

Numerical Computation of Minimal Polynomial Bases: A Generalized Resultant Approach

E.N. Antoniou*, A.I.G. Vardulakis and S. Vologiannidis
Aristotle University of Thessaloniki
Department of Mathematics
Faculty of Sciences
54006, Thessaloniki - Greece
E-mail: (antoniou, avardula, svol)@math.auth.gr

March 23, 2005

Abstract

We propose a new algorithm for the computation of a minimal polynomial basis of the left kernel of a given polynomial matrix $F(s)$. The proposed method exploits the structure of the left null space of generalized Wolovich or Sylvester resultants to compute row polynomial vectors that form a minimal polynomial basis of left kernel of the given polynomial matrix. The entire procedure can be implemented using only orthogonal transformations of constant matrices and results to a minimal basis with orthonormal coefficients.

Keywords: polynomial matrices; Minimal polynomial basis; Matrix fraction description.

AMS classification: 15A33, 15A23, 65F30

1. Introduction

The problem of determination of a minimal polynomial basis of a rational vector space (see [8]) is the starting point of many control analysis, synthesis and design techniques based on the “polynomial matrix approach” [22], [6], [15], [21]. Given a rational transfer function matrix $P(s)$ it is usually required to determine left or right coprime polynomial matrix fractional representations (factorizations) of $P(s)$ of the form $P(s) = D_L^{-1}(s)N_L(s) = N_R(s)D_R^{-1}(s)$. Moreover, in many applications apart from the coprimeness requirement of the above factorizations, it is often desirable to have factorizations where either the denominator matrices $D_L(s), D_R(s)$ or the compound matrices $E(s) := [D_L(s), N_L(s)]$, $F(s) := [N_R^T(s), -D_R^T(s)]^T$ have respectively

*Corresponding Author. Supported by the Greek State Scholarships Foundation (IKY) (Postdoctoral research grant, Contract Number 411/2003-04)

minimal row or column degrees, i.e. they are row or column proper (reduced). Classical examples of such applications are the denominator assignment problem (see [23], [6], [7], [15], [16], [1]) and the determination of a minimal realization (see [22], [21], [20]) of a MIMO rational transfer function, where a minimal in the above sense and coprime factorization of the plant is required. Furthermore, even the problem of row or column reduction of a polynomial matrix itself can be solved using minimal polynomial bases computation techniques as described in [4] and [18].

The classical approach (see [22], [11]) to the problem of finding a minimal polynomial basis of a rational vector space, starts from a given, possibly non-minimal, polynomial basis and in the sequel applying polynomial matrix techniques (extraction of greatest common divisors and unimodular transformations for row/column reduction) one can obtain the desired minimal basis. However, such implementations are known to suffer of serious numerical problems and thus they are not recommended for real-life applications. A numerically reliable alternative to the classical approach has been presented in [3]. The method presented in [3] utilizes the "pencil approach" by applying generalized Schur decomposition on the block companion form of the polynomial matrix, which in turn allows the computation of a minimal polynomial basis of the original matrix. A second alternative appears in [19], where the computation of the minimal basis is accomplished via the Padé approximants of the polynomial matrices involved. Our approach to the problem is comparable to the techniques presented in [12], [13], [14] and [17] where the computation of minimal polynomial bases of matrix pencils is considered and to the one in [2] where the structure of Sylvester resultant matrices is being utilized.

The problem of computation of a minimal basis can be stated as follows. Given a full column rank (over $\mathbb{R}(s)$) polynomial matrix $F(s) \in \mathbb{R}^{(p+m) \times m}[s]$ determine a left unimodular, row proper matrix $E(s) \in \mathbb{R}^{p \times (p+m)}[s]$ such that

$$E(s)F(s) = 0.$$

Then $E(s)$ is a minimal polynomial basis of the left kernel of $F(s)$. Our approach to the problem exploits the structure of the generalized Sylvester or Wolovich resultant (see [5], [1]) of the polynomial matrix $F(s)$. Notice that the methods presented in [3], [19], [12], [13] and [2] deal with the dual problem, i.e. determination of a minimal basis of the right kernel of $E(s)$.

The outline of the paper is as follows. In section 2 we present the necessary mathematical background and notation as well as some known results regarding the structure of generalized resultants. Section 3 presents the main results of the paper along with the proposed algorithm for the computation of minimal polynomial bases. In section 4 we discuss the numerical properties of the proposed algorithm, while in section 5 we provide illustrative examples for the method. Finally, in section 6 we summarize and draw our conclusions.

2. Mathematical background

In the following \mathbb{R} , \mathbb{C} , $\mathbb{R}(s)$, $\mathbb{R}[s]$, $\mathbb{R}_{pr}(s)$, $\mathbb{R}_{po}(s)$ are respectively the fields of *real numbers*, *complex numbers*, *real rational functions*, the rings of *polynomials*, *proper rational* and *strictly proper rational* functions all with coefficients in \mathbb{R} and indeterminate s . For a set \mathbb{F} , $\mathbb{F}^{p \times m}$ denotes the set of $p \times m$ matrices with entries in \mathbb{F} . \mathbb{N}^+ is the set of positive integers. The symbols $rank_{\mathbb{F}}(\cdot)$, $ker_{\mathbb{F}}(\cdot)$ and $Im_{\mathbb{F}}(\cdot)$ denote respectively the rank, right kernel (null space) and image (column span) of the matrix in brackets over the field \mathbb{F} . Furthermore in certain

cases we may use the symbols $\ker_{\mathbb{F}}^L(\cdot)$ and $\text{Im}_{\mathbb{F}}^L(\cdot)$ to denote the left kernel and row span of the corresponding matrix over \mathbb{F} . In case \mathbb{F} is omitted in one of these symbols \mathbb{R} is implied. If $m \in \mathbb{N}^+$ then \mathbf{m} denotes the set $\{1, 2, \dots, m\}$.

A polynomial matrix $T(s) \in \mathbb{R}^{p \times m}[s]$ will be called left (resp. right) unimodular iff $\text{rank}T(s_0) = p$ (resp. $\text{rank}T(s_0) = m$) for every $s_0 \in \mathbb{C}$, or equivalently iff $T(s)$ has no zeros in \mathbb{C} . When $T(s)$ is a square polynomial matrix then $T(s)$ will be called unimodular iff $\text{rank}T(s_0) = p = m$ for every $s_0 \in \mathbb{C}$.

