U. Helmke
Universität Regensburg
Fakultät Mathematik
8400 Regensburg 1, West Germany

and

P. A. Fuhrmann
Department of Mathematics
Ben-Gurion University of the Negev
Beer Sheva, Israel

Submitted by Jan C. Willems

ABSTRACT

We survey the theory of Bezoutians with a special emphasis on its relation to system theoretic problems. Some instances are the connections with realization theory, in particular signature symmetric realizations, the Cauchy index, stability, and the characterization of output feedback invariants. We describe canonical forms and invariants for the action of static output feedback on scalar linear systems of McMillan degree n. Previous results on this subject are obtained in a new and unified way, by making use of only a few elementary properties of Bezout matrices. As new results we obtain a minimal complete set of $2n-2$ independent invariants, an explicit example of a continuous canonical form for the case of odd McMillan degree, and finally a canonical form which induces a cell decomposition of the quotient space for output feedback.

1. INTRODUCTION

The Bezoutian is a rather special quadratic form which is defined for an arbitrary pair of polynomials. It was introduced in the middle of the 19th century and appears in the context of the theory of equations as a basic tool in the study of location of zeros for real and complex polynomials, in stability

LINEAR ALGEBRA AND ITS APPLICATIONS 122/123/124:1039–1097 (1989) 1039

655 Avenue of the Americas, New York, NY 10010 0024-3795/89/$3.50

theory as well as in classical elimination theory. The naming of Bezoutians goes back to Sylvester (1853), and it has been used to great advantage among others by Sylvester, by Cayley, and in a very powerful way by Hermite (1856). In spite of the fact that it is an extremely useful tool, the Bezoutian is hardly mentioned any more in general algebra texts, due to the trend in mathematics towards more abstraction. Van der Waerden (1931), Lang (1965), MacLane and Birkhoff (1967), and Hoffman and Kunze (1971) are all cases in point. In fact, even Gantmacher's classic treatise (1959) avoids the use of Bezoutians, for no obvious reason. Thus the study and use of Bezoutians has been relegated to special topic areas like the study of Toeplitz and Hankel matrix inversion, as in Heinig and Rost (1984), or to applied areas and in particular the area of system theory, especially stability theory. Some other important uses are in the analysis of determinantal representations of rational curves. For this connection we refer to the important paper by N. Kravitsky (1980) and to the work of Vinnikov (1986).

One reason for the decline of the Bezoutian in algebra texts may be the general decline of matrix theoretic methods and the quest for basis free statements. With this in mind we will stress the fact, proved in Fuhrmann (1981), that the Bezoutian is a matrix representation of a special module homomorphism over the ring of polynomials. Thus the study of the Bezoutian and its properties can be lifted out of its formal computational context and done in a mostly basis free way.

Our purpose in this paper is to restore the Bezoutian to a central place in both linear algebra and system theory by stressing its various applications. The uses of the Bezoutian in the system theoretic context are multifold. As the nonsingularity of the Bezoutian of two scalar polynomials is a criterion for coprimeness, the form makes direct contact with realization theory through controllability and observability tests. When we turn to questions concerning the Cauchy index and the signature of Hankel forms associated with rational functions, the associated Bezoutian gives equivalent results. However, the Bezoutian is much more amenable to computations. The reason for this is that the Bezoutian is linear in its arguments. This, coupled with the Euclidean algorithm, opens up the possibility of recursive computations, clarifying the relation to continued fractions and the establishing of recursive procedures for Hankel and Toeplitz matrix inversions. Maybe the most widely used classical area of application of the Bezoutian is the stability theory of an nth order homogeneous linear differential equation. Moreover the Bezoutian is related to geometric concepts as (A, B) invariant subspaces. With this we have not exhausted its uses in system theory.

Of particular interest to control theorists is a remarkable property of the Bezoutian, one which we will emphasize and exploit in this paper: It gives a complete set of invariants for the action of the static output feedback group on scalar transfer functions. This basic fact is our starting point, from which we will derive and extend, in a unified way, previous results on the output feedback problem. We will restrict ourselves completely to the simplest possible case, namely to the action of the full output feedback group

$$\frac{p(z)}{q(z)} \mapsto \frac{\alpha p(z)}{q(z) + \gamma p(z)}, \qquad \alpha \neq 0, \quad \gamma \in F,$$

on the space $\mathrm{Rat}(n)$ of scalar linear systems $g(z) = p(z)/q(z)$ of McMillan degree n.

A complete set of invariants for the restricted output feedback action

$$g(z) \mapsto \frac{g(z)}{1 + \gamma g(z)}, \qquad g \in \mathrm{Rat}(n),$$

was first given by Yannakoudakis (1981). In Byrnes and Crouch (1985) the authors give, based on classical algebraic geometry, necessary and sufficient conditions for equivalence under the full output feedback group. They also prove, over **R**, that there exists a *continuous* canonical form for output feedback if and only if the McMillan degree n is odd. However, no example of such a canonical form is given.

All these results are shown only for the scalar case. The multivariable case is more complicated and seems to require at least certain genericity assumptions. Hinrichsen and Prätzel-Wolters (1984) have constructed a quasicanonical form for output feedback of multivariable systems, which defines an honest canonical form on a certain generic subclass of linear systems. Similarly, Byrnes and Helton (1986) describe a different canonical form for output feedback, which is based on the matrix cross ratio and which is also defined only for a certain generic subclass of systems.

The paper is organized as follows. In Section 2 we briefly survey the basic facts concerning polynomial models. Section 3 reviews the basic facts about Bezoutians, stressing the polynomial model point of view. In Section 5 we study representation theorems for Bezoutians, relate them to Hankel matrices, and derive some classic connections between the two and their minors. The Kravitsky (1980) identity is an easy corollary of these representations. A connection with the Hurwitz determinants is also established.

In Section 8 we prove Heinig and Rost's (1984) result, Theorem 8.1, and obtain as corollaries several basic results of Yannakoudakis and of Byrnes and Crouch. A minimal set of $2n - 2$ complete (projective) invariants for output feedback is given. In the last section we construct an explicit example of a continuous canonical form for output feedback for the case of odd McMillan degree. Finally we describe a different canonical form, using continued fraction representations. This canonical form has some nice geometric properties, e.g., it leads to a cell decomposition of the quotient space of $\mathrm{Rat}(n)$ under the output feedback action.

With these applications towards the output feedback problem the potential of the Bezoutian has by no means been exhausted. Some topics where the Bezoutian point of view can be expected to be fruitful are:

(1) The output feedback problem in the multivariable case.

(2) The relation of the Bezoutians to Löwner matrices and the Cauchy interpolation problem.

(3) The classification of binary forms.

(4) The analysis of trace forms in algebra and number theory.

We wish to add that the classical Bezoutian has been extended to the multivariable setting by Anderson and Jury (1976). In this connection we point out the following references: Fuhrmann (1983), Wimmer (1989a, 1989b), Wimmer and Pták (1985), and Lerer and Tismenetsky (1982).

2. POLYNOMIAL MODELS

Throughout the appear we will denote by F an arbitrary commutative field. It might be identified later with the real number field R. By $F[z]$ we denote the ring of polynomials over F, by $F((z^{-1}))$ the set of truncated Laurent series in z^{-1}, and by $F[[z^{-1}]]$ and $z^{-1}F[[z^{-1}]]$ the set of all formal power series in z^{-1} and the set of those power series with vanishing constant term, respectively. Let π_+ and π_- be the projections of $F((z^{-1}))$ onto $F[z]$ and $z^{-1}F[[z^{-1}]]$ respectively. Since $F((z^{-1})) = F[z] \oplus z^{-1}F[[z^{-1}]]$, they are complementary projections. Also, $z^{-1}F[[z^{-1}]]$ is isomorphic to $F((z^{-1}))/F[z]$, which is an $F[z]$-module with the module action given by

$$z \cdot h = S_- h = \pi_- zh. \tag{1}$$

Similarly we define

$$S_+ f = zf \qquad \text{for} \quad f \in F[z]. \tag{2}$$

Given a monic polynomial q of degree n, we define a projection π_q in $F[z]$

by

$$\pi_q f = q\pi_- q^{-1} f \qquad \text{for} \quad f \in F[z]. \tag{3}$$

We define the *polynomial model* associated with q as the space

$$X_q = \operatorname{Im} \pi_q \tag{4}$$

endowed with the module structure induced by the shift map defined through

$$S_q f = \pi_q S_+ f \qquad \text{for} \quad f \in F[z], \tag{5}$$

and the *rational model* as the space

$$X^q = \operatorname{Im} \pi^q, \tag{6}$$

where π^q is the projection in $z^{-1}F[[z^{-1}]]$ defined by

$$\pi^q h = \pi_- q^{-1} \pi_+ qh \qquad \text{for} \quad h \in z^{-1}F[[z^{-1}]]. \tag{7}$$

X^q is a submodule of $z^{-1}F[[z^{-1}]]$ with the module structure given by

$$S^q h = S_- h \qquad \text{for} \quad h \in X^q. \tag{8}$$

The two models X_q and X^q associated with the polynomial q are isomorphic, the isomorphism being given by the map $\rho_q : X^q \to X_q$ defined by

$$\rho_q h = qh \qquad \text{for} \quad h \in X^q, \tag{9}$$

i.e. we have $\rho_q S^q = S_q \rho_q$.

A map Z in X_q commutes with S_q if and only if $Z = p(S_q)$ for some polynomial $p \in F[z]$ and $p(S_q)$ is invertible if and only if p and q are coprime. We define a pairing of elements of $F((z^{-1}))$ as follows: for

$$f(z) = \sum_{j=-\infty}^{n_f} f_j z^j$$

and

$$g(z) = \sum_{j=-\infty}^{n_g} g_j z^j$$

let

$$[f, g] = \sum_{j=-\infty}^{\infty} f_{-j-1} g_j. \tag{10}$$

Clearly, since both series are truncated, the sum in (10) is well defined. In terms of this pairing we can make the following identification; see Fuhrmann (1981). The dual of $F[z]$ as a linear space is $z^{-1}F[[z^{-1}]]$. Now, given a nonzero polynomial q the module X_q is isomorphic to $F[z]/qF[z]$. If we define, for a subset M of $F((z^{-1}))$, $M^{\perp}$ by

$$M^{\perp} = \left\{ g \in F\left((z^{-1})\right) \middle| [f, g] = 0 \text{ for all } f \in M \right\}, \tag{11}$$

then in particular $F[z]^{\perp} = F[z]$ and $(qF[z])^{\perp} = X^q$. Since, in general, $(X/M)^* \simeq M^{\perp}$, we have

$$X_q^* = (F[z]/qF[z])^* \simeq [qF[z]]^{\perp} = X^q. \tag{12}$$

But in turn we have $X^q \simeq X_q$, and so X_q^* can be identified with X_q. This can be made more concrete through the use of bilinear form

$$\langle f, g \rangle = \left[q^{-1}f, g \right]. \tag{13}$$

Relative to this bilinear form we have the important relation

$$S_q^* = S_q, \tag{14}$$

so that S_q is self-adjoint.

Let X be a finite dimensional vector space over the field F, and let X^* be its dual space under the pairing $\langle \ , \ \rangle$. Let $\{e_1, \dots, e_n\}$ be a basis for X. Then the set of vectors $\{f_1, \dots, f_n\}$ in X^* is called the *dual basis* if

$$\langle e_i, f_j \rangle = \delta_{ij}, \qquad 1 \leqslant i, j \leqslant n. \tag{15}$$

Let X_q be the polynomial model associated with the polynomial $q(z) = z^n + q_{n-1}z^{n-1} + \cdots + q_0$. The elements of X_q are all polynomials of degree $\leqslant n-1$. We consider the following very natural bases in X_q. The subset of X_q given by $B_{st} = \{f_1, \dots, f_n\}$, where

$$f_i(z) = z^{i-1}, \qquad i = 1, \dots, n, \tag{16}$$

is a basis for X_q. We will refer to this as the *standard basis*.

Given the polynomial q as above, we define

$$e_i(z) = \pi_+ z^{-i} q = q_i + q_{i+1}z + \cdots + z^n, \qquad i = 1, \dots, n. \tag{17}$$

and call the set $B_{co} = \{e_1, \dots, e_n\}$ is the *control basis* of X_q. The important fact about this pair of bases is that relative to the bilinear form $\langle \ , \ \rangle$ of Equation (13) the standard and control bases are dual to each other. In particular since $S_q^* = S_q$, we have $p(S_q)^* = p(S_q^*) = p(S_q)$, and so $p(S_q)$ is a self-adjoint operator in the indefinite metric $\langle \ , \ \rangle$. Thus the matrix representation of $p(S_q)$ relative to any dual pair of bases is symmetric.

The following theorem summarizes the most important properties of linear maps that commute with S_q.

THEOREM 2.1. *Let q be a monic polynomial of degree n, and let $S_q: X_q \to X_q$ be defined by Equation* (5). *Then*:

(i) *The map S_q is cyclic.*

(ii) *Let $Z_q: X_q \to X_q$ be any map commuting with S_q, i.e. any map satisfying*

$$ZS_q = S_q Z. \tag{18}$$

Then there exists a unique polynomial of degree $< n$ such that

$$Z = p(S_q). \tag{19}$$

(iii) *Let r be the g.c.d. of p and q, i.e. $p = rp_1$ and $q = rq_1$ with p_1, q_1 coprime. Then*

$$\operatorname{Ker} p(S_q) = p_1 X_r, \tag{20}$$

and

$$\operatorname{Im} p(S_q) = rX_{q_1}. \tag{21}$$

(iv) *The map $p(S_q)$ is invertible if and only if p and q are coprime.*

(v) *If p and q are coprime, let $a, b \in F[z]$ be solutions of the Bezout equation*

$$p(z)a(z) + q(z)b(z) = 1. \tag{22}$$

Then the inverse of $p(S_q)$ is $a(S_q)$. The polynomial a is uniquely determined provided we require that the condition $\deg a < \deg q$ *be satisfied.*

Proof. We prove only (v). From (22) it follows that

$$p(S_q)a(S_q) + b(S_q)q(S_q) = I, \tag{23}$$

and as $q(S_q) = 0$, we have

$$p(S_q)a(S_q) = I. \tag{24}$$

Note that the polynomials a and b solving Equation (22) can be found using the Euclidean algorithm. We will return to this later in Section 9. ∎

We note, for later use, that

$$\tilde{C}_q = [S_q]_{\mathrm{st}}^{\mathrm{st}} = \begin{pmatrix} 0 & \cdots & 0 & -q_0 \\ 1 & \cdots & 0 & -q_1 \\ & \ddots & \vdots & \vdots \\ & & 1 & -q_{n-1} \end{pmatrix} \tag{25}$$

and

$$C_q = [S_q]_{\mathrm{co}}^{\mathrm{co}} = \begin{pmatrix} 0 & 1 & & \\ & & \ddots & \\ & & & 1 \\ -q_0 & & \cdots & -q_{n-1} \end{pmatrix}, \tag{26}$$

i.e., we obtain the companion matrices as matrix representations.

