Generalized Bezoutian and Sylvester Matrices in Multivariable Linear Control

BRIAN D.O. ANDERSON, FELLOW, IEEE, AND
E.I. JURY, FELLOW, IEEE

Abstract—**Generalized Bezoutian and Sylvester matrices are defined and discussed in this short paper. The relationship between these two forms of matrices is established. It is shown that the McMillan degree of a real rational function can be ascertained by checking the rank of either one of these generalized matrices formed using a polynomial matrix fraction decomposition of the prescribed transfer function matrix. Earlier established results by Rowe and Munro are obtained as a special case. Several theorems related to the rank testing and other properties of the generalized matrices are discussed and various research problems are listed in the conclusion.**

I. Introduction

Recent work in linear multivariable control has started to draw heavily again on frequency domain ideas. The book of Rosenbrock [1] has been an important stimulus; more recently, the book of Wolovich [2] has appeared. This short paper studies problems which arise when one generalizes to the multivariable case the notion of transfer function descriptions in which the numerator and denominator have a common factor; such problems are described in [1] and [2].

Our approach is to work with generalizations of two classical tools used in studying the question of whether a prescribed pair of polynomials is relatively prime. These tools are Bezoutian and Sylvester resultant matrices. Our main results involve the relating of the McMillan degree of a real rational transfer function matrix to the rank of either a generalized Bezoutian or Sylvester matrix. Though such forms of generalized Sylvester resultants [3], [4] and Bezoutians [5] have appeared in the literature, we believe our formulations are more general, with the earlier established results obtainable, as shown in the paper, as special cases. Of course, standard Sylvester and Bezoutian matrices have been applied to a number of linear system theory problems, especially some involving stability, see, e.g., the recent survey [6] and text [7]; the novelty here lies in the generalization to the multivariable situation.

This short paper is structured as follows. In Section II, we introduce the notions of generalized Bezoutian forms and matrices, and, with the aid of a formula connecting the generalized Bezoutian matrix with a truncated Hankel matrix, establish the main result of the section linking the rank of a generalized Bezoutian matrix with the degree of the transfer function matrix with which it is associated.

In Section III, the generalized Sylvester matrix is introduced. A formula is presented, linking it with the generalized Bezoutian matrix of Section II, and allowing the derivation of a result linking the rank of a generalized Sylvester matrix with the degree of the associated transfer function matrix. The fact that the generalized Sylvester matrix is structured with a left triangle of zeros implies that algorithms for checking its rank may be computationally very attractive, as compared with those for checking the rank of a matrix (such as a generalized Bezoutian matrix) with no special distribution of zero elements, even though the latter may have lower dimension.

In Section IV, we note some important problems arising from the earlier contents of the short paper.

Manuscript received December 23, 1975; revised April 19, 1976. Paper recommended by J. B. Pearson, Chairman of the IEEE S-CS Linear Systems Committee. This work was supported by the Australian Research Grants Committee and the Joint Services Electronics Program under Contract F44620-71-C-0087.

B.D.O. Anderson is with the Department of Electrical Engineering, University of Newcastle, N.S.W., Australia.

E. I. Jury is with the Department of Electrical Engineering, University of Newcastle, N.S.W., Australia, on leave from the Department of Electrical Engineering and Computer Sciences and the Electronics Research Laboratory, University of California, Berkeley, CA 94720.

II. Generalized Bezoutian Matrices and Forms

Let $W(z)$ be a $q\times r$ matrix of real rational functions of z. Suppose that two matrix fraction decompositions of $W(z)$ are known:

$$W(z)=A^{-1}(z)B(z)=D(z)C^{-1}(z). \tag{1}$$

Then the *generalized Bezoutian form* associated with the quadruple $\{A,B,C,D\}$ is

$$\Gamma(x,y)=\frac{1}{x-y}[A(x)D(y)-B(x)C(y)]. \tag{2}$$

Remark 2.1: $\Gamma(x,y)$ is integral in x,y.

Remark 2.2: One could conceive of forming a generalized Bezoutian matrix from any quadruple $\{A,B,C,D\}$ of matrices, not necessarily linked by the relation (1). For $\Gamma(x,y)$ to be integral in x,y it is necessary and sufficient that $A(x)D(x)=B(x)C(x)$. However, the relevance of having A or C singular[1] is not fully understood. Henceforth, use of the notation $\{A,B,C,D\}$ will be understood to imply that (1) holds.

Remark 2.3: Let $V(z)=W(z)+J$, so that $\{A,B+AJ,C,D+JC\}$ defines a generalized Bezoutian form $\Gamma_V(x,y)$ associated with $V(z)$. Then it is easily checked that $\Gamma_V=\Gamma$. Similarly, for some constant F, let $U(z)=W(z)\ [I+FW(z)]^{-1}$. Then $\{A+BF,B,C+FD,D\}$ defines a generalized Bezoutian form $\Gamma_U(x,y)$ with $\Gamma_U=\Gamma$.

Remark 2.4: Suppose that $W(z)$ is invertible, so that $\{B,A,D,C\}$ defines a generalized Bezoutian form $\Gamma_{W^{-1}}(x,y)$ associated with $W^{-1}(z)$. Then it is easily checked that $\Gamma_{W^{-1}}=-\Gamma_W$.

There are clearly an infinite number of generalized Bezoutian forms associated with a prescribed $W(z)$. The following result provides some structure to the class of forms.

Proposition 2.1: For given $W(z)$, there exists a generalized Bezoutian form $\Gamma_1(x,y)$ such that all other generalized Bezoutian forms can be written as

$$\Gamma(x,y)=E(x)\Gamma_1(x,y)F(y) \tag{3}$$

for nonsingular polynomial E, F.

