Greatest Common Divisors via Generalized Sylvester and Bezout Matrices

R. R. BITMEAD, S.-Y. KUNG, B. D. O. ANDERSON, FELLOW, IEEE, AND T. KAILATH, FELLOW, IEEE

***Abstract*—We present new methods for computing the greatest common right divisor of polynomial matrices. These methods involve the recently studied generalized Sylvester and generalized Bezoutian resultant matrices, which require no polynomial operations. They can provide a row proper greatest common right divisor, test for coprimeness and calculate dual dynamical indices.**

The generalized resultant matrices are developments of the scalar Sylvester and Bezoutian resultants and many of the familiar properties of these latter matrices are demonstrated to have analogs with the properties of the generalized resultant matrices for matrix polynomials.

I. Introduction

Greatest common divisors (gcd's) of polynomial matrices play an important part in the theory and application of general differential systems as studied extensively by Rosenbrock [1], [2], Wolovich [3], and others. For example, they are useful in 1) obtaining irreducible matrix-fraction descriptions (and hence minimal state-space realizations) of transfer-function matrices, 2) studying decoupling zeros and uncontrollable and unobservable modes of given systems, and 3) obtaining the pole-zero structure of given multivariable systems.

Most of the system-theory literature in this area has focused on the somewhat more restricted problem of devising tests for the coprimeness of matrix polynomials—see, e.g., [4]–[11], or in obtaining irreducible MFD's by more direct methods—see, e.g., [12]–[14]. These methods can in principle often also lead to a gcd, as we explain now. First note that [15] a *greatest common right divisor*[1] (gcrd) of two polynomial matrices $C(s)$ and $D(s)$, having the same number of columns, is any polynomial matrix $R(s)$ such that 1) $R(s)$ is a right divisor of $\{C(s), D(s)\}$, i.e.,

$$C(s)=\bar{C}(s)R(s), D(s)=\bar{D}(s)R(s)$$

Manuscript received January 13, 1978; revised August 28, 1978. Paper recommended by M. Sain, Chairman of the Linear Systems Committee. This work was supported by the U.S. Army Research Office under Grant DAAG29-77-C-0042, by the Australian Research Grants Committee, National Science Foundation under the U.S.-Australian Cooperative Science Program, by the Air Force Office of Scientific Research, Air Force Systems Command, under Contract AF44-620-74-C-0068, and by the National Science Foundation under Grant ENG75-10533.

R. R. Bitmead and B. D. O. Anderson are with the Department of Electrical Engineering, University of Newcastle, New South Wales, Australia.

S.-Y. Kung is with the Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90007.

T. Kailath is with the Information Systems Laboratory, Stanford University, Stanford, CA 94305.

[1]Similar definitions and results apply to greatest common left divisors (gcld's), so that we shall confine our discussions to gcrd's.

0018-9286/78/1200-1043$00.75 ©1978 IEEE

for some polynomial matrices $\{\bar{C}(s),\bar{D}(s)\}$; and 2) it is divisible by any other right divisor, say $R_1(s)$, of $\{C(s),D(s)\}$, i.e., $R(s)=M(s)R_1(s)$, for some polynomial matrix $M(s)$.

gcrd's are clearly not unique, since they can differ by *unimodular* factors, i.e., polynomial matrices with constant (nonzero) determinants. For nondegeneracy, we shall also require that the matrix $F(s)=[D'(s)C'(s)]'$ has full rank (for almost all s) because otherwise we could have gcrd's of arbitrarily high degree. In system theory, this condition is often assured by having one of the matrices, say $C(s)$, be square and nonsingular. For example, the pair $\{C(s),D(s)\}$ often arises as a so-called right MFD (matrix-fraction description) of a matrix transfer function $H(s)$,

$$H(s)=D(s)C^{-1}(s).$$

A pair $\{\bar{C}(s),\bar{D}(s)\}$ is said to be right coprime if it has only unimodular right divisors. If $\bar{C}(s)$ is nonsingular, $\bar{D}(s)\bar{C}^{-1}(s)$ is said to be an irreducible right MFD.