A polynomial matrix $X(s) \in \mathbb{R}^{p \times m}[s]$ ($p \geq m$) is called column proper or column reduced iff its highest column degree coefficient matrix, denoted by X^{hc} , which is formed by the coefficients of the highest powers of s in each column of $X(s)$, has full column rank. The column degrees of $X(s)$ are usually denoted by $\deg_{ci} X(s)$, $i \in \mathbf{m}$. Respectively $Y(s) \in \mathbb{R}^{p \times m}[s]$ ($p \leq m$) is called row proper or row reduced iff $Y^\top(s)$ is column proper and the row degrees of $Y(s)$ are denoted by $\deg_{ri} Y(s)$, $i \in \mathbf{p}$.

Let $F(s) \in \mathbb{R}^{(p+m) \times m}[s]$ with $\text{rank}_{\mathbb{R}(s)} F(s) = m$ and $E(s) \in \mathbb{R}^{p \times (p+m)}[s]$ with $\text{rank}_{\mathbb{R}(s)} E(s) = p$ be polynomial matrices such that

$$E(s)F(s) = 0. \quad (2.1)$$

When (2.1) is satisfied and $E(s)$ is row proper and left unimodular, $E(s)$ is a minimal polynomial basis [8] of the (rational vector space spanning the) left kernel of $F(s)$ and the row degrees $\deg_{ri} E(s) =: \mu_i$, $i \in \mathbf{p}$ of $E(s)$ are the *invariant minimal row indices* of the left kernel of $F(s)$ or simply the *left minimal indices* of $F(s)$. Similarly when (2.1) is satisfied with $F(s)$ column proper and right unimodular, $F(s)$ is a minimal polynomial basis of the (rational vector space spanning the) right kernel of $E(s)$ and the column degrees $\deg_{ci} F(s) =: \nu_i$, $i \in \mathbf{m}$ of $F(s)$ are the *invariant minimal column indices* of the right kernel of $E(s)$ or simply the *right minimal indices* of $E(s)$.

Given a polynomial matrix $F(s) \in \mathbb{R}^{(p+m) \times m}[s]$ and $k, a, b \in \mathbb{N}^+$ we introduce the matrices

$$S_{a,b}(s) := [I_a, sI_a, \dots, s^{b-1}I_a]^\top \in \mathbb{R}^{ba \times a}[s], \quad (2.2)$$

$$X_k(s) := S_{p+m,k}(s)F(s) \in \mathbb{R}^{(p+m)k \times m}[s]. \quad (2.3)$$

Formula (2.3) is essentially the basis for the construction of generalized resultants. Let $F(s) = F_0 + sF_1 + \dots + s^q F_q$, $F_i \in \mathbb{R}^{(p+m) \times m}$ and write

$$X_k(s) = R_k S_{m,q+k}(s), \quad (2.4)$$

where R_k

$$R_k := \begin{bmatrix} F_0 & F_1 & \dots & F_q & 0 & \dots & 0 \\ 0 & F_0 & F_1 & \dots & F_q & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \dots & \ddots & 0 \\ 0 & \dots & 0 & F_0 & F_1 & \dots & F_q \end{bmatrix} \in \mathbb{R}^{(p+m)k \times m(q+k)}. \quad (2.5)$$

The matrix R_k is known [5] as the *Generalized Sylvester Resultant* of $F(s)$.

Let $\nu_i = \deg_{ci} F(s)$, $i \in \mathbf{m}$ be the column degrees of $F(s)$. Similarly to [22] (page 242) $X_k(s)$ can be written as

$$X_k(s) = M_{ek} \text{block diag}\{S_{1,\nu_i+k}(s)\}_{i \in \mathbf{m}}, \quad (2.6)$$

where $M_{ek} \in \mathbb{R}^{(m+p)k \times (mk + \sum_{i=1}^m \nu_i)}$. The matrix M_{ek} is defined [1] as the *Generalized Wolovich Resultant* of $F(s)$.

Write $F(s) = [f_1(s), f_2(s), \dots, f_m(s)]$ where $f_i(s) = f_{i0} + sf_{i1} + \dots + s^{\nu_i} f_{ik_i} \in \mathbb{R}^{(m+p) \times 1}[s]$, $i \in \mathbf{m}$ are the columns of $F(s)$. Then it is easy to see that

$$M_{ek} = [R_k^1, R_k^2, \dots, R_k^m], \quad (2.7)$$

where

$$R_k^i = \begin{bmatrix} f_{i0} & f_{i1} & \dots & f_{ik_i} & 0 & \dots & 0 \\ 0 & f_{i0} & f_{i1} & \dots & f_{ik_i} & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \dots & \ddots & \vdots \\ 0 & 0 & \dots & f_{i0} & f_{i1} & \dots & f_{ik_i} \end{bmatrix} \in \mathbb{R}^{(m+p)k \times (\nu_i + k)} \quad i \in \mathbf{m},$$

is the generalized Sylvester resultant of the column $f_i(s)$, $i \in \mathbf{m}$ of $F(s)$. It is easy to see that the two types of generalized resultants are related through

$$R_k = [M_{ek}, 0_{(m+p)k, b}] P_k, \quad (2.8)$$

where $P_k \in \mathbb{R}^{m(q+k) \times m(q+k)}$ is a column permutation matrix. The fact that R_k contains at least $b = mq - \sum_{i=1}^m \nu_i$, where $q = \max_{i \in \mathbf{m}} \{\nu_i\}$, zero columns has been observed in [23]. The following result will be very useful in the sequel

Theorem 2.1. *Let $E(s) \in \mathbb{R}^{p \times (p+m)}[s]$ be a minimal polynomial basis for the left kernel of $F(s)$ as in (2.1) and let $\mu_i = \deg_{r_i} E(s)$, $i \in \mathbf{p}$ be the invariant minimal row indices of the left kernel of $F(s)$. Then*

$$\text{rank} R_k = \text{rank} M_{ek} = (p+m)k - v_k, \quad (2.9)$$

where $v_k = \sum_{i: \mu_i < k} (k - \mu_i) = \dim \ker^L M_{ek} = \dim \ker^L R_k$.

