllability matrix of a system whose inner behaviour is given by  $D$ , and  $C(D)$  is a compression of the shift operator (defined on vector polynomials) to the remainder subspace corresponding to  $D$ , with characteristic polynomial equal to  $\det D$ . A factorization formula for finite-rank block Hankel matrices is proved.

The generalization of the finite companion matrix  $C(D)$  permits to construct new linearizations of nonmonic matrix polynomials. These linearizations have considerably smaller dimension than the standard ones. As a consequence, any system of linear difference or differential equations with constant coefficients can be transformed into a first order system of dimension  $n = \det D$ .

**Keywords :** matrix polynomial, companion matrix, difference equations  
**AMS(MOS) Classification :** Primary : 15A21, Secondary : 26C99,39A10.

# Extending the notions of companion and infinite companion to matrix polynomials

Marc Van Barel\*

Katholieke Universiteit Leuven, Department of Computer Science  
Celestijnenlaan 200A, B-3001 Heverlee, Belgium

Vlastimil Pták† and Zdeněk Vavřín†

Czech Academy of Sciences, Institute of Mathematics  
Žitná 25, 115 67 Praha, The Czech Republic

July 1997

## Abstract

An extended infinite companion matrix  $\tilde{C}_\infty(D)$  and an infinite companion matrix  $C_\infty(D)$  for a (nonmonic in general) matrix polynomial  $D$  is introduced and the finite companion matrix  $C(D)$  is generalized to the nonmonic case. These matrices generalize all properties of the infinite and finite companion (Frobenius) matrix corresponding to a scalar polynomial. In particular,  $C_\infty(D)$  is a controllability matrix of a system whose inner behaviour is given by  $D$ , and  $C(D)$  is a compression of the shift operator (defined on vector polynomials) to the remainder subspace corresponding to  $D$ , with characteristic polynomial equal to  $\det D$ . A factorization formula for finite-rank block Hankel matrices is proved.

The generalization of the finite companion matrix  $C(D)$  permits to construct new linearizations of nonmonic matrix polynomials. These linearizations have considerably smaller dimension than the standard ones. As a consequence, any system of linear difference or differential equations with constant coefficients can be transformed into a first order system of dimension  $n = \det D$ .

## 1 INTRODUCTION

The well-known and important notion of the companion (sometimes called also Frobenius) matrix dates back to Kronecker. Given a monic polynomial  $d$  of degree  $n$ ,  $d(z) = d_0 + d_1z + \dots + d_{n-1}z^{n-1} + z^n$ , the companion matrix  $C(d)$  of  $d$  is defined via

$$C(d) = \begin{bmatrix} 0 & \cdots & 0 & -d_0 \\ 1 & & 0 & -d_1 \\ & \ddots & & \vdots \\ 0 & & 1 & -d_{n-1} \end{bmatrix}.$$

Besides its classical fields of application, both the matrix and its related forms occur in the mathematical theory of linear systems when looking for a “canonical” form of realization of these systems.

---

\* This research has been partially supported by the grant of the “Fonds voor Wetenschappelijk Onderzoek-Vlaanderen” (FWO-VI) under contract number G.0278.97

† This research has been partially supported by the grant of the Academy of Sciences of the Czech Republic No. 1030701

The notion of the infinite companion matrix of a scalar polynomial introduced by the second-named author arose in a natural manner in the course of his study of an extremal problem in spectral theory [9]. Although the matrix was originally intended as a technical tool only, it soon became evident that it has interesting connections with what is known today as dilation theory, in particular that it is nothing more than a matrix representation  $P$  of a projection operator which establishes the lifting relation between the classical companion matrix  $C(d)$  of the polynomial  $d$  and the matrix representation  $S$  of the shift operator on the space of all polynomials

$$C(d)P = PS. \quad (1)$$

Equivalently,  $C(d)$  represents a compression of the shift operator to the space of polynomials of degree at most  $n - 1$ .

The idea of using compressions of the shift in interpolation problems appears first in the important paper of D. Sarason [12] and forms one of the starting points of dilation theory developed later to such a perfection by Szökefalvi – Nagy and Foias. Recall the definition of the *infinite companion matrix for scalar polynomials*.

The infinite companion matrix of the polynomial  $d$ , denoted by  $C_\infty(d)$ , is the  $n \times \infty$  matrix  $P = [t_{jk}]_{j=0}^{n-1}{}_{k=0}^\infty$  which may be described by each of the following equivalent conditions (see [11]):

- (1) The column of index  $r$  consists of the coefficients of the polynomial obtained as remainder upon dividing  $s^r$  by  $d(s)$ —in other words,  $s^r - \sum_{0 \leq j \leq n-1} t_{jr} s^j$  is divisible by  $d(s)$ .
- (2) The  $j$ th row is the solution of the linear homogeneous difference equation with characteristic polynomial  $d$  and initial conditions  $t_{jj} = 1$  and  $t_{jk} = 0$  for  $0 \leq k \leq n - 1$ ,  $k \neq j$ .
- (3) For each  $r$  the submatrix of  $P$  consisting of the  $n$  consecutive columns  $r, r + 1, \dots, r + n - 1$  equals  $C(d)^r$ .
- (4) The matrix  $P$  satisfies the intertwining relation  $PS = C(d)P$  and the initial conditions  $t_{jj} = 1$  and  $t_{jk} = 0$  for  $0 \leq k \leq n - 1$ ,  $k \neq j$ .
- (5) The transpose of the matrix  $P$  represents the operator which assigns to each vector  $(v_0, \dots, v_{n-1})^T$  the infinite sequence satisfying the linear homogeneous difference equation with characteristic polynomial  $d$  and the initial conditions  $v_0, \dots, v_{n-1}$ .

The connections between the above five conditions are interesting. The intertwining condition (4) which says that the shift is a lifting of  $C(d)$  is essential. Iterating this relation, we obtain

$$C(d)^r P = PS^r;$$

this immediately implies (3) and represents the matrix  $C_\infty(d)$  in the form of a controllability matrix: the  $r$ th column equals  $C(d)^r e_0$ , where  $e_0 = (1, 0, \dots, 0)^T$ .

In the definition of the infinite companion matrix of a polynomial  $d$  the conditions above are formulated in terms of the coefficients of the polynomial. Explicit formulae for the entries of the matrix expressed in terms of the zeros of  $d$  are given in [10].

In view of the many important connections of these considerations with difference equations, model operators, Hankel operators, convergence of iterative processes, linear control systems etc., it is desirable to extend the study of the infinite companion matrix to matrix polynomials.

In a paper of M. Van Barel and A. Bultheel [13] on the minimal partial realization problem the authors used a certain factorization of the corresponding infinite block Hankel matrix. One of the factors turns out to be a particular case of the matrix to be defined here i.e. in the case when the matrix polynomial is in the canonical form introduced by S. Beghelli and R. Guidorzi [3, 2]. In this case it is also possible to construct the finite companion matrix from the infinite one.

The algorithm considered in [13] leads to a matrix polynomial in the canonical form defined by Beghelli and Guidorzi [3, 2]. The corresponding finite companion matrix to be defined here can easily be extracted from the infinite one.

When looking for a general approach to the infinite companion matrix, an immediate reformulation of condition (1) yields a natural solution: The matrix  $C_\infty(d)$  is the matrix representation of the “remainder operator”  $R(d)$  given by  $R(d)p = r$  if  $p = dq + r$ ,  $\deg r < \deg d$ .

The generalized notion of division by a nonsingular matrix polynomial  $D$  (nonmonic in general) is well-known (see, e.g., [8]) and defines the “remainder operator”  $R(D)$ . It is important to note that this operator has been studied by P. Fuhrmann in [5]. In contrast to [5], we are focusing on the matrices representing the division operator and on their properties analogous to properties of the scalar infinite companion matrix. Our natural choice of basis is such that any vector polynomial  $p(s) = \sum p_i s^i \in \mathbf{F}^q[s]$  is represented by the infinite column vector  $\hat{p} = \text{col}_{i=0}^\infty [p_i]$ . (We put  $p_i = 0$  for  $i > \deg p$ .)

Considering the operator  $R(D)$  as a mapping on the space  $\mathbf{F}^q[s]$ , our definition results into the s.c. extended infinite companion matrix  $\tilde{C}_\infty(D)$  of dimension  $\infty \times \infty$ . It is necessary to impose a special restriction on the matrix polynomial  $D$ , more specifically, to assume  $D$  row reduced, if a definition of an infinite companion matrix of dimension  $n \times \infty$  ( $n = \deg \det D$ ) like in the scalar case is required. As a generalization of (1), the finite companion matrix  $C(D)$  should now fulfil the lifting relation with the matrix  $S^q$  which represents the operator of multiplication by  $s$  on the space  $\mathbf{F}^q[s]$  (this operator will be denoted by  $\mathbf{S}$  and called (block) shift):

$$C(D)C_\infty(D) = C_\infty(D)S^q.$$

This implies that  $C(D)$  is nothing more than the matrix representation of multiplication by  $s$  modulo  $D$  on the s.c. remainder space  $\mathcal{R}(D)$ . See Definitions 3.1 and 6.2 for details.

The purely algebraic background of dilation theory was outlined first by P. Fuhrmann in his paper [5]. The theory of the infinite companion matrix in its scalar form [11] as well as in its present more general form represents an algebraic analog of ideas used in dilation theory. Consequently our work, although different in form – since we concentrate on explicit matrix forms of the operators – is close in spirit to the work of P. Fuhrmann.

The ideas as well as the formulae obtained in the present paper will be applied to set up new explicit expressions for solutions of systems of linear difference equations with constant coefficients; the authors intend to describe these in the paper [13], in preparation.