Suppose now that we have an MFD $D(s)C^{-1}(s)$ for which by one of the methods of [12]–[14] we have found an irreducible description $\bar{D}(s)\bar{C}^{-1}(s)$. Then a gcrd of $\{C(s),D(s)\}$ can be found as $\bar{C}^{-1}(s)C(s)$. However, this method is not only indirect, it will generally require direct symbolic manipulation of polynomial matrices, which may not be very convenient for computer evaluation. The same criticism applies to the standard method of MacDuffee [15], [16], which finds a gcd as a polynomial combination

$$R(s)=P(s)C(s)+Q(s)D(s).$$

In special cases, e.g., when $F(s)$ is column reduced, the Euclidean algorithm could be used to find a suitable pair $\{P(s),Q(s)\}$. In this paper we shall first describe how a gcrd can be determined from elementary operations on a constant real matrix—a so-called *generalized Sylvester resultant matrix* (Section II). We shall show that certain ranks associated with this matrix yield information on the so-called *dual dynamical indices* [17]–[19] of the matrix $[C'(s)\ D'(s)]'$ and thereby also yield a new test for coprimeness [19], [20]. This rank information can be determined by using numerically stable orthogonal transformations (cf. [14], [18]), but doing this will destroy some information about the gcrd's; to preserve the gcrd's we have to use unimodular rather than orthogonal transformations. We shall show how to do this while also exploiting the special shift-invariant Toeplitz structure of the generalized Sylvester resultant to reduce the number of computations [19] (Section III).

In the latter part of the paper we show how gcrd's and minimal indices can also be obtained from the closely related generalized Bezout resultant of [20].

II. The Generalized Sylvester Resultant

Two scalar polynomials $d(s)$ and $c(s)$ are coprime if and only if there exist no nontrivial polynomials $a(s)$ and $b(s)$ with degree less than the degree of $c(s)$ and $d(s)$ such that $a(s)d(s)+b(s)c(s)=0$. Equating coefficients leads to a special matrix, known as the Sylvester resultant matrix of $c(s)$ and $d(s)$, whose nonsingularity implies the coprimeness of $c(s)$ and $d(s)$. As noted in [19], this approach also extends nicely to the matrix case. First, we can show that two matrix polynomials[2] $C(s)$ and $D(s)$ will be right coprime if and only if there exists an irreducible pair of row-reduced (row-proper)[3] polynomial matrices $[A(s)\ B(s)]$ such that $A(s)$ has the same determinantal degree as $C(s)$ and

$$A(s)D(s)+B(s)C(s)=0. \tag{2.1}$$

Suppose now that $A(s)=\sum_{i=0}^{l-1}A_is^{l-1-i}$ while $\{B(s),C(s),D(s)\}$ have similar expressions and are of degrees $l-1$, m and m, respectively. Then by equating the coefficients of the various powers of s in (2.1), we shall obtain the equations

$$[A_0,B_0,A_1,B_1,\cdots,A_{k-1},B_{k-1}]S_k=0, \qquad k=1,2,\ldots \tag{2.2}$$

[2]We assume that $\{C(s),D(s)\}$ have the same number of columns and that $[C'(s)\ D'(s)]'$ has full column rank.

[3]By this we mean that the matrix, whose ith row is composed of the coefficients of the highest power of s in the ith row of the polynomial matrix, is full rank. Other characterizations also exist, see e.g., [3].

where

$$S_k=\begin{bmatrix} D_0 & D_1 & \cdots & D_m & 0 & \cdots & 0\\ C_0 & C_1 & & C_m & 0 & & 0\\ 0 & D_0 & & D_{m-1} & D_m & & 0\\ 0 & & & C_{m-1} & C_m & & 0\\ \vdots & & & & & & \\ 0 & \cdots & 0 & D_0 & D_1 & & D_m\\ 0 & & 0 & C_0 & C_1 & \cdots & C_m \end{bmatrix}, \quad 2k \text{ block rows} \tag{2.3}$$

is called the *generalized Sylvester resultant* of C and D of order k [19]. An inconsequentially different form was obtained in [20] by a different argument based on generalized Bezout resultants (other somewhat more restricted forms of matrix Sylvester resultants have been given by Rosenbrock [4], Rowe [6] and Wolovich [3, p. 234]).

Equation (2.2) suggests then, that in dealing with questions of coprimeness, it will be useful to examine the left null space of S_k, for increasing k.

In doing this, it will be useful to note that elementary row operations on the matrix S_k correspond in an obvious way to elementary operations on the rows of the polynomial matrix

$$F(s)=[D'(s)\quad C'(s)]'.$$

If we define $E(s)=[A(s)\quad B(s)]$, with $\{A(s),B(s)\}$ *coprime* and $E(s)$ *row-reduced* then the relation (2.1) can be rewritten as

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

Forney has shown [17] that the constraints on $E(s)$ imply the following.

1) Any polynomial row vector $p(s)$ that is orthogonal to $F(s)$ has the form

$$p(s)=\sum_{i=1}^{q}w_i(s)E_i(s)=w(s)E(s)$$

for some polynomial row vector $w(s)$. (The E_i are rows of E.)

2) If $p(s)=w(s)E(s)$, then

$$\deg p=\max_{\{i:w_i\neq 0\}}[\deg w_i+\nu_i]$$

where ν_i = the degree of the ith row of $E(s)$.