Proof. The rank formula (2.9) for the generalized Sylvester resultant first appeared in [5], while the corresponding result for the generalized Wolovich resultant has been established in [1]. Furthermore, the fact that $\text{rank} R_k = \text{rank} M_{ek}$ becomes obvious in view of equation (2.8). ■

3. Computation of Minimal Polynomial Bases

It is evident from the last result of the above section that the orders of the left minimal indices of a polynomial matrix $F(s)$ are closely related to the structure of the generalized Sylvester or Wolovich resultants. Furthermore, the following result shows the connection between the coefficients of a minimal polynomial basis of the left kernel of $F(s)$ and a basis of the left kernel of either R_k or M_{ek} .

Theorem 3.1. *Let $E(s)$ be a minimal polynomial basis for the left kernel of $F(s)$ as in (2.1). Let $\mu_i = \deg_{r_i} E(s)$, $i \in \mathbf{p}$ be the invariant minimal row indices of the left kernel of $F(s)$, and denote by a_k the number of rows of $E(s)$ with $\mu_i = k$. Then*

$$\ker^L R_k = \ker^L M_{ek} = \text{Im}^L L_k, \quad (3.1)$$

where $L_k \in \mathbb{R}^{v_k \times k(p+m)}$, is defined by

$$\text{block diag}\{S_{1,k-\mu_i}(s)\}_{i:\mu_i < k} E_k(s) = L_k S_{p+m,k}(s), \quad (3.2)$$

and $E_k(s) \in \mathbb{R}^{\gamma_k \times (p+m)}$ is a polynomial matrix that consists of all $\gamma_k = \sum_{i=0}^{k-1} a_i$ rows of $E(s)$ with row degrees satisfying $\mu_i < k$.

Proof. Since $E_k(s)$ consists of rows of $E(s)$ satisfying $\mu_i = \deg_{r_i} E(s) < k$, in view of (2.1) it is easy to see that

$$E_k(s)F(s) = 0, \quad (3.3)$$

for every $s \in \mathbb{C}$. Postmultiplying (3.2) by $F(s)$ and using (3.3) gives

$$L_k X_k(s) = 0,$$

with $X_k(s)$ defined in (2.3). Now using respectively (2.4) and (2.6), we get

$$L_k R_k S_{m,q+k}(s) = 0 \text{ and } L_k M_{ek} \text{block diag}\{S_{1,\nu_i+k}(s)\}_{i \in \mathbf{m}} = 0,$$

for every $s \in \mathbb{C}$. Thus

$$L_k R_k = 0 \text{ and } L_k M_{ek} = 0,$$

which proves that $\text{Im}^L L_k \subset \ker^L R_k$ and $\text{Im}^L L_k \subset \ker^L M_{ek}$. Furthermore it is easy to see that L_k has full row rank since the existence of a (constant) row vector $\bar{w}^\top \in \mathbb{R}^{1 \times v_k}$ s.t. $\bar{w}^\top L_k = 0$, would imply (via eq. 3.2) existence of a polynomial vector $w^\top(s) \in \mathbb{R}^{1 \times \gamma_k}[s]$ satisfying $w^\top(s)E_k(s) = 0$, which contradicts the fact that $E(s)$ consists of linearly independent polynomial row vectors. Thus

$$\dim \text{Im}^L L_k = \sum_{i:\mu_i < k} (k - \mu_i) = \dim \ker^L M_{ek} = \dim \ker^L R_k,$$

which completes the proof. ■

Our aim is to propose a method for the determination of a minimal polynomial basis for the left kernel of $F(s)$. As it will be shown in the sequel this can be done via numerical computations on successive generalized Sylvester or Wolovich resultants of the polynomial matrix $F(s)$. The key idea is that if we already know a part of the minimal polynomial basis $E(s)$ of the left kernel of $F(s)$, corresponding to rows with row degrees less than k , then we can easily determine linearly independent polynomial row vectors with degree exactly equal to k , that belong to the left kernel of $F(s)$.

Recall that $E_k(s) \in \mathbb{R}^{\gamma_k \times (p+m)}[s]$ is the matrix defined in theorem 3.1, i.e. it is a part of the minimal polynomial basis $E(s)$ of the left kernel of $F(s)$ that contains only those rows of $E(s)$ with $\mu_i = \deg_{r_i} E(s) < k$. For $k = 1, 2, 3, \dots$ we define the sequence of rational vector spaces

$$\mathcal{F}_k = \text{Im}_{\mathbb{R}(s)}^L E_k(s). \quad (3.4)$$

It is easy to see that

$$\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \dots \subseteq \mathcal{F}_{\mu+1} = \ker_{\mathbb{R}(s)}^L F(s), \quad (3.5)$$

where $\mu = \max_{i \in \mathbf{p}} \{\mu_i\}$, while obviously

$$\dim_{\mathbb{R}(s)} \mathcal{F}_k = \gamma_k. \quad (3.6)$$

Theorem 3.2. Let $E_k(s)$ be a minimal polynomial basis of \mathcal{F}_k . Define the $(v_k + \gamma_k) \times (p + m)(k + 1)$ matrix \bar{L}_{k+1} from the relation

$$\text{block diag}\{S_{1,k-\mu_i+1}(s)\}_{i \in \gamma_k} E_k(s) = \bar{L}_{k+1} S_{p+m,k+1}(s), \quad (3.7)$$

and let $\bar{N}_{k+1} \in \mathbb{R}^{a_k \times (p+m)(k+1)}$ be such that the $(a_k + v_k + \gamma_k) \times (p + m)(k + 1)$ compound matrix $\tilde{L}_{k+1} := [\bar{L}_{k+1}^\top, \bar{N}_{k+1}^\top]^\top$ satisfies

$$\text{rank} \tilde{L}_{k+1} = v_{k+1} \text{ and } \tilde{L}_{k+1} M_{e(k+1)} = 0, \quad (3.8)$$

i.e. such that \tilde{L}_{k+1} is a basis of $\ker^L M_{e(k+1)}$. Then the rows of the polynomial matrix¹

$$\tilde{E}_{k+1}(s) := \begin{bmatrix} E_k(s) \\ N_{k+1}(s) \end{bmatrix},$$

form a minimal polynomial basis of \mathcal{F}_{k+1} where $N_{k+1}(s) := \bar{N}_{k+1} S_{p+m,k+1}(s) \in \mathbb{R}^{a_k \times (p+m)}[s]$.