The contents of our present paper is the following: In Section 3 we introduce the extended infinite companion matrix  $\tilde{C}_\infty(D)$ . Section 4 contains some necessary facts on matrix polynomials. In Section 5, we introduce the infinite companion matrix  $C_\infty(D)$  for row reduced  $D$  and we also describe the connection of the canonical form of  $D$  by Beghelli, Guidorzi and of the Gauss-Jordan elimination for the matrix  $\tilde{C}_\infty(D)$ . In Section 6 the definition of the finite companion matrix  $C(D)$  for any row reduced matrix polynomial  $D$  is given and both the spectrum and the Jordan chains are described. In Section 7, a row reduced matrix polynomial is “linearized”, i.e., transformed into a monic matrix polynomial of degree 1. By this linearization a linear homogeneous system of difference, resp. differential equations with constant coefficients the characteristic polynomial of which is  $D(s)$  is transformed into a larger monic system of first order. Section 8 contains a formula connecting the matrix  $C_\infty(D)$  with the scalar infinite companion matrix  $C_\infty(d)$  where  $d = \det D$ . In Section 9, it is shown that any block Hankel matrix  $H$  of finite rank can be written as the product of the transpose of an infinite companion matrix, a submatrix of  $H$  and, in general, another infinite companion matrix.

## 2 NOTATION AND PRELIMINARIES

In the sequel, the following notation will be used:

- $\mathbf{F}$  denotes an arbitrary (finite or infinite) field.

- $\mathbf{F}^q[s]$  denotes the set of the vector polynomials, i.e., the polynomials with coefficients belonging to  $\mathbf{F}^q$ .
- Similarly  $\mathbf{F}^{p \times q}[s]$  denotes the set of all  $p \times q$  matrix polynomials.
- $\mathbf{F}^q(s)$  denotes the set of all vector rational functions.
- The coefficients of a matrix polynomial  $D \in \mathbf{F}^{p \times q}[s]$  are denoted by  $D_k$ , i.e.,  $D(s) = \sum_k D_k s^k$  with  $D_k = 0$  for  $k > \deg D$ .
- Stacking vector: We often need stacking vectors of coefficients of polynomial vectors. We indicate these stacking vectors by a hat. For example, if  $a$  is a polynomial vector, then by  $\hat{a}$  we mean the infinite column vector obtained by stacking the coefficients of the polynomial vector  $a$  and completing by zeros. Thus the value of  $a$  in the point  $s$  equals  $a(s) = [I_q \ s I_q \ s^2 I_q \ \dots] \hat{a}$ . If  $b$  is a vector whose elements are strictly proper rational functions, the stacking vector  $\hat{b}$  is defined so that  $b(s) = [s^{-1} I_q \ s^{-2} I_q \ \dots] \hat{b}$ . Similarly, if we have a sequence of vectors  $\{y_i\}$  having finite length, the stacking vector is denoted by  $\hat{y} = [y_0^T, y_1^T, \dots]^T$ .
- $S$  denotes the “shift operator” on the space  $\mathbf{F}^q[s]$ , given by  $Sp(s) = sp(s)$  for any  $p(s) \in \mathbf{F}^q[s]$ .
- $S = [ \delta_{i,j+1} ]_{i,j=0}^{\infty}$  denotes the infinite “shift matrix”. Then  $S^q = [ \delta_{i,j+q} ]_{i,j=0}^{\infty}$  is the infinite “block shift matrix”.
- In the sequel, the matrix polynomial  $D$  is always assumed to be nonsingular, i.e.,  $\det D$  is a nonzero polynomial which we shall denote by  $d$ .
- If  $\mathcal{I}$  is an index set and  $M$  an arbitrary matrix then we introduce the matrix  $M^{(\mathcal{I})}$ , resp.  $M_{(\mathcal{I})}$ , as the submatrix of  $M$  formed by columns, resp. rows of  $M$  with indices belonging to  $\mathcal{I}$ .
- Given two matrices  $A = (a_{ij})$  and  $B$ , their Kronecker (tensor) product is defined as

$$A \otimes B = \begin{bmatrix} a_{00}B & a_{01}B & \dots \\ a_{10}B & a_{11}B & \dots \\ \vdots & & \end{bmatrix}$$

(see [1], page 415).

### 3 EXTENDED INFINITE COMPANION MATRIX FOR MATRIX POLYNOMIALS

Note that, for a scalar polynomial  $d$  of degree  $n$ ,  $C_\infty(d)$  is an  $n \times \infty$  matrix. It represents the operator  $R(d)$  considered as an operator mapping  $\mathbf{F}[s]$  onto  $\mathbf{F}_{n-1}[s]$  (the “remainder space”, consisting of all polynomials of degree less than  $n$ ). In the matrix case, we have to impose a special restriction on the matrix polynomial  $D$  to obtain an  $n \times \infty$  matrix ( $n = \text{rank } R(D)$ ) if we use the monomial basis  $\mathcal{B}$  defined in Definition 3.2. The reason is that the range of  $R(D)$ , though being  $n$ -dimensional where  $n = \deg \det D$ , is not necessarily spanned by  $n$  vectors of the basis  $\mathcal{B}$ . In this section we introduce the  $\infty \times \infty$  s.c. “extended” infinite companion matrix. The number of its nonzero rows is finite but greater than  $n$  in general. Section 4 deals with the case of row reduced matrix polynomials  $D$ ; in that case, the “remainder space” is spanned by  $n$  elements of the basis  $\mathcal{B}$  and the corresponding  $n$  nonzero rows of  $\tilde{C}_\infty(D)$  form the  $n \times \infty$  infinite companion matrix  $C_\infty(D)$  which is a complete analogy with the scalar case.

Let us start by recalling the generalized definition for division by any square nonsingular (nonmonic in general) matrix polynomial  $D$  (see [8], Theorem 6.3-15).

**Theorem 3.1 (Division theorem for matrix polynomials)** *Let  $D$  be any nonsingular matrix polynomial and let*

$$\mathcal{R}(D) = \{r \in \mathbf{F}^q[s] \mid D^{-1}r \text{ is strictly proper}\}.$$

*Then the space  $\mathbf{F}^q[s]$  may be written as a direct sum of  $D\mathbf{F}^q[s]$  and of  $\mathcal{R}(D)$ .*

**proof.** For any  $p \in \mathbf{F}^q[s]$ , decompose the rational vector function  $D^{-1}p$  into the polynomial and strictly proper rational part and then premultiply both parts by  $D$  to get the desired (unique) decomposition.  $\square$

We shall use the notation  $\Pi_-$  for the projection operator defined on  $\mathbf{F}^q(s)$  which assigns to any rational function its strictly proper part.

**Definition 3.1 (Remainder operator, remainder space)** *Given any nonsingular  $D \in \mathbf{F}^{q \times q}[s]$ , we introduce the remainder operator  $R(D)$  on  $\mathbf{F}^q[s]$  as the projection such that  $\text{Ran}(R(D)) = \mathcal{R}(D)$  and  $\text{Ker}(R(D)) = D\mathbf{F}^q[s]$ . We call the space  $\mathcal{R}(D)$  the remainder space of the matrix polynomial  $D$ .*

**Corollary 3.1** *For any nonsingular  $D \in \mathbf{F}^{q \times q}[s]$  and any  $p \in \mathbf{F}^q[s]$ ,*

$$R(D)p = D\Pi_- D^{-1}p. \quad (2)$$

**Definition 3.2 (Basis)** *In the space  $\mathbf{F}^q[s]$  we choose the basis*

$$\mathcal{B} = \{b_t\}_0^\infty \quad \text{where } b_t(s) = e_k s^j \quad \text{if } t = jq + k, 0 \leq k < q.$$

*In this basis the vector polynomial  $p(s) = \sum p_i s^i$  is represented by the stacking vector  $\hat{p}$ .*

**Definition 3.3 (Extended infinite companion matrix)** *Assume that  $D$  is nonsingular. The  $\infty \times \infty$  matrix representation of the operator  $R(D)$  with respect to the basis  $\mathcal{B}$  will be denoted by  $\tilde{C}_\infty(D)$  and will be called the extended infinite companion matrix of the matrix polynomial  $D$ .*

The matrix  $\tilde{C}_\infty(D)$  has a block structure which corresponds to the block structure of the basis  $\mathcal{B}$ ,

$$\tilde{C}_\infty(D) = [C_{j,r}]_{j,r=0}^\infty$$

where the blocks  $C_{j,r}$  are square of order  $q$ . For any vector polynomial  $p(s) = \sum_r p_r s^r$ ,  $p_r \in \mathbf{F}^q$ , the action of  $R(D)$  onto  $p(s)$  can be written in the compact block form:

$$(R(D)p)(s) = \sum_j \left( \sum_r C_{j,r} p_r \right) s^j.$$

The following theorem relates the matrix  $\tilde{C}_\infty(D)$  to the special block Hankel matrix corresponding to the matrix rational function  $D^{-1}(s)$  and is nothing more than a restatement of Corollary 3.1 in matrix form.

**Theorem 3.2** *For any nonsingular matrix polynomial  $D(s) = \sum_k D_k s^k$  (we set  $D_k = 0$  if  $k > \deg D(s)$ ),*

$$\tilde{C}_\infty(D) = \Delta(D)H(D^{-1})$$

*where  $\Delta(D) = [D_{1+i+j}]_{i,j=0}^\infty$ .*

**proof.**

The Hankel matrix  $H(ND^{-1})$  corresponding to the matrix rational function  $ND^{-1}$  is the matrix representation of the Hankel operator which maps any polynomial  $p \in \mathbf{F}^q[s]$  onto the function  $v = \Pi_- ND^{-1}p$ . In the particular case of  $H(D^{-1})$  we put  $N = I$  to get

$v = \Pi_- D^{-1} p$ . If  $p(s) = \sum p_i s^i$ ,  $v(s) = \sum_0^\infty v_j s^{-j-1}$  and if we use the notation for stacking vectors (see Section 2), we can write

$$\hat{v} = H(D^{-1})\hat{p}$$

and it is easy to verify that

$$\Delta(D)\hat{v}$$

are the coefficients of the vector polynomial

$$D\Pi_- D^{-1} p = R(D)p.$$

□

## 4 ROW AND COLUMN REDUCED MATRIX POLYNOMIALS, THE CANONICAL FORM

To study the properties of the matrix  $\tilde{C}_\infty(D)$ , we shall need some basic definitions concerning matrix polynomials and their special forms.

**Definition 4.1 (Degree, row/column degree, highest row/column degree coefficient)**

If we consider  $D(s)$  as an element of  $\mathbb{F}[s]^{q \times q}$ ,  $D(s) = [d_{ij}(s)]$  then  $\deg D(s) = \max_{ij} \deg d_{ij}(s)$ . The row degrees are  $h_i = \max_j \deg d_{ij}(s)$ ,  $i = 0, \dots, q-1$ . The highest row degree coefficient is the  $q \times q$  matrix  $D_{hrdc} = [a_{ij}]$  where  $d_{ij}(s) = a_{ij} s^{h_i} + O(s^{h_i-1})$ ,  $s \rightarrow \infty$ .