Proof: Let $\{A^1,B^1,C^1,D^1\}$ be such that (A^1,B^1) are relatively left prime, and (C^1,D^1) are relatively right prime. Given an arbitrary $\{A,B,C,D\}$, we then have $A=EA^1$, $B=EB^1$, $C=C^1F, D=D^1F$ for some nonsingular polynomials E, F; see [2]. Equation (3) follows immediately from (2).

Remark 2.5: $\Gamma_1(x,y)$ is unique up to left and right multiplication by unimodular $E(x)$ and $F(y)$, respectively. One could determine a canonical $\hat{\Gamma}_1(x,y)$ from $\Gamma_1(x,y)$, for example, by choosing $E(x)$ and $F(y)$ so that $E(x)\Gamma_1(x,0)$ and $\Gamma_1(0,y)F(y)$ were canonical (e.g., in Hermite form [8]).

Remark 2.6: The Bezoutian forms associated with $\{A^1,B^1,C^1,D^1\}$ and $\{A=EA^1,B=EB^1,C=C^1F,D=D^1F\}$ are related by (3) even if the relative primeness conditions do not hold.

Hitherto, no assumption has been made that $W(\infty)$ is finite. However, to keep further results tidy, we shall henceforth assume that $W(\infty)<\infty$, except where an explicit assumption to the contrary is made.

The following result is a minor extension of a result established in [2], the extension resting in the consideration of the case $W(\infty)$ nonzero. The proof will be omitted.

Lemma 2.1: Let $W(z)=D(z)C^{-1}(z)$, with $C(z)=C_0z^m+\cdots+C_m$, $C_0\neq 0$, and $D(z)=D_0z^l+\cdots+D_l$, $D_0\neq 0$. Then $W(\infty)<\infty$ implies $l\leqslant m$ and $W(\infty)=0$ implies $l<m$.

This lemma allows us to write

$$\Gamma(x,y)=\sum_{i=0}^{n}\sum_{j=0}^{m}\Gamma_{ij}x^{i-1}y^{j-1} \tag{4}$$

[1] A square polynomial matrix $A(z)$ is nonsingular if for almost all values of z it is nonsingular or, equivalently, if its determinant is not the zero polynomial. Otherwise it is singular.

where the highest degree in $A(z)$, $C(z)$ is, respectively, n, m; we then define the *generalized Bezoutian matrix* Δ as

$$\Delta=\begin{bmatrix}\Gamma_{11} & \Gamma_{12} & \cdots & \Gamma_{1m}\\ \Gamma_{21} & \Gamma_{22} & \cdots & \Gamma_{2m}\\ \vdots & \vdots & & \vdots\\ \Gamma_{n1} & \Gamma_{n2} & \cdots & \Gamma_{nm}\end{bmatrix}. \tag{5}$$

Suppose that $A(z)=A_0z^n+\cdots+A_n$, $B(z)=B_0z^n+\cdots+B_n$, $C(z)=C_0z^m+\cdots+C_m$, and $D(z)=D_0z^m+\cdots+D_m$. It is not hard to check that

$$\Gamma_{ij}=\sum_{k>0}(A_{n-i-k}D_{m-j+1+k}-B_{n-i-k}C_{m-j+1+k}) \tag{6a}$$

$$=\sum_{k>0}(B_{n-i+k+1}C_{m-j-k}-A_{n-i+k+1}D_{m-j-k}) \tag{6b}$$

where $A_p=0$ for p outside $[0,n]$, etc. Equations (6) probably provide the easiest method for computing entries of Δ. Note that it is not, in general, true that Δ is symmetric, or that $\Gamma_{ij}=\Gamma_{ji}$, in contrast to the situation where $W(z)$ is scalar.

As a parallel to Proposition 2.1, we have the following result.

Proposition 2.2: Let Δ_1 correspond to $\{A^1,B^1,C^1,D^1\}$ and Δ to $\{A,B,C,D\}$ where $A=EA^1, B=EB^1, C=C^1F, D=D^1F$ for nonsingular polynomial matrices $E(z)=E_0z^k+\cdots+E_k$ and $F(z)=F_0z^l+\cdots+F_l$. Then for zero blocks of appropriate dimensions

$$\begin{bmatrix}\Delta & 0\\ 0 & 0\end{bmatrix}=\begin{bmatrix}E_k & 0 & \cdots & 0\\ E_{k-1} & E_k & & \\ \vdots & \vdots & \ddots & \\ \vdots & \vdots & & E_k\\ E_0 & E_1 & & \vdots\\ 0 & E_0 & & \vdots\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & E_0\end{bmatrix}\cdot\Delta_1\begin{bmatrix}F_l & F_{l-1} & \cdots & F_0 & 0 & \cdots & 0\\ 0 & F_l & \cdots & F_1 & F_0 & \cdots & 0\\ \vdots & \vdots & \ddots & & & \ddots & \\ 0 & 0 & & F_l & & & F_0\end{bmatrix} \tag{7}$$

and Δ and Δ_1 have the same rank.

Remark 2.7: The proposition has most significance when (A^1, B^1) and (C^1, D^1) are relatively left and right prime; it follows, for example, that all generalized Bezoutian matrices have the same rank.