The indices $\{\nu_i\}$ characterize the left null space of $F(s)$ [21]. When $\{C(s),D(s)\}$ form a right MFD of a matrix transfer function $H(s)=D(s)C^{-1}(s)$, then the $\{\nu_i\}$ have been called the *dual dynamical indices* of $H(s)$ by Forney [17]: when $H(s)$ is proper they coincide with the observability indices of any minimal state-space realization of $H(s)$. Moreover, the order of any such minimal realization obeys $n_{\min}=\sum\nu_i$. The following result relates the $\{\nu_i\}$ to the generalized Sylvester resultant and has immediate application to the determination of the dual dynamical indices.

Theorem 1 [19]: Given the $q\times r$ transfer function matrix $H=DC^{-1}$ with dual dynamical indices ν_i, we have

$$\operatorname{rank} S_k=(r+1)k-\sum_{\{i:\nu_i<k\}}(k-\nu_i) \tag{2.4}$$

where S_k is the generalized Sylvester resultant matrix of order k of C and D.

Consequently, denoting rank S_k by r_k and the dimension of the row null space of S_k by n_k, the number α_k of dual dynamical indices of value k is given by the formula

$$\begin{aligned}\alpha_k&=(r_k-r_{k-1})-(r_{k+1}-r_k) \qquad k=1,2,\cdots\\ &=(n_{k+1}-n_k)-(n_k-n_{k-1}),\end{aligned} \tag{2.5}$$

initialized by $r_0=n_0=0$, $\alpha_0=r+q-r_1=n_1$.

Proof: Define

$$\mathcal{N}_k \equiv \{w: wS_k = 0, \text{ where } w \text{ is a } k(q+r) \text{ row vector}\}$$

$$U_k \equiv \{v(s): v(s)F(s) = 0 \text{ and } \deg v(s) < k\}$$

$$V_k \equiv \begin{array}{l}\{v(s): v(s) = x(s)E(s) \text{ where } x(s) \text{ is any} \\ \text{polynomial row vector with } \deg x_i < k - \nu_i\}.\end{array}$$

By properties 1) and 2) of $E(s)$, we can assert that $V_k^{\perp} = U_k$ and it is clear that $\mathcal{N}_k$ is isomorphic to U_k. Since dim V_k is clearly equal to $\Sigma_{\{i:\nu_i<k\}}(k-\nu_i)$ this is also the dimension of $\mathcal{N}_k$. Then, noting that dim $\mathcal{N}_k$ + rank $S_k = (r+q)k$ establishes (2.4). The other results follow simply. □

The spanning property of the rows of $E(s)$ then shows that the dimension of the null space of S_k increases uniformly with k once k has surpassed the maximum dual dynamical index. Since $E(s)$ yields an irreducible left MFD of the transfer function matrix, Theorem 1 and the facts that $n_{\min}$ equals the sum of the dual dynamical indices, and that the pair $\{C,D\}$ is right coprime for proper H if and only if deg det C equals $n_{\min}$, yield the following result.

Corollary 1. [19], [20]: With the same hypothesis as Theorem 1, let ν be the least integer for which

$$\text{rank } S_{\nu+1} - \text{rank } S_\nu \leq r$$

(actually it must equal r). Then there exists a left MFD of $H(s)$ of degree ν, but none of degree less than ν. Furthermore,

$$\text{rank } S_{\nu+\mu} = r(\nu+\mu) + n_{\min}$$

for all integers $\mu > 0$. And if H is proper, then for $n > \nu$, $\{C,D\}$ are right coprime if and only if

$$\text{rank } S_n = rn + \deg\det C.$$

The rank information needed for the above calculations can be obtained by numerically stable orthogonal transformations (cf. [14], [18]). However, to calculate a gcrd from the $\{S_k\}$ we have to use elementary (or unimodular) transformations and we now describe an efficient way of doing this.

III. gcrd's via the Sylvester Resultant

The efficiency arises from exploiting the shift-invariant structure of the block-Toeplitz matrices $\{S_k, k=1,2,\cdots\}$. We can find the rank of any matrix by reducing the matrix to row echelon form using elementary row operations. The rank of the matrix is then the number of nonzero rows in its row echelon form and the rows of the echelon form span the row space of the original matrix.

So consider the matrix S_k. If the first block row has been reduced to echelon form, the shift invariance allows us to replace every lower block row by the echelon form of the first block row (shifted to the right) as an intermediate step to the echelon reduction of S_k. This follows from the spanning property of the echelon form together with the shift invariance.

Now, reduce the second block still further using the first block row, allowing only those elementary operations that add rows from the first block row into the second block row or that add only within the second block. When this is completed the first two block rows, taken together, are in row echelon form (with some row orders permuted) as is the first block row itself. We replace all block rows lower than the second by the second block row of the echelon form, suitably shifted, and proceed with the reduction of the third block row using the first two, and so on. We note that zero rows may be removed as they occur, since all information is available from the nonzero rows.