Next we specialize to the case that the polynomial q has n simple roots. Thus

$$q(z)=\prod_{i=1}^{n}(z-\lambda_i)$$

and $\lambda_i \neq \lambda_j$. Now, as $(S_q f)(z)=zf(z)-\rho q(z)$ for some ρ, it follows that α is an eigenvalue of S_q, and f an eigenfunction, if and only if $q(\alpha)=0$ and $f(z)=\rho q(z)/(z-\alpha)$, i.e., α is equal to one of the λ_i.

Clearly $\{p_i(z)=q(z)/(z-\lambda_i)|i=1,\dots,n\}$ is a set of n linearly independent functions in X_q and hence constitutes a basis. We call this the *spectral basis* and denote it by B_{sp}. Obviously

$$p_i(z)=\frac{q(z)}{z-\lambda_i}=\prod_{j\neq i}(z-\lambda_j). \tag{27}$$

Finally we introduce the *interpolation basis* B_{in} in X_q. Let $\pi_1,\dots,\pi_n \in X_q$ be defined by the requirement

$$\pi_i(\lambda_j)=\delta_{ij}, \qquad i,j=1,\dots,n. \tag{28}$$

A simple calculation leads to

$$\pi_i(z)=\frac{p_i(z)}{p_i(\lambda_i)}.$$

Thus $B_{\mathrm{in}}=\{\pi_1,\dots,\pi_n\}$.

Now for an arbitrary polynomial f in X_q we have

$$\langle f,p_j\rangle=\left[q^{-1}f,p_j\right]=\left[f,q^{-1}p_j\right]=\left[f,(z-\lambda_j)^{-1}\right]=f(\lambda_j). \tag{29}$$

In particular

$$\langle \pi_i,p_j\rangle=\pi_i(\lambda_j)=\delta_{ij}. \tag{30}$$

So B_{in} is in fact the dual basis of B_{sp}.

The usage of the term interpolation is justified by the fact that $f(z)=\sum_{i=1}^{n}c_i\pi_i(z)$ is the unique polynomial solution, of degree $\leqslant n-1$, of the interpolation problem

$$f(\lambda_i) = c_i, \qquad i = 1, \dots, n. \tag{31}$$

This is just the *Lagrange interpolation* problem.

Note that for every f in X_q we have the expansion

$$f(z) = \sum_{i=1}^{n} f(\lambda_i) \pi_i(z) \tag{32}$$

and in particular

$$z^k = \sum_{i=1}^{n} \lambda_i^k \pi_i(z). \tag{33}$$

We define the Van der Monde matrix $V = V(\lambda_1, \dots, \lambda_n)$ by

$$V(\lambda_1, \dots, \lambda_n) = \begin{pmatrix} 1 & \cdots & 1 \\ \lambda_1 & \cdots & \lambda_n \\ \vdots & & \vdots \\ \lambda_1^{n-1} & \cdots & \lambda_n^{n-1} \end{pmatrix}. \tag{34}$$

Now Equation (33) can be written as

$$[I]_{\mathrm{st}}^{\mathrm{in}} = \tilde{V} \tag{35}$$

and so, by duality,

$$[I]_{\mathrm{sp}}^{\mathrm{co}} = V. \tag{36}$$

Incidentally the representation (36) of $V(\lambda_1, \dots, \lambda_n)$ is another proof of the nonsingularity of the Van der Monde matrix.

An easy corollary is the well-known diagonalization, by similarity, of the companion matrices.

COROLLARY 2.1. *Let q be as above, and C_q the companion matrix of* (26). *Then for $V = V(\lambda_1, \dots, \lambda_n)$ we have*

$$V^{-1} C_q V = \Lambda \tag{37}$$

and $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$.

Proof. From the equality $S_q I = IS_q$ we get

$$[S_q]_{\text{co}}^{\text{co}}[I]_{\text{sp}}^{\text{co}} = [I]_{\text{sp}}^{\text{co}}[S_q]_{\text{sp}}^{\text{sp}}, \tag{38}$$

or

$$C_q V = V\Lambda. \qquad \blacksquare \tag{39}$$

3. BEZOUTIANS

A quadratic form is associated with any symmetric matrix. We will focus now on quadratic forms induced by polynomials and rational functions. Specifically we will focus on Bezout and Hankel forms because of their connection to system theoretic problems like stability and symmetric realization theory.

Let $p(z) = \sum_{i=0}^{n} p_i z^i$ and $q(z) = \sum_{i=0}^{n} q_i z^i$ be two polynomials, and let z and w be two (generally noncommuting) variables. Then

$$\begin{aligned} p(z)q(w) - q(z)p(w) &= \sum_{i=0}^{n}\sum_{j=0}^{n} p_i q_j (z^i w^j - z^j w^i) \\ &= \sum_{0 \leqslant i < j \leqslant n} (p_i q_j - q_i p_j)(z^i w^j - z^j w^i). \end{aligned} \tag{40}$$

Observe now that

$$z^i w^j - z^j w^i = \sum_{\nu=0}^{j-i-1} z^{i+\nu}(w - z) w^{j-\nu-1}.$$

Thus Equation (40) can be rewritten as

$$p(z)q(w) - q(z)p(w) = \sum_{0 \leqslant i < j \leqslant n} (p_i q_j - p_j q_i) \sum_{\nu=0}^{j-i-1} z^{i+\nu}(w - z) w^{j-\nu-1} \tag{41}$$

or

$$q(z)p(w) - p(z)q(w) = \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij} z^{i-1}(z - w) w^{j-1}. \tag{42}$$

The last equality is obtained by changing the order of summation and properly defining the coefficients b_{ij}.

DEFINITION 3.1. The matrix (b_{ij}) defined by Equation (42) is called the *Bezout form* associated with the polynomials q and p, or just the *Bezoutian* of q and p, and is denoted by $B(q, p)$. If $g = p/q$ is the unique irreducible representation of a rational function g, with q monic, then we will write $B(g)$ for $B(q, p)$. $B(g)$ will be called the Bezoutian of g.

For a study of the functorial properties of the map $g \mapsto B(g)$ we refer to Helmke (1987). It has been noted by Kravitsky that Equation (42) holds even when z and w are a pair of noncommuting variables. This observation is extremely useful in the derivation of representation theorems for Bezoutians.

Note that $B(q, p)$ defines a bilinear form on F^n by

$$B(q,p)(\xi,\eta) = \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij}\xi_i\eta_j \qquad \text{for} \quad \xi,\eta \in F^n.$$

No distinction will be made between the matrix and the bilinear form.

The following theorem summarizes the elementary properties of the Bezoutian.

THEOREM 3.1. *Let* $q, p \in F[z]$ *with* $\max(\deg q, \deg p) \leqslant n$. *Then:*

(i) $B(q, p)$ *is a symmetric matrix.*
(ii) $B(q, p)$ *is linear in* q *and* p.
(iii) $B(p, q) = -B(q, p)$.

The following theorem is of central importance in that it reduces the study of Bezoutians to that of intertwining maps of the form $p(S_q)$. These are easier to handle and yield information on Bezoutians. Moreover, since the homomorphisms of polynomial models generalize to the multivariable case, we can extend the theory of Bezoutians to that context. Thus the analysis of the Bezoutian is reduced to the study of the map $p(S_q)$, which is much easier.

THEOREM 3.2. *Let* $p, q \in F[z]$, *with* $\deg p \leqslant \deg q$. *Then the Bezoutian* $B = B(q, p)$ *of* q *and* p *satisfies*

$$B(q,p) = \left[p(S_q)\right]_{\mathrm{co}}^{\mathrm{st}}. \tag{43}$$

Proof. Because of its importance we give two different proofs of the theorem. Our starting point is Equation (42). We note that in this form the equality holds for any pair of noncommuting variables. In particular we will choose for z and w linear maps.

We assume now that $\deg p \leqslant \deg q = n$ and choose a monic $r \in F[z]$ such that $\deg r = \deg q$ and r and q are coprime. This is always possible; in fact $r(z) = q(z)+1$ is such a polynomial. Furthermore let us substitute in Equation (42) $z = S_r$ and $w = S_q$. Obviously the polynomial models X_q and X_r are equal as sets, though they carry different module structures. Since $q(S_q) = 0$, it follows that

$$q(S_r)p(S_q) = \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}S_r^{i-1}(S_r - S_q)S_q^{j-1}. \tag{44}$$

Next we note that for every polynomial f of degree $\leqslant n-2$ we have

$$(S_r f)(z) = (S_q f)(z) = zf(z),$$

whereas

$$S_q z^{n-1} = \pi_q z^n = z^n - q(z) = -\sum_{i=0}^{n-1} q_i z^i.$$

Therefore we have

$$(S_r - S_q)z^i = \begin{cases} 0 & \text{for } i = 0,\dots,n-2, \\ q - r & \text{for } i = n-1. \end{cases} \tag{45}$$

Note that, since both q and r are monic, $q - r$ is of degree $\leqslant n-1$. It follows that we have the following matrix representation:

$$[S_r - S_q]_{\text{st}}^{\text{st}} = \begin{pmatrix} 0 & \cdots & 0 & q_0 - r_0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & q_{n-1} - r_{n-1} \end{pmatrix}$$

$$= \begin{pmatrix} q_0 - r_0 \\ \vdots \\ q_{n-1} - r_{n-1} \end{pmatrix} (0 \quad \cdots \quad 0 \quad 1).$$

Apply the map $q(S_r)$ to the polynomial 1 to obtain

$$q(S_r)1 = \pi_r q \cdot 1 = q - r.$$

Given two polynomials $b, c \in X_q$ we define the rank one operator $b \otimes \tilde{c}$ by

$$(b \otimes \tilde{c}) f = \langle f, c \rangle b. \tag{46}$$

Note that in terms of coordinates relative to the dual pair of bases we can write

$$\langle f, c \rangle = [\tilde{c}]^{\mathrm{st}} [f]^{\mathrm{co}}. \tag{47}$$

Hence from Equation (44) it follows that, given an arbitrary $f \in X_q$,

$$q(\mathrm{S}_r) p(\mathrm{S}_q) f = \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij} \mathrm{S}_r^{i-1} \big(q(\mathrm{S}_r) 1 \otimes \tilde{1} \big) \mathrm{S}_q^{j-1}.$$

But as

$$\begin{aligned} \mathrm{S}_r^{i-1} \big(q(\mathrm{S}_r) 1 \otimes \tilde{1} \big) \mathrm{S}_q^{j-1} f &= \langle \mathrm{S}_q^{j-1} f, 1 \rangle \mathrm{S}_r^{i-1} q(\mathrm{S}_r) 1 \\ &= \langle f, \mathrm{S}_q^{j-1} 1 \rangle q(\mathrm{S}_r) \mathrm{S}_r^{i-1} 1 \\ &= \langle f, z^{j-1} \rangle q(\mathrm{S}_r) z^{i-1}, \end{aligned}$$

we get, using the fact that $q(\mathrm{S}_r)$ is invertible, the equality

$$p(\mathrm{S}_q) f = \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij} \langle f, z^{j-1} \rangle z^{i-1}. \tag{48}$$

Hence

$$\big[p(\mathrm{S}_q) \big]_{\mathrm{co}}^{\mathrm{st}} [f]^{\mathrm{co}} = \big[p(\mathrm{S}_q) f \big]_{\mathrm{co}}^{\mathrm{st}} = \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij} [z^{i-1}]^{\mathrm{st}} \Big[\widetilde{z^{j-1}} \Big]^{\mathrm{st}} [f]^{\mathrm{co}},$$

or

$$\big[p(\mathrm{S}_q) \big]_{\mathrm{co}}^{\mathrm{st}} = \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij} E_{ij}, \tag{49}$$

where

$$E_{ij} = [z^{i-1}]^{\mathrm{st}} \Big[\widetilde{z^{j-1}} \Big]^{\mathrm{st}} \tag{50}$$

is the matrix whose i, j entry is 1 and all other entries zero. Thus the representation (43) follows.

For another proof note that

$$\pi_+ w^{-k}(z-w)w^{j-1} = \begin{cases} 0 & \text{if} \quad j<k, \\ -1 & \text{if} \quad j=k, \\ (z-w)w^{j-k-1} & \text{if} \quad j>k. \end{cases} \tag{51}$$

So from Equation (42) it follows that

$$q(z)\pi_+ w^{-k}p(w) - p(z)\pi_+ w^{-k}q(w)\big|_{w=z} = \sum_{i=1}^{n} b_{ik} z^{i-1}, \tag{52}$$

so that, with $\{e_1, \dots, e_n\}$ the control basis of X_q and p_k defined by

$$p_k(z) = \pi_+ z^{-k} p(z), \tag{53}$$

we have

$$q(z)p_k(z) - p(z)e_k(z) = \sum_{i=1}^{n} b_{ik} z^{i-1}. \tag{54}$$

Applying the projection π_q, we obtain

$$\pi_q p e_k = p(S_q)e_k = \sum_{i=1}^{n} b_{ik} z^{i-1}, \tag{55}$$

which, from the definition of a matrix representation, completes the proof. ■

Corollary 3.1. *Let $p, q \in F[z]$, with $\deg p < \deg q$. Then the last row and column of the Bezoutian $B(q,p)$ consist of the coefficients of p.*

Proof. By Equation (55) and the fact that the last element of the control basis satisfies $e_n(z) = 1$, we have

$$\sum_{i=1}^{n} b_{in} z^{i-1} = (\pi_q p e_n)(z) = (\pi_q p)(z) = p(z) = \sum_{i=0}^{n-1} p_i z^i, \tag{56}$$

i.e.

$$b_{in} = p_{i-1}, \qquad i = 1, \ldots, n. \tag{57}$$

The statement for rows follows from the symmetry of the Bezoutian. ∎

The next theorem presents some well-known facts concerning Bezoutians. The interest in the polynomial model approach is the characterization of the Bezoutian as a matrix representation of an intertwining map. This leads to a conceptual theory of Bezoutians, in contrast to the purely computational approach that has been prevalent in their study for a long time.

THEOREM 3.3. *Given two polynomials $p, q \in F[z]$, then*

(i) *$B(q, p)$ is invertible if and only if q and p are coprime;*
(ii) *$\dim(\operatorname{Ker} B(q, p))$ is equal to the degree of the g.c.d. of q and p.*

Proof. Part (i) follows from Theorem 2.1(iv) and Theorem 3.2. Similarly part (ii) follows from Theorem 2.1(iii). ∎

As an immediate consequence we derive some well-known results concerning Bezoutians. Previous proofs were all computational.

