On Symmetric Rational Transfer Functions

Paul A. Fuhrmann

Ben Gurion University of the Negev
Beer Sheva, Israel

Submitted by R. W. Brockett

ABSTRACT
A detailed study of symmetric transfer functions is presented.

1. INTRODUCTION

The object of this paper is a detailed study of symmetric transfer functions. One of the central themes of system theory is the study of the interrelation between the external properties of a system, be they given in terms of input-output relations, weighting patterns, or transfer functions, and the internal properties of the system, namely the properties of the realizations of the system.

Thus it is natural to expect that given some external symmetry properties of a transfer function, these properties have their counterpart in some realizations possessing internal symmetry properties. This is certainly not a new subject; some previous work is that of Youla and Tissi [62], Brockett and Skoog [12], and Brockett [7–9].

Indeed, for the special type of symmetry considered here, namely symmetry under transposition of real rational transfer matrices, the existence of a signature symmetric realization has been known for some time. In the special case of a scalar transfer function the signature of the signature matrix in a signature symmetric realization has been shown to be equal to the signature of the Hankel matrix induced by the transfer function g; to the signature of the Bezoutian of the polynomials p and q, where g = p/q; and to the Cauchy index of g.

It was to be expected that this circle of ideas, which in the scalar case is mostly classical (as the references to the work of Hermite [39] and Hurwitz [41] indicate), would have corresponding multivariable generalizations. This
indeed began recently to happen. Anderson and Jury [1] defined a multivariable Bezoutian form which is a generalization of the classical scalar Bezoutian. Bezoutians have for long played a role in the theory of equations, giving alternative criteria for coprimeness of polynomials, as well as in control and system theory. In this context they appear in algebraic stability theory, again a classical topic whose study goes back to Hermite [39]. Another instance is the appearance of Bezoutians as matrices intertwining some special realizations, as in Kailath [40, p. 144] or Casti [15, p. 99]. That this is no coincidence will hopefully be made clear in this paper.

First and foremost the paper’s intention is to tell a story, and tell it in a unified way. As such it contains some results which are old and some which are new. The unifying theme is the use of the theory of polynomial models. Thus we re-prove some classical results, such as the resultant theorem, the Chinese remainder theorem, the Hermite-Hurwitz theorem, and the theorem of Frobenius on the representation of an arbitrary square matrix as the product of two symmetric ones, not only to make the exposition self-contained, but also to familiarize the reader with polynomial models in relation to some known theorems as well as to provide motivation and intuition for the passage to the more difficult multivariable case. Thus naturally the original contributions are concentrated in the latter part of the paper. Specifically, the results on the generalized Bezoutians as block matrix representations, the multivariable version of the Chinese remainder theorem, the matrix partial fraction decompositions, and especially the approach to signature symmetric realizations by way of the equivalence relation between symmetric transfer functions and the corresponding canonical form seem all to be new.

We begin in the next section with a short introduction to polynomial models and their use in realization theory. This theory has been developed recently by Fuhrmann [22–26], Emre and Hautus [19], Fuhrmann and Willems [29, 30], and Khargonekar and Emre [47], and it seems ideally suited for the unification of several current approaches to linear system theory. In Section 3 we collect some useful information on bases and dual bases in the context of polynomial models. This sheds light on some well-known matrix relations whose previous proof was mostly computational.

Next, in Section 4, we pass to the study of multivariable, or generalized, Bezoutians in the context of the theory of polynomial models. The emphasis will be on interpreting the Bezoutian as a (block) matrix representation of an intertwining map. Since intertwining maps for polynomial models have been exhaustively studied and are well understood, many results on generalized Bezoutians can be easily derived from this study. That the setting is a natural one is indicated by the comparison with the difficulties encountered in Householder [40] or Datta [16], where in some instances the analysis of special cases comes instead of a proof.
In Section 5 we use polynomial models to rederive the scalar Hermite-Hurwitz theorem. We end that section by giving a simplified treatment of the Frobenius result for computing the signature of a Hankel matrix. This we do by the use of Bezoutian which can be more directly related to the Euclidean algorithm. In this connection the papers of Kalman [45] and Gragg and Lindquist [37] are relevant.

We pass, in Section 6, to the explicit construction of a signature symmetric realization of a scalar transfer function \( g = \frac{p}{q} \). The point we try to make is that, while the representation \( \frac{p}{q} \) gives all the i/o information, it is not a good encoding for the symmetry, trivial in this case. A more appropriate representation would be one of the Rosenbrock type, namely

\[
g = (re)q^{-1}r + s,
\]

where the signature information is carried in the polynomial \( e \). The existence of such representations and the corresponding factorization of \( p = er^2 \mod q \), will be given in complete detail with an eye to the multivariable generalizations. The method uses partial fraction decompositions, which localizes the problem. The local pieces are put together to obtain a global factorization by the Chinese remainder theorem, which is an interpolation result.

From this analysis it is clear that it would be profitable to develop multivariable analogs for the tools used in the scalar case. These are topics of independent interest. We begin, in Section 7, by proving a multivariable generalization of the Chinese remainder theorem. In the next section we develop a partial fraction decomposition with matrix fractions. Special attention is given to the implications of the symmetry property of the rational function.

With this machinery ready we return, in Section 9, to the construction of signature symmetric realizations of real symmetric rational transfer functions. This is done by applying the same line of reasoning as in the scalar case, i.e. local representations that exhibit the symmetry, which are interpolated to obtain a global signature symmetric realization. The computations are adapted from Bitmead and Anderson [3], the basic paper on the matrix Cauchy index and a work which greatly influenced the research reported in this paper. While we prefer not to adopt the Bitmead-Anderson definition of the matrix Cauchy index, the ideas are very similar. The local representations exhibit certain canonical forms generalizing the local scalar Cauchy index. Gohberg, Lancaster and Rodman [36] use the name signature characteristic for what is essentially the full signature information carried by a symmetric polynomial matrix, and we will follow their definition.
The next section, a short one, is devoted to a generalization of a theorem of Frobenius [21]. The proof uses polynomial models in a very compact way. The interest in this theorem of Frobenius is that it links the study of self-adjoint operators in (finite dimensional) indefinite metric spaces with the study of symmetric polynomial matrices. Once this identification is made, the whole machinery of signature symmetric realizations and the Cauchy characteristic can be applied either to the spectral analysis of a symmetric nonsingular polynomial matrix (not necessarily monic) or alternatively to the reduction of a self-adjoint operator in an indefinite metric space to canonical form. This in turn is shown to be equivalent to the analysis, going back to Kronecker [1868] and Weierstrass [1868], of the simultaneous reduction by congruence of two real symmetric matrices, of which one is assumed nonsingular, to canonical form.

Some work closely related to the point of view taken in this paper, especially in regard to signature symmetric realizations and indefinite metric spaces, can be found in Wimmer [60, 61].

The methods developed and used in this paper depend strongly on a bilinear form defined on the space of truncated vector Laurent series. In fact, if we choose a skew-symmetric bilinear form instead, then we get a parallel theory dealing with the analysis of real rational transfer functions possessing Hamiltonian symmetry, i.e. for which \( G(-z) = G(z) \), and canonical forms for Hamiltonian maps in symplectic spaces. This will be the subject of a forthcoming paper.

2. POLYNOMIAL MODELS AND REALIZATIONS

Let \( F \) be an arbitrary field. We let \( F[z] \) denote the ring of polynomials over \( F \), \( F((z^{-1})) \) the set of truncated Laurent series in \( z^{-1} \), and \( F[[z^{-1}]] \) 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 \( \pi_+ \) and \( \pi_- \) 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. Analogously we define the spaces \( F^m[z] \), \( F^{n \times m}[z] \), etc.

Given a nonsingular \( D \) in \( F^{n \times n}[z] \) we define a projection \( \pi_D \) in \( F^m[z] \) by

\[
\pi_D f = D \pi_- D^{-1} f
\]  

and let

\[
X_D = \text{Range} \; \pi_D.
\]
A polynomial vector \( f \) belongs to \( X_D \) if and only if \( D^{-1}f \) is strictly proper. \( X_D \) becomes an \( F[z] \)-module if we let
\[
p \cdot f = \pi_D(pf)
\]
for all \( p \in F[z] \) and \( f \in X_D \). We denote by \( S_D \) the map
\[
S_Df = \pi_Dzf.
\]
We call the module \( X_D \) a polynomial model.