This procedure produces a "shifted row echelon form," E_k, that has the same block structure as S_k. We note that r_k equals the number of nonzero rows in E_k, and $r_{k+1} - r_k$ equals the number of nonzero rows in the final block of E_{k+1}. We claim (see Theorem 1 below) that the final block row of E_{k+1}, for k greater than the maximum dual dynamical index, defines the coefficients of a gcrd.

An example will clarify the procedure.

Example: As an illustration of the above algorithm we consider the transfer function

$$H(s) = D(s)C^{-1}(s) = \begin{bmatrix} 2s+1 & s^2+1 \\ s^2+2s+1 & s^2+2s \end{bmatrix}\begin{bmatrix} 2s^2+3s+5 & s^3+4s+1 \\ s^2+s-1 & s^2+s-1 \end{bmatrix}^{-1}. \quad (3.1)$$

We form

$$S_1 = \begin{bmatrix} 0 & 0 & 0 & 1 & 2 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 2 & 2 & 1 & 0 \\ 0 & 1 & 2 & 0 & 3 & 4 & 5 & 2 \\ 0 & 0 & 1 & 1 & 1 & 1 & -1 & -1 \end{bmatrix} \quad (3.2)$$

and reduce it to row echelon form E_1, before extending it by E_1 shifted (note that rank S_1 = rank E_1 = 4). We have

$$\left[\begin{array}{cccccccccc} 0 & 1 & 2 & 0 & 3 & 4 & 5 & 2 & 0 & 0 \\ \cdot & \cdot & 1 & 1 & 1 & 1 & -1 & -1 & 0 & 0 \\ \cdot & \cdot & \cdot & 1 & 2 & 0 & 1 & 1 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & 1 & 1 & 2 & 1 & 0 & 0 \\ \hline \cdot & \cdot & \cdot & 1 & 2 & 0 & 3 & 4 & 5 & 2 \\ \cdot & \cdot & \cdot & \cdot & 1 & 1 & 1 & 1 & -1 & -1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 2 & 0 & 1 & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 1 & 2 & 1 \end{array}\right]. \quad (3.3)$$

Now reduce this matrix to "shifted row echelon" form, E_2, by subtracting multiples of the first four rows from the last four rows and working within the last four rows. One row becomes zero, and as a result is deleted, leaving three rows in the last block row of E_2. Then extend E_2 by the nonzero rows left from the last block shifted two places to the right. (Note that rank S_2 = 7.)

$$\left[\begin{array}{cccccccccccc} 0 & 1 & 2 & 0 & 3 & 4 & 5 & 2 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & 1 & 1 & 1 & -1 & -1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & 2 & 0 & 1 & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & 1 & 2 & 1 & \cdot & \cdot & \cdot & \cdot \\ \hline \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 2 & 3 & 5 & 2 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 1 & 0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 2 & 0 & 1 & 1 & \cdot & \cdot \\ \hline \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 2 & 3 & 5 & 2 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 1 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 2 & 0 & 1 & 1 \end{array}\right] \quad (3.4)$$

and reduce this to shifted row echelon form

$$E_3 = \left[\begin{array}{cccccccccccc} 0 & 1 & 2 & 0 & 3 & 4 & 5 & 2 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & 1 & 1 & 1 & -1 & -1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & 2 & 0 & 1 & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 & 1 & 2 & 1 & \cdot & \cdot & \cdot & \cdot \\ \hline \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 2 & 3 & 5 & 2 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 1 & 0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 2 & 0 & 1 & 1 & \cdot & \cdot \\ \hline \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 2 & 3 & 5 & 2 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 1 & 0 \end{array}\right]. \quad (3.5)$$

Since rank S_3 − rank S_2 = 2 = dim $C(s)$, we stop here and the final two rows, as we shall prove, give a gcrd $R(s)$ of $C(s)$ and $D(s)$ as

$$R(s) = \begin{bmatrix} 2s+5 & 3s+2 \\ 1 & s \end{bmatrix}. \quad (3.6)$$

We may obtain a minimal right MFD $\{\bar{C}(s), \bar{D}(s)\}$ as follows:

$$\bar{D}(s) = D(s)R^{-1}(s) = \begin{bmatrix} 0.5 & s-1.5 \\ 0.5s & -0.5s+1 \end{bmatrix};$$

$$\bar{C}(s) = \begin{bmatrix} 0.5s-1 & s^2-1.5s \\ 0.5s-0.5 & -0.5s+1.5 \end{bmatrix}.$$

The sequence of ranks $\{r_k\}$ of the generalized Sylvester matrices for the example is $\{4,7,9\}$ which, in view of Theorem 1, immediately allows us to state that $\alpha_0 = 0$, $\alpha_1 = \alpha_2 = 1$, $\alpha_3 = \alpha_4 = \cdots = 0$. Thus, the dual dynamical indices are $\{1,2\}$.