Proof: We have, with A^1 and C^1 of degree n and m,

$$[I\ xI\ \cdots\ x^{n-1+k}I]\begin{bmatrix}\Delta & 0\\ 0 & 0\end{bmatrix}[I\ yI\ \cdots\ y^{m-1+k}]'=\Gamma(x,y)$$

where the size of the zero blocks depends on the degrees of EA^1 and C^1F, which are no greater than $n+k$ and $m+l$. Also,

$$\Gamma(x,y)=E(x)\Gamma_1(x,y)F(y)$$
$$=E(x)[I\ xI\ \cdots\ x^{n-1}I]\Delta_1[I\ yI\ \cdots\ y^{m-1}I]'F(y).$$

Now it is easily checked by direct calculation that

$$E(x)[I\ xI\ \cdots\ x^{n-1}I]=[I\ xI\ \cdots\ x^{n-1+k}I]\begin{bmatrix}E_k & 0 & \cdots & 0\\ E_{k-1} & E_k & & \\ \vdots & & \ddots & \\ \vdots & & & E_k\\ E_0 & E_1 & & \\ 0 & E_0 & & \\ \vdots & \vdots & \ddots & \\ 0 & 0 & & E_0\end{bmatrix}$$

with a similar relation for $[I\ yI\ \cdots\ y^{m-1}I]'F(y)$. Equation (7) is immediate.

That Δ and Δ_1 have the same rank is an immediate consequence of the following lemma.

Lemma 2.2: Let $E(z)=E_0z^k+\cdots+E_k$ be a $p\times p$ matrix polynomial. Then $E(z)$ is nonsingular if and only if

$$\mathcal{E}=\begin{bmatrix}E_k & 0 & \cdots & 0\\ E_{k-1} & E_k & \cdots & 0\\ \vdots & & \ddots & \\ \vdots & & & E_k\\ E_0 & E_1 & & \vdots\\ 0 & E_0 & & \vdots\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & & E_0\end{bmatrix}$$

with n block columns has full rank for all n.

Proof: Consider a vector x with n block subvectors each of dimension p; thus $x=[x_{n-1}'\ \cdots\ x_0']'$.

Also, let $y=[y_{n+k-1}'\ \cdots\ y_0']'$ have $(n+k)$ block subvectors of dimension p. Then $y=\mathcal{E}_x$ if and only if $Y(z)=E(z)X(z)$ where $X(z)=x_0z^{n-1}+\cdots+x_n$ and $y(z)=y_0z^{n+k-1}+\cdots+y_{n+k-1}$. The result is immediate.

Having established that all generalized Bezoutian matrices have the same rank, we identify this rank in the following main result of this section.

Theorem 2.1: Let Δ be any generalized Bezoutian matrix associated with a real rational $q\times r$ $W(z)$ with $W(\infty)<\infty$. Then $\operatorname{rank}\Delta=\delta[W(z)]$. [Here, $\delta[W(z)]$ denotes the McMillan degree of $W(z)$.]

The proof of the result hinges on the following lemma, which seems of independent interest. A scalar version of the lemma appears in [9].

Lemma 2.3: With notation as pointed out earlier, suppose that $W(\infty)=0$ and $W(z)=W_0z^{-1}+W_1z^{-2}+W_2z^{-3}+\cdots$. Let H_{nm} denote the truncated block Hankel matrix

$$H_{nm}=\begin{bmatrix}W_0 & W_1 & \cdots & W_{m-1}\\ W_1 & W_2 & \cdots & W_m\\ \vdots & \vdots & & \vdots\\ W_{n-1} & W_n & \cdots & W_{n+m-1}\end{bmatrix}.$$

Then

$$\Delta = \begin{bmatrix} A_{n-1} & A_{n-2} & \cdots & A_1 & A_0 \\ A_{n-2} & A_{n-3} & \cdots & A_0 & 0 \\ d & & & 0 & 0 \\ \cdot & & \ddots & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ A_0 & 0 & \cdots & 0 & 0 \end{bmatrix} H_{nm} \begin{bmatrix} C_{m-1} & C_{m-2} & \cdots & C_1 & C_0 \\ C_{m-2} & C_{m-3} & \cdots & C_0 & 0 \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \ddots & \cdot & \cdot \\ C_0 & 0 & \cdots & 0 & 0 \end{bmatrix} \tag{8}$$

Proof: Observe that

$$[x^{-1}I\ x^{-2}I\ \cdots] \begin{bmatrix} W_0 & W_1 & W_2 & \cdots \\ W_1 & W_2 & \cdots & \cdots \\ W_2 & \cdot & & \\ \vdots & \vdots & & \end{bmatrix} \begin{bmatrix} y^{-1}I \\ y^{-2}I \\ \cdot \\ \vdots \end{bmatrix}$$

$$= \frac{1}{xy}[W_0 + W_1(x^{-1}+y^{-1}) + W_2(x^{-2}+x^{-1}y^{-1}+y^{-2}) + \cdots]$$

$$= \frac{1}{xy}\frac{1}{y^{-1}-x^{-1}}[W_0(y^{-1}-x^{-1}) + W_1(y^{-2}-x^{-2})$$

$$+ W_2(y^{-3}-x^{-3}) + \cdots]$$

$$= \frac{1}{x-y}[W(y) - W(x)]$$

$$= \frac{1}{x-y}[D(y)C^{-1}(y) - A^{-1}(x)B(x)].$$

Therefore,

$$A(x)[x^{-1}I\ x^{-2}I\ \cdots]H[y^{-1}I\ y^{-2}I\ \cdots]'C(y)$$

$$= \frac{1}{x-y}[A(x)D(y) - B(x)C(y)] = \Gamma(x,y).$$

Here, H is the (untruncated) Hankel matrix associated with $W(z)$. Now write $A(x)x^{-i} = \hat{A}^i(x) + \tilde{A}^i(x)$ where $\hat{A}^i(x)$ is polynomial in x, and $\tilde{A}^i(x)$ is polynomial in x^{-1} with no constant term. Using the fact that $\Gamma(x,y)$ is integral, we have