The column degrees and the highest column degree coefficient  $D_{hcdc}$  are defined analogously.

**Definition 4.2** We say that  $D(s)$  is row reduced if  $D_{hrdc}$  is nonsingular. Similarly, we say that  $D(s)$  is column reduced if  $D_{hcdc}$  is nonsingular.

**Lemma 4.1** (See, e.g., [8]) The matrix polynomial  $D(s)$  is row reduced if and only if  $\sum h_i = \deg \det D$ .

Observe that any row reduced (column reduced)  $D$  is nonsingular.

**Definition 4.3** We introduce the relation of right equivalence  $\sim_R$  by

$$D \sim_R D' \quad \text{if and only if} \quad D' = DU$$

for some unimodular  $U$ . We shall use the notation  $\mathcal{U}_R(D)$  for the class of polynomials right equivalent to  $D$ .

Similarly we can introduce the left equivalence  $\sim_L$  and the class  $\mathcal{U}_L(D)$ .

**Lemma 4.2** (REMAINDER OPERATORS FOR EQUIVALENT MATRIX POLYNOMIALS) Suppose that  $D, D'$  are nonsingular.

(1) If  $D' = DM$  for some matrix polynomial  $M$  then  $\text{Ker } R(D') \subset \text{Ker } R(D)$ .

(2) If  $D' \in \mathcal{U}_R(D)$ , i.e.,  $D' = DU$  for some unimodular  $U$  then

- (a)  $\text{Ker } R(D') = \text{Ker } R(D)$ ,
- (b)  $R(D') = R(D')R(D)$  and

$$\tilde{C}_\infty(D') = \tilde{C}_\infty(D')\tilde{C}_\infty(D),$$

(c)  $\text{rank } R(D') = \text{rank } R(D)$  and  $R(D')|_{\mathcal{R}(D)}$  is an isomorphism between  $\mathcal{R}(D)$  and  $\mathcal{R}(D')$ ,

(d) span rows  $\tilde{C}_\infty(D') = \text{span rows } \tilde{C}_\infty(D)$ .

(3) If  $D' \in \mathcal{U}_L(D)$ , i.e.,  $D' = UD$  for some unimodular  $U$  then  $\mathcal{R}(D') = U\mathcal{R}(D)$ .

**proof.**

(1)  $\text{Ker } R(D') = DM\mathbf{F}^q[s] \subset D\mathbf{F}^q[s] = \text{Ker } R(D)$ .

(2) (a) follows by a double use of (1) since  $D = D'U^{-1}$  where  $U^{-1}$  is a matrix polynomial.

(b) and (c) This is a consequence of a more general and well-known property;

For any two projections  $P_1, P_2$  such that  $\text{Ker } P_1 = \text{Ker } P_2$  the following holds:  $P_1 = P_1P_2$  and, moreover,  $P_1|_{\text{Ran } P_2}$  is an isomorphism between  $\text{Ran } P_2$  and  $\text{Ran } P_1$ .

**proof.** The range of  $I - P_i$  being  $\text{Ker } P_i$ , we have  $P_1(I - P_2) = 0$  and  $P_2(I - P_1) = 0$ , i.e.,

$$P_1P_2 = P_1, \quad P_2P_1 = P_2.$$

Denote by  $T$  the restriction to  $\text{Ran } P_2$  of the operator  $P_1$ . The identity  $P_1P_2 = P_1$  shows that  $T$  maps  $\text{Ran } P_2$  onto  $\text{Ran } P_1$ , the identity  $P_2P_1 = P_2$  implies that  $T$  is injective.  $\square$

(d) follows from the fact that

$$\text{Ker } \tilde{C}_\infty(D') = \text{Ker } \tilde{C}_\infty(D).$$

(3)  $p \in \mathcal{R}(D)$  if and only if  $D^{-1}p$  is strictly proper, i.e., if and only if  $(UD)^{-1}Up$  is strictly proper which is if and only if  $Up \in \mathcal{R}(D')$ .  $\square$

**Lemma 4.3 (Special elements in the class  $\mathcal{U}_R(D)$ )** Assume  $D$  nonsingular.

- (1) Each class of equivalence  $\mathcal{U}_R(D)$  contains at least one row reduced element. The sequences of row degrees  $(h_i)$  of such elements can be different in general.
- (2) Each class of equivalence  $\mathcal{U}_R(D)$  contains at least one column reduced element. If  $D'$  and  $D''$  are two column reduced elements of  $\mathcal{U}_R(D)$  and if  $(k'_i)$  and  $(k''_i)$  are the corresponding sequences of their column degrees then  $(k''_i)$  is just a permutation of  $(k'_i)$ .
- (3) Each class of equivalence  $\mathcal{U}_R(D)$  contains at least one column reduced element with an upper triangular highest column degree coefficient. All these elements have the same sequence of column degrees  $(k_i)$  (including the order).
- (4) Each class of equivalence  $\mathcal{U}_R(D)$  contains exactly one element which is both column reduced with an upper triangular highest column degree coefficient and row reduced with the identity matrix as its highest row degree coefficient.

**Definition 4.4** For each  $D$ , we shall denote by  $D_c$  the unique element of the class  $\mathcal{U}_R(D)$  which possesses the properties sub (4) in the preceding lemma. Following S. Beghelli and R. Guidorzi [2], the matrix polynomial  $D_c$  will be called the canonical form of  $D$  and the matrix polynomial  $D$  will be said to be in the (input-output) canonical form if  $D = D_c$ .

In the proof of Lemma 4.3 we include an algorithm to construct the input-output canonical form of any nonsingular  $D$ . In fact this algorithm is very similar to the one described in [2] which transforms not one but two polynomial matrices.

**proof.** Let us show the uniqueness of the set of column degrees in (2). Take two column reduced matrix polynomials

$$\begin{aligned} D' &= DU', & U' \text{ unimodular and} \\ D'' &= DU'', & U'' \text{ unimodular} \end{aligned}$$

with column degrees  $k'_i$  and  $k''_i$ . Hence,

$$D' = D''U$$

where  $U = (U'')^{-1}U'$  is unimodular. We shall use the expressions:

$$D''(s) = (D''_{hcdc} + D''_{csp}(s))E(k'')$$

and

$$D'(s) = (D'_{hcdc} + D'_{csp}(s))E(k')$$

where  $D''_{csp}(s)$  and  $D'_{csp}(s)$  are strictly proper rational matrices and  $E(l) = \text{diag}(s^{l_i})$  and  $l \in \mathbf{Z}^q$ . By assumption, both  $D''_{hcdc}$  and  $D'_{hcdc}$  are nonsingular. Now

$$(D'')^{-1}(s) = E(-k'')((D''_{hcdc})^{-1} + \Delta_{sp}(s))$$

for a strictly proper rational  $\Delta_{sp}(s)$ . We can now express

$$U(s) = (D'')^{-1}(s)D'(s) = E(-k'')(U_N + U_{sp}(s))E(k')$$

where  $U_N = (D''_{hcdc})^{-1}D'_{hcdc}$  is a constant nonsingular matrix and  $U_{sp}(s)$  is strictly proper rational. We can always find a permutation  $\pi$  with corresponding permutation matrix  $P$  such that  $U_N P$  has all diagonal entries different from zero. Comparing diagonal entries in

$$U(s)P = E(-k'')(U_N P + U_{sp}(s)P)E(\pi k'),$$

which are polynomials on the left, we get that

$$(\pi k')_i \geq k''_i.$$

Since  $\sum_i (\pi k')_i = \deg \det D'(s) = \deg \det D''(s) = \sum_i k''_i$ , we conclude  $(\pi k')_i = k''_i$ .

The uniqueness in (3) can be proved by similar arguments.

Let us prove (4):

Consider the following elementary polynomial transformation matrices:

- $E_{i,j}$  = the identity matrix in which column  $i$  and column  $j$  are interchanged;
- $E_{i,j}(q(s))$  = the identity matrix in which column  $j$  is replaced by column  $j$  plus  $q(s) \in F[s]$  times column  $i$ ;
- $E_i(k)$  = the identity matrix where column  $i$  is multiplied by  $k \in F \setminus \{0\}$ .

Suppose we have a nonsingular polynomial  $q \times q$  matrix  $D(s)$ . The coefficient of degree  $k$  of element  $d_{i,j}(s)$  is denoted by  $d_{i,j,k}$ . The degrees of the columns of  $D(s)$  are denoted by  $k_1, \dots, k_q$ .

First of all, to make a nonsingular polynomial matrix column reduced, i.e., where the highest degree coefficients of the columns are linearly independent, we can use the following algorithm:

- while not column reduced do

$$\begin{aligned} & - \text{Consider the set of linear equations} \\ & \begin{bmatrix} d_{1,1,k_1} & \cdots & d_{1,q,k_q} \\ \vdots & & \vdots \\ d_{q,1,k_1} & \cdots & d_{q,q,k_q} \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_q \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \\ & \text{which always has a nontrivial solution } c_i. \end{aligned}$$

- Define  $l$  such that  $l = \max\{k_i | c_i \neq 0\}$ .
- Replace  $D$  by  $D\Pi_{i=1, i \neq l}^q E_{i,l}(-c_i/c_l s^{k_l - k_i})$ .
- By this transformation, the degree of column  $l$  of  $D$  has been decreased.

Now we are going to make  $D$  column reduced with the highest degree coefficients forming a unit upper triangular matrix.

- for  $i = q, q-1, \dots, 1$  do
  - Define  $l$  such that  $l = \min\{k_j | 1 \leq j \leq q, d_{i,j,k_l} \neq 0\}$ .
  - Replace  $D$  by  $DE_{l,i}$ .
  - Replace  $D$  by  $DE_i(1/d_{i,i,k_i})$ .
  - for  $j = 1, 2, \dots, i-1$ 
    - \* Replace  $D$  by  $DE_{l,j}(-d_{i,j,k_j} s^{k_j - k_i})$ .