We may note that the row operations to find null vectors of S_3 were

$$\begin{bmatrix} 0 & 0 \vdots 0 & 0 \vdots 1 & -1 \vdots 0 & 1 \vdots 0 & 3 \vdots -1 & -2 \\ 0 & 2 \vdots 0 & -2 \vdots 1 & -5 \vdots 0 & 3 \vdots 2 & 3 \vdots -1 & 0 \end{bmatrix} S_3 = 0.$$

Accordingly, the matrix on the left above represents the coefficients matrix of $E(s)$, from which $\{A(s), B(s)\}$ and an irreducible left MFD $-A^{-1}(s)B(s)$ can be constructed.

The justification of the gcrd calculation is provided by the following result.

Theorem 2. [19]: Let C and D be two polynomial matrices (with the same number of columns) and let S_k be the corresponding generalized Sylvester matrix of order k, with E_k the shifted row echelon form of S_k as derived via the algorithm. Then, provided k is greater than the maximum row degree ν of any dual basis, the nonzero rows of the last block row of E_k are the coefficients of a gcrd G_k of C and D.

Proof: The key step is to show that the polynomial matrix whose coefficients make up the final block row of E_k is a unimodular multiple of $F(s)$. Then Theorem 1 will show that for $k > \nu$, we will have a collection of m nonzero rows that spans the row space of F. This is clearly a gcrd. The proof of the first statement above is involved and the reader is referred to [18] and [19] (which does have some typographical errors). □

We remark that our method for gcrd evaluation is a modified form of Gaussian elimination, with only partial pivoting. Because of this restriction, complete numerical stability of the algorithm cannot be assured. However, this seems to be a fundamental limitation, and we are not aware of any better method of finding a gcrd. We may note that a Smith form equivalent of the gcrd can be found in a more stable way—using orthogonal transformations—by the methods of [14]; these methods can also be used to find the dual dynamical indices.

It should be pointed out that our method is not restricted to polynomials with real coefficients but can be applied to coefficients from any field. In particular for finite fields, the present algorithm will probably be quite well-behaved from a numerical point of view.

To continue with our results here, we remark that having found a gcrd, a test for coprimeness of the given polynomials is to check that the gcrd is unimodular. This check can be performed in several ways, e.g., by computing its determinant. Perhaps a simpler method is to reduce the gcrd to row-reduced form (by elementary row operations)—then unimodularity will be equivalent to all the row degrees being zero. It turns out that one way of obtaining a row-reduced gcrd is just to continue the echelon-form reduction.

Theorem 3. Let C, D, S_k, E_k, G_k and ν be as in Theorem 2. Then for some constant $k_1 \geq \nu$ the gcrd G_k derived from the final block row of the shifted row echelon form of S_k is the same for all $k \geq k_1$ and is a row proper gcrd of C and D.

This constant, k_1, has been reached when all r nonzero rows of the last block row of E_k have pivot indices (column numbers of the leading nonzero elements) in different residue classes mod r.

Proof: We note the following properties.

1) The rows of a row proper gcrd G have minimal degree in the class of all matrices which generate the left ideal generated by C and D. (This is seen by considering degree det G.)

2) There exists a number k_1 such that $G = MC + ND$ where M and N are polynomial matrices of degree k_1 or less.

3) The row degrees of the gcrd G_k are nonincreasing with k. (This is a consequence of the algorithm.)

4) If there is a row degree reduction between the minimal row degree generator derivable from S_k and that derivable from S_{k+1}, then it will occur as a row degree reduction from G_k to G_{k+1}. (This follows because the rows of the echelon form E_k, as a whole, represent a set with maximal ordered pivot indices of all possible linear combinations of rows of S_k and should any change occur in these indices between E_k and E_{k+1} it must occur in E_{k+1} because of the nature of the algorithm.)

The occurrence of the row proper gcrd for $k = k_1$ is established by properties 1), 2), and 4). The constancy for $k > k_1$ is proven by 3), and the final property of the pivot indices is seen to be implicit in 3) and 4). □

IV. The Generalized Bezoutian Resultant Matrix

A. Resultant and gcrd Calculation

In this section we establish a connection between the row echelon forms of the generalized Sylvester and the generalized Bezoutian matrix of [20], which will show that the generalized Bezoutian matrix mirrors many of the properties of the generalized Sylvester matrix.

Consider a quadruple of polynomial matrices $\{A(s), B(s), C(s), D(s)\}$, related by $AD - BC = 0$, where A and B have q rows, C and D have r columns.