Similarly, given a nonsingular polynomial matrix \( D \), we define a projection map \( \pi^D \) in \( z^{-1}F^m[[z^{-1}]] \) by
\[
\pi^Dh = \pi D^{-1} \pi_+ Dh
\]
and let
\[
X^D = \text{Range } \pi^D.
\]

We define an \( F[z] \)-module structure on \( X^D \) by letting
\[
p \cdot h = \pi_- ph \quad \text{for } h \in X^D, \ p \in F[z].
\]
We will denote by \( S^D \) the map in \( X^D \) given by
\[
S^Dh = \pi_- zh.
\]

In \( F^m((z^{-1})) \) we define a bilinear form by
\[
[f, g] = \sum_{j=-\infty}^{\infty} \tilde{f}_j g_{j-1},
\]
where \( f(z) = \sum_{j=-\infty}^{\infty} f_j z^j \), \( g(z) = \sum_{j=-\infty}^{\infty} g_j z^j \). We have easily that \((F^m[z])^{-1} = F^m[z]\). Moreover, since we can identify \((F^m[z])^* \) with \( z^{-1}F^m[[z^{-1}]] \) and easily establish that \((DF^m[z])^{-1} = \tilde{X}_D^D \), we have the identification of \( X^*_D \) with \( X^D \). Now \( X^D \) and \( \tilde{X}_D \) are isomorphic, and so we can identify \( X^*_D \) with \( \tilde{X}_D \) under the pairing
\[
\langle f, g \rangle = [D^{-1}f, g]
\]
for \( f \in X_D \) and \( g \in \tilde{X}_D \).
In the special case that the underlying field is the complex field \( \mathbb{C} \) we modify the definition of duality slightly. Thus we assume, given \( x, y \in \mathbb{C}^n \), that \( (x, y) \) is their usual inner product given by \( (x, y) = \sum x_i \overline{y}_i \). Then we replace (2.9) by

\[
[f, g] = \sum_j (f_j, g_{-j-1}). \tag{2.11}
\]

Consequently in many results derived in the sequel on matrix representations the Hermitian adjoint has to replace the transposed matrix. Mostly we will omit the details.

For a full analysis of duality in the context of polynomial models we refer to Fuhrmann [27]. We quote next from Fuhrmann [22].

**Theorem 2.1.** Given two polynomial models \( X_D \) and \( X_{D_i} \), corresponding to nonsingular polynomial matrices \( D \) and \( D_i \) in \( F^{n \times n}[z] \) and \( F^{n_i \times n_i}[z] \) respectively, then a map \( Z: X_D \to X_{D_i} \) is an \( F[z] \)-homomorphism if and only if there exist polynomial matrices \( M \) and \( N \) such that

\[
MD = D_iN \quad \tag{2.12}
\]

and

\[
Zf = \pi_{D_i} Mf. \quad \tag{2.13}
\]

Moreover \( Z \) is surjective if and only if \( M \) and \( D_i \) are left coprime, and injective if and only if \( N \) and \( D \) are right coprime.

Notice that (2.7) implies

\[
\tilde{N} \tilde{D}_i = \tilde{D} \tilde{M}, \quad \tag{2.14}
\]

and \( W : X_{\tilde{D}_i} \to X_{\tilde{D}} \) given by

\[
Wg = \pi_{\tilde{D}} \tilde{N} \tilde{g} \quad \tag{2.15}
\]

is also a module homomorphism. It is easily checked that actually \( W = Z^* \) where \( Z^* \) is defined as the unique map for which

\[
\langle Zf, g \rangle = \langle f, Z^* g \rangle \quad \tag{2.16}
\]
for all \( f \in X_D \) and \( g \in X_{D_1} \). We note also that \( Z: X_D \to X_{D_1} \) is an \( F[z] \)-homomorphism if and only if
\[
ZS_D = S_{D_1}Z,
\]
(2.17)
i.e., if and only if \( Z \) is an intertwining map for \( S_D \) and \( S_{D_1} \).

Up to this point the discussion has been purely module theoretic. We proceed now to make contact with realization theory. Thus let \( G \) be a \( p \times m \) strictly proper transfer function admitting what we refer to as a Rosenbrock type representation, namely one of the form
\[
G(z) = V(z)T(z)^{-1}U(z) + W(z)
\]
(2.18)
where \( V, T, U, \) and \( W \) are polynomial matrices of appropriate sizes and \( T \) is assumed nonsingular. The following theorem has been proved in Fuhrmann [22, 24].

**Theorem 2.2.** Let \( G \) have the representation (2.18). Then with the state space \( X_T \) the system \((A, B, C)\) defined by
\[
\begin{align*}
A &= S_T, \\
Bx &= \pi_T x \\text{for} &\quad x \in F^m, \\
Ce &= (VT^{-1}f)^{-1} \\text{for} &\quad f \in X_T
\end{align*}
\]
(2.19)
is a realization of \( G \). This realization is reachable if and only if \( T \) and \( U \) are left coprime, observable if and only if \( V \) and \( T \) are right coprime.

We will call this the realization associated with the representation (2.18).

Let us consider now an observable pair \((A, C)\) and the corresponding state-output transfer function \( C(zI - A)^{-1} \). By observability \( zI - A \) and \( C \) are right coprime. Let \( D^{-1}H \) be a left coprime factorization of \( C(zI - A)^{-1} \). The following lemma has been proved by Hautus and Heymann [38] and Wimmer [59].

**Lemma 2.3.** (a) Let \( D(z)^{-1}H(z) \) be a left coprime factorization of \( C(zI - A)^{-1} \) with \((A, C)\) observable. Then the columns of \( H \) form a basis for \( X_D \).

(b) Given a polynomial matrix \( N \), then \( D^{-1}N \) is strictly proper if and only if there exists a constant matrix \( B \) such that
\[
N(z) = H(z)B.
\]
(2.20)
Proof. We consider the realizations associated with the coprime factorizations

\[ C(zI - A)^{-1} = D(z)^{-1}H(z). \]

The first one has \( F^n \) as both input and state space, the input map being the identity map, which is of course both injective and surjective. This means that in the realization associated with \( D^{-1}H \), which is isomorphic to the previous one, the map

\[ x \to H(z)x \]

is both injective and surjective. This proves (a).

To prove (b) we note that the sufficiency of (2.20) for \( D^{-1}N \) to be strictly proper is trivial. Conversely, assume \( D^{-1}N \) is strictly proper. Then so is \( D^{-1}Nx \) for any constant vector \( x \). This implies that \( Nx \in X_D \) and so, as the columns of \( H \) form a basis for \( X_D \), for some vector \( b_x \) we have \( Nx = Hb_x \). By letting \( x \) vary through a set of basis elements and by linear extension, there exists a linear map \( B \) such that (2.20) holds.

We remark that Lemma 2.3 has a dual version which can be obtained easily by transposing matrices. We omit the details.

3. BASES, INTERPOLATION, AND QUADRATIC FORMS

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

\[ \langle e_i, f_j \rangle = \delta_{ij}, \quad 1 \leq i, j \leq n. \quad (3.1) \]

Given \( q \in F[z] \) with \( q(z) = z^n \cdot q_{n-1}z^{n-1} + \cdots + q_0 \), then the elements of \( X_q \) are all polynomials of degree \( \leq n - 1 \). In particular the subset of \( X_q \) given by \( \mathfrak{B}_0 = \{ f_1, \ldots, f_n \} \), where

\[ f_i(z) = z^{i-1}, \quad i = 1, \ldots, n, \quad (3.2) \]

is a basis for \( X_q \). We will refer to this as the standard basis of \( X_q \) and denote it by \( \mathfrak{B}_0 \). Since

\[ S_q z^i = \begin{cases} z^{i+1} & \text{if } 0 \leq i < n - 1, \\ - (q_0 + \cdots + q_{n-1}z^{n-1}) & \text{if } i = n - 1, \end{cases} \quad (3.3) \]
the matrix representation of \( S_q \) relative to the basis \( \mathcal{B}_0 \), is

\[
\begin{pmatrix}
0 & \cdots & 0 & -q_0 \\
1 & & & -q_1 \\
& \ddots & & \vdots \\
& & 1 & -q_{n-1}
\end{pmatrix}
\]

i.e., it is the \textit{companion matrix} of \( q \).