The last step makes  $D$  also row reduced where the highest degree coefficients of the rows of  $D$  form the identity matrix.

- for  $j = 1, 2, \dots, q$  do
  - for  $d = k_j, k_{j-1}, \dots, k_{\min}$  do
    - \* for  $i = q, q-1, \dots, 1$  do
      - if  $i \neq j$  and  $d \geq k_i$  and  $D_{i,j,d} \neq 0$  then
        - $\gamma = d - k_i$ .
        - Replace  $D$  by  $DE_{i,j}(-D_{i,j,d} s^\gamma)$ .

Note that all transformations we used, are elementary polynomial, i.e., they are unimodular. Hence, their product is unimodular, too.

Now we prove the uniqueness of this element: Suppose we have two matrices  $D'(s), D''(s) \in \mathcal{U}_R(D)$  having the given properties. We know by Lemma 4.3(3) that  $k(D') = k(D'')$  and the properties of the canonical form imply that  $h(D') = k(D') = k(D'') = h(D'')$ .

Because  $D'(s), D''(s) \in \mathcal{U}_R(D)$ , we can write

$$\begin{aligned} D'(s) &= D(s)U'(s), \\ D''(s) &= D(s)U''(s) \end{aligned}$$

or

$$\begin{aligned} (D'')^{-1}(s)D'(s) &= (U'')^{-1}(s)U'(s) \\ &= U(s) \end{aligned}$$

with  $U(s)$  unimodular. Now, using the row reducedness of  $D'(s)$  and  $D''(s)$ ,

$$(D''_{hrdc} + D''_{rpr}(s))^{-1} E^{-1}(k'') E(k') (D'_{hrdc} + D'_{rpr}(s)) = U(s).$$

Because  $E^{-1}(k'') E(k') = I_q$  and  $D''_{hrdc} = D'_{hrdc} = I_q$ , we get

$$I_q + D_{rpr}(s) = U(s)$$

with  $D_{rpr}(s)$  proper rational. Because  $U(s)$  is polynomial, we get  $U(s) = I_q$ . Hence,  $D'(s) = D''(s)$ .

Now, existence of special elements by (1)–(3) is a consequence of (4). □

## 5 INFINITE COMPANION MATRIX FOR ROW REDUCED MATRIX POLYNOMIALS

In this section, we study the additional properties of the extended infinite companion matrix  $\tilde{C}_\infty(D)$  when the matrix polynomial  $D$  is row reduced. The remainder subspace is a coordinate subspace with respect to the basis  $\mathcal{B}$ , i.e., a subspace spanned by  $n$  elements of the basis  $\mathcal{B}$  with  $n = \deg \det D$ . This makes possible to restrict the infinite companion matrix to an  $n \times \infty$  matrix.

**Lemma 5.1** *Let  $r \in \mathbf{F}^q[s]$  and  $D$  row reduced with row degrees  $h_i$  be given. Then  $D^{-1}r$  is strictly proper rational if and only if  $\deg r_i < h_i$ ,  $i = 0, \dots, q-1$ .*

This lemma can be found in [8, Lemma 6.3-11].

**Definition 5.1 (Index set)** *For any nonsingular  $D(s)$  with row degrees  $h_i$ , we introduce the index set  $\mathcal{I}(D)$  by*

$$\mathcal{I}(D) = \{t \mid t = jq + i, 0 \leq j < h_i, 0 \leq i < q\}.$$

**Corollary 5.1 (Remainder space for a row reduced  $D$ )** *Let  $D$  be row reduced with row degrees  $h_i$ . Then the remainder space  $\mathcal{R}(D)$  is spanned by  $\{e_i s^j, 0 \leq j < h_i\} = \{b_t, t \in \mathcal{I}(D)\}$ .*

**Lemma 5.2** *If  $D$  is nonsingular then*

$$\text{rank } \tilde{C}_\infty(D) = \text{rank } \mathcal{R}(D) = \deg \det D = n.$$

*In particular, if  $D$  is row reduced then the rows of  $\tilde{C}_\infty(D)$  with indices in  $\mathcal{I}(D)$  are linearly independent and all other rows are zero.*

**proof.** If  $D$  is row reduced, the assertion follows from Corollary 5.1. In the general case use Lemma 4.3, (2)(c).  $\square$

**Definition 5.2 (Infinite companion matrix for a row reduced matrix polynomial)** *If  $D$  is row reduced then we introduce the infinite companion matrix  $C_\infty(D)$  as the matrix representation of the operator  $R(D) : \mathbf{F}^q[s] \rightarrow \mathcal{R}(D)$  with respect to the basis  $\{b_t \mid t \in \mathcal{I}(D)\}$  in  $\mathcal{R}(D)$ . Equivalently,  $C_\infty(D) = \tilde{C}_\infty(D)_{(\mathcal{I}(D))}$ .*

Note that  $C_\infty(D)^{(\mathcal{I}(D))} = I_n$  because  $R(D)|_{\mathcal{R}(D)}$  is the  $n$ -dimensional identity operator.

If  $D$  is not only row reduced but also in the canonical form, the infinite companion matrix  $C_\infty(D)$  has additional structure. To describe this structure, we need the following terminology.

**Definition 5.3 (Row-echelon form)** *We say that a matrix  $M = [m_{ik}]$  has row-echelon form if the following two conditions are satisfied:*

- (i) *If  $j_i = \min\{k, m_{ik} \neq 0\}$  then the  $j_i$ th column of  $M$  equals the canonical vector  $e_i = \text{col}(\delta_{ij})$  for  $i = 1, 2, \dots$*
- (ii) *The sequence  $\{j_i\}$  is monotone increasing.*

**Remark 5.1** *If  $D_c$  is a canonical matrix polynomial then the infinite companion matrix  $C_\infty(D_c)$  has the row-echelon form.*

**Lemma 5.3** *Let  $M$  be a nonzero  $k \times l$  matrix ( $k, l \leq \infty$ ). Then there is a unique positive integer  $r$  (or infinity) and unique matrices  $M_0, G$  of dimensions  $k \times r, r \times l$  such that  $G$  has the row echelon form, the columns of  $M_0$  are linearly independent and*

$$M = M_0 G.$$

**Definition 5.4** *The unique matrix  $G$  defined by the previous lemma will be called the row-echelon form of  $M$ .*

**proof.** A method to factorize  $M$  as  $M = M_0 G$  with  $G$  in row echelon form is the following. Because  $M$  is nonzero, we can take for the first column of  $M_0$  the first nonzero column of  $M$ . The second column of  $M_0$  is the next column of  $M$  which together with the first column of  $M_0$  forms a linearly independent set . . . . For any  $k$ , the  $k$ th column of  $G$  has only a finite number of nonzero elements and these elements give the coefficients of the finite linear combination of the columns of  $M_0$  to get the  $k$ th column of  $M$ . It is clear that  $G$  has the row echelon form and that this factorization is unique. □

**Remark 5.2** *We emphasize that in general the product of two infinite matrices is not defined. In Lemma 5.3 we use the special form of  $G$  which implies that any of its columns has only a finite number of nonzero entries.*

**Lemma 5.4** *Let  $B$  be a linear operator and  $P$  be a projection, both defined on the same linear space  $\mathcal{L}$ . Suppose that*

$$\text{Ker } P \subset \text{Ker } B.$$

*Then*

$$B = BP.$$

**proof.** For any element  $x \in \mathcal{L}$  we can write

$$Bx = B(Px + (I - P)x).$$

Since  $(I - P)x \in \text{Ker } P \subset \text{Ker } B$ , it holds

$$B(I - P) = 0$$

and the proof is complete. □

**Theorem 5.5** *Let  $M$  be a  $k \times \infty$  matrix and  $D$  a nonsingular polynomial matrix. Then the following properties are equivalent:*

(i)  $M\hat{p} = 0$  for all  $p \in DF^q[s]$

(ii)  $M = M\tilde{C}_\infty(D)$

(iii) There is a  $k \times \infty$  matrix  $P$  such that  $M = P\tilde{C}_\infty(D)$ .

**proof.** (i)  $\rightarrow$  (ii)

$M$  represents an operator  $T$  defined on  $F^q[s]$  such that  $\text{Ker } T \supset \text{Ker } R(D)$ . By the preceding lemma,  $T = TR(D)$ . In matrix representation, this is (ii).

(ii)  $\rightarrow$  (iii)

Evident.

(iii)  $\rightarrow$  (i)

Evident since  $\tilde{C}_\infty(D)\hat{p} = 0$  for all  $p \in DF^q[s]$ . □

**Theorem 5.6** *Under the assumptions of Theorem 5.5 and denoting  $n = \deg \det D$ , the following are equivalent:*

- (i)  $M\hat{p} = 0$  if and only if  $p \in D\mathbf{F}^q[s]$
- (ii)  $M = M\tilde{C}_\infty(D)$  and  $\text{rank } M = n$
- (iii) There is an  $\infty \times \infty$  matrix  $P$  such that  $M = P\tilde{C}_\infty(D)$  and  $\text{rank } M = n$ .
- (iv) The row-echelon form of  $M$  is  $C_\infty(D_c)$  where  $D_c$  is the canonical form of  $D$ .

**proof.**

(i)  $\rightarrow$  (ii) Let  $T$  be the operator on  $\mathbf{F}^q[s]$  the matrix representation of which is  $M$ . By (i),  $\text{Ker } T = D\mathbf{F}^q[s] = \text{Ker } R(D)$ . By Theorem 5.5,  $M = M\tilde{C}_\infty(D)$ . The equality of the kernels implies the equality of ranks:  $\text{rank } T = \text{rank } R(D) = n$  so that also  $\text{rank } M = n$ .

(ii)  $\rightarrow$  (iii) Evident.

(iii)  $\rightarrow$  (iv) It is

$$\begin{aligned}\tilde{C}_\infty(D) &= \tilde{C}_\infty(D)\tilde{C}_\infty(D_c) \\ &= \tilde{C}_\infty(D)^{(T)}C_\infty(D_c)\end{aligned}$$

so that

$$M = P\tilde{C}_\infty(D)^{(T)}C_\infty(D_c). \quad (3)$$

Now  $\text{rank } M = n$  so that the matrix  $M_0 = P\tilde{C}_\infty(D)^{(T)}$  has rank at least  $n$ . Since this matrix has just  $n$  columns, they have to be linearly independent. Moreover, we know that  $C_\infty(D_c)$  has the row echelon form. It follows that (3) gives the unique decomposition described in Lemma 5.3. The property  $M_0 = M^{(T)}$  follows from the fact that  $C_\infty(D_c)^{(T)} = I_n$ .