The generalized Bezoutian form associated with $\{A, B, C, D\}$ is

$$\Gamma(x,y) = \frac{1}{x-y}[A(x)D(y) - B(x)C(y)] = \sum_{i=1}^{n} \sum_{j=1}^{m} \Gamma_{ij} x^{i-1} y^{j-1}$$

where n is the highest power of s in $[A\ B]$ and m that of $[C\ D]$. We then define the generalized Bezoutian matrix Δ as (Γ_{ij}).

A situation where such a quadruple of matrices might often arise is as obvious left and right MFD's of a given transfer function matrix $H(s)$ (see, e.g., [22], [23]). For example, if $H(s)$ is given as a matrix of scalar rational functions then we could take the denominator polynomial to be $a(s)I$, where $a(s)$ is the least common multiple of all denominators.

The following results are drawn from [20].

Lemma 1: Let Δ be a generalized Bezoutian matrix associated with the quadruple $\{A, B, C, D\}$, then rank $\Delta = n_{\min}[A\ B] = n_{\min}\begin{bmatrix} C \\ D \end{bmatrix}$. Consequently, either pair of matrices is coprime if and only if the highest degree minor has degree rank Δ.

Lemma 2: Let $\{A(s), B(s)\}$ and $\{C(s), D(s)\}$ be matrix pairs such that $AD - BC = 0$, let S_k be the generalized Sylvester resultant matrix of order k associated with $\{C(s), D(s)\}$ and let Δ be the generalized Bezoutian matrix associated with $\{A, B, C, D\}$. Then for all $k \geq n$, the degree of $[A\ B]$, S_k is row equivalent to

$$\left.\begin{matrix} k \\ \text{block} \\ \text{rows} \end{matrix}\right\{ \left[\begin{array}{cccccccc} C_0 & C_1 & \cdots & C_m & 0 & \cdots & & 0 \\ 0 & C_0 & \cdots & C_{m-1} & C_m & \cdots & & 0 \\ \vdots & & & & & & & \vdots \\ 0 & \cdots & 0 & C_0 & | \; C_1 & & \cdots & C_m \\ \hline 0_{nq \times kr} & & & & | & & \tilde{\Delta} & \end{array}\right] \tag{4.1}$$

where $\tilde{\Delta}$ is Δ with its block columns in reverse order.

From Lemma 2 and Theorem 2, we may expect that the coefficients of a gcrd are contained in the echelon form of the matrix in (4.1). But there is a difficulty in exploiting the idea, however; we still need a scheme to choose the required rows of the echelon form, a shift no longer being available to guide the choice. We shall look first at the regular case to see how the difficulty is overcome.

The regular case corresponds to the highest degree coefficient matrix of $[C'D']'$ having full rank, and in the remainder of this section we assume, without loss of generality, that if $[C'D']'$ is regular, then C is regular and there is an associated transfer function $H(s) = D(s)C^{-1}(s)$.

Theorem 4. [23]: With the same hypothesis as Lemma 2, and $C(s)$ regular, suppose the row echelon form of the following matrix is constructed:

$$\begin{bmatrix} C_0 & C_1 & \cdots & C_m \\ 0_{nq \times r} & & \tilde{\Delta} & \end{bmatrix}. \tag{4.2}$$

Then the coefficients of the rows of a row proper gcrd of $\{C, D\}$ are given by those rows of the echelon form whose pivot indices are the maximum in each residue class mod r.

Proof: The row equivalence of Lemma 2 together with the regularity of $C(s)$ imply that the echelon form of (4.2) contains the $r + n_{\min}$ rows of greatest pivot index from the echelon form of S_k and that the first r rows of the echelon form have pivot indices in differing residue classes mod r. Theorem 3 establishes the result. □

Next we consider nonregular $C(s)$. We choose some number σ such that $C(\sigma)$ is nonsingular—since $C(s)$ is a nonsingular polynomial matrix almost every σ will do—and construct the polynomial matrix $\hat{C}(s) = C(s+\sigma) = \hat{C}_0 s^m + \hat{C}_1 s^{m-1} + \cdots + \hat{C}_m$. These coefficients are simply related. Similarly we construct $\hat{D}(s) = D(s+\sigma), \hat{A}(s), \hat{B}(s)$ for the same σ.