Proof. Postmultiplying (3.7) by $F(s)$ and taking into account (2.6) and the fact that $E_k(s)F(s) = 0$ for every $s \in \mathbb{C}$, it is easily seen that

$$\bar{L}_{k+1} M_{e(k+1)} = 0, \quad (3.9)$$

while the rows of \bar{L}_{k+1} are linearly independent. We seek to find linearly independent row vectors that together with the rows of \bar{L}_{k+1} form a basis of $\ker^L M_{e(k+1)}$. According to theorem 2.1 $\dim \ker^L M_{e(k+1)} = v_{k+1}$ which compared to the number of rows of \bar{L}_{k+1} shows that we need another a_k linearly independent vectors to form a complete basis of $\ker^L M_{e(k+1)}$. Assume we determine a $a_k \times (p + m)(k + 1)$ full row rank matrix \bar{N}_{k+1} such that

$$\bar{N}_{k+1} M_{e(k+1)} = 0, \quad (3.10)$$

with rows linearly independent to those of \bar{L}_{k+1} , i.e. such that

$$\text{rank} \tilde{L}_{k+1} = v_{k+1}. \quad (3.11)$$

Obviously the rows of the compound matrix in the above equation form a basis for the left kernel of $M_{e(k+1)}$. It is easy to verify that the rows of the polynomial matrix $N_{k+1}(s)$ will satisfy

$$N_{k+1}(s)F(s) = 0.$$

Furthermore the polynomial rows of $N_{k+1}(s)$ will have degrees exactly k , since if there exists a row of $N_{k+1}(s)$ with $\deg_{s r_i} N_{k+1}(s) < k$, the corresponding row of \bar{N}_{k+1} would be a linear combination of the rows of \bar{L}_{k+1} , which contradicts (3.11).

It is easy to see that there exists a row permutation matrix P such that

$$P \bar{L}_{k+1} = \begin{bmatrix} L_k & 0 \\ X_k & E_k^{hr} \end{bmatrix}, \quad (3.12)$$

¹If $a_k = 0$ then obviously $\tilde{E}_{k+1}(s) = E_k(s)$.

where $L_k \in \mathbb{R}^{v_k \times k(p+m)}$ is a basis of $\ker^L M_{e_k}$ as defined in (3.2), X_k is a constant matrix, and E_k^{hr} is the highest row coefficient matrix of $E_k(s)$. Accordingly partition \bar{N}_{k+1} as follows

$$\bar{N}_{k+1} = [Y_k, N_{k+1}^{hr}],$$

where $Y_k \in \mathbb{R}^{a_k \times (p+m)k}$ and $N_{k+1}^{hr} \in \mathbb{R}^{a_k \times (p+m)k}$ is the highest row coefficient matrix of $N_{k+1}(s)$. We shall prove that $\tilde{E}_{k+1}(s)$ is row proper or equivalently that the highest row degree coefficient matrix of $\tilde{E}_{k+1}(s)$ has full row rank. Obviously

$$\tilde{E}_{k+1}^{hr} = \begin{bmatrix} E_k^{hr} \\ N_{k+1}^{hr} \end{bmatrix}.$$

Assume that $\tilde{E}_{k+1}(s)$ is not row proper. Then there exists a row vector $[a^\top, b^\top]$ such that

$$[a^\top, b^\top] \begin{bmatrix} E_k^{hr} \\ N_{k+1}^{hr} \end{bmatrix} = 0. \quad (3.13)$$

Combining equations (3.9), (3.10) and (3.12) we obtain

$$\begin{bmatrix} L_k & 0 \\ X_k & E_k^{hr} \\ Y_k & N_{k+1}^{hr} \end{bmatrix} M_{e(k+1)} = 0,$$

while premultiplying the above equation by $[0, a^\top, b^\top]$, with a^\top, b^\top chosen as in (3.13) we get

$$[a^\top, b^\top] \begin{bmatrix} X_k \\ Y_k \end{bmatrix} M_{e_k} = 0.$$

Now since L_k is a basis of the left kernel of M_{e_k} there exists a row vector c^\top such that

$$[a^\top, b^\top] \begin{bmatrix} X_k \\ Y_k \end{bmatrix} = c^\top L_k.$$

It is easy to verify that

$$[-c^\top, a^\top, b^\top] \begin{bmatrix} P & 0 \\ 0 & I_{a_k} \end{bmatrix} \begin{bmatrix} \bar{L}_{k+1} \\ \bar{N}_{k+1} \end{bmatrix} = 0,$$

which contradicts (3.11). Thus $\tilde{E}_{k+1}(s)$ is row proper and thus has full row rank over $\mathbb{R}(s)$. Hence the rows of $\tilde{E}_{k+1}(s)$ form a basis of the rational vector space \mathcal{F}_{k+1} . Furthermore $\tilde{E}_{k+1}(s)$, as row proper, has no zeros at $s = \infty$ [21] (Corollary 3.100, page 144). It remains to show that $\tilde{E}_{k+1}(s)$ has no finite zeros. Consider the rational vector space \mathcal{F}_{k+1} and its minimal polynomial basis formed by the rows of $\tilde{E}_{k+1}(s)$. The row orders $\mu_i = \deg_{r_i} \tilde{E}_{k+1}(s)$ are the minimal invariant indices of \mathcal{F}_{k+1} and denote by $ord\mathcal{F}_{k+1}$ the (Forney invariant) minimal order of \mathcal{F}_{k+1} , which in our case is