(iv)  $\rightarrow$  (i) Evident.  $\square$

## 6 FINITE COMPANION MATRIX FOR A ROW REDUCED MATRIX POLYNOMIAL

In this section we give a generalization of the notion “finite companion matrix” (or “Frobenius matrix”) to matrix polynomials. We introduce it as the matrix representation of the operator on  $\mathcal{R}(D)$  of multiplication by  $s$  modulo  $D(s)$ , i.e. the compressed shift. Note that our particular choice of basis leads to the particular restriction of the definition to row reduced matrix polynomials only.

**Definition 6.1 (Operator  $S(D)$ )** *For any nonsingular  $D$  the operator  $S(D) : \mathcal{R}(D) \rightarrow \mathcal{R}(D)$  is defined by*

$$S(D) = R(D)\mathbf{S}|_{\mathcal{R}(D)}.$$

If  $D(s)$  is row reduced, the matrix of  $S(D)$  w.r.t. the basis  $\{b_t \mid t \in \mathcal{I}(D)\}$  equals

$$(C_\infty(D)S^q)^{(\mathcal{I}(D))}$$

where  $S$  is the “infinite shift matrix”,  $S = [\delta_{i,j+1}]_{i,j=0}^\infty$ .

**Definition 6.2 (Finite companion matrix for a row reduced matrix polynomial)**

*The  $n \times n$  matrix  $C(D)$  defined as*

$$C(D) = (C_\infty(D)S^q)^{(\mathcal{I}(D))}$$

*is called the finite companion matrix of  $D$ .*

Because  $\forall p \in \mathbb{F}^q[s]$

$$R(D)(sp(s)) = S(D)R(D)p(s)$$

holds, the following theorem is true.

**Theorem 6.1 (Intertwining with the shift)** *For any row reduced  $D(s)$  the equality*

$$C_\infty(D)S^q = C(D)C_\infty(D)$$

*holds.*

Similar to the scalar case, this relation implies the following property of the infinite and finite companion matrices of a row reduced matrix polynomial.

**Corollary 6.1** *Let  $D$  be row reduced.*

(1)  $C_\infty(D) = [C_0, C(D)C_0, C^2(D)C_0, \dots] = \text{row}_{j=0}^\infty(C^j(D)C_0)$  where  $C_0$  is the submatrix of  $C_\infty(D)$  given by its first  $q$  columns.

(2)  $(C_\infty(D)S^q)^{(X(D))} = C(D)^j, j = 0, 1, \dots$

Corollary 6.1 shows that  $C_\infty(D)$  has the form of a controllability matrix. From linear system theory, we know that if  $D' \sim_R D$  or  $D' \sim_L D$  are both row reduced then  $C(D)$  and  $C(D')$  are similar (see, e.g., [5, Theorem 4.8]). More precisely, it holds

**Lemma 6.2 (SIMILARITY OF FINITE COMPANION MATRICES FOR EQUIVALENT MATRIX POLYNOMIALS)**

*Let  $D' \sim_R D$  be both row reduced. Denote  $M_1 = C_\infty(D)^{(X')}$ ,  $M_2 = C_\infty(D')^{(X)}$ . Then*

$$M_1 M_2 = I_n$$

*and*

$$C(D) = M_1 C(D') M_2 \tag{4}$$

*so that the companion matrices  $C(D)$ ,  $C(D')$  are similar.*

**proof.** Since  $\text{Ker } R(D) = \text{Ker } R(D')$  by Lemma 4.2(2)(a), we have  $R(D) = R(D)R(D')$  using Lemma 5.4 so that, in matrix form,

$$\tilde{C}_\infty(D) = \tilde{C}_\infty(D)\tilde{C}_\infty(D')$$

whence

$$C_\infty(D) = C_\infty(D)^{(X')} C_\infty(D'). \tag{5}$$

Hence,

$$C_\infty(D)^{(X)} = I_n = C_\infty(D)^{(X')} C_\infty(D')^{(X)} = M_1 M_2.$$

By symmetry,

$$C_\infty(D') = C_\infty(D')^{(X)} C_\infty(D). \tag{6}$$

Using (5), (6) and Theorem 6.1, we have

$$\begin{aligned} C(D)C_\infty(D) &= C_\infty(D)S^q = C_\infty(D)^{(X')} C_\infty(D')S^q = \\ &= C_\infty(D)^{(X')} C(D')C_\infty(D') = C_\infty(D)^{(X')} C(D')C_\infty(D')^{(X)} C_\infty(D). \end{aligned}$$

Since the rows of  $C_\infty(D)$  are linearly independent, the identity (4) follows.  $\square$

In the following, we shall need the concept of “left congruence”.

**Definition 6.3** *For any nonsingular  $D$ , we write  $p_1 \sim p_2 \text{ mod}_L D(s)$  if  $p_1 - p_2 = Du$  for some vector polynomial  $u$ .*

Now we are going to study the characteristic values and characteristic vectors of the finite companion matrix.

**Theorem 6.3** *If  $D(s)$  is row reduced then*

$$d(C(D)) = 0$$

with  $d = \det D$ .

**proof.**

The matrix  $d(C(D))$  represents the operator  $d(S(D))$  on  $\mathcal{R}(D)$ . By definition of  $S(D)$ ,

$$d(S(D)) = R(D)d(S)|_{\mathcal{R}(D)}$$

so that for all  $p \in \mathcal{R}(D)$

$$d(S(D))p(s) = R(D)d(s)p(s).$$

However,  $d(s)p(s) = D(s)D_{adj}(s)p(s)$  and, consequently,

$$d(S(D))p(s) = 0.$$

This means that

$$d(C(D)) = 0.$$

□

**Theorem 6.4** *The number  $\lambda_0$  is an eigenvalue of  $S(D)$  if and only if  $D(\lambda_0)$  is singular, in other words  $d(\lambda_0) = 0$ .*

*If  $\lambda_0$  is an eigenvalue of  $S(D)$  then the corresponding set of eigenvectors consists of all polynomials of the form*

$$(s - \lambda_0)^{-1}D(s)q$$

where  $0 \neq q \in \text{Ker } D(\lambda_0)$ .

**proof.** If  $\lambda_0$  is an eigenvalue and  $p$  the corresponding eigenvector of  $S(D)$  then  $(s - \lambda_0)p(s)$  equals  $D(s)q(s)$  for some vector polynomial  $q(s)$ . Since  $p \in \mathcal{R}(D)(D)$ , we have  $D^{-1}(s)p(s) = (s - \lambda_0)^{-1}q(s)$  strictly proper so that  $q(s)$  is a constant vector. It cannot be the zero vector ( $p(s) \neq 0$ ) and, consequently,  $D(\lambda_0)$  is singular. The converse is immediate. □

**Definition 6.4 (Jordan chains of a matrix polynomial)** *Let  $D \in \mathbb{F}^{q \times q}[s]$ . The sequence of vectors  $\{x_i\}_0^{k-1}$  with  $x_i \in \mathbb{F}^q[s]$  and  $x_0 \neq 0$  is called a Jordan chain of length  $k$  for the matrix polynomial  $D$  corresponding to the eigenvalue  $\lambda_0 \in \mathbb{F}$  if  $D(s)x(s)$  is divisible by  $(s - \lambda_0)^k$  where*

$$x(s) = x_0 + x_1(s - \lambda_0) + x_2(s - \lambda_0)^2 + \cdots + x_{k-1}(s - \lambda_0)^{k-1}.$$

*The constant vector  $x_0$  is called an eigenvector of  $D(s)$  corresponding to the eigenvalue  $\lambda_0$  while the other vectors  $x_1, \dots, x_{k-1}$  are called generalized eigenvectors.*

We also recall the classical notion of a Jordan chain of vectors for a linear operator. If  $T$  is a linear operator on a vector space and  $\lambda_0$  is given, a sequence  $v_0, \dots, v_{k-1}$ ,  $v_0 \neq 0$  of vectors is said to be a Jordan chain of length  $k$  for  $T$  corresponding to  $\lambda_0$  if

$$(T - \lambda_0 I)v_0 = 0, (T - \lambda_0 I)v_1 = v_0, \dots, (T - \lambda_0 I)v_{k-1} = v_{k-2}.$$

The definition for matrices is analogous. Observe that  $v_0, \dots, v_{k-1}$  is a Jordan chain for a matrix  $T$  if and only if it is a Jordan chain in the sense of Definition 6.4 for the linear matrix polynomial  $T(s) = sI - T$ .

For more information about Jordan chains of matrix polynomials, we refer the interested reader to the book of Gohberg, Lancaster and Rodman [7].

**Theorem 6.5** (*Jordan chains of  $D$* )

Suppose  $D$  is row reduced. To each  $\lambda_0$  which is an eigenvalue of  $D(s)$  (hence also an eigenvalue of  $S(D)$ ), there is a one-to-one correspondence of the corresponding Jordan chains for  $D(s)$  and for  $S(D)$ .

If  $q^{(0)}, \dots, q^{(k-1)}$  is a chain of length  $k$  for  $D(s)$  then introduce the polynomials  $p_0(s), \dots, p_{k-1}(s)$  by

$$p_i(s) = \frac{1}{(s - \lambda_0)^{i+1}} D(s) q_i(s)$$

where

$$q_i(s) = q^{(0)} + q^{(1)}(s - \lambda_0) + \dots + q^{(i)}(s - \lambda_0)^i.$$

The polynomials  $p_i(s)$  belong to  $\mathcal{R}(D)$  and form a Jordan chain for  $S(D)$  of length  $k$ .

Conversally, if  $p_0(s), \dots, p_{k-1}(s)$  is a Jordan chain for  $S(D)$  then  $(S(D) - \lambda_0)^k p_{k-1}(s) = 0$  so that there exists a polynomial  $q$  such that  $(s - \lambda_0)^k p_{k-1}(s) = D(s)q(s)$ . If

$$q(s) = \sum q^{(i)}(s - \lambda_0)^i$$

is its expansion then the first  $k$  coefficients  $q^{(0)}, \dots, q^{(k-1)}$  form a Jordan chain for  $D(s)$ .