It is of considerable interest to characterize the elements of the dual basis. Since we have the identification \( X_q^* = X_q \) under the pairing \((2.10)\), the dual basis elements are also polynomials of degree \(< n \). Given the polynomial \( q \) as above, we define

\[
e_i(z) = \pi_+ z^{-i} q = q_i + q_{i+1} z + \cdots + z^{n-i}, \quad i = 1, \ldots, n.
\]

**Theorem 3.1.** The set \( \mathcal{B}_c = \{e_1, \ldots, e_n\} \) is the dual basis to the basis \( \mathcal{B}_0 = \{1, z, \ldots, z^{n-1}\} \), relative to the bilinear form \( \langle , \rangle \) introduced in \((2.10)\).

**Proof.** We have

\[
\langle e_i, f_j \rangle = \left[ q^{-1} e_i, f_j \right] = \left[ q^{-1} \pi_+ z^{-i} q, z^{j-1} \right]
\]

\[
= \left[ \pi_- q^{-1} \pi_+ qz^{-i}, z^{j-1} \right]
\]

\[
= \left[ \pi_- q z^{-i}, z^{j-1} \right] = \left[ z^{-i}, \pi_+ z^{j-1} \right]
\]

\[
= \left[ z^{-i}, z^{j-1} \right] = \delta_{ij}.
\]

It follows from the general study of duality that the matrix representation of \( S_q \) with respect to the basis \( \mathcal{B}_c \) is

\[
\begin{pmatrix}
0 & 1 \\
\vdots & \ddots \\
- q_0 & - q_1 & \cdots & - q_{n-1}
\end{pmatrix}
\]

which is the control form. For this reason \( \mathcal{B}_c \) is referred to as the \textit{control
basis. Of course the matrix representation (3.6) can be verified directly by observing that

\[
S_q e_i = \begin{cases} 
  e_{i-1} - q_{j-1} e_n & \text{if } i = 2, \ldots, n \\
  -q_0 e_n & \text{if } i = 1.
\end{cases}
\]

(3.7)

As it is obvious that

\[ IS_q = S_q I, \]

(3.8)

the next corollary follows trivially.

**Corollary 3.2.** If $\mathcal{B}_0, \mathcal{B}_c$ are the standard and control bases of $X_q$, then

\[
[I]^{\mathcal{B}_0 \mathcal{B}_c} [S_q]^{\mathcal{B}_0 \mathcal{B}_c} = [S_q]^{\mathcal{B}_0 \mathcal{B}_c} [I]^{\mathcal{B}_0 \mathcal{B}_c}.
\]

(3.9)

But

\[
[I]^{\mathcal{B}_0 \mathcal{B}_c} = \begin{pmatrix}
  q_1 & \cdots & & q_{n-1} & 1 \\
  \vdots & \ddots & \ddots & \vdots & \vdots \\
  \vdots & & \ddots & \vdots & \vdots \\
  q_{n-1} & \cdots & & 0 & \\
  1 & & & & \end{pmatrix},
\]

(3.10)

so it follows that this Hankel matrix intertwines the companion matrices (3.4) and (3.6). This result appears in Taussky [54], Barnett [2], Langer [51], Gohberg et al. [34, 35], and Kailath [44].

Consider now the special case in which $q$ has only simple zeros, i.e.,

\[ q(z) = \prod (z - a_i) \] with $a_i \neq a_j$ for $i \neq j$. Let

\[
d_i(z) = \frac{q(z)}{z - a_i} = \prod_{j \neq i} (z - a_j).
\]

(3.11)

**Lemma 3.3.** The $d_i$ are eigenfunctions of $S_q$ corresponding to the eigenvalues $a_i$. Conversely, every eigenfunction of $S_q$ is a multiple of a $d_i$. 
Proof. With $d_i$ as above we have

$$(S_q - a_i)d_i = \pi_a(z - a_i)d_i = \pi_a(z - a_i)q(z - a_i)^{-1} = \pi_a q = 0.$$  

Conversely, let $f$ be an eigenfunction corresponding to the eigenvalue $a$. Then $S_q f - af$ or $\pi_a(z - a)f = 0$, which means that for some polynomial $h$ we have $(z - a)f(z) = q(z)h(z)$. Since $(z - a)f(z)/q(z)$ is a proper rational function, it follows that necessarily $h$ is a constant, say $c$. Thus $f(z) = cq(z) (z - a)^{-1}$, and as $f$ is a polynomial we must have $h$ is a constant, say $c$. Thus $f(z) = cq(z) (z - a)^{-1} = cd_i(z)$, which concludes the proof.

The next lemma characterizes the dual base.

**Lemma 3.4.** Let $q(z) = \prod_{i=1}^n (z - a_i)$, and assume $a_i = a_j$ if and only if $i = j$. Then $\{v_i, \ldots, v_n\}$, with $v_i(z) = d_i(z)/d_i(a_i)$ where $d_i$ is defined by (3.11), is a basis of $X_q$, and its dual basis $\{v_i^*, \ldots, v_n^*\}$ is given by $v_i^* = d_i$.

Proof. By our assumptions and Lemma 3.4, the $d_i$ are eigenfunctions of $S_q$ corresponding to different eigenvalues, and hence linearly independent. Since $\dim X_q = n$, they form a basis. Now

$$\langle d_i, d_j \rangle = [q^{-1}d_i, d_j] = [d_i, q^{-1}d_j]$$

$$= [d_i, q^{-1}q(z - a_j)^{-1}] = [d_i, (z - a_j)^{-1}] = d_i(a_j).$$

Obviously $d_i(a_j) = 0$ for $i \neq j$ and is different from zero if $i = j$. Thus

$$\left\langle \frac{d_i}{d_i(a_i)}, d_j \right\rangle = \delta_{ij},$$

which proves the lemma.

Let us compute the matrix representation of $S_q$ relative to the spectral basis $\{d_i/d_i(a_i)\}$, which we will refer to as the spectral basis or alternatively as the interpolation basis of $X_q$, and will denote by $\mathcal{B}_s$. Since

$$S_q d_i = a_i d_i,$$

also

$$S_q(d_i(a_i)^{-1}d_i) = a_i(d_i(a_i)^{-1}d_i).$$
and hence

$$\left[ S_q \right]_{\mathbb{V}_i} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$ 

By duality, or by reference to (3.13), we have also

$$\left[ S_q \right]_{\mathbb{V}_i} = \left[ S_q \right]_{\mathbb{V}_i}^*.$$ 

As the polynomials $d_i(z)/d_i(a_i)$ (which are just the Lagrange interpolation polynomials) are a basis of $X_q$, we have, for $f \in X_q$, that

$$\sum_{j=1}^{n} c_j d_j(z) = f(z).$$

Evaluating this equality at $z = a_i$, we obtain

$$f(a_i) = \sum_{j=1}^{n} \frac{c_j d_j(z)}{d_j(a_i)} = c_i.$$ 

This explains the reference to interpolation. In fact $f$, defined by (3.14), is the unique solution to the problem of finding a polynomial of degree $< n$ satisfying $f(a_i) = c_i$, i.e., which interpolates the value $c_i$ at $a_i$, $i = 1, \ldots, n$. Another suggestive way of writing (3.14) is

$$X_q = \sum d_j X_{z-a_j}.$$ 

Yet another restatement is the following. Given prescribed remainders modulo $q_i(z) = z - a_i$ (namely constants $c_i$), then there exists a unique polynomial $f$ of degree $< n$ such that $f \mod (z - a_i)$ is $c_i$, $i = 1, \ldots, n$. This is a special case of the Chinese remainder theorem; see for example Lang [50], Newman [52]. This theorem we prove next in the context of polynomial models.