Now we notice that, since $C(\sigma)$ is nonsingular so is $\hat{C}_m$ and consequently the matrix

$$\bar{C}(s) = s^m \hat{C}(s^{-1}) = \hat{C}_m s^m + \hat{C}_{m-1} s^{m-1} + \cdots + \hat{C}_0 \tag{4.3}$$

is a regular polynomial matrix. Also, construct $\bar{A}(s) = s^n \hat{A}(s^{-1}), \bar{B}(s) = s^n \hat{B}(s^{-1}), \bar{D}(s) = s^m \hat{D}(s^{-1})$. Thus, we have a quadruple of polynomial matrices $\{\bar{A}, \bar{B}, \bar{C}, \bar{D}\}$—associated with the rational matrix $H[(s+\sigma)^{-1}]$—with $\bar{C}$ regular. By Theorem 4, we may find a row-proper gcrd of $\{\bar{C}, \bar{D}\}$ by using the generalized Bezoutian matrix associated with $\{\bar{A}, \bar{B}, \bar{C}, \bar{D}\}$.

Our next problem is: Given a row proper gcrd of the polynomial matrices $\bar{C}$ and $\bar{D}$, how do we get a gcrd of $\hat{C}$ and $\hat{D}$? We answer this simply in the following theorem.

Theorem 5: Suppose we have four polynomial matrices $\hat{C}, \hat{D}, \bar{C}, \bar{D}$ related as follows:

$$\hat{C}(s) = s^m \bar{C}(s^{-1}) \qquad \hat{D}(s) = s^m \bar{D}(s^{-1})$$

where degree $\{\hat{C}, \hat{D}\} = m$ and $\hat{C}(0)$ is nonsingular.

Let $\bar{R}(s)$ be a row proper gcrd of $\{\bar{C}, \bar{D}\}$ with row degrees $n_1, n_2, \cdots, n_r$; then

$$\hat{R}(s) = \text{diag}\{s^{n_1}, s^{n_2}, \cdots, s^{n_r}\} \bar{R}(s^{-1})$$

is a gcrd (not necessarily row proper) of $\{\hat{C}, \hat{D}\}$. Hence, $\{\hat{C}, \hat{D}\}$ are right coprime if and only if $\{\bar{C}, \bar{D}\}$ are also.

Proof: Since $\bar{R}$ is a row proper gcrd of $\{\bar{C}, \bar{D}\}$ we have $\bar{C} = \bar{K}\bar{R}, \bar{D} = \bar{H}\bar{R}$ for right coprime polynomial matrices $\bar{K}, \bar{H}$. Forney's predictable degree property may then be used to show that $\hat{R}$ is a right divisor of both $\hat{C}$ and $\hat{D}$, and that the coprimeness of the associated matrices $\hat{K}, \hat{H}$ then follows, using the fact that $\hat{C}(0)$ has full rank. Thus $\hat{R}$ is indeed a gcrd of $\{\hat{C}, \hat{D}\}$. □

Given a gcrd $\hat{R}(s)$ of $\{\hat{C}, \hat{D}\}$ we may easily revert to a gcrd R of $\{C, D\}$ by shifting the origin back to $s=0$ from $s=\sigma$.

We remark here that there is a simple relationship between the Bezoutian matrix, $\hat{\Delta}$, associated with $\{\hat{A}, \hat{B}, \hat{C}, \hat{D}\}$ and that, $\bar{\Delta}$, associated with $\{\bar{A}, \bar{B}, \bar{C}, \bar{D}\}$. It may be seen easily by examining the generalized Bezoutian forms that $\hat{\Delta}$ equals $-\bar{\Delta}$ with block rows and block columns reversed. Since row operations do not affect the ultimate row echelon form, we really need only worry about the column reordering.

B. Dual Dynamical Indices Calculation

As with the generalized Sylvester resultant, the dual dynamical indices can also be obtained from the generalized Bezoutian matrix.

Theorem 6: Let $\bar{\Delta}$ be the generalized Bezoutian matrix associated with a left pair $\{\bar{A}(s), \bar{B}(s)\}$, with greatest common left divisor nonsingular at $s=0$,[4] and with any right pair. Then denoting by ρ_i the rank of the submatrix of $\bar{\Delta}$ formed from the first i block rows,

$$\alpha_k = (\rho_k - \rho_{k-1}) - (\rho_{k+1} - \rho_k) \tag{4.4}$$

where α_k is the number of dual dynamical indices (observability indices) of value k. [Note the parallel with (2.5).]

[4]As before, we may ensure this by moving the origin to $s=\sigma$.

Proof: The proof of Theorem 6 is involved and will only be outlined here. By considering the zero rows and the ranks $\bar{\Delta}_i$ of the generalized Bezoutian matrix associated with a left MFD which is left coprime and row proper, the result is easily established as $\bar{\Delta}$ then has only $n_{\min}$ nonzero rows. The extension to nonrow proper left MFD is achieved by examination of the relationship between this generalized Bezoutian matrix and that of the row proper case. Provided the gcld is nonsingular at $s=0$, rank $\bar{\Delta}_i$ does not change. □

V. Concluding Remarks

We have shown how a gcrd of two matrix polynomials can be computed by elementary operations on two real constant (i.e., nonpolynomial) matrices—the generalized Sylvester and Bezout matrices. No comparable methods seem to be available in the literature, though we may note that our method for the Sylvester resultant generalized a little-known method of Laidacker [24] (also cited in [7] and [25]) for finding the gcrd of two scalar polynomials. Another method sometimes used in the scalar case is Trudi's method [26], which, however, requires evaluation of polynomial determinants and does not seem to have a useful generalization to the matrix case.