**proof.**

If  $q^{(0)}, \dots, q^{(k-1)}$  is a Jordan chain of length  $k$  for  $D(s)$ , consider the polynomials  $q_i$ ,

$$q_i(s) = q^{(0)} + \dots + q^{(i)}(s - \lambda_0)^i \quad \text{for } i = 0, \dots, k-1.$$

There exist polynomials  $p_i(s)$  such that

$$D(s)q_i(s) = (s - \lambda_0)^{i+1}p_i(s).$$

The identity

$$D(s)^{-1}p_i(s) = \frac{q_i(s)}{(s - \lambda_0)^{i+1}}$$

shows that  $D(s)^{-1}p_i(s)$  is strictly proper rational so that  $p_i \in \mathcal{R}(D)$ . Let us show that  $p_0(s), \dots, p_{k-1}(s)$  is a Jordan chain for the operator  $S(D)$ . Since  $p_0 \in \mathcal{R}(D)$  we have

$$\begin{aligned} (S(D) - \lambda_0 I)p_0(s) &= (R(D)S - \lambda_0 I)p_0(s) = R(D)(S - \lambda_0 I)p_0(s) \\ &= R(D)(s - \lambda_0)p_0(s) = R(D)D(s)q_0(s) = 0. \end{aligned}$$

For  $i > 0$

$$\begin{aligned} (S - \lambda_0 I)^{i+1}p_i(s) &= D(s)q_i(s) = D(s)q_{i-1}(s) + D(s)q^{(i)}(s - \lambda_0)^i \\ &= (s - \lambda_0)^i p_{i-1}(s) + D(s)q^{(i)}(s - \lambda_0)^i \end{aligned}$$

whence

$$(S - \lambda_0 I)p_i(s) = (s - \lambda_0)p_i(s) = p_{i-1}(s) + D(s)q^{(i)}.$$

Since the  $p_j$  belong to  $\mathcal{R}(D)$  we obtain

$$(S(D) - \lambda_0 I)p_i(s) = p_{i-1}(s).$$

In the same way the proof can be given in the other direction. □

The important computation of the characteristic polynomial of  $C(D)$  is left to the following section (Theorem 7.4) since it follows easily from the linearization equality.

## 7 LINEARIZATION OF ARBITRARY ROW REDUCED MATRIX POLYNOMIALS

In the study of linear differential as well as difference equations with constant coefficients, the fact that an equation of the  $n$ th order can be easily replaced by a system of  $n$  equations of the first order (and vice-versa) is often useful. This fact has a narrow connection with the transformation, called *linearization*, of the higher degree polynomial into a linear matrix pencil.

Take a differential equation of the  $n$ th order with constant coefficients

$$y^{(n)} + d_{n-1}y^{(n-1)} + \dots + d_1y' + d_0y = \varphi. \quad (7)$$

The most transparent transformation into a system of  $n$  differential equations of the first order is given by the substitution

$$\begin{aligned} y_0 &= y \\ y_1 &= y' \\ &\vdots \\ y_{n-1} &= y^{(n-1)}. \end{aligned}$$

Then the following system is obtained.

$$\begin{aligned} y_0' &= y_1 \\ y_1' &= y_2 \\ &\vdots \\ y_{n-2}' &= y_{n-1} \\ y_{n-1}' &= -d_{n-1}y_{n-1} - \dots - d_0y_0 + \varphi. \end{aligned}$$

Note that, in matrix notation, this system can be written as

$$(y_0', \dots, y_{n-1}') - (y_0, \dots, y_{n-1})C(d) = (0, \dots, 0, \varphi). \quad (8)$$

The characteristic polynomial of the equation (7) is  $d(s)$  and the characteristic polynomial of the system (8) is

$$sI_n - C(d). \quad (9)$$

Their mutual connection is given by

$$[sI_n - C(d)]F(s) = E(s) \begin{bmatrix} d(s) & 0 \\ 0 & I_{n-1} \end{bmatrix} \quad (10)$$

where

$$E(s) = \begin{bmatrix} 1 & -s & & \dots & 0 \\ 0 & 1 & -s & & \\ \vdots & & \ddots & \ddots & \\ & & & -s & \\ 0 & \dots & & & 1 \end{bmatrix}, \quad F(s) = \begin{bmatrix} b_1(s) & -1 & & \dots & 0 \\ b_2(s) & 0 & -1 & & \\ b_3(s) & \vdots & & \ddots & \\ \vdots & & & & -1 \\ b_n(s) & 0 & & \dots & 0 \end{bmatrix}$$

and the polynomials  $b_i$  are given by

$$b_i(s) = d_i + \dots + d_n s^{n-i}.$$

The linear matrix polynomial  $sI_n - [C(d)]^T$  is called a *linearization* of  $d(s)$ .

Generalization of this situation for a  $q \times q$  system of differential equations the characteristic polynomial of which is a  $q \times q$  matrix polynomial  $D(s)$  was also studied. Corresponding results can be found in [8, Lemma 6.3-20] and [7, Theorem 1.1].

If  $D(s)$  is monic then all the  $q$  equations of the system have the same order, say  $K$ . It is reasonable to expect that the “dimension” of the problem is  $Kq$  and this number really coincides with the dimension of the linearization by [8] and [7]. However, the approach taken there is such that the same dimension  $Kq$  is obtained when linearizing a system with a nonmonic characteristic polynomial  $D(s)$ . The solution is a linear matrix pencil  $As - B$  such that the coefficient  $A$  has defect equal to  $Kq - n$ ,  $n$  being the degree of  $\det D(s)$ . This means that besides the  $n$  characteristic values of  $D(s)$ , there are  $Kq - n$  additional characteristic values at infinity.

Our paper offers a different approach to the nonmonic case.

Since we shall need the assumption that  $D(s)$  is row reduced, we first emphasize that it is easy to meet this assumption by an easy pretransformation of the system. Actually, by equivalent row transformations we can get any characteristic polynomial  $D' \in U_L(D)$ . By Lemma 4.3 we can get both row and column reduced characteristic polynomials.

For  $D(s)$  row reduced, the number  $n = \deg \det D(s)$  is just equal to the sum of the orders of all the  $q$  equations. The expected dimension of this problem is, of course,  $n$  and we show here that the *monic* linear matrix pencil  $sI_n - C(D)$  (where  $C(D)$  is our finite companion matrix) is a *natural linearization* having the mentioned *minimal* dimension  $n$ .

The practical aspect of the problem is to find the substitution transforming a system of equations of a higher order into the first order system with characteristic polynomial  $sI_n - C(D)$  (like in (8) for the scalar case). Since this transformation exists, the characteristic values of the corresponding systems have to coincide.

This fact implies a natural question whether a stronger connection of the characteristic polynomials of both systems can be found. Such a connection does exist and performs the second and more theoretical aspect of linearization: The characteristic polynomial of the higher order system can be extended by an identity matrix (of a proper dimension) and then transformed by equivalent transformations (i.e., by pre- and/or postmultiplication by unimodular matrices) into the corresponding linear matrix pencil.

Following [8] and [7], we shall use the name linearization in a narrower sense for this equivalence relation. Our Definition 7.1 is different from the definition used there if the characteristic polynomial  $D(s)$  is not monic.

Let us start with the first aspect: mutual transformation of systems of differential (and difference) equations of higher orders into systems having more equations but of the first order.

### Transformation of difference and differential systems

Let us write for short  $\mathcal{I} = \mathcal{I}(D)$ . If  $D$  is row reduced, denote by  $z(0) < z(1) < \dots < z(n-1)$  the elements of  $\mathcal{I}$  arranged in increasing order. Further let us introduce the notation  $D(S^q)$  for the (block) matrix

$$D(S^q) = \sum_i S^i \otimes D_i$$

where  $\otimes$  denotes the Kronecker product.

With this notation, it is easy to describe the row space of the matrix  $\tilde{C}_\infty(D)$ .

**Lemma 7.1** *For any nonsingular  $D$  and any infinite row vector  $v^T$ , the vector  $v^T$  belongs to the row space of  $\tilde{C}_\infty(D)$  if and only if*

$$v^T D(S^q) = 0.$$

**proof.**

Take  $k = 1$  and  $M = v^T$  in Theorem 5.5. Note that  $p \in DF^q[s]$  if and only if  $\hat{p}$  belongs to the column space of the matrix  $D(S^q)$ . Consequently,

$$v^T D(S^q) = 0 \quad \Leftrightarrow \quad v^T \hat{p} = 0, \quad \forall p \in DF^q[s].$$

Hence, (by Theorem 5.5, (i)→(ii))

$$v^T = v^T \tilde{C}_\infty(D)$$

which means that  $v^T$  is in the row space of  $\tilde{C}_\infty(D)$ . Conversely, if  $v^T$  is in the row space of  $\tilde{C}_\infty(D)$  then, using (iii)→(i) of Theorem 5.5,  $v^T \hat{p} = 0$ ,  $\forall p \in D\mathbf{F}^q[s]$ . Therefore,  $v^T D(S^q) = 0$ .  $\square$

**Theorem 7.2** *Let  $D$  be a row reduced polynomial matrix of dimension  $q$ , set  $n = \deg \det D$ . Then there is a one-to-one correspondence between the solutions of the  $q$ -dimensional system corresponding to  $D$ ,*

$$\sum_0^K D_j^T v^{(j+k)} = 0, \quad k = 0, 1, 2, \dots \quad (11)$$

and the solutions of the  $n$ -dimensional system corresponding to the linear polynomial  $sI_n - C(D)^T$ ,

$$w^{(l+1)} - [C(D)]^T w^{(l)} = 0, \quad l = 0, 1, 2, \dots \quad (12)$$

More precisely, this correspondence is given by the mapping  $T : F^{q \times \infty} \rightarrow F^{n \times \infty}$ ,

$$T\{v^{(k)}\} = \{w^{(l)}\}$$

if

$$w^{(l)} = \hat{v}_{(I+lq)}$$

( $\hat{v}$  is the stacking vector), i.e.,

$$w_j^{(l)} = v_i^{(k+l)} \text{ where } i + kq = z(j).$$

**proof.** We note first that  $\{v^{(k)}\}$  is a solution to (11) if and only if  $(\hat{v})^T D(S^q) = 0$  which is (by the preceding lemma) if and only if  $(\hat{v})^T$  belongs to the row space of  $C_\infty(D)$ . For  $D$  row reduced, we have  $C_\infty(D)^{(I)} = I_n$  and, consequently,  $\hat{v}^T$  has the form

$$(\hat{v})^T = [(\hat{v})^T]^{(I)} C_\infty(D).$$

Denote  $[(\hat{v})^T]^{(I)}$  by  $c^T$ . Now

$$\begin{aligned} (w^{(l)})^T &= ((\hat{v})^T)^{(I+lq)} = \\ &= c^T C_\infty(D)^{(I+lq)} = c^T C(D)^l \end{aligned}$$

(by Theorem 6.1) or  $[C(D)^l]^T c = w^{(l)}$ . Consequently,  $w^{(l+1)} = [C(D)]^T w^{(l)}$  and  $w^{(0)} = c$ .