**Theorem 3.5** (Chinese remainder theorem). Let $q_i \in F[z]$ be pairwise coprime polynomials, and let $q = q_1 \cdots q_s$. Then given polynomials $a_i$ such
that \( \deg a_i < \deg q_i \), there exists a unique polynomial \( f \) such that \( \deg f < \deg q \) and \( f \mod q_i = a_i \).

**Proof.** By the mutual coprimeness of the \( q_j \) we have the direct sum, or spectral, decomposition

\[
X_q = d_1 X_{q_1} \oplus \cdots \oplus d_s X_{q_s}.
\]

(3.15)

For details see Fuhrmann and Willems [30]. Also note that \( S_q d_i a = d_i (S_{q_i} a) \).

Let now \( \deg a_i < \deg q_i \); then \( \pi_{q_i} a_i = a_i \), i.e., \( a_i \in X_{q_i} \). Define \( f \) by

\[
f = \sum_{j=1}^{n} d_i \left( d_j (S_{q_j})^{-1} a_j \right).
\]

(3.16)

Then clearly \( f \in X_q \), i.e., \( \deg f < \deg q \), and

\[
\pi_{q_i} f = \pi_{q_i} \sum_{j=1}^{n} d_j \left( d_j (S_{q_j})^{-1} a_j \right)
\]

\[
= \sum_{j=1}^{n} d_j (S_{q_j}) d_j (S_{q_j})^{-1} a_j = a_i,
\]

as, by the divisibility of \( d_j \) by \( q_i \) for \( i \neq j \), we have \( d_j (S_{q_j}) = 0 \) for \( i \neq j \). Note that the invertibility of \( d_j (S_{q_j}) \) follows from Theorem 2.1 by the coprimeness of \( q_i \) and \( d_j \). This proves the existence part. Suppose now \( f \in X_q \) and \( f \mod q_i = 0 \) for all \( i \). By the mutual coprimeness of the \( q_i \) it follows that \( f \) is divisible by \( q \), and hence necessarily (as \( \deg f < \deg q \)) \( f = 0 \), which proves uniqueness.

It is interesting to observe that in this simple proof coprimeness is used in two different ways. First one uses it to obtain the direct sum representation (3.15), and then to insure the invertibility of \( d_i (S_{q_i}) \). Since both results have a multivariable generalization, one expects also a multivariable version of the theorem. This is indeed the use, and we will return to it in Section 7.

Let us focus now on a special case of the spectral decomposition (3.15).

**Lemma 3.6.** Let \( p, q \in F[z] \). Then \( p \) and \( q \) are coprime if and only if

\[
X_{pq} = qX_p + pX_q.
\]

(3.17)

**Proof.** Assume \( p \) and \( q \) are coprime. Then the equality (3.17) follows from Theorem 2.13 in Fuhrmann and Willems [30], but it can easily be
proved directly. By the Euclidean algorithm there exist polynomials $a$ and $b$ in $F[z]$ for which $ap + bq = 1$. This means that any polynomial $f$ can be written as $f = ap + bq$. This implies
\[
\pi_{pq} f = p \pi_q a + q \pi_p b,
\]
and hence the inclusion
\[
X_{pq} \subseteq pX_q + qX_p
\]
follows. The inverse inclusion holds always, as $pX_q$ and $qX_p$ are submodules of $X_{pq}$.

Conversely, assume (3.17); then $1 = pa + qb$ with $a \in X_q$, $b \in X_p$. This implies the coprimeness of $p$ and $q$. 

The previous lemma yields as a direct corollary the classical Sylvester resultant test for the coprimeness of two polynomials. Given two polynomials $p(z) = p_0 + \cdots + p_m z^m$ and $q(z) = q_0 + \cdots + q_n z^n$, we let
\[
R(p, q) = \begin{pmatrix}
   p_0 & \cdots & p_m & 0 \\
   \vdots & \ddots & \vdots & \vdots \\
   p_0 & \cdots & p_m & 0 \\
   q_0 & \cdots & q_n & 0 \\
   \vdots & \cdots & \vdots & \vdots \\
   q_0 & \cdots & q_n & 0
\end{pmatrix},
\]
which we call the \textit{Sylvester resultant matrix}. We have then

\textbf{Theorem 3.7.} Let $p$ and $q$ be given as above. Then $p$ and $q$ are coprime if and only if $\det R(p, q) \neq 0$.

\textbf{Proof.} The sets $\{p(z)z^i : i = 0, \ldots, n - 1\}$ and $\{q(z)z^i : i = 0, \ldots, m - 1\}$ are bases for $pX_q$ and $qX_p$ respectively. Thus $p$ and $q$ are coprime if and only if their union is a basis for $X_{pq}$. Expressing this in terms of the polynomial coefficients, which are the coordinates relative to the standard basis of $X_{pq}$, yields the result. 

Let us analyze now the case of multiple zeros; the field is assumed to be $R$. If $q(z) = z^m$, then $\mathcal{B} = \{1, z, \ldots, z^{m-1}\}$ is a basis for $X_q$ and its dual basis is
SYMMETRIC RATIONAL TRANSFER FUNCTIONS

181

given by \( \mathfrak{B}^* = (z^{m-1}, \ldots, z, 1) \). Clearly in this case

\[
[I]_{\mathfrak{B}^*} = \begin{pmatrix}
0 & 0 & \cdots & 0 & 1 \\
0 & 0 & \cdots & 1 \\
& & \ddots & \cdots & \ddots \\
0 & 1 \\
1
\end{pmatrix}.
\] (3.18)

In the case of a power of a quadratic polynomial the situation is not much more complicated. Let \( q(z) = (z^2 + 1)^m \); then obviously

\[
\mathfrak{B} = \{1, z, (z^2 + 1), z(z^2 + 1), \ldots, (z^2 + 1)^{m-1}, z(z^2 + 1)^{m-1}\}
\]
is a basis for \( X_q \) and its dual basis is given by

\[
\mathfrak{B}^* = \{z(z^2 + 1)^{m-1}, (z^2 + 1)^{m-1}, \ldots, z, 1\}.
\]

This follows from the following easy computation, with \( 0 \leq k, l \leq 1 \):

\[
\langle (z^2 + 1)^i z^k, (z^2 + 1)^l z^i \rangle = \left[ (z^2 + 1)^{m-i} (z^2 + 1)^{i+l} z^{k+l}, 1 \right]
\]

\[
= \begin{cases}
0 & \text{if } i+j \neq m-1, \\
0 & \text{if } i+j = m-1 \text{ and } k+l \neq 1, \\
1 & \text{if } i+j = m-1 \text{ and } k+l = 1.
\end{cases}
\]

Again it is clear that (3.18) holds where the matrix on the right is of size \( 2m \times 2m \).

The following lemma gives a useful computational rule. Define a map \( T_a \) in \( F^m((z^{-1})) \) by

\[
(T_a f)(z) = f(z - a).
\] (3.19)

**Lemma 3.8.** Given \( f, g \in F^m((z^{-1})) \), then

(a) \( [T_a f, g] = [f, T_{-a} g] \),

(b) \( [T_a f, T_a g] = [f, g] \).

**Proof.** A direct computation shows that

\[
\left[ (z-a)^i, 1 \right] = \begin{cases}
0, & i = -1, \\
1, & i = -1.
\end{cases}
\]
and more generally

\[
[(z-a)^j, (z-a)^k] = \begin{cases} 
0, & j+k \neq -1, \\
1, & j+k = -1.
\end{cases}
\]

This implies, for \(f(z) = \sum_j f_j z^j\) and \(g(z) = \sum_k g_k z^k\),

\[
[T_a f, T_a g] = \left[ \sum_j f_j (z-a)^j, \sum_k g_k (z-a)^k \right] \\
= \sum_{j,k} \bar{g}_k f_j \left[ (z-a)^j, (z-a)^k \right] \\
- \sum_j \bar{g}_{-j-1} f_j = [f, g],
\]

which proves (b). (a) follows by

\[
[T_a f, g] = [T_{-a} T_a f, T_{-a} g] = [f, T_{-a} g].
\]

Using this lemma, we can state

**LEMMA 3.9.**

(a) If \(q(z) = (z-a)^m\), then

\[
\mathcal{B} = \{1, z-a, \ldots, (z-a)^{m-1}\} \tag{3.20}
\]

and

\[
\mathcal{B}^* = \{(z-a)^{m-1}, \ldots, 1\} \tag{3.21}
\]

are a pair of dual bases for \(X_{(z-a)^m}\).