COROLLARY 3.2. *Let $p, q \in F[z]$, with $\deg p \leqslant \deg q$. Assume p has a factorization $p = p_1 p_2$. Let $q(z) = \sum_{i=0}^{n} q_i z^i$ be monic, i.e. $q_n = 1$. Let C_q be the companion matrix of q, and let K be the matrix*

$$K = \begin{pmatrix} q_1 & \cdot & \cdot & \cdot & \cdot & q_{n-1} & 1 \\ \cdot & & & & \cdot & 1 & \\ \cdot & & & \cdot & \cdot & & \\ \cdot & & \cdot & \cdot & & & \\ \cdot & \cdot & \cdot & & & & \\ q_{n-1} & 1 & & & & & \\ 1 & & & & & & \end{pmatrix} \tag{58}$$

Then

$$B(q, p_1 p_2) = B(q, p_1) p_2(\tilde{C}_q) = p_1(C_q) B(q, p_2), \tag{59}$$

$$B(p, q) = K p(\tilde{C}_q) = p(C_q) K, \tag{60}$$

$$B(p, q)\tilde{C}_q = C_q B(p, q). \tag{61}$$

Note that the matrix K in Equation (58) satisfies $K = B(q,1)$ and is both a Bezoutian and a Hankel matrix.

Proof. Note that for the standard and control bases of X_q we have

$$C_q = [S_q]_{\mathrm{st}}^{\mathrm{st}} \quad \text{and} \quad \tilde{C}_q = [S_q]_{\mathrm{co}}^{\mathrm{co}}. \tag{62}$$

(59): This follows from the equality

$$p(S_q) = p_1(S_q)p_2(S_q) \tag{63}$$

and the fact that

$$\begin{aligned} B(q, p_1p_2) &= [(p_1p_2)(S_q)]_{\mathrm{co}}^{\mathrm{st}} = [p_1(S_q)p_2(S_q)]_{\mathrm{co}}^{\mathrm{st}} \\ &= [p_1(S_q)]_{\mathrm{co}}^{\mathrm{st}}[p_2(S_q)]_{\mathrm{co}}^{\mathrm{co}} \\ &= [p_1(S_q)]_{\mathrm{st}}^{\mathrm{st}}[p_2(S_q)]_{\mathrm{co}}^{\mathrm{st}}. \end{aligned} \tag{64}$$

(60): This follows from (59) for the trivial factorization $p = p \cdot 1$.
(61): From the commutativity of S_q and $p(S_q)$ it follows that

$$[p(S_q)]_{\mathrm{co}}^{\mathrm{st}}[S_q]_{\mathrm{co}}^{\mathrm{co}} = [S_q]_{\mathrm{st}}^{\mathrm{st}}[p(S_q)]_{\mathrm{co}}^{\mathrm{st}}. \tag{65}$$

We note that

$$[S_q]_{\mathrm{co}}^{\mathrm{co}} = \left([S_q]_{\mathrm{st}}^{\mathrm{st}}\right)^{\tilde{}}. \qquad \blacksquare \tag{66}$$

The representation (60) for the Bezoutian is sometimes referred to as the *Barnett factorization.*

The next result, concerning diagonalization of Bezoutians by congruence transformations, is apparently well known. It is used in Datta (1978). In Heinig and Rost (1984) it is credited to Lander. In this connection we refer also to Furhmann and Datta (1987).

THEOREM 3.4. *Let $q(z)$ be a monic nth degree polynomial having n simple zeros $\lambda_1, \dots, \lambda_n$, and let p be a polynomial of degree $\leqslant n$. Then the*

Bezoutian $B(q, p)$ *satisfies the following identity*:

$$\tilde{V}B(q,p)V = R, \tag{67}$$

where R *is the diagonal matrix* $\mathrm{diag}(r_1,\dots,r_n)$ *and*

$$r_i = p(\lambda_i)q_i(\lambda_i) = p(\lambda_i)q'(\lambda_i). \tag{68}$$

Proof. The trivial operator identity

$$Ip(S_q)I = p(S_q) \tag{69}$$

implies the matrix equality

$$[I]^{\mathrm{in}}_{\mathrm{st}}[p(S_q)]^{\mathrm{st}}_{\mathrm{co}}[I]^{\mathrm{co}}_{\mathrm{sp}} = [p(S_q)]^{\mathrm{in}}_{\mathrm{sp}}. \tag{70}$$

As $S_q q_i = \lambda_i q_i$, it follows that

$$p(S_q)p_i = p(\lambda_i)p_i = p(\lambda_i)q_i(\lambda_i)\pi_i. \tag{71}$$

Now $q_i(\lambda_i) = \prod_{j\neq i}(\lambda_i - \lambda_j)$, but $q'(z) = \sum_{i=1}^n q_i(z)$, and hence

$$q'(\lambda_i) = \prod_{j\neq i}(\lambda_i - \lambda_j) = q_i(\lambda_i). \tag{72}$$

Thus

$$R = [p(S_q)]^{\mathrm{in}}_{\mathrm{sp}}, \tag{73}$$

and the result follows. ■

4. REALIZATION

Let $g(z)$ be a (not necessarily proper) rational function. We say that a quadruple of linear maps (E, A, B, C) is a *realization* of g if

$$G(z) = C(zE - A)^{-1}B; \tag{74}$$

it is a *minimal realization* if E and A are of minimal possible dimension.

A pencil $zE - A$ is called *nonsingular* if E and A are square matrices and $\det(zE - A)$ is not identically zero.

Two pencils $zE_1 - A_1$ and $zE_2 - A_2$ are called strictly equivalent if there exist nonsingular matrices P and R such that

$$P(zE_1 - A_1)R = zE_2 - A_2. \tag{75}$$

Clearly this is an equivalence relation. If E_1 is nonsingular, so is E_2, and in that case $PE_1(zI - E_1^{-1}A_1)RE_2^{-1} = zI - A_2E_2^{-1}$, so $(PE_1)^{-1}A = RE_2^{-1}$. Let $PE_1 = S$; then

$$S(zI - E_1^{-1}A_1)S^{-1} = zI - A_2E_2^{-1}. \tag{76}$$

It follows that E_1^{-1} and E_2^{-2} are similar.

The *McMillan degree* of g, denoted by $\delta(g)$, is defined as the dimension of a minimal realization.

THEOREM 4.1. *Let g be rational and $g = p/q$ with p and q coprime. Then* $\delta(g) = \operatorname{rank} B(q, p)$.

REMARK. The advantage of using the Bezoutian is that it does not require the properness of g.

Proof. Let $g = p/q$ with p and q coprime. In that case $\operatorname{rank} B(q, p) = n$. Let $p = aq + r$ with $\deg r < \deg rq$. So $p/q = a + r/q$ with r/q strictly proper, a a polynomial, and clearly r and q coprime. Taking minimal realizations of a and r/q of degrees m and $n - m$, respectively, we have

$$a(z) = \tilde{c}_\infty(I - zA_\infty)^{-1}b_\infty \tag{77}$$

and

$$g_-(z) = \frac{r(z)}{q(z)} = \tilde{c}(zI - A)^{-1}b. \tag{78}$$

Putting the two realizations together, we have

$$\begin{aligned} g(z) &= \begin{pmatrix} \tilde{c} & \tilde{c}_\infty \end{pmatrix} \begin{pmatrix} zI - A & 0 \\ 0 & I - zA_\infty \end{pmatrix}^{-1} \begin{pmatrix} b \\ b_\infty \end{pmatrix} \\ &= \tilde{c}(zI - A)^{-1}b + \tilde{c}_\infty(I - zA_\infty)^{-1}b_\infty \\ &= g_- + g_+. \end{aligned} \tag{79}$$

So $\delta(g) \leqslant n = \operatorname{rank} B(q, p)$.

Conversely, assume $\delta(g)=n$. Then $g(z)=\tilde{c}(zE-A)^{-1}b$ with E, A $n\times n$ matrices. Let $g=p/q$, with p and q coprime. We may, without loss of generality, assume the equality (79) holds. Let A_∞ be $m\times m$; then A is $(n-m)\times(n-m)$, and $g(z)=a(z)+r(z)/q(z)$ with $\deg q=n-m$. Thus $qg=qa+r=p$, and hence $g=p/q$ and p,q are coprime. From this it follows that rank $B(q,p)=n$. ∎

5. REPRESENTATION OF BEZOUTIANS

The relation (42) leads easily to some interesting representation formulas for the Bezoutian. These formulas were obtained by Kravitsky (1980) and Pták (1984). They are also implicit in Trench (1965). In this connection see also Fuhrmann (1986).

For a polynomial a of degree n we define the *reverse polynomial* $a^\#$ by

$$a^\#(z)=a(z^{-1})z^n. \tag{80}$$

THEOREM 5.1. *Let $p(z)=\sum_{i=0}^n p_i z^i$ and $q(z)=\sum_{i=0}^n z_i z^i$ be polynomials of degree n. Then the Bezoutian has the following representations:*

$$\begin{aligned}B(q,p)&=\left[p(\tilde{S})q^\#(S)-q(\tilde{S})p^\#(S)\right]J\\&=J\left[p(S)q^\#(\tilde{S})-q(S)p^\#(\tilde{S})\right]\\&=-J\left[p^\#(\tilde{S})q(S)-q^\#(\tilde{S})p(S)\right]\\&=-\left[p^\#(S)q(\tilde{S})-q^\#(S)p(\tilde{S})\right]J.\end{aligned} \tag{81}$$

Proof. We define the $n\times n$ shift matrix S by

$$S=\begin{pmatrix}0&1&0&\cdots&0&0\\ \cdot&\cdot&1&\cdots&0&0\\ \cdot&&\cdot&\ddots&\vdots&\vdots\\ \cdot&&&\cdot&1&0\\ \cdot&&&&\cdot&1\\ 0&\cdot&\cdot&\cdot&\cdot&\end{pmatrix}, \tag{82}$$

and the $n \times n$ transposition matrix J by

$$J = \begin{pmatrix} 0 & \cdot & \cdot & \cdot & \cdot & 0 & 1 \\ 0 & \cdot & \cdot & \cdot & \cdot & 1 & 0 \\ & \cdot & & & \cdot & \cdot & \cdot \\ & \cdot & & \cdot & \cdot & & \cdot \\ & \cdot & \cdot & \cdot & & & \cdot \\ 0 & 1 & \cdot & \cdot & \cdot & \cdot & 0 \\ 1 & 0 & \cdot & \cdot & \cdot & \cdot & 0 \end{pmatrix}. \tag{83}$$

Note also that for an arbitrary polynomial a we have

$$Ja(S)J = a(\tilde{S}). \tag{84}$$

From Equation (42) we easily obtain

$$\begin{aligned} p(z)q^{\sharp}(w) - q(z)p^{\sharp}(w) &= p(z)q(w^{-1})w^n - q(z)p(w^{-1})w^n \\ &= \left\{ p(z)q(w^{-1}) - q(z)p(w^{-1}) \right\} w^n \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij} z^{i-1}(z - w^{-1}) w^{-j+1} w^n \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij} z^{i-1}(zw - 1) w^{n-j}. \end{aligned} \tag{85}$$

In this identity we substitute now $z = \tilde{S}$ and $w = S$. Thus we get for the central term

$$(\tilde{S}S - I) = -\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}, \tag{86}$$

and

$$\tilde{S}^{i-1}(\tilde{S}S - I)S^{n-j} \tag{87}$$

is the matrix whose only nonzero term is -1 in the $i, n-j$ position. This implies the identity

$$B(q,p) = \left\{ p(\tilde{S})q^{\sharp}(S) - q(\tilde{S})p^{\sharp}(S) \right\} J. \tag{88}$$

The other identities are similarly derived. ■

From the representation (88) of the Bezoutian we obtain, by expanding the matrices, the *Gohberg-Semencul* (1972) formula

$$\begin{aligned}
B(q,p) &= \left(p(\tilde{S})q^{\sharp}(S) - q(\tilde{S})p^{\sharp}(S)\right)J \\
&= \left[\begin{pmatrix} p_0 & & 0 \\ \vdots & \ddots & \\ p_{n-1} & \cdots & p_0 \end{pmatrix}\begin{pmatrix} q_n & \cdots & q_1 \\ & \ddots & \vdots \\ 0 & & q_n \end{pmatrix}\right. \\
&\quad \left. - \begin{pmatrix} q_0 & & \\ \vdots & \ddots & \\ q_{n-1} & \cdots & q_0 \end{pmatrix}\begin{pmatrix} 0 & p_{n-1} & \cdots & p_1 \\ & & \ddots & \vdots \\ & & & p_{n-1} \end{pmatrix}\right]J \\
&= \left[\begin{pmatrix} p_0 & & 0 \\ \vdots & \ddots & \\ p_{n-1} & \cdots & p_0 \end{pmatrix}\begin{pmatrix} q_1 & \cdots & q_n \\ \vdots & ⋰ & \\ q_n & & 0 \end{pmatrix}\right. \\
&\quad \left. - \begin{pmatrix} q_0 & & 0 \\ \vdots & \ddots & \\ q_{n-1} & \cdots & q_0 \end{pmatrix}\begin{pmatrix} p_1 & \cdots & p_{n-1} & 0 \\ \vdots & ⋰ & & \\ p_{n-1} & & & \\ 0 & & & 0 \end{pmatrix}\right].
\end{aligned} \tag{89}$$

Given two polynomials p and q of degree n, we let their *Sylvester resultant* $\operatorname{Res}(p,q)$ be defined by

$$\operatorname{Res}(p,q) = \begin{pmatrix} p_0 & \cdots & p_{n-1} & p_n & & \\ & \ddots & \vdots & \vdots & \ddots & \\ & & p_0 & p_1 & \cdots & p_n \\ q_0 & \cdots & q_{n-1} & q_n & & \\ & \ddots & \vdots & \vdots & \ddots & \\ & & q_0 & q_1 & \cdots & q_n \end{pmatrix}. \tag{90}$$

It is well known that the resultant $\operatorname{Res}(p,q)$ is nonsingular if and only if p and q are coprime. Equation (90) can be rewritten as the 2×2 block matrix

$$\operatorname{Res}(p,q) = \begin{pmatrix} p(S) & p^{\sharp}(\tilde{S}) \\ q(S) & q^{\sharp}(\tilde{S}) \end{pmatrix}, \tag{91}$$

where the reverse polynomials $p^{\sharp}$ and $q^{\sharp}$ are defined in Equation (80).