$$[\hat{A}^1(x)\ \hat{A}^2(x)\ \cdots\ \hat{A}^n(x)]H_{nm}[\hat{C}^{1\prime}(y)\ \cdots\ \hat{C}^{m\prime}(y)]' = \Gamma(x,y)$$

or

$$[I\ xI \cdots x^{n-1}I] \begin{bmatrix} A_{n-1} & A_{n-2} & \cdots & A_0 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & \ddots & \cdot \\ \cdot & \cdot & & \cdot \\ A_1 & A_0 & & \cdot \\ A_0 & 0 & \cdots & 0 \end{bmatrix}$$

$$\cdot H_{nm} \begin{bmatrix} C_{m-1} & C_{m-2} & \cdots & C_1 & C_0 \\ C_{m-2} & C_{m-3} & \cdots & C_0 & 0 \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \ddots & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ C_0 & 0 & \cdots & 0 & 0 \end{bmatrix} \begin{bmatrix} I \\ yI \\ \cdot \\ \cdot \\ \cdot \\ y^{m-1}I \end{bmatrix}$$

$$= [I\ xI\ \cdots\ x^{n-1}I]\Delta[I\ yI\ \cdots\ y^{m-1}I]'.$$

Equation (8) is immediate.

Proof of Theorem 2.1: Suppose first that $W(\infty)=0$. Because $W(z) = A^{-1}(z)B(z)$ with A of degree n, $A(z)[W_0z^{-1} + W_1z^{-2} + \cdots]$ is polynomial, which in turn can be seen to imply that the block rows of H are linear combinations of the first n block rows; similarly, the block columns are linear combinations of the first m block columns. Therefore, $\delta[W(z)] = \operatorname{rank} H = \operatorname{rank} H_{nm}$, the first equality being standard [10].

There exist matrix fraction decompositions of $W(z)$ with A_0 and C_0 nonsingular. (An example is to be found in Remark 2.8.) For such decompositions it follows from (8) that $\operatorname{rank}\Delta = \operatorname{rank} H_{nm} = \delta[W(z)]$. By Proposition 2.2, it follows that $\operatorname{rank}\Delta = \delta[W(z)]$ for all $\{A,B,C,D\}$. In case $W(\infty)\neq 0$ but $W(\infty)<\infty$, Remark 2.3 extends the proof of the theorem to this case.

Remark 2.8: A special case of this theorem has been obtained by Munro [5] who, with $W(z) = (V_1z^{n-1} + \cdots + V_n)(z^n + a_1z^{n-1} + \cdots + a_n]^{-1}$, takes $A(z) = (z^n + a_1z^{n-1} + \cdots + a_n)I_q$ and $C(z) = (z^n + a_1z^{n-1} + \cdots + a_n)I_r$. In this case, it is easily checked using (2) that $\Gamma(x,y) = \Gamma(y,x)$, i.e., $\Gamma_{ij} = \Gamma_{ji}$.

Remark 2.9: Let $W(z)$ be symmetric, with a left matrix fraction decomposition (A,B) and thus right matrix fraction decomposition (A',B'). Then $\{A,B,A',B'\}$ determines a Δ which by (8) is symmetric, with the same rank and signature as the symmetric matrix H_{nn}.

Remark 2.10: Theorem 2.1 extends to improper $W(z)$ in the following way.[2] First, suppose $W(z)$ is square. Then, for some J, $X(z) = [W(z) + J]^{-1}$ is proper, and has degree equal to that of $W(z)$. By Remarks 2.3 and 2.4, $\Gamma_X = -\Gamma_W$ and by the theorem $\delta[X] = \operatorname{rank}\Delta_X$. So $\delta[W] = \operatorname{rank}\Delta$. In case $W(z)$ is not square, suppose that $q<r$. Set

$$\hat{W}(z) = \begin{bmatrix} W(z) \\ 0_{r-q\times r} \end{bmatrix}, \quad \hat{A}(z) = A(z)\oplus I_{r-q}, \quad \hat{B}(z) = \begin{bmatrix} B(z) \\ 0_{r-q\times r} \end{bmatrix}$$

$$\hat{C}(z) = C(z), \quad \hat{D}(z) = \begin{bmatrix} D(z) \\ 0_{r-q\times r} \end{bmatrix}.$$

Then one can check that

$$\hat{\Gamma} = \begin{bmatrix} \Gamma \\ 0_{r-q\times r} \end{bmatrix}$$

while $\delta[\hat{W}] = \delta[W]$. This argument shows that the result for square W implies the result for nonsquare W. A second way to see this result is as follows. As shown in [11], with $W(z)$ improper, for almost all constant F, one has $U(z) = W(z)[I + FW(z)]^{-1}$ proper; also, $\delta[U(z)] = \delta[W(z)]$. By Remark 2.3, one can select matrix fraction descriptions causing $\Gamma_U = \Gamma_W$. Application of Theorem 2.1 shows rank $\Gamma_W = \delta[W(z)]$

Remark 2.11: By imposing constraints on $\{A,B,C,D\}$ one can obtain some fine structure in Δ which may be useful in applications. For example, if A and C are, respectively, row and column proper [2], Δ contains a number of zero rows and columns, while if also the pairs (A,B) and (C,D) are left and right prime, Δ has precisely $\delta[W(z)]$ nonzero rows and columns. The positions of these zero rows and columns define and are defined by the row and column indices of A and C.

Remark 2.12: The pairs (A,B) and (C,D) are relatively left and right prime, respectively, if and only if $\delta[W(z)] = \operatorname{degree}[\det A(z)] = \operatorname{degree}[\det C(z)]$. Accordingly, Theorem 2.1 provides a tool for checking the relative primeness of the two pairs.