(b) If \(q(z) = [(z-a)^2 + b^2]^m\), then

\[
\mathcal{B} = \{1, z-a, [(z-a)^2 + b^2], (z-a)[(z-a)^2 + b^2], \ldots, \\
\times [(z-a)^2 + b^2]^{m-1}, (z-a)[(z-a)^2 + b^2]^{m-1}\} \tag{3.22}
\]

and

\[
\mathcal{B}^* = \{(z-a)[(z-a)^2 + b^2]^{m-1}, [(z-a)^2 + b^2]^{m-1}, \ldots, z-a, 1\} \tag{3.23}
\]

are a pair of dual bases for \(X_{[(z-a)^2 + b^2]^m}\).
We note that the matrix representation of $S_q$ relative to the basis $\mathfrak{B}$ of (3.20) is

\[
[S_q]^{\mathfrak{B}} = \begin{pmatrix}
 a & a & \cdots & a \\
 1 & a & \cdots & 1 \\
 & & \ddots & & \\
 & & & 1 & a
\end{pmatrix},
\] (3.24)

i.e., it is the corresponding Jordan block. Analogously, in case (b) we have, with respect to the basis $\mathfrak{B}$ of (3.22), the matrix representation

\[
[S_q]^{\mathfrak{B}} = \begin{pmatrix}
 a & -b^2 & & \\
 1 & a & 0 & \\
 0 & 1 & a & -b^2 \\
 0 & 0 & 1 & a & 0 \\
 & & & \ddots & \\
 & & & & \ddots \\
 0 & 1 & a & -b^2 & \\
 0 & 0 & 1 & a & 
\end{pmatrix},
\] (3.25)

which is essentially the real Jordan canonical form corresponding to $[(z - a)^2 + b^2]^m$.

Let us consider now the general scalar case where

\[
q(z) = q_1(z) \cdots q_i(z)
\] (3.26)

is a factorization of $q$ into powers of irreducible coprime polynomials. Thus $q_i(z) = (z - a_i)^{m_i}$ or $q_i(z) = [(z - a_i)^2 + b_i^2]^{m_i}$. Define, as previously,

\[
d_i = q/q_i.
\] (3.27)

We have then the spectral decomposition of $X_q$ given by

\[
X_q = d_1 X_{q_1} \oplus \cdots \oplus d_i X_{q_i}.
\] (3.28)

In the previous lemma we constructed a pair of dual bases for $X_q$, and we would like to use these in the construction of a pair of dual bases for $X_q$.

We note first that the direct sum decomposition (3.28) is also an orthogonal direct sum decomposition relative to the indefinite metric in $X_q$. Indeed,
for \( i \neq j \) if \( f d_j X_q \) and \( g d_j X_q \), then \( f = d_i f_i, g = d_j g_j \) and

\[
\langle f, g \rangle = \langle d_i f_i, d_j g_j \rangle
\]

\[
= \left[ q^{-1} d_i f_i, d_j g_j \right] = \left[ d_j q^{-1} d_i f_i, g_j \right] = 0,
\]
as \( q \) divides \( d_i d_j \), or equivalently \( d_j q^{-1} d_i \) is a polynomial.

Let now \( \{v_i^{(k)}\}_{i=1} \) be a basis of \( X_{a_k} \) which satisfies

\[
\langle v_i^{(k)}, v_j^{(k)} \rangle \delta_{i, (m_k + 1 - j)},
\]
i.e., the dual basis is \( \{v_{m_k + 1 - j}\}_{i=1} \). The existence of such bases has been proved in Lemma 3.9. Define now

\[
\mathcal{B} = \left\{ d_k \left( d_k \left( S_{k} \right)^{-1} \right) v_i^{(k)} : k = 1, \ldots, s, i = 1, \ldots, m_k \right\}. \tag{3.29}
\]

Since \( d_k \left( S_{k} \right) \) is, by the coprimeness of \( d_k \) and \( q_k \) an invertible map in \( X_{q} \), then this is clearly a basis in \( X_{q} \).

**Theorem 3.10.** The basis \( \mathcal{B}^* \) dual to the basis \( \mathcal{B} \) of (3.29) is

\[
\mathcal{B}^* = \left\{ d_k v_i^{(k)} : k = 1, \ldots, s, i = 1, \ldots, m_k \right\}. \tag{3.30}
\]

**Proof.** By the orthogonality of the invariant subspaces \( d_i X_q \), it suffices to prove

\[
\langle d_k \left( d_k \left( S_{k} \right)^{-1} v_{m_k + 1 - j}^{(k)} \right), d_k v_i^{(k)} \rangle = \delta_{ij}.
\]

Indeed,

\[
\langle d_k \left( d_k \left( S_{k} \right)^{-1} v_{m_k + 1 - j}^{(k)} \right), d_k v_i^{(k)} \rangle_{X_q} = \left[ q^{-1} d_k \left( d_k \left( S_{k} \right)^{-1} v_{m_k + 1 - j}^{(k)} \right), d_k v_i^{(k)} \right]
\]

\[
= \left[ q_k^{-1} d_k \left( d_k \left( S_{k} \right)^{-1} v_{m_k + 1 - j}^{(k)} \right), v_i^{(k)} \right]
\]

\[
= \langle d_k \left( S_{k} \right) d_k \left( S_{k} \right)^{-1} v_{m_k + 1 - j}^{(k)} v_i^{(k)} \rangle_{X_{a_k}}
\]

\[
= \langle v_{m_k + 1 - j}^{(k)} v_i^{(k)} \rangle_{X_{a_k}} = \delta_{ij}. \]

\[ \blacksquare \]
We proceed next to introduce the notion of a self-dual map. Given a (not necessarily finite dimensional) vector space $X$ and its dual space $X^*$, then a map $Z: X \rightarrow X^*$ is self-dual if

$$\langle Zx, y \rangle = \langle x, Zy \rangle$$

for all $x, y \in X$.

Given a self-dual map $Z: X \rightarrow X^*$, we have associated with it a quadratic form $\langle Zx, x \rangle$. For any choice of basis in $X$, say $(e_1, e_2, \ldots)$, and with $x = \sum x_i e_i$, we have $\langle Zx, x \rangle = \sum \sum Z_{ij} x_i x_j$, where $(Z_{ij})$ is the matrix representation relative to the basis $(e_1, e_2, \ldots)$ and its dual basis in $X^*$. Obviously $(Z_{ij})$ is symmetric.

In case the field is real or complex, the signature of $Z$, denoted by $\sigma(Z)$, is defined as the difference between the number of positive and negative eigenvalues of $Z$. By the Sylvester inertia theorem (see Gantmacher [31]), the signature is well defined, i.e., it is independent of the choice of basis.

4. ON GENERALIZED BEZOUTIANS

Bezoutians, quadratic forms associated with a pair of polynomials, have been used for a long time in the theory of equations, giving criteria for coprimeness of a pair of polynomials, and in the analysis of location of zeros of polynomials, especially in stability theory, as in Hermite's work [39]. Recently the classical concept of a Bezoutian has been generalized by Anderson and Jury [1]. It is interesting to note that Bezoutians appear naturally in the polynomial analysis of $(A, B)$-invariant subspaces; see Fuhrmann and Willems [30]. The main point is that Bezoutians are special (block) matrix representations of intertwining maps. That they can be used to good purpose will be seen in the next section, where they will be used to compute the signature of a Hankel matrix, simplifying a long-standing result of Frobenius [20].