$$ord\mathcal{F}_{k+1} = \sum_{i:\mu_i < k+1} \mu_i.$$

The rows of $\tilde{E}_{k+1}(s)$ span also the rational vector space \mathcal{F}_{k+1} . It is known ([21], page 137) that if $\mathcal{Z}\{\tilde{E}_{k+1}(s)\}$ is the total number of (finite and infinite) zeros, $\delta_M\{\tilde{E}_{k+1}(s)\}$ is the McMillan degree of $\tilde{E}_{k+1}(s)$, and $ord\mathcal{F}_{k+1}$ is the (Forney invariant [8]) minimal order of the rational vector space spanned by the rows of $\tilde{E}_{k+1}(s)$ then

$$\delta_M\{\tilde{E}_{k+1}(s)\} = \mathcal{Z}\{\tilde{E}_{k+1}(s)\} + ord\mathcal{F}_{k+1},$$

but $\tilde{E}_{k+1}(s)$ is row proper and thus its McMillan degree is equal to the sum of its row indices, which by construction coincides with $\sum_{i:\mu_i < k+1} \mu_i = ord\mathcal{F}_{k+1}$. Thus $\mathcal{Z}\{\tilde{E}_{k+1}(s)\} = 0$ which establishes the fact that $\tilde{E}_{k+1}(s)$ has no finite zeros. Thus the polynomial matrix $\tilde{E}_{k+1}(s)$ is a row proper and left unimodular, i.e. a minimal polynomial basis of \mathcal{F}_{k+1} . ■

The above theorem essentially allows us to determine successively a minimal polynomial basis of $\ker_{\mathbb{R}(s)}^L F(s)$. Starting with $k = 0$ one can determine a minimal polynomial basis of \mathcal{F}_1 , i.e. the part of the minimal polynomial basis of $\ker_{\mathbb{R}(s)}^L F(s)$ with row indices $\mu_i = 0$. Using this part of the polynomial basis and applying again the procedure of theorem 3.2 for $k = 1$, we determine a minimal polynomial basis of \mathcal{F}_2 . The entire procedure can be repeated until we have a minimal polynomial basis consisting of p row vectors.

In order to obtain numerically stable results one can use singular value decomposition to obtain orthonormal bases of the kernels of constant matrices involved. Furthermore, the rows of \bar{N}_{k+1} can be chosen not only to be linearly independent to those of \bar{L}_{k+1} , but orthogonal to each one of them. This can be done by computing an orthonormal basis of the left kernel of $[M_{e(k+1)}, \bar{L}_{k+1}^\top]$. The coefficients of a minimal polynomial basis computed this way will form a set of orthonormal vectors, i.e. $E_{k-1}E_{k-1}^\top = I_p$.

The entire procedure can be summarized in the following algorithm:

- **Step 1.** Compute an orthonormal basis \bar{N}_1 of $\ker^L M_{e1}$, and set $E_1 = \bar{N}_1$
- **Step 2.** Set $k = 2$
- **Step 3.** Using (3.7) compute \bar{L}_k for $E_{k-1}(s) = E_{k-1}S_{p+m,k}$
- **Step 4.** Determine an orthonormal basis \bar{N}_k of $\ker^L [M_{ek}, \bar{L}_k^\top]$ and set $E_k = \begin{bmatrix} E_{k-1} & | & 0 \\ \bar{N}_k & & \end{bmatrix}$
- **Step 5.** Set $k = k + 1$
- **Step 6.** If $\{\# \text{ of rows } E_{k-1}\} < p$ go to Step 3
- **Step 7.** The minimal polynomial basis is given by $E_{k-1}S_{p+m,k-1}(s)$

Notice that the above procedure can be applied even if the matrix $F(s)$ has not full column rank over $\mathbb{R}(s)$. Assuming that $rank_{\mathbb{R}(s)} F(s) = r < m$, we can modify step 6 so that the loop stops if $\{\# \text{ of rows } E_{k-1}\} = p + m - r$, since obviously $p + m - r$ is the dimension of the left kernel of $F(s)$. In case r is unknown, we can still use the proposed algorithm by leaving the loop running until k reaches mq , since mq is known to be the upper bound for the maximal left minimal index μ , but with a significant overhead in computational cost (see next section for more details).

Obviously the proposed algorithm can be easily modified to compute right minimal polynomial bases, by simply transposing the polynomial matrix whose right null space is to be determined. Finally, notice that throughout the above analysis we have used the generalized Wolovich resultant because in general it has less columns than the corresponding generalized Sylvester resultant (see (2.8)). However, the left null space structure of both resultants is identical and the proposed algorithm can be implemented using either.

4. Numerical Considerations

The proposed algorithm requires successive determination of orthonormal bases of left kernels of the matrices $[M_{ek}, \bar{L}_k^\top]$ for each $k = 1, 2, 3, \dots$. The most reliable method to obtain orthonormal bases of null spaces is undoubtedly singular value decomposition (SVD) (see for instance [9]). Thus the computational complexity at each step of the algorithm is about $O(n^3 k^3)$ (where for ease of notation we use $n := p + m$ for $F(s) \in \mathbb{R}^{(p+m) \times m}[s]$). Applying standard SVD implementations (such as Golub-Reinsch SVD or R-SVD) at each step would result in a relatively high computational cost, since the SVD computed in step k cannot be reused for the next iteration. However, recently a fast and backward stable algorithm for updating the SVD when rows or columns are appended to a matrix, appeared in [10]. The cost of each update is quadratic to matrix dimensions. Applying this technique at each step, could effectively reduce the total cost of our algorithm up to the k th step, to $O(n^3 k^3)$. Taking into account that the iterations will continue until $k = \mu + 1$, where μ is the maximal left minimal index of $F(s)$, we can conclude that the cost of the proposed algorithm is about $O(\mu^3 n^3)$.

Comparing this to the complexity of the algorithm in [3] which is about $O(q^3 n^3)$, where q is the maximum degree of s in $F(s)$, we can see that our implementation can be more efficient if $\mu < q$. On the other hand the upper bound for μ is mq , so the complexity of our algorithm can get as high as $O(m^3 q^3 n^3)$. However, this upper bound will only be reached in extreme cases where $F(s)$ has only one left minimal index of order greater than zero and no finite and infinite elementary divisors (see example 5.3). In general when $F(s)$ has finite zeros (elementary divisors) and/or is not column reduced or even better if $p \approx m$, it is expected that $\mu \leq q$. Still, in bad cases where generalized resultants tend to become very large, one may employ sparse or structured matrix techniques to improve the efficiency of the algorithm.