Example 2.1: The transfer function matrix

$$W(z) = \begin{bmatrix} \dfrac{1}{z+1} & \dfrac{2}{z-2} \\ \dfrac{2}{z-2} & 0 \end{bmatrix}$$

has matrix fraction descriptions defined by

$$A(z) = C(z) = \begin{bmatrix} z^2-z-2 & 0 \\ 0 & z-2 \end{bmatrix} \quad B'(z) = D(z) = \begin{bmatrix} z-2 & 2 \\ 2z+2 & 0 \end{bmatrix}.$$

[2] We are grateful to Prof. T. Kailath for suggesting the possibility of this extension (see also Remark 3.7).

Then

$$
\begin{aligned}
\Gamma(x,y)&=\frac{1}{x-y}\begin{bmatrix}(x^2-x-2)(y-2)-(x-2)(y^2-y-2) & 2(x^2-x-2)-(2x+2)(y-2)\\ (x-2)(2y+2)-2(y^2-y-2) & 0\end{bmatrix}\\
&=\begin{bmatrix}xy-2(x+y)+4 & 2x+2\\ 2y+2 & 0\end{bmatrix}\\
&=[I\ \ xI]\left[\begin{array}{cc:cc} 4 & 2 & -2 & 0\\ 2 & 2 & 2 & 0\\ \hdashline -2 & 2 & 0 & 0\\ 0 & 0 & 0 & 0\end{array}\right]\begin{bmatrix}I\\ \\ yI\end{bmatrix}.
\end{aligned}
$$

Evidently, $\operatorname{rank}\Delta=3$, and it is easy to check that $\delta[W(z)]=3$.

III. Generalized Sylvester Matrices

We work with a real rational $q\times r$ matrix $W(z)$ with matrix fraction decomposition as in (1), with $A(z)$, $C(z)$ of degree n, m and with $W(\infty)=0$. The results to follow can be adjusted to cover the case of $0\neq W(\infty)<\infty$.

We define for each integer p a *generalized Sylvester matrix*[3] $S^p(C,D)$ with $2p-1$ block rows by

$$
S^p(C,D)
$$

$$
=\begin{bmatrix}
C_0 & C_1 & & & C_m & 0 & \cdot & \cdot & 0\\
0 & C_0 & & & & \cdot & & & \\
\cdot & 0 & \cdot & & & \cdot & & & \\
\cdot & & \cdot & & C_2 & \cdot & \cdot & C_m & 0\\
\cdot & & & C_0 & C_1 & \cdot & \cdot & \cdot & C_m\\
\cdot & & & 0 & D_1 & \cdot & \cdot & \cdot & D_m\\
\cdot & & & D_1 & D_2 & \cdot & \cdot & D_m & 0\\
\cdot & & \cdot & & & & \cdot & & \cdot\\
\cdot & & \cdot & & & \cdot & & & \cdot\\
 & & & & \cdot & & & & \cdot\\
0 & D_1 & & & \cdot & & & & \cdot\\
D_1 & D_2 & \cdots & D_m & 0 & \cdot & \cdot & \cdot & 0
\end{bmatrix}
\begin{array}{l}
\left.\begin{array}{l}\\ \\ \\ \\ \\ \end{array}\right\}p-1 \text{ block rows}\\
\left.\begin{array}{l}\\ \\ \\ \\ \\ \\ \\ \end{array}\right\}p \text{ block rows.}
\end{array}
\tag{9}
$$

Remark 3.1: As one would expect, one can define a generalized Sylvester matrix $\bar{S}^p(A,B)$ and obtain results akin to those for $S^p(C,D)$. In view of the fact that (A',B') defines a right matrix fraction decomposition for $W'(z)$, we might define $\bar{S}^p(A,B)=S^p(A',B')$. To eliminate transposes, however, it is tidier to take $\bar{S}^p(A,B)=[S^p(A',B')]'$. This matrix has an upper triangle of zeros, and triangular blocks of zeros in the bottom left and right corners.

By direct verification using definition (9), we have the following lemma.

Lemma 3.1: Let $F(z)=F_0z^l+\cdots+F_l$ be a nonsingular matrix. Then

$$
[0\ \vdots\ S^p(C^1F,D^1F)]
$$

$$
=S^p(C^1,D^1)\begin{bmatrix}
F_0 & F_1 & \cdots & F_l & 0 & \cdots & 0\\
0 & F_0 & \cdots & F_{l-1} & F_l & \cdots & 0\\
\cdot & \cdot & & & & & \\
\cdot & \cdot & \ddots & & & \ddots & \\
\cdot & \cdot & & & & & \\
0 & 0 & & F_0 & & & F_l
\end{bmatrix}. \tag{10}
$$

(The zero block appears in the left side of (10) if and only if the polynomial C^1F has degree less than $m+l$.)

Remark 3.2: Using Lemma 2.2, we see that

$$
\operatorname{rank}S^p(C^1F,D^1F)=\operatorname{rank}S^p(C^1,D^1) \tag{11}
$$

for all nonsingular $F(z)$. Since any pair (C,D) for which $W(z)=D(z)C^{-1}(z)$ is of the form (C^1F,D^1F) for a relatively prime (C^1,D^1), it follows that $\operatorname{rank}S^p(C,D)$ is the same for all right matrix fraction decompositions of $W(z)$.

We now link the notions of the generalized Bezoutian and Sylvester matrices.