On the other hand, if  $\{w^{(l)}\}$  is any solution to (12) such that

$$w^{(0)} = c$$

then

$$[w^{(l)}]^T = c^T C(D)^l$$

and if we set  $(\hat{v})^T = c^T C_\infty(D)$ , it holds

$$[w^{(l)}]^T = [(\hat{v})^T]^{(I+lq)}$$

so that

$$w^{(l)} = [\hat{v}]_{(I+lq)}.$$

Consequently, the solution  $\{v^{(k)}\}$  of (11) with the initial conditions  $v_i^{(k)} = c_j$  (for  $j = 0, \dots, n-1$  and  $i, k$  such that  $i+kq = z(j)$ ) is the unique solution such that  $T\{v^{(k)}\} = \{w^{(l)}\}$ .  $\square$

The case of differential equations is completely analogous. The solutions of the systems

$$\sum_{j=0}^K D_j^T v^{(j)}(x) = 0 \quad (13)$$

and

$$w'(x) = [C(D)]^T w(x) \quad (14)$$

can be put into a one-to-one correspondence by the equality

$$w_j^{(l)} = v_i^{(k+l)} \text{ where } i + kq = z(j).$$

Now we turn to the second, theoretical point of view.

**Definition 7.1** A linear monic matrix polynomial  $sI - A$  is called a linearization of the (not necessarily monic) matrix polynomial  $D(s)$  if there are unimodular matrices  $E(s)$  and  $F(s)$  and a positive integer  $r$  such that

$$sI - A = E(s) \begin{bmatrix} D(s) & 0 \\ 0 & I_r \end{bmatrix} F(s). \quad (15)$$

We start by formulating the main results and then we continue by giving the proofs.

**Theorem 7.3** Let  $D(s) = \sum_{k=0}^K D_k s^k$  be a row reduced not necessarily monic matrix polynomial, let  $n = \deg \det D$ . Then the monic linear matrix polynomial  $sI - C(D)$  is a linearization of  $D$ . More precisely, it holds

$$[sI - C(D)]F(s) = E(s)\Delta(s) \quad (16)$$

where  $\Delta(s) = \text{diag}(D(s), I_{n-q})$  for unimodular matrix polynomials  $E(s)$ ,  $F(s)$  given by

$$F(s) = [B(s) - (S^q)^*] \begin{matrix} (T) \\ (T) \end{matrix},$$

$B(s)$  being the matrix

$$B(s) = \begin{bmatrix} B_1(s) \\ B_2(s) \\ \vdots \\ B_K(s) \\ 0 \\ \vdots \end{bmatrix} [ I \quad 0 \quad \dots \quad 0 \quad \dots ]$$

where

$$B_j(s) = D_j + D_{j+1}s + \dots + D_K s^{K-j},$$

and

$$E(s) = [I - s(S^q)^*] \begin{matrix} (T) \\ (T) \end{matrix}.$$

As a corollary, we compute the characteristic polynomial of the finite companion matrix.

**Theorem 7.4**

$$\det(sI - C(D)) \det D_{hrdc} = \det D(s).$$

**proof.**

Compute the determinants of all matrices in the above matrix identity (16). It holds  $\det \Delta(s) = \det D(s)$  and  $\det[I - s(S^q)^*]_{(T)}^{(T)} = 1$  so that the degree of the determinant on the r.h.s. equals  $n = \deg \det[sI - C(D)]$ . Consequently, the matrix  $[B(s) - (S^q)^*]_{(T)}^{(T)}$  is unimodular and

$$\det[sI - C(D)]c = \det D(s)$$

with  $c$  a nonzero constant. We know that the highest coefficient in  $\det D(s)$  is  $D_{hrdc}$  so that the constant  $c$  is equal to  $D_{hrdc}$ .  $\square$

The proof of the linearization theorem is based on a simple operator identity for the shift operator on the space of vector polynomials. This identity was already used in the case of scalar polynomials [11, page 90] and is restated here because it is of independent interest. We use the symbol  $P_0$  for the operator which assigns to any vector polynomial  $p = \sum p_i s^i \in \mathbf{F}^q[s]$  the constant vector polynomial  $p_0 \in \mathbf{F}^q[s]$ .

$$(rI - \mathbf{S}) \sum_0^{K-1} r^j (\mathbf{S}^*)^{j+1} = P_0 \sum_0^{K-1} r^j (\mathbf{S}^*)^j + r^K (\mathbf{S}^*)^K - I; \quad (17)$$

when restricted to the subspace of all polynomials of degree not exceeding  $K$ , it assumes the following form

$$(rI - \mathbf{S}) \sum_0^{K-1} r^j (\mathbf{S}^*)^{j+1} = P_0 \sum_0^K r^j (\mathbf{S}^*)^j - I \quad (18)$$

because, in that case,  $r^K (\mathbf{S}^*)^K = P_0 r^K (\mathbf{S}^*)^K$ . To prove the identity (17), take any  $r \in \mathbf{F}$ , write down the sum of the geometric series

$$(r\mathbf{S}^* - I) \sum_0^{K-1} r^j (\mathbf{S}^*)^j = r^K (\mathbf{S}^*)^K - I \quad (19)$$

and replace, on the left hand side, the identity by  $\mathbf{S}\mathbf{S}^* + P_0$  so that

$$r\mathbf{S}^* - I = (rI - \mathbf{S})\mathbf{S}^* - P_0.$$

For any (constant) matrix  $A \in \mathbf{F}^{q \times q}$  we denote by  $M(A)$  the operator of multiplication by  $A$  on the space  $\mathbf{F}^q[s]$ , i.e.  $M(A)p(s) = Ap(s)$ . Note that the matrix of  $M(A)$  with respect to the basis  $\mathcal{B}$  is the block diagonal matrix  $\text{diag}(A, A, A, \dots)$ .

We introduce the *matrix polynomial function of the shift operator* by

$$D(\mathbf{S}) = \sum M(D_i) \mathbf{S}^i.$$

The next step consists in composing both sides of the identity (18) with the operator  $D(\mathbf{S})P_0$ . Let us precompute the sum on the l.h.s.,

$$\sum_{j=0}^{K-1} r^j (\mathbf{S}^*)^{j+1} D(\mathbf{S})P_0 = \sum_{j=0}^{K-1} r^j (\mathbf{S}^*)^{j+1} \sum_{i=0}^K M(D_i) \mathbf{S}^i P_0.$$

Using the fact that multiplication operators commute with the shifts and reducing the sums by means of the equalities  $\mathbf{S}^* \mathbf{S} = I$  and  $\mathbf{S}^* P_0 = 0$ , the expression can be written as

$$\sum_{i=0}^K M(D_i) \sum_{0 \leq j \leq i-1} r^j \mathbf{S}^{i-j-1} P_0.$$

Writing  $k$  for  $i - j - 1$  so that  $i = k + 1 + j$ , this sum appears in the form

$$\sum_{k=0}^{K-1} \left( \sum_{j=0}^{K-k-1} r^j M(D_{k+1+j}) \right) \mathbf{S}^k P_0 = \sum_{k=0}^{K-1} M(B_{k+1}(r)) \mathbf{S}^k P_0.$$

The r.h.s. of (18) composed by  $D(\mathbf{S})P_0$  gives

$$P_0 \sum_0^K r^j (\mathbf{S}^*)^j D(\mathbf{S})P_0 - D(\mathbf{S})P_0 = \sum_0^K r^j M(D_j)P_0 - D(\mathbf{S})P_0 = \\ M(D(r))P_0 - D(\mathbf{S})P_0.$$

In this manner, we obtain the following identity

$$(rI - \mathbf{S}) \sum_0^{K-1} M(B_{k+1}(r)) \mathbf{S}^k P_0 = M(D(r))P_0 - D(\mathbf{S})P_0. \quad (20)$$

Using (20) and the identities  $(rI - \mathbf{S})\mathbf{S}^* = -(I - r\mathbf{S}^*)\mathbf{S}\mathbf{S}^*$  and  $\mathbf{S}^*M(D(r))P_0 = 0$ , we obtain

$$(rI - \mathbf{S}) \left( \sum_0^{K-1} M(B_{k+1}(r)) \mathbf{S}^k P_0 - \mathbf{S}^* \right) = M(D(r))P_0 - D(\mathbf{S})P_0 - (rI - \mathbf{S})\mathbf{S}^* = \\ (I - r\mathbf{S}^*)M(D(r))P_0 - D(\mathbf{S})P_0 + (I - r\mathbf{S}^*)\mathbf{S}\mathbf{S}^* = \\ (I - r\mathbf{S}^*)(M(D(r))P_0 + \mathbf{S}\mathbf{S}^*) - D(\mathbf{S})P_0.$$

We introduce the notation  $\mathbf{B}(r)$  for the operator  $\sum_0^{K-1} M(B_{k+1}(r)) \mathbf{S}^k P_0$ .