Based on the preceding, Kravitsky's result can be stated as

THEOREM 5.2. *Let* $p(z)=\sum_{i=0}^{n}p_iz^i$ *and* $q(z)=\sum_{i=0}^{n}q_iz^i$ *be polynomials of degree* n. *Then*

$$\widetilde{\operatorname{Res}(p,q)}\begin{pmatrix}0 & J\\ -J & 0\end{pmatrix}\operatorname{Res}(p,q)=\begin{pmatrix}0 & B(p,q)\\ -B(p,q) & 0\end{pmatrix}. \tag{92}$$

Proof. By expanding the left side of (92) we have

$$\begin{pmatrix}p(\tilde{S}) & q(\tilde{S})\\ p^{\sharp}(S) & q^{\sharp}(S)\end{pmatrix}\begin{pmatrix}0 & J\\ -J & 0\end{pmatrix}\begin{pmatrix}p(S) & p^{\sharp}(\tilde{S})\\ q(S) & q^{\sharp}(\tilde{S})\end{pmatrix}$$

$$=\begin{pmatrix}p(\tilde{S})Jq(S)-q(\tilde{S})Jp(S) & p(\tilde{S})Jq^{\sharp}(\tilde{S})-q(\tilde{S})Jp^{\sharp}(\tilde{S})\\ p^{\sharp}(S)Jq(S)-q^{\sharp}(S)Jp(S) & p^{\sharp}(S)Jq^{\sharp}(\tilde{S})-q^{\sharp}(S)Jp^{\sharp}(\tilde{S})\end{pmatrix}.$$

Now

$$p(\tilde{S})Jq(S)-q(\tilde{S})Jp(S)=J[p(S)q(S)-q(S)p(S)]=0.$$

Similarly

$$p^{\sharp}(S)Jq^{\sharp}(\tilde{S})-q^{\sharp}(S)Jp^{\sharp}(\tilde{S})=\left[p^{\sharp}(S)q^{\sharp}(S)-q^{\sharp}(S)p^{\sharp}(S)\right]J=0.$$

In the same way

$$p(\tilde{S})Jq^{\sharp}(\tilde{S})-q(\tilde{S})Jp^{\sharp}(\tilde{S})=J\left[p(S)q^{\sharp}(\tilde{S})-q(S)p^{\sharp}(\tilde{S})\right]=B(q,p).$$

Finally

$$p^{\sharp}(S)Jq(S)-q^{\sharp}(S)Jp(S)=J\left[p(\tilde{S})q^{\sharp}(S)-q^{\sharp}(\tilde{S})p(S)\right]=B(q,p).$$

This proves the theorem. ∎

As an immediate corollary we obtain the following

COROLLARY 5.1. *Let* p *and* q *be polynomials of degree* n. *Then*

$$|\det\operatorname{Res}(p,q)|=|\det B(p,q)|. \tag{93}$$

In particular the nonsingularity of either matrix implies that of the other. As *is also well known these conditions are equivalent to the coprimeness of the polynomials p and q.*

6. BEZOUT AND HANKEL MATRICES

In this section we explore some of the close relations between Bezout and Hankel matrices.

DEFINITION 6.1. Let $g = p/q$ be a strictly proper rational function. The *Hankel map* $H_g: F[z] \to z^{-1}F[[z^{-1}]]$ induced by g is defined by

$$H_g f = \pi_- gf \qquad \text{for} \quad f \in F[z]. \tag{94}$$

If we assume p and q in the representation $g = p/q$ are coprime, then H_g is not invertible, as it has a large kernel, given by

$$\operatorname{Ker} H_g = qF[z] \tag{95}$$

and a small image, given by

$$\operatorname{Im} H_g = X^q. \tag{96}$$

However, if we define a map $\overline{H}: X_q \to X^q$ by

$$\overline{H}f = H_g f \qquad \text{for} \quad f \in X_q, \tag{97}$$

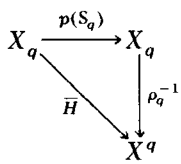
then, because

$$F[z] = X_q \oplus qF[z], \tag{98}$$

$\overline{H}$ becomes an invertible map.

The relation of the map $\overline{H}$ to the intertwining map $p(S_q)$ is given by the following.

THEOREM 6.1. *Let the map* $\rho_q: X^q \to X_q$ *be defined by Equation* (9), *and let* $\overline{H}$ *be defined as above. Then the following diagram is commutative*:



Proof. We compute, for $f \in X_q$,

$$\rho^{-1}p(S_q)f = q^{-1}\pi_q pf = q^{-1}q\pi_- q^{-1}pf = \pi_- gf = H_g f = \overline{H}f. \quad \blacksquare \qquad (100)$$

From the previous diagram we obtain a result of Lander (1974) characterizing the inverse of finite, nonsingular Hankel matrices as Bezoutians. We will consider *minimal rational extensions* of the sequence $g_1, \dots, g_{2n-1}$, i.e. strictly proper rational functions of the form

$$g(z) = \frac{p(z)}{q(z)} = \sum_{i=1}^{\infty} g_i z^{-i}, \qquad (101)$$

with p and q coprime and q of minimal degree.

THEOREM 6.2. *Let H be the Hankel matrix*

$$H = \begin{pmatrix} g_1 & \cdots & g_n \\ \vdots & & \vdots \\ g_n & \cdots & g_{2n-1} \end{pmatrix}, \qquad (102)$$

which is assumed to be nonsingular. Let $g = p/q$ be any minimal rational extension of the sequence $g_1, \dots, g_{2n-1}$ with p and q coprime. Let a be the unique polynomial of degree $< n$ satisfying the Bezout equation (22). *Then we have*

$$H^{-1} = B(q, a). \qquad (103)$$

Proof. Since g is a minimal rational extension of the sequence $g_1, \dots, g_{2n-1}$, we must have $\deg q = n$. Let us write

$$q(z) = z^n + q_{n-1}z^{n-1} + \cdots + q_0. \qquad (104)$$

From Equation (100) we obtain

$$\overline{H} = \rho^{-1}p(S_q). \qquad (105)$$

If a is the polynomial appearing in the Bezout identity (22), then clearly we obtain the important relationship

$$\overline{H}a(S_q) = \rho_q^{-1}. \tag{106}$$

The proof is completed by taking the right matrix representations in Equation (106). We consider the standard and control bases in X_q and B_{rc}, the basis of X^q obtained as the image of B_{co} under the isomorphism ρ_q^{-1}, i.e. $B_{\mathrm{rc}} = \{q^{-1}e_1, \dots, q^{-1}e_n\}$. Obviously we have

$$\left[\rho_q^{-1}\right]_{\mathrm{co}}^{\mathrm{rc}} = I. \tag{107}$$

From Equation (106) it follows therefore that

$$[\overline{H}]_{\mathrm{st}}^{\mathrm{rc}}\left[a(S_q)\right]_{\mathrm{co}}^{\mathrm{st}} = I. \tag{108}$$

Now, by Theorem 3.2,

$$\left[a(S_q)\right]_{\mathrm{co}}^{\mathrm{st}} = B(q, a). \tag{109}$$

To complete the proof we will show that $[\overline{H}]_{\mathrm{st}}^{\mathrm{rc}}$ is equal to the Hankel matrix H in (102). Indeed, if h_{ij} is the i, j element of $[\overline{H}]_{\mathrm{st}}^{\mathrm{rc}}$, then we have, by the definition of a matrix representation of a linear transformation, that

$$\overline{H}z^{j-1} = \pi_- pq^{-1}z^{j-1} = \sum_{i=1}^{n} \frac{h_{ij}e_i}{q}. \tag{110}$$

So

$$\sum_{i=1}^{n} h_{ij}e_i = q\pi_- q^{-1}pz^{j-1} = \pi_q pz^{j-1}. \tag{111}$$

Using the fact that B_{co} is the dual basis of B_{st} under the pairing (13), we have

$$\begin{aligned} h_{kj} &= \sum_{i=1}^{n} h_{ij}\langle e_i, z^{k-1}\rangle = \langle \pi_q pz^{j-1}, z^{k-1}\rangle \\ &= \left[q^{-1}q\pi_- q^{-1}pz^{j-1}, z^{k-1}\right] = \left[\pi_- q^{-1}pz^{j-1}, z^{k-1}\right] \\ &= \left[gz^{j-1}, z^{k-1}\right] = \left[g, z^{j+k-2}\right] = g_{j+k-1}. \qquad \blacksquare \end{aligned} \tag{112}$$

Thus we have not only shown that the inverse of a Hankel matrix, if it exists, is a Bezoutian matrix, but we have also identified the corresponding polynomials. While there are many minimal rational extensions, they all naturally lead to the same inverse for H.

Given two coprime polynomials p and q with $\deg p \leqslant \deg q = n$, it goes back to Hermite (1856) that

$$\sigma(H_{p/q}) = \sigma(B(q,p)), \tag{113}$$

where σ denotes signature. For this, and results related to realization theory and the Hermite-Hurwitz theorem, see Fuhrmann (1983). Now, by Sylvester's law of inertia, it follows that the quadratic forms $H_{p/q}$ and $B(q,p)$ are congruent. It is of interest to find such a specific congruence relation, and this we proceed to do. In a slightly different formulation the following results appear in Krein and Naimark (1936). In their paper the connection between Bezout and Hankel forms is credited to Hermite himself.

THEOREM 6.3. *Let $g = p/q$ with p and q coprime polynomials satisfying $\deg p < \deg q = n$. Then, if $g(z) = \sum_{i=1}^{\infty} g_i z^{-i}$, we have*

$$\begin{pmatrix} g_1 & \cdots & g_n \\ \vdots & & \vdots \\ g_n & \cdots & g_{2n-1} \end{pmatrix} = \begin{pmatrix} & & & 1 \\ & & \cdot^{\cdot^{\cdot}} & \psi_1 \\ & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \vdots \\ 1 & \psi_1 & \cdots & \psi_{n-1} \end{pmatrix} \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix} \begin{pmatrix} & & & 1 \\ & & \cdot^{\cdot^{\cdot}} & \psi_1 \\ & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \vdots \\ 1 & \psi_1 & \cdots & \psi_{n-1} \end{pmatrix}, \tag{114}$$

$$\begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix} = \begin{pmatrix} q_1 & \cdots & q_{n-1} & 1 \\ \vdots & \cdot^{\cdot^{\cdot}} & 1 & \\ q_{n-1} & \cdot^{\cdot^{\cdot}} & & \\ 1 & & & \end{pmatrix} \times \begin{pmatrix} g_1 & \cdots & g_n \\ \vdots & & \vdots \\ g_n & \cdots & g_{2n-1} \end{pmatrix} \begin{pmatrix} q_1 & \cdots & q_{n-1} & 1 \\ \vdots & \cdot^{\cdot^{\cdot}} & 1 & \\ q_{n-1} & \cdot^{\cdot^{\cdot}} & & \\ 1 & & & \end{pmatrix}. \tag{115}$$

Note that equation (115) can be written more compactly as

$$B(q,p) = B(q,1)H_{p/q}B(q,1). \tag{116}$$

Proof. To this end recall the diagram (99) and also the definition of the bases B_{st}, B_{co}, and B_{rc}. Clearly the operator equality

$$\bar{H} = \rho_q^{-1}p(S_q) \tag{117}$$

implies a variety of matrix representations, depending on the choice of bases. In particular it implies

$$[\bar{H}]_{st}^{rc} = [\rho_q^{-1}]_{st}^{rc}[p(S_q)]_{co}^{st}[I]_{st}^{co} = [\rho_q^{-1}]_{co}^{rc}[I]_{st}^{co}[p(S_q)]_{co}^{st}[I]_{st}^{co}. \tag{118}$$

Now it was proved in Theorem 6.2 that

$$[\bar{H}]_{st}^{rc} = H_n = \begin{pmatrix} g_1 & \cdots & g_n \\ \vdots & & \vdots \\ g_n & \cdots & g_{2n-1} \end{pmatrix}. \tag{119}$$

Obviously $[\rho_q^{-1}]_{co}^{rc} = I$ and, using matrix representations for dual maps,

$$[I]_{st}^{rc} = [\tilde{I}]_{st}^{co} = [\tilde{I}]_{co}^{st}. \tag{120}$$

Here we used the fact that B_{st} and B_{co} are dual bases. So $[I]_{co}^{st}$ is symmetric. In fact it is a Hankel matrix, easily computed to be

$$[I]_{co}^{st} = \begin{pmatrix} q_1 & \cdot & \cdot & \cdot & q_{n-1} & 1 \\ \cdot & & & \cdot & \cdot & \\ \cdot & & \cdot & \cdot & & \\ \cdot & \cdot & \cdot & & & \\ q_{n-1} & \cdot & & & & \\ 1 & & & & & \end{pmatrix}, \tag{121}$$

and so actually

$$[I]_{st}^{co} = \widetilde{[I]_{st}^{co}} = R. \tag{122}$$

Finally, from the identification of the Bezoutian as the matrix representation of an intertwining map for S_q

$$B(q,p) = [p(S_q)]_{co}^{st}, \tag{123}$$

we obtain the equality

$$H_n = \tilde{R}B(q,p)R \tag{124}$$

with R defined by (122). ■

COROLLARY 6.1. *Let $g = p/q$ with p and q coprime polynomials satisfying $\deg p < \deg q = n$. If $g(z) = \sum_{i=1}^{\infty} g_i z^{-i}$ and*

$$H_n = \begin{pmatrix} g_1 & \cdots & g_n \\ \vdots & & \vdots \\ g_n & \cdots & g_{2n-1} \end{pmatrix}, \tag{125}$$

then

$$\det H = \det B(q,p). \tag{126}$$

Proof. Clearly $\det R = (-1)^{n(n-1)/2}$ and so $(\det R)^2 = (-1)^{n(n-1)} = 1$, and hence (126) follows from (124). Since obviously

$$[I]_{st}^{co}[I]_{co}^{st} = [I]_{co}^{co} = I, \tag{127}$$

we have

$$[I]_{st}^{co} = \left([I]_{co}^{st}\right)^{-1}. \quad \blacksquare \tag{128}$$

The matrix $[I]_{st}^{co}$ has a polynomial characterization.

LEMMA 6.1. *Let $q(z) = z^n + q_{n-1}z^{n-1} + \cdots + q_0$, and B_{co}, B_{st} be the control and standard bases respectively of X_q. Then*

$$R = [I]_{st}^{co} = \begin{pmatrix} & & & & 1 \\ & & & \cdot & \psi_1 \\ & & \cdot & \cdot & \vdots \\ & \cdot & \cdot & & \vdots \\ 1 & \psi_1 & \cdot & \cdot & \psi_{n-1} \end{pmatrix}, \tag{129}$$

where, for $\phi(z) = q^\sharp(z) = z^n q(z^{-1})$, $\psi(z) = \psi_0 + \cdots + \psi_{n-1}z^{n-1}$ is the unique solution, of degree $< n$, of the Bezout equation

$$\phi(z)\psi(z) + z^n\sigma(z) = 1. \tag{130}$$

With J the transposition matrix defined in (82) we compute

$$J[I]_{st}^{co} = \left([I]_{co}^{st}J\right)^{-1} = \left(\phi(S_{z^n})\right)^{-1}. \tag{131}$$

However, if we consider the map S_{z^n}, then

$$\begin{pmatrix} 1 & q_{n-1} & \cdot & \cdot & q_1 \\ & \cdot & \cdot & & \cdot \\ & & \cdot & \cdot & \vdots \\ & & & \cdot & q_{n-1} \\ & & & & 1 \end{pmatrix} = q^\sharp(S_{z^n}) = \phi(S_{z^n}), \tag{132}$$

and its inverse is given by $\psi(S_{z^n})$, where ψ solves the Bezout equation (130). This completes the proof. ■

The matrix identities appearing in Theorem 6.3 are interesting also as they include in them some other identities for the principal minors of H_n. We state it in the following way.