Let $F$ be an arbitrary field, and let $G$ be a $p \times m$ strictly proper transfer function, i.e., $G$ is a rational element in $z^{-1}F^{p \times m}[[z^{-1}]]$. Let

$$G(z) = T(z)^{-1}U(z) = N(z)D(z)^{-1}$$

be two matrix fraction representations of $G$, where no coprimeness assumptions are made. The generalized Bezoutian associated with the quadruple $(T, U, N, D)$ is

$$\Gamma(z, w) = \frac{T(z)N(w) - U(z)D(w)}{z - w}.$$
Clearly, since (4.1) implies
\[ T(z)N(z) = U(z)D(z), \]
it follows that \( \Gamma(z, w) \) is a polynomial in the variables \( z, w \). Thus \( \Gamma(z, w) \) can be written as
\[ \Gamma(z, w) = \sum_i \Gamma_{ij} z^{i-1} w^{j-1} \quad (4.4) \]
with \( \Gamma_{ij} \in F^{p \times m} \). We saw in the previous section how a matrix polynomial relation of the form (4.3) is linked to module homomorphisms and consequently to intertwining maps.

Let now \( T, U, N, D \) be as in (4.1), and let \( \Gamma \) be the associated generalized Bezoutian. Define the map \( z : X_D \to X_T \) by
\[ Zf = \pi_T U f \quad \text{for} \quad f \in X_D. \quad (4.5) \]

**Lemma 4.1.** For each vector \( x \in F^m \) and each \( j \) we have \( \sum_i \Gamma_{ij} x^{i-1} w^{j-1} \in X_T \).

**Proof.** Since \( \Gamma(z, w)x = \sum_i \sum_j \Gamma_{ij} x^{i-1} w^{j-1} \), it follows that
\[ \pi_+ w^{-(k-1)} \left( \sum_i \sum_j \Gamma_{ij} x^{i-1} w^{j-1} \right) \bigg|_{w=0} = \sum_i \Gamma_{ik} x w^{i-1}. \quad (4.6) \]
Let on the other hand
\[ g(z) = \frac{(T(z)N(w) - U(z)D(w))x}{z - w}. \quad (4.7) \]
Now clearly for each \( w \in F \) we have \( g \in X_T \) as
\[ T^{-1} g = \frac{N(w)x}{z - w} - \frac{T(z)^{-1} U(z)D(w)x}{z - w}, \]
and this shows that \( T^{-1} g \) is strictly proper in the variable \( z \). Now \( g \) can be written as \( g(z) = \sum g_j(w) z^j \), which is in \( X_T \) for all \( w \). In particular it follows that
\[ \pi_+ w^{-1} g \big|_{w=0} = \sum_j \left( g_j(w) - g_j(0) \right) w^{-1} z^j \big|_{w=0} \in X_T, \]
and hence, by induction, also
\[ \pi_+ w^{-(k-1)}g|_{w=0} \in X_T. \]

Assume now that in the matrix fraction \( G = ND^{-1} \) the polynomial matrix 
\( D(s) = D_0 + D_1 z + \cdots + D_s z^s \) is row proper. In this case it has been shown in Fuhrmann \[22\] that \( X_D \) admits the control representation
\[ X_D = \left\{ \sum_j E_j(z)x_j : x \in F^m \right\}. \quad (4.8) \]

Here \( E_j \) are the polynomial matrices defined by
\[ E_j = \pi_+ z^{-j}D \quad (4.9) \]
or
\[ E_j(z) = D_j + D_{j+1} z + \cdots + D_s z^{s-j}. \]

While the \( x_j \) are not uniquely determined, the decomposition \( f(z) = \sum E_j(z)x_j \) of \( f \) into components in the direct sum \( X_D = W_1 \oplus \cdots \oplus W_s \), where
\[ W_j = \{ E_j(z)x : x \in F^m \}, \]
is unique.

We can now identify the generalized Bezoutian as a block matrix representation of an intertwining map relative to two direct sum decompositions of \( X_D \) and \( X_T \).

**Theorem 4.2.** Let \( G \) be a \( p \times m \) strictly proper transfer function admitting the matrix fraction representations
\[ G = T^{-1}U = ND^{-1}, \quad (4.10) \]
and assume that \( D \) is column proper. Let \( Z : X_D \to X_T \) be defined by (4.5). Then
\[ Z(E_k x) = \sum_i \Gamma_{ik} x x^{-i}. \quad (4.11) \]
In other words, the generalized Bezoutian associated with \((T, U, N, D)\) is the block matrix representation of \(Z\) relative to the direct sum decompositions of \(X_D\) given by

\[X_D = W_1 \oplus \cdots \oplus W_s\]  \hspace{1cm} (4.12)

with \(W_i = (E_i(z)x : x \in F^m)\), where \(E_i\) is defined by (4.9), and

\[X_T = V_1 \oplus \cdots \oplus V_s\]  \hspace{1cm} (4.13)

with \(V_i = (z^{i-1}x : x \in F^m)\).

Proof. We clearly have \((T(z)N(w) - U(z)D(w))x = \sum \sum \Gamma_{ij} x^{j-i-1}(z - w)w^{j-1}\). Fixing \(z\), we apply the operator \(\pi_+ w^{-k}\) to both sides and obtain

\[T(z)\pi_+ w^{-k}N(w)x - U(z)\pi_+ w^{-k}D(w)x = \sum \sum \Gamma_{ij} x^{j-i-1}\pi_+ w^{-k}(z - w)w^{j-1}.\]

Now

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

Substituting \(w = z\) in the previous equality, we obtain

\[T(z)\pi_+ z^{-k}N(z)x - U(z)\pi_+ z^{-k}D(z)x = - \sum \Gamma_{ik} x^{i-1}.\]

Applying next the projection \(\pi_T\), we have

\[\pi_T T(z)\pi_+ z^{-k}N(z)x = 0\]

and so

\[-\pi_T U\pi_+ z^{-k}D(z)x = -\pi_T UE_jx = - \sum \Gamma_{ik} x^{i-1},\]

which proves the theorem.
Our identification of the Bezoutian as a matrix representation of an intertwining map easily produces some interesting corollaries. In this connection see also Fuhrmann [28]. Given the matrix fraction representations (4.1), the map $Z : X_D \to X_T$ defined by (4.5) is an intertwining map for $S_D$ and $S_T$, i.e.

$$ZS_D = S_TZ.$$  \hspace{1cm} (4.14)

Given a choice of basis $\mathcal{A}$ in $X_D$ and $\mathcal{A}_1$ in $X_T$, then we have the following matrix equality:

$$\begin{bmatrix} Z \end{bmatrix}_{\mathcal{A}_1} \begin{bmatrix} S_D \end{bmatrix}_{\mathcal{A}} = \begin{bmatrix} S_T \end{bmatrix}_{\mathcal{A}_1} \begin{bmatrix} Z \end{bmatrix}_{\mathcal{A}}.$$  \hspace{1cm} (4.15)

In particular let us specialize to the scalar case. Thus let $p$ and $q$ be polynomials with $\deg p \leq \deg q$, and let $g = p/q$ be a rational function. Let $q(z) = z^n + q_{n-1}z^{n-1} + \cdots + q_0$, and let

$$B(q, p)(z, w) = \frac{q(z)p(w) - p(z)q(w)}{z - w} = \sum b_{ij}z^{i-1}w^{j-1} \hspace{1cm} (4.16)$$

We will denote by $B$ also the matrix $(b_{ij})$. The elements of the space $X_q$ are polynomials of degree $\leq n-1$. We denote by $\mathcal{A}_0$ the standard basis of $X_q$, i.e. $\mathcal{A} = \{1, z, \ldots, z^{n-1}\}$. The control basis will be denoted by $\mathcal{A}_C = (e_1, \ldots, e_n)$, where $e_i(z) = z^{n-i} + \cdots + q_{i+1}z + q_i$.

Of course as a consequence of Theorem 4.2 we know that

$$B(q, p) = \begin{bmatrix} Z \end{bmatrix}_{\mathcal{A}_0},$$  \hspace{1cm} (4.17)

where $Z : X_q \to X_q$ is given by

$$Zf = \pi_q pf = p(S_q)f.$$  \hspace{1cm} (4.18)

The next result, a trivial generalization of Corollary 3.2, has been previously obtained by Taussky [54] and Barnett [2].