Composing this identity to the left by  $R(D)$  and using the identity  $R(D)D(\mathbf{S})P_0 = 0$ , we have

$$R(D)(rI - \mathbf{S})(\mathbf{B}(r) - \mathbf{S}^*) = R(D)(I - r\mathbf{S}^*)(M(D(r))P_0 + \mathbf{S}\mathbf{S}^*).$$

If we consider the restriction to the subspace  $\mathcal{R}(D)$  and if we join an additional condition that  $R(D)M(D(r))P_0 = M(D(r))P_0$  on  $\mathcal{R}(D)$ , we can omit  $R(D)$  on the r.h.s. and we have to replace the shift on the l.h.s. by  $S(D)$ . In this way, we obtain the following theorem the matrix representation of which is Theorem 7.3 above:

**Theorem 7.5** *Suppose that  $D$  is row reduced and, moreover, that  $R(D)M(D(r))P_0 = M(D(r))P_0$  restricted to  $\mathcal{R}(D)$ .*

*Then*

$$[rI - S(D)][\mathbf{B}(r) - \mathbf{S}^*] = [I - r\mathbf{S}^*][M(D(r))P_0 + \mathbf{S}\mathbf{S}^*]$$

*if the operators on both sides are restricted to the space  $\mathcal{R}(D)$ .*

**Remark 7.1** *Note that the condition  $R(D)M(D(r))P_0 = M(D(r))P_0$  on  $\mathcal{R}(D)$  holds, e.g., in the case when all row degrees of  $D$  are positive.*

## 8 CONNECTION BETWEEN THE INFINITE COMPANION MATRIX FOR $D$ AND THE SCALAR INFINITE COMPANION MATRIX FOR $\det D$

Let  $D$  be any row reduced matrix polynomial in  $\mathbb{F}^{q \times q}[s]$  and let  $d = \det D$ . (Recall that  $n = \deg d = \sum \text{row degrees of } D$ .)

Note that

$$dI_q = DD_{\text{adj}}$$

and use Lemma (4.2)(1) to get

$$\text{Ker } R(dI_q) \subset \text{Ker } R(D).$$

From this fact it can be shown that

$$\text{span rows } C_\infty(dI_q) \supset \text{span rows } C_\infty(D).$$

This implies that

$$C_\infty(D) = MC_\infty(dI_q)$$

for an  $n \times nq$  matrix  $M$ . Since division by  $dI_q$  is equivalent to scalar division by  $d$  in each component, we can write

$$C_\infty(dI_q) = C_\infty(d) \otimes I_q.$$

Since the first  $n$  columns of  $C_\infty(d)$  form the identity matrix  $I_n$ , we conclude that the first  $nq$  columns of the matrix  $C_\infty(d) \otimes I_q = C_\infty(dI_q)$  form the identity matrix of order  $nq$ . After having noticed this fact, we can easily find the explicit form of the matrix  $M$  and formulate the following theorem.

**Theorem 8.1** *For any row reduced matrix polynomial  $D \in \mathbf{F}^{q \times q}[s]$ ,*

$$C_\infty(D) = M(C_\infty(d) \otimes I_q)$$

where

$$M = [C_0, C(D)C_0, C(D)^2C_0, \dots, C(D)^{n-1}C_0]$$

and the matrix  $C_0$  is formed by the first  $q$  columns of  $C_\infty(D)$ .

In other words,  $M$  is formed by the first  $nq$  columns of  $C_\infty(D)$ .

## 9 FACTORIZATION OF INFINITE BLOCK HANKEL MATRICES HAVING FINITE RANK

It is a well-known fact that any scalar infinite Hankel matrix having finite rank corresponds to a (strictly proper) rational function,  $H = H(\omega) = (\omega_{i+j})$  where  $\omega(s) = q(s)/f(s) = \sum_{i=0}^{\infty} \omega_i s^{-i-1}$  [6, p. 207]. This matrix has a well-known factorization

$$H = V_f M V_f^T$$

where  $V_f$  is the (in general confluent) Vandermonde matrix, corresponding to the zeros of  $f$  and  $M$  is an (in general block) diagonal matrix. There is also another factorization given by Pták in terms of the infinite companion matrix,

$$H = [C_\infty(f)]^T H_0 C_\infty(f)$$

where  $H_0$  is the leading  $n \times n$  submatrix of  $H$ .

The generalization of the first formula for block Hankel matrices is studied (even for finite block Hankel matrices) in [4]. Here we give a block generalization of the factorization using infinite companion matrices.

Let  $\Omega(s) = \sum \Omega_i s^{-i-1} \in \mathbf{F}^{p \times q}(s)$  and  $H(\Omega) = [\Omega_{i+j}]$ .

Choose one right and one left polynomial matrix fraction description of  $\Omega$  (not necessarily coprime):

$$\Omega = N_R D_R^{-1} = D_L^{-1} N_L.$$

Recall that the Hankel operator  $\mathbf{H} = \mathbf{H}(\Omega)$  corresponding to the matrix rational function  $\Omega$  is the operator which maps any  $p \in \mathbf{F}^q[s]$  onto the strictly proper rational function

$$v = \Pi_- \Omega p.$$

In terms of expansions in powers of  $s$ , the Hankel operator  $\mathbf{H}$  is represented by the Hankel matrix  $H$ . Evidently

$$\text{Ker } \mathbf{H} \supset D_R \mathbf{F}^q[s],$$

i.e.  $\text{Ker } R(D_R) \subset \text{Ker } \mathbf{H}$ . We can apply Theorem 5.5 to corresponding matrix representations to obtain

$$H = H \tilde{C}_\infty(D_R). \quad (21)$$

Using the same reasoning for  $H^T$ , we can write

$$H^T = H^T \tilde{C}_\infty(D_L^T). \quad (22)$$

Substituting (22) into (21), we get the following theorem.

**Theorem 9.1 (Factorization of block Hankel matrices)**

(1) Using the notation introduced above,

$$H = [\tilde{C}_\infty(D_L^T)]^T H \tilde{C}_\infty(D_R).$$

(2) If we moreover suppose that  $D_R, D_L^T$  are row reduced then

$$H = [C_\infty(D_L^T)]^T H_{rs} C_\infty(D_R)$$

where  $H_{rs}$  is the  $r \times s$  submatrix of  $H$  given by

$$H_{rs} = H_{\begin{smallmatrix} \mathcal{I}_R \\ \mathcal{I}_L \end{smallmatrix}}$$

where  $\mathcal{I}_R = \mathcal{I}(D_R)$ ,  $\mathcal{I}_L = \mathcal{I}(D_L^T)$  and  $r = |\mathcal{I}_R| = \deg \det D_R$ ,  $s = |\mathcal{I}_L| = \deg \det D_L$ .

(3) If  $D_R, D_L^T$  are row reduced and, moreover, both decompositions  $\Omega = N_R D_R^{-1}$  and  $\Omega = D_L^{-1} N_L$  are coprime then

$$H = C_\infty(D_L^T) H_n C_\infty(D_R)$$

where  $H_n = H_{\begin{smallmatrix} \mathcal{I}_R \\ \mathcal{I}_L \end{smallmatrix}}$  is a square nonsingular submatrix of  $H$  of order

$$n = \deg \det D_R = \deg \det D_L = \text{rank } H.$$

The fact that  $n = \deg \det D_R = \deg \det D_L = \text{rank } H$  under the assumptions of item (3) can be found e.g. in [8], page 442.

**Remark 9.1** Let  $\Omega = N_R D_R^{-1}$  be any right coprime polynomial matrix fraction decomposition. Then the unique row echelon form of the Hankel matrix  $H$  is the infinite companion matrix

$$\begin{bmatrix} C_\infty(D_c) \\ 0 \end{bmatrix}$$

where  $D_c$  is the canonical form of  $D_R$ .

In conclusion we note that the intertwining relation

$$(S^p)^T H = H S^q \quad (23)$$

together with the factorization equalities give the following:

$$\begin{aligned} [C_\infty(D_L^T)]^T [C(D_L^T)]^T H_n C_\infty(D_R) = \\ [C_\infty(D_L^T)]^T H_n C(D_R) C_\infty(D_R) \end{aligned}$$

and, consequently,

$$[C(D_L^T)]^T H_n = H_n C(D_R).$$

This is compression of the relation (23). This relation shows also that  $[C(D_L^T)]^T$  and  $C(D_R)$  are similar while  $H_n$  is the corresponding transformation matrix.

## References

- [1] S. Barnett. *Polynomials and linear control systems*, volume 77 of *Pure and Applied Mathematics*. Marcel Dekker, Inc., New York, 1983.
- [2] S. Beghelli and R. Guidorzi. A new input-output canonical form for multivariable systems. *IEEE Transactions on Automatic Control*, AC-21:692–696, 1976. Addendum: *IEEE Transactions on Automatic Control* AC-23:93, 1978.
- [3] S. Beghelli and R. Guidorzi. A new realization algorithm for linear multivariable systems. *Ricerche di Automatica*, 7(2-3):164–181, December 1976.
- [4] S. Feldmann and G. Heinig. Vandermonde factorization and canonical representations of block-Hankel matrices. *Linear Algebra and Its Applications*, 241:247–278, 1996.
- [5] P.A. Fuhrmann. Algebraic system theory: An analyst’s point of view. *Journal Franklin Institute*, 301(6):521–540, June 1976.
- [6] F.R. Gantmacher. *The theory of matrices, Volume II*. Chelsea, New York, 1959.
- [7] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix polynomials*. Computer Science and Applied Mathematics. Academic Press, 1982.
- [8] T. Kailath. *Linear Systems*. Prentice-Hall, 1980.
- [9] V. Pták. Spectral radius, norms and iterates and the critical exponent. *Linear Algebra and Its Applications*, 1:245–260, 1968.
- [10] V. Pták. An infinite companion matrix. *Commentationes Mathematicae Universitatis Carolinae*, 19(3):447–458, 1978.
- [11] V. Pták. The infinite companion matrix. *Linear Algebra and Its Applications*, 166:65–95, 1992.
- [12] D. Sarason. Generalized interpolation in  $H^\infty$ . *Transactions of the American Mathematical Society*, 127:179–203, 1967.
- [13] M. Van Barel and A. Bultheel. An algebraic method to solve the minimal partial realization problem for matrix sequences. In C.I. Byrnes, C.F. Martin, and R.E. Saeks, editors, *Linear Circuits, Systems and Signal Processing: Theory and applications*, pages 93–100. Elsevier Science Publishers B.V. (North-Holland), 1988. Proceedings of The International Symposium on the Mathematical Theory of Networks and Systems (MTNS87), USA, June 15-19, 1987.