COROLLARY 6.2. *Under the previous assumption we have, for $k \leqslant n$,*

$$\begin{pmatrix} g_1 & \cdots & g_k \\ \vdots & & \vdots \\ g_k & \cdots & g_{2k-1} \end{pmatrix} = \begin{pmatrix} & & & 1 \\ & & \cdot^{\cdot^{\cdot}} & \psi_1 \\ & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \vdots \\ 1 & \psi_1 & \cdots & \psi_{k-1} \end{pmatrix}$$

$$\times \begin{pmatrix} b_{n-k+1\,n-k+1} & \cdots & b_{n-k+1\,n} \\ \vdots & & \vdots \\ b_{n\,n-k+1} & \cdots & b_{nn} \end{pmatrix} \begin{pmatrix} & & & 1 \\ & & \cdot^{\cdot^{\cdot}} & \psi_1 \\ & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \vdots \\ 1 & \psi_1 & \cdots & \psi_{k-1} \end{pmatrix}. \tag{133}$$

In particular we obtain

COROLLARY 6.3. *The rank of the Bezoutian of the polynomials q and p is equal to the rank of the Hankel matrix $H_{p/q}$ and hence to the order of the largest nonsingular principal minor, when starting from the lower right hand corner.*

Proof. Equation (133) implies the following equality:

$$\det\left(g_{i+j-1}\right)_{i,j=1}^{k} = \det\left(b_{ij}\right)_{i,j=n-k+1}^{n}. \qquad ■ \tag{134}$$

At this point we recall the connection, given in Equation (126), between the determinant of the Bezoutian and that of the resultant of the polynomials q and p. In fact this can not only be made more precise, but we can find relations between the minors of the resultant and thosé of the corresponding

Hankel and Bezout matrices. We state this as

THEOREM 6.4. *Let q be monic of degree n, and p of degree $\leqslant n-1$. Then the lower $k \times k$ right hand corner of the Bezoutian $B(q,p)$ is given by*

$$
\begin{pmatrix}
b_{n-k+1\, n-k+1} & \cdots & b_{n-k+1\, n} \\
\vdots & & \vdots \\
b_{n\, n-k+1} & \cdots & b_{nn}
\end{pmatrix}
$$

$$
= \begin{pmatrix}
p_{n-k} & \cdots & p_0 & & \\
\vdots & & & \ddots & \\
\vdots & & & & p_0 \\
\vdots & & & & \vdots \\
p_{n-k+1} & \cdots & \cdots & \cdots & p_{n-1}
\end{pmatrix}
\begin{pmatrix}
q_{n-k+1} & \cdots & \cdots & q_{n-1} & 1 \\
\vdots & & \cdot^{\cdot^{\cdot}} & 1 & \\
\vdots & \cdot^{\cdot^{\cdot}} & & & \\
q_{n-1} & 1 & & & \\
1 & & & &
\end{pmatrix}
$$

$$
- \begin{pmatrix}
q_{n-k} & \cdots & q_0 & & \\
\vdots & \ddots & & \ddots & \\
\vdots & & \ddots & & q_0 \\
\vdots & & & \ddots & \vdots \\
q_{n-1} & \cdots & \cdots & \cdots & q_{n-k}
\end{pmatrix}
\begin{pmatrix}
p_{n-k+1} & \cdots & \cdots & p_{n-1} & 0 \\
\vdots & & \cdot^{\cdot^{\cdot}} & 0 & \\
\vdots & \cdot^{\cdot^{\cdot}} & & & \\
p_{n-1} & 0 & & & \\
0 & & & &
\end{pmatrix}. \tag{135}
$$

Moreover we have

$$
\det\left(g_{i+j}\right)_{i,j=1}^{n} = \det\left(b_{ij}\right)_{i,j=n-k+1}^{n}
$$

$$
= \begin{vmatrix}
p_{n-k} & q_{n-k} & & & & & & \\
\vdots & \vdots & p_{n-k} & q_{n-k} & & & & \\
\vdots & \vdots & \vdots & \vdots & \ddots & & & \\
\vdots & \vdots & \vdots & \vdots & & \ddots & & \\
\vdots & \vdots & \vdots & \vdots & & & p_{n-k} & q_{n-k} \\
p_{n-1} & q_{n-1} & \vdots & \vdots & & & \vdots & \vdots \\
0 & 1 & p_{n-1} & q_{n-1} & & & \vdots & \vdots \\
 & & 0 & 1 & \ddots & & \vdots & \vdots \\
 & & & & \ddots & \ddots & \vdots & \vdots \\
 & & & & & & p_{n-1} & q_{n-1} \\
 & & & & & & 0 & 1
\end{vmatrix}. \tag{136}
$$

Before proving the theorem let us take note of the following result; see Halmos (1958).

LEMMA 6.2. *Let A, B, C, and D be square matrices such that C and D commute. Then*

$$\det\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(AD - BC). \tag{137}$$

Proof. Without loss of a generality, let $\det D \neq 0$. Since clearly

$$\begin{pmatrix} A & B \\ 0 & D \end{pmatrix} = \begin{pmatrix} A - BD^{-1}C & B \\ C & D \end{pmatrix}\begin{pmatrix} I & 0 \\ D^{-1}C & I \end{pmatrix}, \tag{138}$$

we have

$$\det\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(A - BD^{-1}C)\det D$$

$$= \det(A - BCD^{-1})\det D = \det(AD - BC). \quad \blacksquare \tag{139}$$

Proof of Theorem. The resultant of q and p is given by Equation (90) or equivalently by

$$\operatorname{Res}(p,q) = \begin{pmatrix} p(S) & p^{\#}(\tilde{S}) \\ q(S) & q^{\#}(\tilde{S}) \end{pmatrix}. \tag{140}$$

Clearly $q^{\#}(S)$ is invertible, and the matrices $q^{\#}(S)$ and $p^{\#}(S)$ commute. Thus the previous lemma can be invoked to yield

$$\det \operatorname{Res}(p,q) = \det\big(p(\tilde{S})q^{\#}(S) - q(\tilde{S})p^{\#}(S)\big). \tag{141}$$

Hence as, $B(p,q) = (p(\tilde{S})q^{\#}(S) - q(\tilde{S})p^{\#}(S))J$, we have

$$\det B(p,q) = \det \operatorname{Res}(p,q)\det J = \det \operatorname{Res}(p,q)(-1)^{n(n-1)/2} \tag{142}$$

We can use the factor $(-1)^{n(n-1)/2}$ to reorder the columns of the resultant in the following way:

$$
\begin{vmatrix}
p_0 & q_0 & & & & & & \\
\cdot & \cdot & p_0 & q_0 & & & & \\
\cdot & \cdot & \cdot & \cdot & \cdot & & & \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & & \\
\cdot & \cdot & \cdot & \cdot & & \cdot & \cdot & \\
\cdot & \cdot & \cdot & \cdot & & & p_0 & q_0 \\
p_{n-1} & q_{n-1} & \cdot & \cdot & & & \cdot & \cdot \\
 & 1 & p_{n-1} & q_{n-1} & & & \cdot & \cdot \\
 & & \cdot & 1 & \cdot & & \cdot & \cdot \\
 & & & & \cdot & \cdot & \cdot & \cdot \\
 & & & & \cdot & \cdot & \cdot & \cdot \\
 & & & & & 1 & p_{n-1} & q_{n-1} \\
 & & & & & & 0 & 1
\end{vmatrix},
$$

and hence we have

$$
\begin{vmatrix}
g_1 & \cdots & g_n \\
\vdots & & \vdots \\
g_n & \cdots & g_{2n-1}
\end{vmatrix} = \det(b_{ij})
$$

$$
= \begin{vmatrix}
p_0 & q_0 & & & & & & \\
\cdot & \cdot & p_0 & q_0 & & & & \\
\cdot & \cdot & \cdot & \cdot & \cdot & & & \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & & \\
\cdot & \cdot & \cdot & \cdot & & \cdot & \cdot & \\
\cdot & \cdot & \cdot & \cdot & & & p_0 & q_0 \\
p_{n-1} & q_{n-1} & \cdot & \cdot & & & \cdot & \cdot \\
 & 1 & p_{n-1} & \cdot & & & \cdot & \cdot \\
 & & \cdot & 1 & \cdot & & \cdot & \cdot \\
 & & & & \cdot & \cdot & \cdot & \cdot \\
 & & & & \cdot & \cdot & \cdot & \cdot \\
 & & & & & 1 & p_{n-1} & q_{n-1} \\
 & & & & & & 0 & 1
\end{vmatrix}. \tag{143}
$$

From the representation (89) of the Bezoutian it follows that the $k \times k$ lower right hand corner of the Bezoutian has the required representation (135).

Using Lemma 6.2 as before, we obtain the determinantal relations

$$
(-1)^{k(k-1)/2}\det\left(b_{ij}\right)_{i,j=n-k+1}^{n}
$$

$$
= \left|\begin{array}{cccccc}
p_{n-k} & \cdots & p_0 & & q_{n-k} & \cdots & q_0 & \\
\vdots & & & \ddots & \vdots & & & \ddots \\
\vdots & & & p_0 & \vdots & & & q_0 \\
\vdots & & & \vdots & \vdots & & & \vdots \\
p_{n-1} & \cdots & \cdots & p_{n-k} & q_{n-1} & \cdots & \cdots & q_{n-k} \\
0 & p_{n-1} & \cdots & p_{n-k+1} & 1 & \cdots & \cdots & q_{n-k+1} \\
 & & \ddots & \vdots & & \ddots & & \vdots \\
 & & & p_{n-1} & & & & q_{n-1} \\
 & & & 0 & & & & 1
\end{array}\right|
$$

$$
= \left|\begin{array}{cccccccc}
p_{n-k} & q_{n-k} & & & & & & \\
\vdots & \vdots & p_{n-k} & q_{n-k} & & & & \\
\vdots & \vdots & \vdots & \vdots & \ddots & & & \\
\vdots & \vdots & \vdots & \vdots & & \ddots & p_{n-k} & q_{n-k} \\
p_{n-1} & q_{n-1} & \vdots & \vdots & & & \vdots & \vdots \\
0 & 1 & p_{n-1} & q_{n-1} & & & \vdots & \vdots \\
 & & 0 & 1 & \ddots & & \vdots & \vdots \\
 & & & & & \ddots & p_{n-1} & q_{n-1} \\
 & & & & & & 0 & 1
\end{array}\right|. \tag{144}
$$

An application of Equation (134) yields the equality (136). From (136) the Hurwitz stability test can easily be obtained. ■

7. PLÜCKER CHARACTERIZATION OF BEZOUTIANS

We give a characterization of an arbitrary $n \times n$ Bezoutian matrix B by means of certain quadratic relations among the entries of B. These quadratic formulas appeared first in Plücker's work in the 19th century, in the description of the Grassmann manifold as a subvariety of a projective space.

Let $G_d(F^{n+1})$, $0 < d < n+1$, denote the Grassmann variety consisting of all d-dimensional linear subspaces of F^{n+1}. For $d = 1$ this specializes to the n-dimensional projective space

$$
\mathbf{P}^n(F) = G_1(F^{n+1}),
$$

and we use the homogenous coordinates

$$[x_0 : x_1 : \cdots : x_n]$$

to describe a point in $\mathbf{P}^n(F)$. Any $V \in G_d(F^{n+1})$ has a basis, formed by the columns of a full rank $(n+1)\times d$ matrix

$$X = \begin{pmatrix} X^0 \\ \vdots \\ X^n \end{pmatrix} \in F^{(n+1)\times d}.$$

For any increasing sequence $I = (i_1, \dots, i_n)$ of integers with

$$0 \leqslant i_1 < i_2 \cdots < i_d \leqslant n,$$

let X^I denote the $d \times d$ submatrix of X formed by the row vectors $X^{i_1}, \dots, X^{i_n}$. Thus one of the $\binom{n+1}{d}$ minors

$$d(I) = d_I(X) = \det X^I$$

will be nonzero. Using the lexicographical ordering on the set of indices I, we obtain a sequence of minors $d_1(X), \dots, d_{\binom{n+1}{d}}(X)$ of X.

Definition 7.1. The *Plücker embedding* of $G_d(F^{n+1})$ is the algebraic map

$$\mathscr{P}: G_d(F^{n+1}) \to \mathbf{P}^{\binom{n+1}{d}-1}(F)$$

which associates to each basis matrix X of $V \in G_d(F^{n+1})$ the unique point of $\mathbf{P}^{\binom{n+1}{d}-1}$ with homogeneous coordinates $[d_1(X), \dots, d_{\binom{n+1}{d}}(X)]$.

It is well-known fact that $\mathscr{P}$ is an algebraic embedding. Thus the image set of $\mathscr{P}$ is isomorphic to $G_d(F^{n+1})$ and is defined by a set of quadratic relations, the *Plücker relations* of $G_d(F^{n+1})$; see Hodge and Pedoe (1968) or Kleiman and Laksov (1972). Instead of considering the Plücker relations for an arbitrary Grassmannian, we assume $d = 2$, which is the case of most interest to us.