**Corollary 4.3.** Let $q(z) = z^n + q_{n-1}z^{n-1} + \cdots + q_0$, and let $C_q$ be the companion matrix of $q$, i.e.

$$C_q = \begin{pmatrix} 0 & -q_0 \\ 1 & -q_1 \\ \vdots & \vdots \\ \vdots & \vdots \\ 1 & -q_{n-1} \end{pmatrix}.$$  \hspace{1cm} (4.19)
Then
\[ B(q, p) = J p(\hat{C}_q) = p(C_q)J, \]
where
\[ J = \begin{pmatrix} q_1 & \cdots & q_{n-1} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ 1 & \cdots & 1 & q_{n-1} \end{pmatrix}. \] (4.20)

Proof. Since \( B(q, p) = [p(S_q)]_{\mathbb{P}_c}^{\mathbb{P}_0} \), it follows that
\begin{align*}
B(q, p) &= [p(S_q)]_{\mathbb{P}_c}^{\mathbb{P}_0} [I]_{\mathbb{P}_c}^{\mathbb{P}_0} - [I]_{\mathbb{P}_c}^{\mathbb{P}_0} [p(S_q)]_{\mathbb{P}_c}^{\mathbb{P}_0} \tag{4.21} \\
&= p \left( [S_q]_{\mathbb{P}_c}^{\mathbb{P}_0} [I]_{\mathbb{P}_c}^{\mathbb{P}_0} - [I]_{\mathbb{P}_c}^{\mathbb{P}_0} [S_q]_{\mathbb{P}_c}^{\mathbb{P}_0} \right). \tag{4.22}
\end{align*}

Now \( [S_q]_{\mathbb{P}_c}^{\mathbb{P}_0} = C_q \), and a simple computation shows that \( [S_q]_{\mathbb{P}_c}^{\mathbb{P}_c} = \hat{C}_q \). Moreover, since \( e_i(z) = z^{n-i} + q_{n-i-1} + \cdots + q_i \), it is clear that \( [I]_{\mathbb{P}_c}^{\mathbb{P}_0} = J \). \( \blacksquare \)

The next corollary was derived by Datta [16] and Gohberg et al. [32].

**Corollary 4.4.** Let \( B(q, p) \) be the Bezoutian of the polynomials \( q \) and \( p \). Then
\[ C_q B(q, p) = B(q, p) \hat{C}_q. \tag{4.23} \]

Proof. Since \( S_q \) commutes with \( p(S_q) \), it follows that
\[ [S_q]_{\mathbb{P}_c}^{\mathbb{P}_0} [p(S_q)]_{\mathbb{P}_c}^{\mathbb{P}_0} = [p(S_q)]_{\mathbb{P}_c}^{\mathbb{P}_0} [S_q]_{\mathbb{P}_c}^{\mathbb{P}_0}. \] (4.24)

It is clear from these examples that the fundamental result on the Bezoutians is given by Theorem 4.2 and it [or say the relation (4.14)] can
provide an abundance of matrix equalities depending on how many specific base choices we are ready to make.

For a transfer function $G$, let us denote by $\delta(G)$ the McMillan degree of $G$. For a definition see Kalman, Falb, and Arbib [46] or Brockett [5]. The following result, due to Anderson and Jury [1], follows as an easy corollary.

**Corollary 4.5.** Let $\Gamma$ be a generalized Bezoutian associated with the matrix fraction representations

$$G = T^{-1}U = ND^{-1}$$

of a strictly proper transfer function $G$. Then $\text{rank } \Gamma = \delta(G)$.

**Proof.** Range $Z$ is the reachability subspace of the realization associated with the factorization $T^{-1}U$ of $G$. Since this is an observable realization, it follows that $\text{rank } \Gamma = \dim \text{Range } Z = \delta(G)$.

Next we study some implications of symmetry as related to signatures and Bezoutians.

Let $G$ be an $m \times m$ strictly proper transfer function which is symmetric, i.e., $\tilde{G}(z) = G(z)$, where $G(z) = \sum G_j z^{-j}$ and $\tilde{G}(z) = \sum \tilde{G}_j z^{-j}$. Assume $G$ has a matrix fraction representation

$$G = Q^{-1}P; \quad \text{(4.25)}$$

then also

$$\tilde{G} = \tilde{P}\tilde{Q}^{-1}. \quad \text{(4.26)}$$

Let now $G$ be a symmetric strictly proper rational $m \times m$ matrix function. Assume that $G$ has a matrix fraction representation of the form $G = Q^{-1}P = \tilde{P}\tilde{Q}^{-1}$.

There are two self-dual maps associated with $G$. The first one is the Hankel map $H_G$ induced by $G$. We have $H_G: R^m[z] \rightarrow z^{-1}R^m[[z^{-1}]]$ defined by

$$H_Gf = \pi_- Gf \quad \text{for } f \in R^m[z]. \quad \text{(4.27)}$$

Since $R^m[z]^* = z^{-1}R^m[[z^{-1}]]$ it follows from the symmetry of $G$ that $H_G$ is a self-dual map. Given $f(z) = \sum f_j z^j$ with $f_j \in R^m$, we have

$$\langle H_Gf, f \rangle = \sum_i \sum_j (G_{i+j-1} f_{i-1})^{-1} f_{j-1} \quad \text{(4.28)}$$

with $G(z) = \sum G_j z^{-j}$. 

The second self-dual map arises out of the equality \( Q^{-1}P = \tilde{P}\tilde{Q}^{-1} \), which yields
\[
P\tilde{Q} = Q\tilde{P},
\]
and consequently a self-dual map \( Z: \tilde{X}_\beta \rightarrow X_\alpha \) given by
\[
Zf = \pi_\alpha Pf \quad \text{for} \quad f \in X_\alpha.
\]

**Theorem 4.6.** Let \( G \) be an \( m \times m \) symmetric strictly proper transfer function, and let (4.25) be a matrix fraction representation, assuming \( Q \) is row proper. Let \( (\Gamma_{ij}) \) be the generalized Bezoutian associated with \((Q, \tilde{P}, P, Q)\). Then \( \Gamma_{ij} = \Gamma_{ji} \), i.e., the Bezoutian is block symmetric.

**Proof.** Since in this case the direct sum decompositions of \( X_\alpha \) and \( X_\beta \) given by (4.12) and (4.13) are dual decompositions, it follows that the corresponding block matrix representation \([Z]\), which by Theorem 4.2 is just the generalized Bezoutian, is block symmetric.

The next theorem relates the quadratic forms that correspond to the two self-adjoint maps \( Z \) and \( H_G \).

**Theorem 4.7.** Let \( G \) be a symmetric \( m \times m \) strictly proper transfer function. Let \( H_G \) be the Hankel map induced by \( G \), let \( Z: \tilde{X}_\beta \rightarrow X_\alpha \) be the map defined by (4.30), and let \( \Gamma \) be the Bezoutian associated with \((Q, \tilde{P}, P, Q)\). Then the signatures of \( H_G \), \( Z \), and \( \Gamma \) are equal.

**Proof.** Let \( f \in \mathbb{R}^m[z] \). Then if \( f \in \tilde{Q}\mathbb{R}^m[z] \), we have \( f = Qg \) and \( H_G f = \pi_- Gf = \pi_- P\tilde{Q}^{-1}Qg = \pi_- \tilde{P}g = 0 \). So it follows that
\[
[H_G f, f] = [H_G \pi_\beta f, \pi_\alpha f] = \langle \pi_\alpha G \pi_\beta f, \pi_\alpha f \rangle - \langle \pi_- \tilde{G} \pi_\beta f, \pi_\alpha f \rangle = [Q^{-1}P\pi_\alpha f, \pi_\beta f] = \langle P\pi_\beta f, \pi_\alpha f \rangle = \langle \pi_\alpha P\pi_\beta f, \pi_\beta f \rangle = \langle Z\pi_\beta f, \pi_\beta f \rangle.
\]
So it follows that \( \sigma(H_G) = \sigma(Z) \). Finally, the Bezoutian \( \Gamma \) is just a block matrix representation of \( Z \), so \( \sigma(\Gamma) = \sigma(Z) \) and the theorem is proved.