References

[1] H. H. Rosenbrock, *State Space and Multivariable Theory*. London: Nelson, 1970.
[2] ——, *Computer-Aided Control System Design*. New York: Academic, 1974.
[3] W. A. Wolovich, *Linear Multivariable Systems*. New York: Spnger-Verlag, 1974.
[4] H. H. Rosenbrock, "Generalized resultant," *Electron. Lett.*, vol. 4, pp. 250–251, 1968.
[5] N. Munro, "Determination of the least order of transfer function matrices," *Proc. IEE*, vol. 118, pp. 1507–1510.
[6] S. Barnett, "Regular polynomial matrices having relatively prime determinants," *Proc. Cambridge Philos. Soc.*, vol. 65, pp. 585–590, 1969.
[7] S. Barnett, *Matrices and Control Theory*. London: Van Nostrand Reinhold, 1971.
[8] A. Rowe, "The generalized resultant matrix," *J. Inst. Math. Appl.*, vol. 9, pp. 390–396, 1972.
[9] E. Emre and L. M. Silverman, "Relatively prime polynomial matrices: Algorithms," in *Proc. IEEE Conf. on Decision and Control*, 1975, Houston, TX.
[10] ——, "New criteria and system theoretic interpretation for relatively prime polynomial matrices," *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 239–242, Apr. 1977.
[11] I. C. Gohberg and L. E. Lerer, "Resultants of matrix polynomials," *Bul. Am. Math. Soc.*, vol. 82, no. 4, pp. 565–567, July 1976.
[12] S.-H. Wang and E. J. Davison, "A minimization algorithm for the design of linear multivariable systems," *IEEE Trans. Automat. Contr.*, vol. AC-18, pp. 220–225, June 1973.
[13] E. J. Davison, W. Gesing, and S.-H. Wang, "An algorithm for obtaining the minimal realization of a linear time-invariant system and determining if a system is stabilizable-detectable," in *Proc. IEEE Conf. on Decision and Control, 1977*, New Orleans, LA, pp. 777–781.
[14] S.-Y. Kung, T. Kailath, and M. Morf, "Fast and stable algorithms for the minimal design problem," in *Proc. 4th IFAC Symp. on Multiv. Technological Syst.*, D. P. Atherton, Ed. London: Pergamon, 1979.
[15] C. C. MacDuffee, *The Theory of Matrices*. New York: Chelsea, 1956.
[16] H. Elliott, "Implementation of computer algorithms related to multivariable systems theory," M.S. dissertation, Div. Eng., Brown Univ., Providence, RI, June 1975.
[17] G. D. Forney, "Minimal bases of rational vector spaces, with application to multivariable linear systems," *SIAM J. Contr.*, vol. 13, pp. 493–520, 1975.
[18] S.-Y. Kung, "Multivariable and multidimensional systems: Analysis and design," Ph.D. dissertation, Dep. Elec. Eng., Stanford Univ., Stanford, CA, June 1977.
[19] S.-Y. Kung, T. Kailath, and M. Morf, "A generalized resultant matrix for polynomial matrices," in *Proc. IEEE Conf. on Decision and Control*, 1976, Florida, pp. 892–895.
[20] B. D. O. Anderson and E. I. Jury, "Generalized Bezoutian and Sylvester matrices in multivariable linear control," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 551–556, Aug. 1976.
[21] J. H. M. Wedderburn, "Lectures on matrices," *Am. Math. Soc. Colloq. Publications*, vol. 17, 1934.
[22] R. R. Bitmead and B. D. O. Anderson, "The matrix Cauchy index: Properties and applications," *J. SIAM App. Math.*, pp. 655–672, Dec. 1977.
[23] B. D. O. Anderson and R. R. Bitmead, "New tools in the matrix fraction description of linear systems," *Int. Conf. on Syst. Sci.*, Patras, Greece, 1976.
[24] M. A. Laidacker, "Another theorem relating Sylvester's matrix and the greatest common divisor," *Math. Mag.*, pp. 126–128, May 1969.
[25] E. I. Jury, *Inners and Stability of Dynamic Systems*. New York: Wiley, 1974.
[26] A. S. Householder, "Bezoutiants, elimination and localization," *SIAM Rev.*, pp. 73–78, Jan. 1970.