THEOREM 7.1. *Let $d(ij)$, $0 \leqslant i < j \leqslant n$, denote the homogeneous coordinates of a point in $\mathbf{P}^{\binom{n+1}{2}-1}(F)$. Then there exist linearly independent vectors*

$$\begin{pmatrix} p_0 \\ \vdots \\ p_n \end{pmatrix} \begin{pmatrix} q_0 \\ \vdots \\ q_n \end{pmatrix} \in F^{(n+1)}$$

with

$$d(ij) = p_i q_j - p_j q_i \qquad \text{for all} \quad 0 \leqslant i < j \leqslant n$$

if and only if for all $0 \leqslant i < j < k < l \leqslant n$

$$d(ij)d(kl) - d(ik)d(jl) + d(il)d(jk) = 0.$$

Proof. The necessity is checked by a brute force computation:

$$\begin{aligned}
d(ij)&d(kl) - d(ik)d(jl) + d(il)d(jk) \\
&= (q_i p_j - p_i q_j)(q_i p_l - p_j q_l) \\
&\quad - (q_i p_i - p_i q_i)(q_j p_l - p_j q_l) \\
&\quad + (q_i p_l - p_i q_l)(q_j p_k - p_j q_k) \\
&= q_i q_k (p_j p_l - p_l p_j) + q_i q_l (p_j p_k - p_k p_j) \\
&\quad + q_i q_j (p_k p_l - p_l p_k) + q_j q_k (p_i p_l - p_l p_i) \\
&\quad + q_j q_j (p_i p_k - p_k p_i) + q_k q_l (p_i p_j - p_j p_i).
\end{aligned}$$

To prove sufficiency, consider a point in $\mathbf{P}^{\binom{n+1}{2}-1}(F)$ whose homogeneous coordinates $d(ij)$, $0 \leqslant i < j \leqslant n$, satisfy the Plücker relations

$$d(ij)d(kl) - d(ik)d(jl) + d(il)d(jk) = 0.$$

Let $d(i_* j_*)$ denote the first nonzero element among the $d(ij)$, ordered lexicographically. Since the $d(ij)$ are homogeneous coordinates, we may assume $d(i_* j_*) = 1$. Now define linearly independent vectors

$$\begin{pmatrix} p_0 \\ \vdots \\ p_n \end{pmatrix}, \begin{pmatrix} q_0 \\ \vdots \\ q_n \end{pmatrix} \in F^{(n+1)}$$

by

$$p_i = \begin{cases} 0, & i < i_*, \\ 1, & i = i_*, \\ 0, & i_* < i \leqslant j_*, \\ -d(j_*, i), & i > j_*, \end{cases}$$

$$q_j = \begin{cases} 0, & j < j_*, \\ 1, & j = j_*, \\ d(i_*, j), & j > j_*. \end{cases}$$

Using the Plücker relations it is easily checked that

$$p_i q_j - p_j q_i = d(ij),$$

and the result follows. ∎

Given the pair of polynomials $p(z) = \sum_{i=0}^{n} p_i z^i$ and $q(z) = \sum_{i=0}^{n} q_i z^i$, then we have

$$q(z)p(w) - p(z)q(w) = \sum_{i=0}^{n} \sum_{j=0}^{n} d(ij) z^i w^j$$

with

$$d(ij) = q_i p_j - p_i q_j \qquad \text{for all} \quad 0 \leqslant i, j \leqslant n.$$

There is a simple formula relating the entries of the Bezoutian to the $d(ij)$.

LEMMA 7.1. *For all* $0 \leqslant i, j \leqslant n$

$$d(ij) = b_{i-1,j} - b_{i,j-1}.$$

Here we assume $b_{ij} = 0$ for $i, j < 0$ or $i, j > n$. Using this simple lemma we can reformulate the Plücker relations for $G_2(F^{n+1})$ as giving a necessary and sufficient condition for an arbitrary symmetric matrix to be Bezoutian.

THEOREM 7.2. *Let $B = (b_{ij}) \in F^{n \times n}$ be a symmetric matrix. Then there exist polynomials $p, q \in F[x]$ with* $\max(\deg q, \deg p) \leqslant n$ *and $B = B(q, p)$ if and only if for all $0 \leqslant i < j < k < l \leqslant n$ we have*

$$\begin{aligned}(b_{i-1,j} - b_{i,j-1})(b_{k-1,l} - b_{k,l-1}) \\ - (b_{i-1,k} - b_{i,k-1})(b_{j-1,l} - b_{j,l-1}) \\ + (b_{i-1,l} - b_{i,l-1})(b_{j-1,k} - b_{j,k-1}) = 0.\end{aligned}$$ ■

Let $B_n(F)$ denote the set of all $n \times n$ Bezoutian matrices, i.e

$$B_n(F) = \{b \in F^{n \times n} | \exists q, p \in F[x], \max(\deg q, \deg p) \leqslant n, \text{ such that } B = B(q, p)\}.$$

We call $B_n(F)$ the Bezout variety.

THEOREM 7.3. *$B_n(F)$ is an irreducible algebraic subvariety of $F^{n \times n}$ of dimension $2n - 1$. The origin $B = 0$ is an isolated singular point of $B_n(F)$; thus $B_n(F) \setminus \{0\}$ is smooth.*

Proof. That $B_n(F)$ is algebraic follows from Theorem 7.2. In fact it implies that $B_n(F)$ is quadric. Since the entries b_{ij} depend linearly upon the Plücker coordinates of the Grassmannian $G_2(F^{n+1})$, embedded in $\mathbf{P}^{\binom{n+1}{2}-1}(F)$, we see that $B_n(F)$ is isomorphic to the affine cone over the Grassmannian $G_2(F^{n+1})$. Thus

$$\dim B_n(F) = 1 + \dim G_2(F^{n+1}) = 2(n-1) + 1 = 2n - 1,$$

and $B_n(F) \setminus \{0\}$ is smooth. The irreducibility of $B_n(F)$ is clear, since $B_n(F)$ is the image of the affine space of pairs of polynomials (q, p) with $\max(\deg q, \deg p) \leqslant n$ under the algebraic mapping $(q, p) \to B(q, p)$. Since that space is irreducible, its image $B_n(F)$ is also irreducible. ■

REMARK. The proof shows more. Consider the $F^* = \mathrm{GL}_1(F)$ action

$$\alpha : F^* \times (B_n(F) \setminus \{0\}) \to B_n(F) \setminus \{0\}$$

defined by

$$(\lambda, B) \mapsto \lambda \cdot B,$$

i.e. by scalar multiplication. This is a free algebraic action of the reductive group $GL_1(F)$, and the quotient variety

$$\tilde{B}_n(F) = (B_n(F) \backslash \{0\}) / GL_1(F)$$

is a projective variety embedded in $\mathbf{P}^{\binom{n+1}{2}-1}$, which is isomorphic, by Theorem 7.3, to the Grassmannian $G_2(F^{n+1})$.

REMARK. There are $\binom{n+1}{4}$ quadratic equations in Theorem 7.2 which describe the Bezoutian matrices as a subvariety of the symmetric $n \times n$ matrices. Since

$$\dim B_n(F) = 2n - 1 > \frac{n(n+1)}{2} - \binom{n+1}{4},$$

this dimension count indicates that the system of quadratic equations for $B_n(F)$, given in Theorem 7.2, is far from being minimal. It seems of interest to find a minimal, independent set of quadratic equations characterizing the Bezoutian matrices.

8. CASCADE AND OUTPUT FEEDBACK EQUIVALENCE

DEFINITION 8.1. Let $g = p/q$ and $\hat{g} = \hat{p}/\hat{q}$ be two rational functions. We say that g and $\hat{g}$ are *cascade equivalent* if for some

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in GL_2(F)$$

we have

$$\hat{g} = \frac{\alpha g + \beta}{\gamma g + \delta}. \tag{145}$$

If g and $\hat{g}$ are proper, then they are *output feedback equivalent* if they are cascade equivalent and $\beta = 0$, i.e.

$$\hat{g} = \frac{\alpha g}{\gamma g + \delta}. \tag{146}$$

If g and $\hat{g}$ are proper, then they are *direct gain output feedback equivalent*, or *static output feedback equivalent*, if they are output feedback equivalent and $\alpha = \delta$, or equivalently

$$\hat{g} = \frac{g}{1 - kg}. \tag{147}$$

The Bezoutian behaves particularly simply under linear changes of variables. In fact from Theorem 3.1 one has

LEMMA 8.1. *For any* $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathrm{GL}_2(F)$ *we have*

$$B(\alpha q + \beta p, \gamma q + \delta p) = (\alpha\delta - \beta\gamma) B(q, p). \tag{148}$$

In particular this shows that the Bezoutian $B(q, p)$ of a transfer function $g = p/q$ is an invariant under the action of direct gain output feedback, i.e. $g \mapsto g/(1 - kg)$.

Given two rational functions g and $\hat{g}$, it is of interest to find conditions on their output feedback equivalence. Let $g(z) = p(z)/q(z)$ be any irreducible representation with q monic. Let $p(z) = \sum_{i=0}^{n-1} p_i z^i$ and $q(z) = \sum_{i=0}^{n} q_i z^i$. Obviously the subspace of F^{n+1} spanned by

$$\begin{pmatrix} p_0 & q_0 \\ \vdots & \vdots \\ p_{n-1} & q_{n-1} \\ 0 & 1 \end{pmatrix}$$

is an output feedback invariant. In particular, up to a constant factor, all 2×2 minors, i.e. all the functions

$$p_i q_j - p_j q_i, \tag{149}$$

are output feedback invariant.

For the case of direct gain output feedback it has been proved by Yannakoudakis (1981) that g and $\hat{g}$ are equivalent if and only if all the functions $p_i q_j - p_j q_i$ agree on g and $\hat{g}$. The following theorem is a formalization of results implicit in Antoulas (1986), Fuhrmann (1985), and Heinig and Rost (1984). It shows that the Bezoutian $B(g)$ of a transfer function g is, up to a scalar factor, a complete invariant for output feedback.

THEOREM 8.1. *Given two arbitrary rational functions g and $\hat{g} \in F(z)$. Then*

(a) *$\hat{g}$ is a cascade equivalent to g if and only if the respective Bezoutians are proportional, that is,*

$$B(\hat{g}) = cB(g) \tag{150}$$

for some nonzero $c \in F$.

(b) *If g and $\hat{g}$ are (strictly) proper, then $\hat{g}$ is output feedback equivalent to g if and only if the respective Bezoutians are proportional, that is, if and only if equality* (150) *holds.*

(c) *If g and $\hat{g}$ are strictly proper, then $\hat{g}$ is direct gain output feedback equivalent to g if and only if the respective Bezoutians are equal, that is, if and only if*

$$B(\hat{g}) = B(g). \tag{151}$$

Proof. Since cascade equivalence is transitive and every nonproper transfer function is clearly cascade equivalent to a strictly proper one, then it clearly suffices to prove part (c). So assume g_1 and g_2 are strictly proper. The necessity of Equation (151) follows from Lemma 8.1. To prove sufficiency assume the equality (151) holds. Note that if $g = p/q$ then

$$\begin{aligned} B(q,p)(z,w) &= \frac{q(z)p(w) - p(z)q(w)}{z-w} = \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij} z^{i-1} w^{j-1} \\ &= \frac{q(z)p(w) - q(z)p(z) + q(z)p(z) - p(z)q(w)}{z-w} \\ &= -q(z)\frac{p(z)-p(w)}{z-w} + p(z)\frac{q(z)-q(w)}{z-w}. \end{aligned}$$

This implies, by taking limits, that

$$\sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij} z^{i+j-2} = p(z)q'(z) - q(z)p'(z). \tag{152}$$

Assume now that $B(g) = B(\hat{g})$. Since, by Corollary 3.1, the numerator polynomial is determined by the last column of the Bezoutian, we have $p = \hat{p}$. Let $h = q - \hat{q}$, and assume without loss of generality that h is not identically zero. Now, by Equation (152),

$$B(q - \hat{q}, p) = 0,$$

or

$$p(z)[\hat{q}'(z)\text{-}q'(z)] - [\hat{q}(z) - q(z)]p'(z) = p(z)h'(z) - h(z)p'(z) = 0.$$

However, this implies that

$$\frac{d}{dz}\left(\frac{p(z)}{h(z)}\right) = 0,$$

or $p(z)/h(z) =$ constant, and hence the equivalence of g and $\hat{g}$. ∎

REMARK. Theorem 8.1 generalizes a previous result of Yannakoudakis (1981). Apparently Yannakoudakis was unaware of the Bezoutian and its properties. He proves in a direct but more complicated way that the quadratic form $B(q,1)H_g B(q,1)$ is a complete invariant for direct gain output feedback, which, in view of (115) in Theorem 6.3, is exactly what is stated in Theorem 8.1.

THEOREM 8.2. *Let $g = p/q$ and $\hat{g} = \hat{p}/\hat{q}$ be two strictly proper rational functions. Then $B(\hat{g}) = B(g)$ if and only if*

$$\hat{p} = p \tag{153}$$

and

$$\hat{p}'\hat{q} - \hat{p}\hat{q}' = p'q - pq', \tag{154}$$

where p' denotes the derivative of p.

Proof. The necessity is obvious by Corollary 3.1 and the identity

$$B(q,p)(z,z) = q'(z)p(z) - q(z)p'(z). \tag{155}$$

To prove sufficiency let us note that Equations (153) and (154) are equivalent to the identity

$$\left(\frac{1}{\hat{g}}\right)' = \left(\frac{1}{g}\right)'. \tag{156}$$

Thus there exists a constant $k \in F$ such that

$$\frac{1}{\hat{g}(z)} = \frac{1}{g(z)} - k, \tag{157}$$

which is equivalent to

$$\hat{g}(z) = \frac{g(z)}{1 - kg(z)}. \quad \blacksquare \tag{158}$$

The *breakaway points* of a real transfer function $g \in \mathbf{R}(z)$ are defined as the real complex roots of

$$p'(z)q(z) - q'(z)p(z) = 0. \tag{159}$$

These points are the branch points of the corresponding root loci; see Willems (1970). Using Theorem 8.2 we immediately obtain another remarkable result of Byrnes and Crouch [1985].

Corollary 8.1. *The real and complex zeros and breakaway points of a strictly proper transfer function $g \in \mathbf{R}(z)$ are a complete set of invariants for static output feedback.*

Now clearly it is of interest to be able to reconstruct the Bezoutian from the set of zero and breakaway points. This in fact can be done, by an application of the division algorithm. So assume we look for $g = p/q$ and we know the polynomials $p(z)$ and $r(z) = p(z)q'(z) - q(z)p'(z)$. Let us write

$$q(z) = a_1(z)p(z) - q_1(z), \qquad \deg q_1 < \deg p. \tag{160}$$

Differentiating, we can write

$$q'(z) = a'_1(z)p(z) + a_1(z)p'(z) - q'(z) \tag{161}$$

and hence

$$\begin{aligned} r(z) &= p(z)q'(z) - q(z)p'(z) \\ &= a_1'(z)p^2(z) + a_1(z)p'(z)p(z) \\ &\quad - q_1'(z)p(z) - a_1(z)p(z)p'(z) + p'(z)q_1(z) \\ &= a_1'(z)p^2(z) - [p(z)q_1'(z) - p'(z)q_1(z)]. \end{aligned} \tag{162}$$

Now we clearly have, inspecting the highest coefficient, that

$$\deg[pq_1' - p'q_1] = \deg p + \deg q_1 - 1 < 2\deg p - 1 = \deg p^2 - 1. \tag{163}$$

Thus Equation (162) is just the division rule applied to r and p^2. It follows that the polynomial $a_1(z)$ is only determined up to an additive constant. Inspecting Equation (160), it follows that q is determined up to a constant multiple of p, i.e. up to constant output feedback.