Proposition 3.1: Let $\{A,B,C,D\}$ be associated with a real rational $W(z)$ with $W(\infty)=0$, and suppose A, C have degree n and m, respectively. Partition $S^n(C,D)$ as

$$
S^n(C,D)=\begin{array}{c}\begin{bmatrix}R & & T\\ S & & U\end{bmatrix}\\ \begin{array}{cc}(n-1) & m\\ \text{block} & \text{block}\\ \text{columns} & \text{columns.}\end{array}\end{array}
\begin{array}{l}(n-1)\text{ block rows}\\ n\text{ block rows}\\ \\ \\ \\ \end{array} \tag{12}
$$

Define

$$
P=\begin{bmatrix}
-B_1 & -B_2 & \cdots & -B_{n-1}\\
0 & -B_1 & \cdots & -B_{n-2}\\
\cdot & & & \cdot\\
\cdot & & \ddots & \cdot\\
\cdot & & & \\
 & & & -B_1\\
0 & & & 0
\end{bmatrix} \tag{13a}
$$

$$
Q=\begin{bmatrix}
A_{n-1} & A_{n-2} & \cdots & A_0\\
A_{n-2} & A_{n-3} & \cdots A_0 & 0\\
\cdot & & & \cdot\\
\cdot & & \cdot & \cdot\\
\cdot & \cdot & & \cdot\\
A_0 & 0 & \cdots & 0
\end{bmatrix} \tag{13b}
$$

$$
J=\begin{bmatrix}
0_r & 0_r & \cdots & & I_r\\
0_r & 0_r & \cdots & I_r & 0_r\\
\cdot & & & \cdot & \cdot\\
\cdot & & \cdot & \cdot & \cdot\\
\cdot & I_r & & \cdot & \cdot\\
I_r & 0_r & \cdots & & 0_r
\end{bmatrix}. \tag{13c}
$$

Then

$$
\begin{bmatrix}I & 0\\ P & Q\end{bmatrix}\begin{bmatrix}R & T\\ S & U\end{bmatrix}=\begin{bmatrix}R & T\\ 0 & \Delta J\end{bmatrix}. \tag{14}
$$

Proof: By direct calculation, using (6) and the fact that $A(z)D(z)=B(z)C(z)$.

Remark 3.3: Proposition 3.1 is a generalization of an old formula applicable to scalar polynomials [7].

[3]As noted in the Introduction, a special case of this matrix was formulated by Rosenbrock [3], who assumed $C(z)$ in (1) to be of the form $c(z)I$, where $c(z)$ is a scalar polynomial.

Using lower case italic letters for the polynomial coefficients, and with $A(z)=C(z)$ and $B(z)=D(z)$ one has

$$\begin{bmatrix} -b_1 & -b_2 & \cdots & -b_{n-1} & a_{n-1} & a_{n-2} & \cdots & a_1 & a_0 \\ 0 & -b_1 & \cdots & -b_{n-2} & a_{n-2} & a_{n-3} & \cdots & a_0 & 0 \\ \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \ddots & \cdot & \cdot & \cdot & \ddots & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot & \cdot & & \cdot & \cdot \\ 0 & 0 & \cdots & -b_1 & a_1 & a_0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & a_0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

$$\cdot \begin{bmatrix} a_0 & a_1 & \cdots & a_n & 0 & \cdots & 0 \\ 0 & a_0 & \cdots & & & & \\ & & \ddots & & & & \\ & & a_0 & a_1 & \cdots & a_n & \\ & & 0 & b_1 & \cdots & b_n & \\ & & b_1 & b_2 & \cdots & 0 & \\ & & \ddots & & & & \\ & b_1 & \cdots & & & & \\ b_1 & b_2 & \cdots & b_n & 0 & \cdots & 0 \end{bmatrix} = \text{Bezoutian with rows in reverse order.}$$

As an almost immediate consequence, we have the following

Proposition 3.2: Let (C,D) be a right matrix fraction decomposition of a real rational $W(z)$ with $W(\infty)=0$ and let n be such that there exists a left matrix fraction decomposition (A,B) with n the degree of $A(z)$. Then $\operatorname{rank} S^n(C,D)=r(n-1)+\delta[W(z)]$.

Proof: First observe there is no loss of generality in taking A_0 nonsingular. For suppose A_0 is singular; let $U(z)$ be a unimodular matrix such that $U(z)A(z)$ is row proper. It is always possible to find such a matrix, see [2], such that with the degree of row i denoted by ν_i, one has $n=\nu_1 \geqslant \nu_2 \geqslant \nu_3 \geqslant \cdots \geqslant \nu_q$. Then with $E(z)=\operatorname{diag}[1, z^{\nu_1-\nu_2}, \cdots z^{\nu_1-\nu_q}]$, $E(z)U(z)A(z)$ has degree n and the coefficient matrix of z^n is nonsingular; of course, (EUA, EUB) constitutes a left matrix fraction decomposition of $W(z)$

Next, let $F(z)$ be such that the coefficient of the highest degree term in $C(z)F(z)$ is nonsingular. We shall show that $\operatorname{rank} S^n(CF,DF)=r(n-1)+\delta[W(z)]$. By Remark 3.2, the theorem will hold.

Since A_0 is nonsingular, Q in (13b) is nonsingular, and from (14) it is evident that $\operatorname{rank} S^n(CF,DF)=\operatorname{rank} R+\operatorname{rank}(\Delta J)$. Since the highest degree term in CF is nonsingular, R is nonsingular from (9) and (12). Also J is nonsingular. Thus, $\operatorname{rank} S^n(CF,DF)=r(n-1)+\operatorname{rank}\Delta=r(n-1)+\delta[W(z)]$. This completes the proof.

In view of the fact that the pair (C,D) is relatively right prime if and only if the degree of $\det C(z)$ is $\delta[W(z)]$, see [2], we have an immediate corollary generalizing results of Rowe [4].