From a numerical stability point of view, each step of the algorithm is stable since it depends on SVD computations. A complete stability analysis of the entire algorithm is hard to be accomplished, since there is not a standard way to define small perturbations for polynomial matrices. However, there are some points in the procedure that are worth mentioning. The actual procedure of computing an orthonormal basis of the left kernel of $Q_k := [M_{ek}, \bar{L}_k^\top]$, involves the computation of the singular value decomposition $Q_k = U_k^\top \Sigma_k V_k$, where U_k, V_k are orthogonal and $\Sigma_k = \text{diag}\{\sigma_i(Q_k), 0\}$ and $\sigma_i(Q_k)$ are the singular values of Q_k . In the sequel the rank of Q_k is determined by choosing r_k such that $\sigma_{r_k}(Q_k) \geq \delta_k > \sigma_{r_k+1}(Q_k)$, where $\delta_k = \|Q_k\|_\infty u$ and $u > 0$ is a (small) number such that δ_k is consistent with the machine precision ([9]). The basis of the left kernel of Q_k is then given by the rows of U_k corresponding to singular values smaller than δ_k . It is easy to see that if \bar{N}_k (using the notation in step 4 of the algorithm) is such a basis then $\|\bar{N}_k M_{ek}\|_2 < \delta_k$. The product $\bar{N}_k M_{ek}$ gives the coefficients of the multiplication of the newly computed rows of the minimal polynomial basis, by $F(s)$, so it is important to keep $\|\bar{N}_k M_{ek}\|_2$ small relatively to the magnitude of M_{ek} . On the other hand it is

easy to see that $\|M_{ek}\|_\infty = \|M_{e1}\|_\infty$, while due to the special structure of \bar{L}_k it can be seen that $\|\bar{L}_k^\top\|_\infty \leq 2p$. To avoid problematic situations, where for example $\|M_{ek}\|_\infty \ll \|Q_k\|_\infty$ which may lead to erroneous computation of r_k , it is necessary to scale M_{ek} in order to "balance" the components of Q_k . Experimental results show that a good practice is to normalize $F(s)$ using $\|M_{e1}\|_\infty$, i.e. setting $\bar{F}(s) = F(s)/\|M_{e1}\|_\infty$. In such a case a quick calculation yields

$$\|\bar{N}_k M_{ek}\|_2 < u \|M_{e1}\|_\infty (2p + 1),$$

i.e. that the coefficients of the product $E(s)F(s)$ will be of magnitude about u times the magnitude of the coefficients of $F(s)$, which is close to zero compared to the magnitude of $F(s)$.

5. Examples

The examples bellow have been computed on an PC, with relative machine precision $EPS = 2^{-52} \simeq 2.22045 * 10^{-16}$.

Example 5.1. Consider the example 5.2 in [3]. Given then transfer function $P(s) = D_L^{-1}(s)N_L(s)$, where

$$D_L(s) = (s + 2)^2(s + 3) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, N_L(s) = \begin{bmatrix} 3s + 8 & 2s^2 + 6s + 2 \\ s^2 + 6s + 2 & 2s^2 + 7s + 8 \end{bmatrix}.$$

We construct the compound matrix $F(s) = [D_L(s), -N_L(s)]^\top$ and compute a minimal basis of the left kernel of $F(s)$. We first notice that $\|M_{e1}\|_\infty = 36$ and normalize $F(s)$ by setting $\bar{F}(s) = F(s)/\|M_{e1}\|_\infty$. Due to lack of space we use less decimal digits in intermediate results but the actual computations carried out with the above mentioned relative machine precision. The normalized form of $F(s)$ is

$$\bar{F}(s) = \begin{bmatrix} 0.028s^3 + 0.194s^2 + 0.444s + 0.333 & 0. \\ 0. & 0.028s^3 + 0.194s^2 + 0.444s + 0.333 \\ -0.0833s - 0.222 & -0.028s^2 - 0.167s - 0.056 \\ -0.056s^2 - 0.167s - 0.056 & -0.083s^2 - 0.194s - 0.222 \end{bmatrix}.$$

For $k = 1$ we compute

$$M_{e1} = \begin{bmatrix} 0.333 & 0.444 & 0.194 & 0.028 & 0. & 0 & 0 & 0 & 0 \\ 0. & 0 & 0 & 0 & 0.333 & 0.444 & 0.194 & 0.028 & 0 \\ -0.222 & -0.083 & 0 & 0 & -0.056 & -0.167 & -0.028 & 0 & 0 \\ -0.056 & -0.167 & -0.056 & 0 & -0.222 & -0.194 & -0.083 & 0 & 0 \end{bmatrix},$$

and calculate its left kernel, which in this case is empty. Thus $E_1 = \emptyset$. For $k = 2$ we compute

$$M_{e2} = \begin{bmatrix} 0.333 & 0.444 & 0.194 & 0.028 & 0 & 0. & 0 & 0 & 0 & 0 \\ 0. & 0 & 0 & 0 & 0 & 0.333 & 0.444 & 0.194 & 0.028 & 0 \\ -0.222 & -0.083 & 0 & 0 & 0 & -0.056 & -0.167 & -0.0278 & 0 & 0 \\ -0.056 & -0.167 & -0.056 & 0 & 0 & -0.222 & -0.194 & -0.083 & 0 & 0 \\ 0 & 0.333 & 0.444 & 0.194 & 0.028 & 0 & 0. & 0 & 0 & 0 \\ 0 & 0. & 0 & 0 & 0 & 0 & 0.333 & 0.444 & 0.194 & 0.028 \\ 0 & -0.222 & -0.083 & 0 & 0 & 0 & -0.056 & -0.167 & -0.028 & 0 \\ 0 & -0.056 & -0.167 & -0.056 & 0 & 0 & -0.222 & -0.194 & -0.083 & 0 \end{bmatrix}.$$