These invariants, i.e. the zeros and breakaway points, are not independent, and the question arises whether one can give a minimal set of invariants for output feedback. To deal with this question we consider the quotient space of all transfer functions of McMillan degree n modulo output feedback. In this connection we refer also to Brockett and Krishnaprasad [1980]. Explicitly, let

$$\mathrm{Rat}(n) = \left\{ g = \frac{p}{q} \in \mathbf{R}(z) \,\middle|\, p, q \text{ coprime}, \deg p < \deg q = n \right\} \tag{164}$$

denote the space of all real strictly proper transfer functions of degree n, and let $\mathscr{F}$ denote the output feedback group, consisting of all invertible $\begin{pmatrix} \alpha & 0 \\ \gamma & \delta \end{pmatrix} \in \mathrm{GL}_2(\mathbf{R})$. Then $\mathscr{F}$ acts on Rat(n) by

$$g \mapsto \frac{\alpha g}{\gamma g + \delta}. \tag{165}$$

Byrnes and Crouch have shown that the quotient space $\mathrm{Rat}(n)/\mathcal{F}$ is a smooth quasiprojective variety of dimension $2n-2$ embedded in the projective space

$$\mathbf{P}^N(\mathbf{R}), \qquad N = \binom{n+1}{2} - 1,$$

by the Plücker embedding. Using the Bezoutian, we can prove a stronger version. First, let us show

LEMMA 8.2. *The quotient space* $\mathrm{Rat}(n)/\mathcal{F}$ *is an analytic manifold of dimension* $2n-2$.

Proof. $\mathcal{F}$ acts algebraically on $\mathrm{Rat}(n)$, and the stabilizer subgroup of each $g \in \mathrm{Rat}(n)$ is equal to the nonzero scalar multiples of the identity. Hence there is only one orbit type, and it remains to show that the graph of the action is closed. Thus let g_k, h_k be a sequence of feedback equivalent systems converging in $\mathrm{Rat}(n)$ to g_∞ and h_∞ respectively. Since g_k, h_k are feedback equivalent, their Bezoutians are proportional, i.e., $B(g_k) = c_k B(h_k)$ for some nonzero c_k. By continuity, $B(g_k)$ and $B(h_k)$ converge to $B(g_\infty)$ and $B(h_\infty)$ respectively. Hence c_k must converge to a nonzero constant c_∞ with $B(g_\infty) = c_\infty B(h_\infty)$. By Theorem 8.2, g_∞ is feedback equivalent to h_∞ and the result follows. ∎

We saw, in Theorem 6.2, that given any transfer function $g \in \mathrm{Rat}(n)$, the inverse of the Bezoutian $B(g)$ is a Hankel matrix with $2n-1$ entries $(g_1, \dots, g_{2n-1})$. Now associate to every $g \in \mathrm{Rat}(n)$ the unique point in the projective space $\mathbf{P}^{2n-2}(\mathbf{R})$ whose homogeneous coordinates are the $2n-1$ entries $(g_1, \dots, g_{2n-1})$ of the Hankel matrix $B(g)^{-1}$. By Theorem 8.1 this defines an embedding $\imath : \mathrm{Rat}(n)/\mathcal{F} \hookrightarrow \mathbf{P}^{2n-2}(\mathbf{R})$ defined by

$$[g] \mapsto [g_1; \dots; g_{2n-1}], \tag{166}$$

which is easily seen to be algebraic. The embedding $\imath$ gives an isomorphism of the quotient space $\mathrm{Rat}(n)/\mathcal{F}$ onto the Zariski open subvariety of $\mathbf{P}^{2n-2}(\mathbf{R})$ defined by nonsingular arbitrary $n \times n$ Hankel matrices. Thus this shows

THEOREM 8.3. $\mathrm{Rat}(n)/\mathcal{F}$ *is a quasiprojective variety, embedded into* $\mathbf{P}^{2n-2}(\mathbf{R})$ *as an open subvariety.*

Of course, by reason of dimension, one cannot embed $\mathrm{Rat}(n)/\mathcal{F}$ into some projective space $\mathbf{P}^k(\mathbf{R})$ of smaller dimension, $k < 2n-2$.

COROLLARY 8.2. *The homogeneous coordinates of the* $2n-1$ *distinct entries of the Hankel matrix* $H = B(g)^{-1}$ *form a minimal and complete set of (projective) invariants for output feedback.*

We call these invariants the *Bezout invariants.*

9. CANONICAL FORMS

Throughout this section we identify the field F with $\mathbf{R}$, the field of real numbers. When dealing with group actions an important topic is the existence or nonexistence of continuous canonical forms. For the output feedback group action on $\mathrm{Rat}(n)$ given by

$$g \mapsto \frac{\alpha g}{\beta + \gamma g}, \tag{167}$$

Byrnes and Crouch (1985) have shown that a globally defined continuous canonical form does exist if and only if the McMillan degree n is odd. Their argument is based on topological obstruction theory and does not allow one to explicitly construct such a continuous canonical form for odd n. As we shall see, the construction of the Bezoutian allows us to explicitly construct such a continuous canonical form. For this purpose we introduce a normalized version of the Bezoutian.

DEFINITION 9.1. Let $g = p/q \in \mathrm{Rat}(n)$, with n odd. The *normalized Bezoutian* is defined as

$$\mathrm{BN}(g) = \mathrm{BN}(q,p) = \frac{1}{\sqrt[n]{\det \mathrm{B}(g)}}\mathrm{B}(g) \tag{168}$$

Obviously we have

$$\mathrm{BN}(\lambda g) = \mathrm{BN}(g) \tag{169}$$

for any nonzero real number λ, and $\det \mathrm{BN}(g) = 1$. By the same argument as in the proof of Theorem 8.2 one shows

THEOREM 9.1. *Assume* n *is odd. Then* $g, h \in \mathrm{Rat}(n)$ *are output feedback equivalent if and only if* $\mathrm{BN}(g) = \mathrm{BN}(h)$.

Clearly, for $g \in \mathrm{Rat}(n)$, the entries of $\mathrm{BN}(g)$ are smooth, semialgebraic functions of g.

Given any normalized Bezoutian $\mathrm{BN} = \mathrm{BN}(g)$, we construct two coprime polynomials $p_{\mathrm{BN}}, q_{\mathrm{BN}}$ as follows. First define

$$p_{\mathrm{BN}}(z) = \sum_{i=0}^{n-1} p_i z^i, \tag{170}$$

where

$$\begin{pmatrix} p_0 \\ \vdots \\ p_{n-1} \end{pmatrix} = \mathrm{BN}\, e_n \tag{171}$$

denotes the last column of BN. To construct q_{BN}, let

$$H = \begin{pmatrix} g_1 & \cdots & g_n \\ \vdots & & \vdots \\ g_n & \cdots & g_{2n-1} \end{pmatrix} \tag{172}$$

denote the inverse of BN. Then

$$q_{\mathrm{BN}}(z) = \begin{vmatrix} g_1 & \cdots & g_{n-1} & g_n & 1 \\ g_2 & \cdots & g_n & g_{n+1} & z \\ \vdots & & \vdots & \vdots & \vdots \\ g_n & \cdots & g_{2n-2} & g_{2n-1} & z^{n-1} \\ g_{n+1} & \cdots & g_{2n-1} & 0 & z^n \end{vmatrix}. \tag{173}$$

Note that, by our normalization (168) and Equation (116), $\det H = 1$.

THEOREM 9.2. *Let n be odd, and $\mathrm{BN} = \mathrm{BN}(g)$ denote the normalized Bezoutian of $g = p/q \in \mathrm{Rat}(n)$. Then the map $\beta: \mathrm{Rat}(n) \to \mathrm{Rat}(n)$ given by*

$$\frac{p}{q} \mapsto \frac{p_{\mathrm{BN}}}{q_{\mathrm{BN}}}, \tag{174}$$

with p_{BN}, q_{BN} *defined by Equations* (171) *and* (173) *respectively, is a smooth, semialgebraic canonical form for output feedback.*

Proof. That β is smooth and semialgebraic is clear by construction and by Theorem 9.1. Obviously the polynomials p_{BN}, q_{BN} are invariant under output feedback. Thus it remains to show that

(a) $\beta(g)=\beta(h)$ implies that g and h are output feedback equivalent;
(b) $\beta(g)$ is output feedback equivalent to g.

It was shown in Fuhrmann (1986a) that

$$q_{BN}(z)=q(z)-\xi p_{BN}(z) \tag{175}$$

for some real number ξ. Hence

$$B(\beta(g))=B(q_{BN}(z),p_{BN}(z))=B(q(z),p_{BN}(z)). \tag{176}$$

By Equation (171) and Corollary 3.1 we have

$$p_{BN}(z)=\frac{1}{\sqrt{\det B(q,p)}}p(z). \tag{177}$$

Hence

$$B(\beta(g))=\mathrm{BN}(g). \tag{178}$$

An application of Theorem 9.1 proves (a) and (b) ■

A second, different canonical form for output feedback is defined using continued fraction expansions. To this end we review the basic facts concerning the continued fraction expansion of rational functions and its relation to the Euclidean algorithm and to partial realizations. We refer to Kalman (1979) and Gragg and Lindquist (1983) for comprehensive expositions of the partial realization problem and its relation to continued fraction expansions.

Let g be a strictly proper rational function, and let $g=p/q$ be an irreducible representation of g with q monic of degree n. We define, using the division rule for polynomials, a sequence of polynomials q_i, and a sequence of nonzero constants β_i and monic polynomials $a_i(z)$, referred to as *atoms*, by

$$q_{-1} = q, \qquad q_0 = p,$$

$$q_{i+1}(z) = a_{i+1}(z)q_i(z) - \beta_i q_{i-1}(z) \tag{179}$$

with $\deg q_{i+1} < \deg q_i$. The procedure ends when q_r is the g.c.d. of p and q. Since p and q are assumed coprime, q_r is a nonzero constant.

The atoms $a_1, \dots, a_r$ are real monic polynomials of degrees $n_1, \dots, n_r$ such that

$$n_1 + \cdots + n_r = n \tag{180}$$

and β_i are nonzero real numbers with

$$\delta_i = \operatorname{sign} \beta_i. \tag{181}$$

In terms of the β_i and the $a_i(z)$, g has the continued fraction representation

$$g(z) = \cfrac{\beta_0}{a_1(z) - \cfrac{\beta_1}{a_2(z) - \cfrac{\beta_2}{a_3(z) - \ddots \cfrac{\beta_{r-2}}{a_{r-1}(z) - \cfrac{\beta_{r-1}}{a_r(z)}}}}} \tag{182}$$

We define now two sequences of polynomials P_k and Q_k by the three term recursion formulas

$$P_{-1} = -1, \qquad P_0 = 0,$$

$$P_{k+1}(z) = a_{k+1}(z)P_k(z) - \beta_k P_{k-1}(z) \tag{183}$$

and

$$Q_{-1} = 0, \qquad Q_0 = 1,$$

$$Q_{k+1}(z) = a_{k+1}(z)Q_k(z) - \beta_k Q_{k-1}(z) \tag{184}$$

The expansion of P_k/Q_k in powers of z^{-1} agrees with that of g up to order $2\sum_{i=1}^{k} \deg a_i + \deg a_{k+1}$; see Gragg and Lindquist (1983). Computing the expression

$$Q_{k+1}P_k - Q_k P_{k+1} \tag{185}$$

and using the previous recursion formulas, we get

$$\begin{aligned} Q_{k+1}P_k - Q_k P_{k+1} &= (a_{k+1}Q_k - \beta_k Q_{k-1})P_k - (a_{k+1}P_k - \beta k_{Pk-1})Q_k \\ &= \beta_k(Q_k P_{k-1} - Q_{k-1}P_k), \end{aligned} \tag{186}$$

and by induction it follows that

$$\begin{aligned} Q_{k+1}P_k - Q_k P_{k+1} &= \beta_k \cdots \beta_0(Q_0 P_{-1} - Q_{-1}P_0) \\ &= -\beta_k \cdots \beta_0. \end{aligned} \tag{187}$$

This implies that for each k, Q_k and P_k are coprime. Moreover the Bezout identities

$$P_k(z)A_k(z) + Q_k(z)B_k(z) = 1 \tag{188}$$

hold with

$$A_k(z) = \frac{Q_{k-1}(z)}{\beta_0 \cdots \beta_{k-1}} \tag{189}$$

and

$$B_k(z) = \frac{-P_{k-1}(z)}{\beta_0 \cdots \beta_{k-1}}. \tag{190}$$

Obviously the Bezout identity (188) implies also the coprimeness of the polynomials A_k and Q_k. Since the Bezoutiant is linear in each of its arguments we can get a recursive formula for $B(A_k, Q_k)$:

$$\begin{aligned} B(A_k, Q_k) &= B\big((\beta_0 \cdots \beta_{k-1})^{-1} Q_{k-1}, Q_k\big) \\ &= B\big((\beta_0 \cdots \beta_{k-1})^{-1} Q_{k-1}, a_k Q_{k-1} - \beta_{k-1} Q_{k-2}\big) \\ &= (\beta_0 \cdots \beta_{k-1})^{-1} Q_{k-1}(z) B(1, a_k) Q_{k-1}(w) \\ &\quad + B\big((\beta_0 \cdots \beta_{k-2})^{-1} Q_{k-2}, Q_{k-1}\big) \\ &= \cdots = \sum_{j=1}^{k} (\beta_0 \cdots \beta_{k-1})^{-1} Q_j(z) B(1, a_{j+1}) Q_j(w). \end{aligned} \tag{191}$$

Thus we have obtained the polynomial relation

$$\begin{aligned} &Q_k(z) A_k(w) - A_k(z) Q_k(w) \\ &\quad = \sum_{j=1}^{k} (\beta_0 \cdots \beta_{k-1})^{-1} Q_j(z) \left[\frac{a_{j+1}(z) - a_{j+1}(w)}{z - w} \right] Q_j(w) \end{aligned} \tag{192}$$

or the equivalent matrix relation

$$B(Q_k, A_k) = \sum_{j=1}^{k} (\beta_0 \cdots \beta_{j-1})^{-1} \tilde{R}_j B(a_{j+1}, 1) R_j, \tag{193}$$

where R_j is the $n_j \times n$ Toeplitz matrix

$$R_j = \begin{pmatrix} \sigma_0^{(j)} & \sigma_1^{(j)} & \cdots & \sigma_{n_j}^{(j)} & & \\ & \ddots & \ddots & & \ddots & \\ & & \sigma_0^{(j)} & \sigma_1^{(j)} & \cdots & \sigma_{n_j}^{(j)} \end{pmatrix} \tag{194}$$

and

$$Q_j(z) = \sum_{j=0}^{n_j} \sigma_\nu^{(j)} z^\nu. \tag{195}$$

We will refer to (192) as the *generalized Christoffel-Darboux formula*. It reduces to the regular Christoffel-Darboux formula [see Akhiezer (1965)] in the generic case, namely when all the atoms a_j have degree one.

Combining Equation (193) with Theorem 6.2, we obtain the following.