Corollary 3.1: Let $(C(z),D(z))$ be a pair of matrices with $C(z)$ nonsingular and $r\times r$, and with $D(z)$ of lower degree than $C(z)$. Let n be such that there exists a left matrix fraction decomposition (A,B) of $W(z)=D(z)C^{-1}(z)$ with A of degree n. Then (C,D) is relatively right prime if and only if $\operatorname{rank} S^n(C,D)=r(n-1)+\operatorname{degree}[\det C]$.

The earlier work of Rowe [4] demanded that the coefficient of the highest degree term of $C(z)$ be nonsingular, and demanded that $A(z)$ be of the form $(z^n+a_1z^{n-1}+\cdots+a_n)I$ for scalar a_i.

The major result of this section in essence eliminates the requirement in the preceding proposition and corollary that an integer n be known.

Theorem 3.1: Let (C,D) be a right matrix fraction decomposition of a real rational $q\times r$ $W(z)$ with $W(\infty)=0$. Let ν be the least integer for which

$$\operatorname{rank} S^{\nu+1}(C,D)-\operatorname{rank} S^{\nu}(C,D)\leqslant r.$$

Then there exists a left matrix fraction decomposition (A,B) of $W(z)$ with A of degree ν, but none with A of degree less than ν. Moreover,

$$\operatorname{rank} S^{\nu+\mu+1}(C,D)-\operatorname{rank} S^{\nu+\mu}(C,D)=r$$

for all integer $\mu\geqslant 0$.

Remark 3.4: The theorem suggests that the ranks of generalized Sylvester matrices of growing size be examined; when the rate of increase stabilizes, this identifies ν. Note that ν, as the lowest possible degree of any $A(z)$ in a left matrix fraction decomposition, and therefore the least possible value of the maximum row index of any row proper $A(z)$, is the maximum Kronecker invariant, or the observability index in any minimal state-space realization of $W(z)$ [2].

Proof: Using Remark 3.2, we can assume with no loss of generality that C_0 is nonsingular. Then, since

$$S^{\nu+1}(C,D)=\begin{bmatrix} C_0 & C_1 & C_2 & \cdots & C_m & 0 & \cdots & 0 \\ 0 & & & & & & & \\ \cdot & & & S^{\nu}(C,D) & & & & \\ \cdot & & & & & & & \\ 0 & & & & & & & \\ D_1 & D_2 & \cdots & D_m & 0 & 0 & \cdots & 0 \end{bmatrix}$$

it is immediate with ν as defined in the theorem hypothesis that $\operatorname{rank} S^{\nu+1}(C,D)=\operatorname{rank} S^{\nu}(C,D)+r$ and that the last block row of $S^{\nu+1}(C,D)$ is a linear combination of the preceding block rows. Thus, for some matrices B_j, A_j [not yet associated with a pair $B(z)$, $A(z)$] we have

$$[D_1\ D_2\ \cdots\ D_m\ 0\ \cdots\ 0]=[B_1\ B_2\ \cdots\ B_\nu\ -A_\nu\ \cdots\ -A_1]\cdot\begin{bmatrix} C_0 & C_1 & \cdots & C_m & 0 & \cdots & 0 \\ 0 & & & & & & \\ \cdot & & & S^{\nu}(C,D) & & & \\ \cdot & & & & & & \\ 0 & & & & & & \end{bmatrix}.$$

Writing these relations out gives

$$D_1=B_1C_0, D_2=B_1C_1+B_2C_0-A_1D_1,\cdots$$

which, if we define

$$A(z)=z^{\nu}+A_1z^{\nu-1}+\cdots \quad B(z)=B_1z^{\nu-1}+\cdots,$$

are equivalent to $A(z)D(z)=B(z)C(z)$, i.e., $A^{-1}B=DC^{-1}=W$. This establishes the first claim of the theorem.

Suppose now that there exists a left matrix fraction decomposition (A,B) with $A(z)$ of degree $\sigma<\nu$. Using Proposition 3.2, we see that $\operatorname{rank} S^{\nu}(C,D)-\operatorname{rank} S^{\sigma}(C,D)=r(\nu-\sigma)$, and this provides a contradiction to the definition of ν.

Finally, $(z^{\mu+1}A, z^{\mu+1}B)$ is a left matrix fraction decomposition of degree $\mu+1+\nu$ for all integer $\mu\geqslant 0$. By Proposition 3.2, we obtain $\operatorname{rank} S^{\mu+1+\nu}(C,D)-\operatorname{rank} S^{\nu}(C,D)=(\mu+1)r$ from which the final claim of the theorem follows.

Remark 3.5: An upper bound on ν is provided by the degree of $\det C(z)$, since there exist (A,B) with $A=I_q \det C(z)$.

Remark 3.6: As a simple application of the ideas of Proposition 3.2 and Theorem 3.1, consider the matrices

$$S^1=[H] \quad S^2=\begin{bmatrix} I & -F \\ 0 & H \\ H & 0 \end{bmatrix}$$

$$S^3=\begin{bmatrix} I & -F & 0 \\ 0 & I & -F \\ 0 & 0 & H \\ 0 & H & 0 \\ H & 0 & 0 \end{bmatrix}$$

and so on, associated with a $q\times r$ $W(z)=H(zI-F)^{-1}$. Then $\operatorname{rank} S^r=r(r-1)+\delta[W(z)]$, and $\operatorname{rank} S^r=r^2$ if and only if $[F,H]$ is a completely observable pair. In case $\operatorname{rank} S^{l+1}-\operatorname{rank} S^l=r$ for some $l<r$, we can conclude complete observability if and only if $\operatorname{rank} S=rl$. Of course, this is not a new result, see, e.g., [1]. Looked at from the Bezoutian point of view, we obtain the following. Suppose that $H(zI-F)^{-1}=[\det(zI-F]^{-1}[H_1z^{n-1}+\cdots+H_n]$. One can check that $\Gamma(x,y)=H_1x^{n-1}+\cdots+H_n$ and so $\delta[W(z)]=\operatorname{rank}\Delta'=\operatorname{rank}[H_n'H_{n-1}'\cdots H_1']$.