Since $E_1 = \emptyset$, $\bar{L}_2 = \emptyset$ so we have to compute the left kernel of M_{e_2}

$$\bar{N}_2 = \begin{bmatrix} -0.343 & -0.514 & -0.343 & -0.686 & 4.227 \cdot 10^{-16} & -2.516 \cdot 10^{-16} & -7.473 \cdot 10^{-16} & -0.171 \end{bmatrix},$$

and set $E_2 = \bar{N}_2$. We proceed for $k = 3$ and compute M_{e_3} and \bar{L}_3

$$M_{e_3} = \begin{bmatrix} 0.33 & 0.44 & 0.19 & 0.03 & 0 & 0 & 0. & 0 & 0 & 0 & 0 & 0 \\ 0. & 0 & 0 & 0 & 0 & 0 & 0.33 & 0.44 & 0.19 & 0.03 & 0 & 0 \\ -0.22 & -0.08 & 0 & 0 & 0 & 0 & -0.06 & -0.17 & -0.03 & 0 & 0 & 0 \\ -0.06 & -0.17 & -0.06 & 0 & 0 & 0 & -0.22 & -0.19 & -0.08 & 0 & 0 & 0 \\ 0 & 0.33 & 0.44 & 0.19 & 0.03 & 0 & 0 & 0. & 0 & 0 & 0 & 0 \\ 0 & 0. & 0 & 0 & 0 & 0 & 0 & 0.33 & 0.44 & 0.19 & 0.03 & 0 \\ 0 & -0.22 & -0.08 & 0 & 0 & 0 & 0 & -0.06 & -0.17 & -0.03 & 0 & 0 \\ 0 & -0.06 & -0.17 & -0.06 & 0 & 0 & 0 & -0.22 & -0.19 & -0.08 & 0 & 0 \\ 0 & 0 & 0.33 & 0.44 & 0.19 & 0.03 & 0 & 0 & 0. & 0 & 0 & 0 \\ 0 & 0 & 0. & 0 & 0 & 0 & 0 & 0 & 0.33 & 0.44 & 0.19 & 0.03 \\ 0 & 0 & -0.22 & -0.08 & 0 & 0 & 0 & 0 & -0.06 & -0.17 & -0.03 & 0 \\ 0 & 0 & -0.06 & -0.17 & -0.06 & 0 & 0 & 0 & -0.22 & -0.19 & -0.08 & 0 \end{bmatrix},$$

$$\bar{L}_3 = \begin{bmatrix} -0.343 & -0.514 & -0.343 & -0.686 & 4.227 \cdot 10^{-16} & -2.516 \cdot 10^{-16} \\ 0 & 0 & 0 & 0 & -0.343 & -0.514 \\ -7.473 \cdot 10^{-16} & -0.171 & 0 & 0 & 0 & 0 \\ -0.343 & -0.686 & 4.227 \cdot 10^{-16} & -2.516 \cdot 10^{-16} & -7.473 \cdot 10^{-16} & -0.171 \end{bmatrix}.$$

Next we compute the left null space of $[M_{e_3}, \bar{L}_3^\top]$ which gives

$$\bar{N}_3 = \begin{bmatrix} 0.312 & -0.092 & 0.535 & -0.272 & -0.076 & 0.110 \\ 0.595 & -0.332 & -6.461 \cdot 10^{-17} & 1.717 \cdot 10^{-16} & 0.224 & -0.038 \end{bmatrix},$$

and set

$$E_3 = \begin{bmatrix} -0.343 & -0.514 & -0.343 & -0.686 & 4.227 \cdot 10^{-16} & -2.516 \cdot 10^{-16} \\ 0.312 & -0.092 & 0.535 & -0.272 & -0.076 & 0.110 \\ -7.473 \cdot 10^{-16} & -0.171 & 0 & 0 & 0 & 0 \\ 0.595 & -0.332 & -6.461 \cdot 10^{-17} & 1.717 \cdot 10^{-16} & 0.224 & -0.038 \end{bmatrix}.$$

We have $p = 2$ rows in E_3 so the loop stops. The (transposed) computed minimal basis of the left kernel of $F(s)$ is then computed by setting $E(s) = E_3 S_{4,3}(s)$, i.e.

$$E^\top(s) = \begin{bmatrix} 8.90294 \times 10^{-17} s - 0.342997 & -2.46895 \times 10^{-16} s^2 - 0.0761616s + 0.311713 \\ 2.32405 \times 10^{-16} s - 0.514496 & -3.18563 \times 10^{-16} s^2 + 0.109531s - 0.091865 \\ -7.15439 \times 10^{-18} s - 0.342997 & 0.223774s^2 + 0.59516s + 0.535487 \\ -0.171499s - 0.685994 & -0.0380808s^2 - 0.332127s - 0.271669 \end{bmatrix},$$

whose (row) partitioning gives the coprime factorization $P(s) = N_R(s)D_R^{-1}(s)$. Notice that $\mu = 2$ which is less than the degree of $F(s)$, $q = 3$.

Example 5.2. Consider the example 5.1 in [3]. Given then transfer function $P(s) = N_R(s)D_R^{-1}(s)$, where

$$N_R(s) = \begin{bmatrix} s^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & s & 0 \\ 0 & 0 & 0 & s \end{bmatrix}, D_R(s) = \begin{bmatrix} 1-s & 0 & 0 & 0 \\ 0 & 1-s & 0 & 0 \\ 0 & -s & 1-s & 0 \\ 0 & 0 & 0 & 1-s \end{bmatrix},$$

and construct the compound matrix $F(s) = [N_R^\top(s), -D_R^\top(s)]^\top$. The minimal basis of the left kernel of $F(s)$ is then given by our algorithm

$$E(s) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & a(s-1) & 0 & 0 & 0 & 0 & -as \\ a(s-1) & 0 & 0 & 0 & 0 & -as^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -bs^2 + cs - b & 0 & 0 & -bs^2 & b(s^2 - s) & 0 & 0 \end{bmatrix},$$

where $a \simeq 0.57735$, $b \simeq 0.333333$ and $c \simeq 0.666667$. The left coprime fractional representation of $P(s) = N_L(s)D_L^{-1}(s)$ can be obtained by appropriately partitioning $E(s)$. Notice that $\mu = 2$ which is equal to $q = 2$.