THEOREM 9.3. *Let H be the nonsingular Hankel matrix* (172), *and let $g = p/q$ be any minimal rational extension of the sequence $g_1, \dots, g_{2n-1}$. Let β_i and $a_i(z)$ be defined by* (179), *and the sequence of polynomials P_k and Q_k by* (183) *and* (184) *respectively. Let the Toeplitz matrix R_j be defined by* (194). *Then*

$$H^{-1}\sum(\beta_0 \cdots \beta_{j-1})^{-1}\tilde{R}_j B(a_{j+1},1)R_j. \tag{196}$$

Suppose $\hat{g} = \hat{p}/\hat{q}$ is output feedback equivalent to $g = p/q$. Then $\hat{p}(z) = \alpha p(z)$, $\alpha \neq 0$, and $\hat{q}(z) = q(z) - kp(z)$. Therefore

$$q_1(z) = a_1(z)p(z) - \beta_0 q(z)$$

implies

$$\begin{aligned}\alpha q_1(z) &= a_1(z)(\alpha p(z)) - \alpha\beta_0 q(z)\\ &= a_1(z)(\alpha p(z)) - \alpha\beta_0[\hat{q}(z) + kp(z)]\\ &= [a_1(z) - \beta_0 k]\hat{p}(z) - (\alpha\beta_0)\hat{q}(z),\end{aligned}$$

and so

$$\hat{\beta}_0 = \alpha\beta_0, \qquad \hat{a}_1(z) = a_1(z)p(z) - \beta_0 k.$$

However, as $q_1/p = \hat{q}_1/\hat{p}$, all other atoms of g and $\hat{g}$ coincide. Thus, changing the transfer function $g(z)$ by output feedback amounts to rescaling β_0 by a nonzero constant and to changing the constant part of $a_1(z)$ by an arbitrary real number, i.e. $a_1(z) \mapsto a_1(z) - \sigma$. In particular, the degrees of n_i of the atoms are output feedback invariant; see Fuhrmann and Krishnaprasad (1986). It follows that any transfer function $g(z)$ as in Equation (182) is output feedback equivalent to a unique transfer function $\hat{g}$ given by

$$\hat{g}(z) = \cfrac{1}{\hat{a}_1(z) - \cfrac{\beta_1}{a_2(z) - \cfrac{\beta_2}{a_3(z) - \ddots \cfrac{\beta_{r-2}}{a_{r-1}(z) - \cfrac{\beta_{r-1}}{a_r(z)}}}}} \tag{197}$$

with $\hat{a}_1(z) = a_1(z) - a_1(0)$.

The map $g \mapsto \hat{g}$ defines a canonical form for output feedback, and by choosing a state space realization for $\hat{g}$ as in Kalman (1979) or Gragg and Lindquist (1983), the following theorem is proved.

THEOREM 9.4. *Let (A, b, c) be a minimal realization of a transfer function g of McMillan degree n. Let $a_i(z)$, β_i be the sequence of atoms of g, and assume that $n_1, \dots, n_r$ are the degrees of the atoms and*

$$a_k(z) = \sum_{i=0}^{n_k - 1} a_i^{(k)} z^i + z^{n_k}. \tag{198}$$

Then (A, b, c) is output feedback equivalent to a unique system of the form $(\hat{A}, \hat{b}, \hat{c})$ of the form

$$\hat{b} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad \hat{c} = (0 \quad \cdots \quad 1 \quad 0 \quad \cdots \quad 0)^{\sim}, \tag{199}$$

then 1 being in the n_1 position, and

$$\hat{A} = \begin{pmatrix} A_{11} & A_{12} & & & \\ A_{21} & A_{22} & \cdot & & \\ & \cdot & \vdots & \vdots & \\ & & & \ddots & A_{r-1\,r} \\ & & & A_{r\,r-1} & A_{rr} \end{pmatrix}, \tag{200}$$

where

$$A_{11}=\begin{pmatrix} 0 & \cdots & 0 & 0 \\ 1 & & & -a_1^{(1)} \\ & \ddots & & \vdots \\ & & 1 & -a_{n_1-1}^{(1)} \end{pmatrix}, \tag{201a}$$

$$A_{ii}=\begin{pmatrix} 0 & \cdots & 0 & -a^{(i)}0 \\ 1 & & & -a_1^{(i)} \\ & \ddots & & \vdots \\ & & 1 & -a_{n_i-1}^{(i)} \end{pmatrix}, \qquad i=2,\dots,r, \tag{201b}$$

$$A_{i+1i}=\begin{pmatrix} 0 & \cdots & 0 & 1 \\ 0 & & & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & 0 & 0 \end{pmatrix}, \tag{202}$$

and $A_{i\,i+1}=\beta_{i-1}A_{i+1i}$.

Proof. Without loss of generality assume $g=p/q$ is already in the canonical continued fraction form given by Equation (197). Let $Q_i(z)$ be the polynomials, associated with q, defined in Equation (184). Then the set

$$B_{\text{or}}=\left\{1,z,\dots,z^{n_1-1},Q_1,zQ_1,\dots,z^{n_2-1}Q_1,\dots,Q_{r-1},\dots,z^{n_{r-1}-1}Q_{r-1}\right\}$$

is clearly a basis for X_q, as it contains one polynomial for each degree between 0 and $n-1$. We call this basis the *orthogonal basis* because of its relation to orthogonal polynomials. In this connection see Gragg (1974) and Fuhrmann (1987).

Next we construct an *associated realization* to $g=p/q$. We choose X_q as the state space and define (A_1, B_1, C_1) through

$$\begin{aligned} A_1 &= S_q \\ B_1\xi &= \xi \qquad \text{for} \quad \xi\in F^m, \\ C_1 f &= \left(pq^{-1}f\right)_{-1} \quad \text{for} \quad f\in X_q. \end{aligned} \tag{203}$$

The realization of g is minimal, by the coprimeness of p and q, [see Fuhrmann (1976)], and its matrix representation with respect to the basis B_{or}

proves the theorem. In the computation of the matrix representation we lean heavily on the recursion formulas (184). ∎

This defines a canonical form for output feedback, which we call the *continued fraction form*. By fixing the block sizes n_i, $i = 1, \dots, r$, and the signs $\delta_i = \operatorname{sign} \beta_{i-1}$, $i = 1, \dots, r$, of the nonzero entries as discrete invariants, it is easy to see that the above canonical form defines a cell decomposition of the orbit space $\mathrm{Rat}(n)/\mathscr{F}$. For this we refer to Furhmann and Krishnaprasad (1986), Helmke, Hinrichsen, and Manthey (1988), Hinrichsen, Manthey, and Prätzel-Wolters (1986), and Manthey (1987) for further details of the continued fraction cell decomposition.

We conclude with an example.

EXAMPLE ($n = 2$). There are only three cells of $\mathrm{Rat}(2)/\mathscr{F}$, parameterized by rational functions

$$g(z) = \frac{1}{z(z+\alpha)}, \qquad \alpha \in \mathbf{R}, \tag{204}$$

$$g(z) = \frac{1}{z + \dfrac{\beta}{z+\gamma}}, \qquad \beta > 0, \quad \gamma \in \mathbf{R}, \tag{205}$$

$$g(z) = \frac{1}{z - \dfrac{\beta}{z+\gamma}}, \qquad \beta > 0, \quad \gamma \in \mathbf{R} \tag{206}$$

$\mathrm{Rat}(2)/\mathscr{F}$ splits into two connected components

$$(\mathrm{Rat}(2,0) \cup \mathrm{Rat}(0,2))/\mathscr{F} \quad \text{and} \quad \mathrm{Rat}(1,1)/\mathscr{F},$$

where the first component is the 2-cell parametrized by (206), while $\mathrm{Rat}(1,1)/\mathscr{F}$ is a Möbius band formed by a 1-cell and a 2-cell. In particular, $\mathrm{Rat}(2)/\mathscr{F}$ is not orientable.

REFERENCES

Akhiezer, N. I. 1965. *The Classical Moment Problem*, Hafner.

Anderson, B. D. O. and Jury, E. I. 1976. Generalized Bezoutian and Sylvester matrices in multivariable linear control, *IEEE Trans. Automat. Control* AC-21:551–556.

Antoulas, A. C. 1986. On recursiveness and related topics in linear systems, *IEEE Trans. Aut. Contr.* AC-31:1121–1135.

Barnett, S. and Lancaster, P. 1980. Some properties of the Bezoutian for polynomial matrices, *Linear Multilinear Algebra* 9:99–110.

Brockett, R. W. and Krishnaprasad, P. S. 1980. A scaling theory for linear systems, *IEEE Trans. Automatic Control* AC-25:197–207.

Byrnes, C. I. and Crouch, P. 1985. Geometric methods for the classification of linear feedback systems, *Systems Control Lett.* 6:239–246.

Byrnes, C. I. and Helton, J. W. 1986. Cascade equivalence of linear systems, *Internat. J. Control* 44:1507–1521.

Datta, B. N. 1978. An elementary proof of the stability criterion of Lienard and Chipart, *Linear Algebra Appl.* 22:89–96.

Fuhrmann, P. A. 1976. Algebraic system theory: An analyst's point of view, *J. Franklin Inst.* 301:521–540.

Furhmann, P. A. 1981. Polynomial models and algebraic stability criteria, in *Proceedings of the Joint Workshop on Feedback and Synthesis of Linear and Nonlinear Systems*, ZIF Bielefeld, June 1981, Lecture Notes in Control and Inform. Sci. 39, Springer-Verlag.

Furhmann, P. A. 1983. On symmetric rational matrix functions, *Lin. Alg. and Appl.*, Special issue on Linear Control Theory, 50:167–250.

Furhmann, P. A. 1985. On the partial realization problem and the recursive inversion of Hankel and Toeplitz matrices, in *Linear Algebra and its Role in System Theory*, Contemp. Math., Amer. Math. Soc.

Fuhrmann, P. A. 1986a. Remarks on the inversion of Hankel matrices, *Linear Algebra Appl.* 81:89–104.

Furhmann, P. A. 1986b. Block Hankel matrix inversion—the polynomial approach, in *Operator Theory and Systems*, Proceeding, Workshop Amsterdam, June 1985.

Furhmann, P. A. 1987. Orthogonal matrix polynomials and system theory, *Rend. Sem. Mat. Univ. Padova*, to appear.

Furhmann, P. A. and Datta, B. N. 1987. On Bezoutians, Van der Monde matrices and the Lienard-Chipart stability criterion, to appear.

Furhmann, P. A. and Krishnaprasad, P. S. 1986. Towards a cell decomposition for rational functions, *IMA J. Math. Control Inform.* 3:137–150.

Gantmacher, F. R. 1959. *The Theory of Matrices*, Chelsea, New York.

Gohberg, I. C. and Semencul, A. A. 1972. On the inversion of finite Toeplitz matrices and their continuous analogs (Russian), *Mat. Issled.*, 7:201–233.

Gragg, W. B. 1974. Matrix interpretations and applications of the continued fraction algorithm, *Rocky Mountain J. Math.* 4:213–225.

Gragg, W. B. and Lindquist, A. 1983. On the partial realization problem, *Linear Algebra Appl.* 50:277–319.

Halmos, P. 1958. *Finite Dimensional Vector Spaces*, Van Nostrand.

Heinig, G. and Rost, K. 1984. *Algebraic Methods for Toeplitz-like Matrices and Operators*, Birkhäuser.

Helmke, U., Hinrichsen, D. and Manthey, W. 1988. A cell decomposition for infinite Hankels of rank $\leqslant n$ and some applications, *Lin. Alg. Appl.* This issue.

Helmke, U. 1987. Rational functions and Bezout forms: A functional correspondence, preprint.

Hermite, C. 1856, Sur le nombre des racines d'une équation algébrique comprise entre des limites donnés, *J. Reine Angew. Math.* 52:39–51.

Hinrichsen, D., Manthey, W. and Prätzel-Wolters, D. 1986. The Bruhat decomposition of finite Hankel matrices, *Systems Control Lett.* 7:173–182.

Hinrichsen, D. and Prätzel-Wolters, D. 1984. A canonical form for static linear output feedback, in *Mathematical Theory of Networks and Systems*, Proceedings Beer Sheva, 1983, Lecture Notes in Control and Inform. Sci. 58, Springer-Verlag.

Hodge, W. V. D. and Pedoe, D. 1968. *Methods of Algebraic Geometry*, Cambridge U.P.

Hoffman, K. and Kunze, R. 1971. *Linear Algebra*, McGraw-Hill.

Kalman, R. E. 1979. On partial realizations, transfer functions and canonical forms, *Acta Polytech. Scand.* 31:9–32.

Kleiman, S. L. and Laksov, D. 1972. Schubert calculus, *Amer. Math. Monthly* 79:1061–1082.

Kravitsky, N. 1980. On the discriminant function of two noncommuting nonselfadjoint operators, *Integral Equations Operator Theory* 3:97–124.

Krein, M. G. and Naimark, M. A. 1936. The method of symmetric and Hermitian forms in the theory of the separation of roots of algebraic equations, English translation in *Linear and Multilinear Algebra* 10:265–308 (1981).

Lander, F. I. 1974. On the discriminant function of two noncommuting nonselfadjoint operators (in Russian), *Mat. Issled.* IX(32):69–87.

Lang, S. 1965. *Algebra*, Addison-Wesley.

Lerer, L. and Tismenetsky, M. 1982. The Bezoutian and the eigenvalue separation problem for matrix polynomials, *Integral Equations Operator Theory* 5:386–445.

MacLane, S. and Birkhoff, G. 1967. *Algebra*, Macmillan, New York.

Manthey, W. 1987. Doctoral Dissertation, Universität Bremen.

Pták, V. 1984. Explicit expressions for Bezoutians, *Linear Algebra Appl.* 59:43–54.

Rosenbrock, H. H. 1968. Relatively prime polynomial matrices, *Electr. Lett.* 4:227–228.

Rowe, A. 1972. The generalized resultant matrix, *J. Inst. Math. Appl.* 9:390–396.

Sylvester, J. 1853. On a theory of the Szygetic relations of two rational integral functions, comprising an application to the theory of Sturm's functions, and that of the greatest algebraic common measure, *Philos. Trans. Roy. Soc. London* 143:407–548.

Trench, W. 1965. An algorithm for the inversion of finite Hankel matrices, *SIAM J. Appl. Math.* 13:1102–1107.

Van der Waerden, B. L. 1931. *Modern Algebra*, Springer, Berlin.

Vinnikov, V. 1986. Self-adjoint determinantal representations of real irreducible cubics, in *Operator Theory and Systems*, Proceedings, Workshop Amsterdam, June 1985.

Willems, J. L. 1970. *Stability Theory of Dynamical Systems*, Wiley.

Wimmer, H. K. 1989a. Polynomial models, Bezoutians and Hankel matrices, to appear.

Wimmer, H. K. 1989b. Generalized Bezoutians and block Hankel matrices, to appear.

Wimmer, H. K. and Pták, V. 1985. On the Bezoutian for polynomial matrices, *Lin. Alg. and Appl.*, 71:167–250.
Yannakoudakis, A. 1981. Invariant algebraic structures in multivariable control, preprint, Lab. d'Automatique de Grenoble.

Received March 1988; final manuscript accepted 15 July 1988