Remark 3.7: Kung, Kailath, and Morf at Stanford University have obtained in an as yet unpublished work, a number of the results of this section by different procedures, which also yield greatest common divisors of the two matrix polynomials and an efficient minimal realization algorithm for the associated transfer function.

Remark 3.8: Remark 2.12 allows the result of Theorem 3.1 to be reformulated as a result on the relative right primeness or otherwise of (C,D).

Example 3.1: The transfer function matrix

$$W(z)=\begin{bmatrix} \frac{1}{z+1} & \frac{2}{z-2} \\ \frac{2}{z-2} & 0 \end{bmatrix}$$

has

$$C(z)=\begin{bmatrix} z^2-z-2 & 0 \\ 0 & z-2 \end{bmatrix} \quad D(z)=\begin{bmatrix} z-2 & 2 \\ 2z+2 & 0 \end{bmatrix}.$$

Then

$$C_0=\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad C_1=\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \quad C_2=\begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$$

$$D_1=\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \quad D_2=\begin{bmatrix} -2 & 2 \\ 2 & 0 \end{bmatrix}.$$

We also have

$$S^1(C,D)=\left[\begin{array}{cc|cc} 1 & 0 & -2 & 2 \\ 0 & 0 & 2 & 0 \end{array}\right]$$

$$\operatorname{rank} S^1(C,D)=2$$

$$S^2(C,D)=\left[\begin{array}{cc|cc|cc} 1 & 0 & -1 & 0 & -2 & 0 \\ 0 & 0 & 0 & 1 & 0 & -2 \\ \hline 0 & 0 & 1 & 0 & -2 & 2 \\ 0 & 0 & 1 & 0 & 2 & 0 \\ \hline 1 & 0 & -2 & 2 & 0 & 0 \\ 1 & 0 & 2 & 0 & 0 & 0 \end{array}\right]$$

$$\operatorname{rank} S^2(C,D)=5.$$

One can check that $\operatorname{rank} S^3(C,D)=7$, so that $\operatorname{rank} S^3(C,D)-\operatorname{rank} S^2(C,D)=2$, or one can use the fact (see Example 2.1 in Section II) that there exists a left matrix fraction decomposition with $A(z)$ of degree 2. By Proposition 3.2, $\delta[W(z)]=5-2(2-1)=3$.

IV. Conclusions

In this short paper we have defined a generalized Bezoutian matrix associated with the left and right matrix fraction decomposition of a prescribed transfer function matrix and obtained several forms of this generalized matrix. We have established various properties of this matrix and in particular noted an easy method for computing its entries. We have also shown that the rank of any generalized Bezoutian matrix is the McMillan degree of the real rational transfer function matrix whose matrix fraction decompositions are used to define the Bezoutian matrix. As a special case of this we have obtained the Munro results.

Furthermore, we have defined the generalized Sylvester matrix and shown its connection with the generalized Bezoutian. The rank of the generalized Sylvester matrix also determines the McMillan degree of the transfer function matrix $W(z)$ with which it is associated. Several properties of this matrix have been discussed and as a special case we have obtained Rowe's results.

A number of problems arise from the preceding material, which we note as follows.

1) Can a method be developed for computing a greatest common divisor of two matrix polynomials which are not relatively prime, using the Sylvester matrix? Such is known in the case of scalar polynomials [7].

2) Can generalized Bezoutian or Sylvester matrices play a useful role in developing stability and positivity tests, again as in the scalar case [7]?

3) Transfer function matrix descriptions of the form $C(z)A^{-1}(z)B(z)$ for polynomials A, B, and C are sometimes used [1]. Can one, via some type of generalized Sylvester matrix reflecting A, B, and C, give a simple formula for $\delta[CA^{-1}B]$?

4) In relation to Remark 3.6, can one work with generalized subresultants rather than the sequence $S^1, S^2, \cdots$? Again, our knowledge of the scalar case suggests it might be possible to trim some columns off S^1, S^2, etc.

References

[1] H. H. Rosenbrock, *State-Space and Multivariable Theory*. London: Nelson, 1970.

[2] W. A. Wolovich, *Linear Multivariable Systems*. New York: Springer, 1974.

[3] H. H. Rosenbrock, "Generalized resultant," *Electron. Lett.*, vol. 4, pp. 250–251, 1968.

[4] A. Rowe, "The generalized resultant matrix," *J. Inst. Math. Appl.*, vol. 9, pp. 390–396, 1972.

[5] N. Munro, "Determination of the least order of transfer-function matrices," *Proc. Inst. Elec. Eng.*, vol. 118, pp. 1507–1510, Oct. 1971.

[6] E. I. Jury, "The theory and applications of the inners," *Proc. IEEE*, vol. 63, pp. 1044–1068, July 1975.

[7] ——, *Inners and Stability of Dynamic Systems*. New York: 1974.

[8] C. C. Macduffee, *The Theory of Matrices*. Berlin: Springer, 1933.

[9] B. D. O. Anderson, "On the computation of the Cauchy index," *Quart. Appl. Math.*, vol. 29, pp. 577–582, Jan. 1972.

[10] R. W. Brockett, *Finite Dimensional Linear Systems*. New York: Wiley, 1970.

[11] R. W. Scott and B. D. O. Anderson, "Comments on 'Conditions for a feedback transfer matrix to be proper,'" this issue, pp. 632–634.