Example 5.3. Consider the matrix

$$F(s) = \begin{bmatrix} -1 & 0 & \cdots & 0 \\ s^q & -1 & \ddots & \vdots \\ 0 & s^q & \ddots & 0 \\ \vdots & \ddots & \ddots & -1 \\ 0 & \cdots & 0 & s^q \end{bmatrix} \in \mathbb{R}^{(m+1) \times m}[s],$$

the minimal polynomial basis for the left kernel of $F(s)$ as computed by the algorithm is

$$E(s) = a[s^{qm}, s^{q(m-1)}, \dots, s^q, 1] \in \mathbb{R}^{1 \times (m+1)}[s],$$

where $a = \frac{1}{\sqrt{m+1}}$. Obviously $\mu = qm$, which is the worst case from a performance point of view. Notice the absence of finite and infinite elementary divisors in $F(s)$ and the fact that $\dim \ker_{\mathbb{R}(s)}^L F(s) = 1$, thus qm is "consumed" in just one left minimal index.

6. Conclusions

In this note we have proposed a resultant based method for the computation of minimal polynomial bases of a polynomial matrix. The algorithm utilizes the left null space structure of successive generalized Wolovich or Sylvester resultants of a polynomial matrix to obtain the coefficients of the minimal polynomial basis of the left kernel of the given polynomial matrix. The entire computation can be accomplished using only orthogonal decompositions and the coefficients of the resulting minimal polynomial basis have the appealing property of being

orthonormal. From a performance point of view our procedure requires about $O(n^3\mu^3)$ floating point operations which is comparable to other approaches ([3], [19]) since in most cases it is expected that the order of the maximal left minimal index μ is close to the degree q of the polynomial matrix itself.

Further research on the subject could address more specific problems like the computation of row or column reduced polynomial matrices using an approach similar to [4] (and the improved version of [18]) or the determination of rank, left minimal indices and greatest common divisors of polynomial matrices.

A test version of the algorithm has been implemented in MathematicaTM 4.2 and is available upon request to anyone interested.

References

- [1] Antoniou E.N., Vardoulakis A.I.G. (2003), *A numerical method for the computation of proper denominator assigning compensators for strictly proper plants*, Proc. of the 11th IEEE Med. Conference on Control and Automation, June 2003, Rhodes, Greece.
- [2] Basilio J.C., Kouvaritakis B., (1997), *An Algorithm for Coprime Matrix Fraction Description using Sylvester Matrices*, Linear Algebra and its Apps, Vol 266, pp. 107-125.
- [3] Beelen T.G., Veltkamp G.W. (1987), *Numerical computation of a coprime factorization of a transfer function matrix*, Systems & Control Letters, Vol. 9, pp. 281-288.
- [4] Beelen T.G., Van Den Hurk G.J., Praagman C., (1988), *A new method for computing a column reduced polynomial matrix*, Systems & Control Letters, Vol. 10, pp. 217-224.
- [5] Bitmead R., Kung S, Anderson B., Kailath T. (1978), *Greatest common divisors via generalized Sylvester and Bezout matrices*, IEEE Trans. Autom. Control, Vol AC-23, No 6, pp. 1043-1047.
- [6] Callier, F.M. and Desoer, C.A. (1982). *Multivariable Feedback Systems*. Springer Verlag. New York.
- [7] Callier, F.M., (2001), *Polynomial equations giving a proper feedback compensator for a strictly proper plant*, Proceedings of the IFAC/IEEE Symposium on System structure and Control, Prague, Czech Republic.
- [8] Forney G.D. (1975), *Minimal bases of rational vector spaces with application to multivariable linear systems*, SIAM J. Contr., Vol. 13, pp. 293-520.
- [9] Golub G., Van Loan C., (1996), *Matrix Computations*, Third Edition, The John Hopkins University Press,
- [10] Gu M., Eisenstat S.C., (1994), *A Stable and Fast Algorithm for Updating the Singular Value Decomposition*, Research Report YALEU/DCS/RR-966, Yale University, New Haven, CT.
- [11] Kailath T., (1980), *Linear Systems*, Englewood Cliffs, NJ: Prentice Hall.

- [12] Karcianas N., (1994), *Minimal Bases of Matrix Pencils: Algebraic Toeplitz Structure and Geometric Properties*, Linear Algebra and its Apps, Vol 205-206, pp. 831-868.
- [13] Karcianas N., Mitrouli M., (2002), *Minimal Bases of Matrix Pencils and Coprime Fraction Descriptions*, IMA Journal of Control and Inf., 19, pp. 245-278.
- [14] Khazanov V. B. , *On Some Properties of the Minimal-in-Degree Irreducible Factorizations of a Rational Matrix*, Journal of Mathematical Sciences, Volume 114, Issue 6, Apr 2003, pp. 1860 - 1862
- [15] Kucera V., (1970), *Discrete Linear Control: The polynomial Equation Approach*, Willey, Cichester, U.K.
- [16] Kucera V., Zagalak P., (1999), *Proper Solutions of polynomial equations*, Proc. of the 14th World Congress of IFAC, London, Elsevier Science, pp. 357-362
- [17] Kublanovskaya V. N., *Rank division algorithms and their applications*, J. Numer. Linear Algebra Appl. 1 (1992), No. 2, pp. 199–213.
- [18] Neven W., Praagman C., (1993), *Column reduction of polynomial matrices*, Linear Algebra and its Apps, Vol 188-189, pp. 569-589.
- [19] Stuchlik-Quere M.P., (1997), *How to compute minimal bases using Pade approximants*, Laboratoire d'Informatique de Paris, Research Report, LIP6 1997/035.
- [20] Vafiadis D., Karcianas N. (1995), *Generalized state-space realizations from matrix fraction descriptions*, IEEE Trans. Automat. Control, Vol. 40 , No. 6, pp. 1134–1137.
- [21] Vardulakis, A.I.G., (1991), *Linear Multivariable Control - Algebraic Analysis and Synthesis Methods*, John Willey & Sons Ltd, New York.
- [22] Wolovich, W.A. (1974), *Linear Multivariable Systems*. Springer Verlag. New York.
- [23] Zhang S.-Y., Chen C.-T. (1983), *Design of Unity Feedback Systems to Achieve Arbitrary Denominator Matrix*, IEEE Trans. Autom. Control, Vol AC-28, No 4, pp. 518-521.