This connection between the signatures of Hankel matrices and Bezoutians leads to some easy derivation of classical algebraic stability criteria. For an account of this we refer to Fuhrmann [28].

5. THE CAUCHY INDEX: SCALAR CASE

Let \( g \) be a rational transfer function with real coefficients. The Cauchy index of \( g(z) \), denoted by \( I_g \), is defined as the number of jumps of \( f \) from \(-\infty\) to \(+\infty\) minus the number of jumps from \(+\infty\) to \(-\infty\) as \( z \) goes from \(-\infty\) to \(+\infty\) through real values.

In this and the next section we will establish some connections between the Cauchy index, the signature of the Hankel map induced by \( g \), and the existence of signature symmetric realizations of \( g \). The central result is the classical result of Hermite [39] and Hurwitz [41]. For a modern geometric proof one can consult Brockett [6], and for a geometric approach to the multivariable case Byrnes [14].

**Theorem 5.1.** Let \( g = p/q \) be a strictly proper rational function with \( p \) and \( q \) coprime. Then

\[
I_g = \sigma(H_g),
\]

where \( \sigma(H_g) \) denotes the signature of the Hankel map induced by \( g \).

**Proof.** Let us analyze first the case where \( q \) is a polynomial with simple real zeros, i.e. \( q(z) = \prod_{i=1}^{n}(z - a_i) \) and \( a_i \neq a_j \) for \( i \neq j \). Let \( d_i(z) = q(z)/(z - a_i) \). Given any polynomial \( u \in X_q \), it has a unique expansion \( u = \sum_{i=1}^{n} u_i d_i \). Then

\[
[H_g u, u] = [\pi_q gu, u] = [\pi_q q^{-1}pu, u] = [q^{-1}q\pi_q q^{-1}pu, u]
\]

\[
= \langle \pi_q pu, u \rangle = \langle p(S_q) u, u \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} u_i u_j \langle p(S_q) d_i, d_j \rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} p(a_i) d_i(a_j) u_i^2,
\]
as \( d_i \) are eigenfunctions of \( S_q \) corresponding to the eigenvalues \( a_i \), and as
\[
\langle d_i, d_j \rangle = d_i(a_i)\delta_{ij}
\]
by (3.12).

From this computation it follows, since \([p(S_q)]_{\mu_0}^\nu = B(q, p)\), that
\[
\sigma(H_\epsilon) = \sigma(B(q, p)) = \sum_{i=1}^{n} \text{sign}(a_i).
\]
(5.2)

On the other hand we have the partial fraction decomposition
\[
g(z) = \frac{p(z)}{q(z)} = \sum_{i=1}^{n} \frac{c_i}{z - a_i}
\]
or
\[
p(z) = \sum_{i=1}^{n} \frac{c_iq(z)}{z - a_i} = \sum_{i=1}^{n} c_id_i(z),
\]
which implies \( p(a_i) = c_id_i(a_i) \), or equivalently that \( c_i = p(a_i)/d_i(a_i) \).

Now obviously
\[
I_a = \sum_{i=1}^{n} \text{sign}(c_i) = \sum_{j=1}^{n} \text{sign} \left( \frac{p(a_i)}{d_i(a_i)} \right),
\]
and as
\[
\text{sign}(p(a_i)d_i(a_i)) = \text{sign} \left( \frac{p(a_i)}{d_i(a_i)} \right),
\]
the equality (5.1) is proved in this case. We pass now to the general case. Let \( q = q_1 \cdots q_s \) be the unique factorization of \( q \) into powers of relatively prime irreducible monic polynomials. As before, we define polynomials \( d_i \) by
\[
d_i(z) = \frac{q(z)}{q_i(z)}.
\]
(5.3)

Since we have the direct sum decomposition
\[
X_q = \sum_{i=1}^{s} d_iX_{q_i},
\]
(5.4)
it follows that each \( f \in X_q \) has a unique representation of the form \( f = \sum_{i=1}^{s} d_i u_i \) with \( u_i \in X_q \). Relative to the indefinite metric of \( X_q \) we have the orthogonality relation

\[
\langle d_i X_{q_i}, d_j X_{q_j} \rangle = 0 \quad \text{for} \quad i \neq j. \tag{5.5}
\]

Indeed, if \( f_i \in X_{q_i} \) and \( g_j \in X_{q_j} \), then

\[
\langle d_i f_i, d_j f_j \rangle = [q^{-1}d_i f_i, d_j f_j] = [d_j q^{-1}d_i f_i, f_j] = 0,
\]
as \( d_i d_j \) is divisible by \( q \) and, by (2.6), \( F[z] \perp F[z] \).

Let \( g = \sum_{i=1}^{s} p_i / q_i \) be the partial fraction decomposition of \( g \). Since the zeros of the \( q_i \) are different, it is clear that

\[
l_g = I_{p/q} = \sum_{i=1}^{s} I_{p_i/q_i}.
\]

Also, as a consequence of the theorem in Fuhrmann [23], it is clear that for the McMillan degree \( \delta(g) \) of \( g \) we have

\[
\delta(g) = \sum_{i=1}^{s} \delta(p_i/q_i),
\]

and hence by a result of Bitmead and Anderson [3] the signatures of the Hankel forms are additive, namely

\[
\sigma(H_g) = \sum_{i=1}^{s} \sigma(H_{p_i/q_i}).
\]

Therefore to prove the Hermite-Hurwitz theorem it suffices to prove it in the case of \( q \) being the power a monic prime. Since we discuss the real case, the primes are of the form \( z - a \) or \( (z - a)^2 + b^2 \), with \( a, b \in \mathbb{R} \).

With the intention of reducing the problem to the cases where \( a = 0 \), \( b = 1 \) we prove the following theorem, due to Brockett and Krishnaprasad [10], which is of independent interest.

**Theorem 5.2.** Let \( g \) be a real rational function. Then the following scaling operations leave the rank and signature of the Hankel map as well as
the Cauchy index invariant:

(i) \( g(z) \to mg(z), m > 0, \)
(ii) \( g(z) \to g(z - a), a \in \mathbb{R}, \)
(iii) \( g(z) \to g(rz), r > 0. \)

Proof. (i) is obvious. To prove the rank invariance let \( g = p/q \) with \( p \) and \( q \) coprime. By the Euclidean algorithm there exist polynomials \( a \) and \( b \) such that \( ap + bq = 1 \). This implies

\[
a(z - a)p(z - a) + b(z - a)q(z - a) = 1
\]
as well as

\[
a(rz)p(rz) + b(rz)q(rz) = 1,
\]
i.e., \( p(z - a), q(z - a) \) are coprime and so are \( p(rz), q(rz) \). Now \( g(z - a) = \frac{p(z - a)}{q(z - a)} \) and \( g(rz) = \frac{p(rz)}{q(rz)} \), which proves the invariance of the McMillan degree, which is the same as the rank of the Hankel map.

Now it is easy to check that, given any polynomial \( u \), we have

\[
[H_g u, u] = [H_{g_a} u_a, u_a],
\]
where \( g_a(z) = g(z - a) \). If we define a map \( R_a : \mathbb{R}[z] \to \mathbb{R}[z] \) by

\[
(R_a u)(z) = u(z - a) = u_a(z),
\]
then \( R_a \) is invertible, \( R_a^{-1} = R_{-a} \), and

\[
[H_g u, u] = [H_{g_a} u_a, u_a] = [H_{g_a} R_a u, R_a u] = [R_a^* H_g R_a u, u],
\]
which shows that

\[
H_g = R_a^* H_{g_a} R_a
\]
and hence that

\[
\sigma(H_g) = \sigma(H_{g_a}),
\]
which proves (ii).

To prove (iii) define, for \( r > 0 \), a map \( P_r : \mathbb{R}[z] \to \mathbb{R}[z] \) by

\[
(P_r u) = u(rz).
\]
Clearly \( P_r \) is invertible and \( P_r^{-1} = P_{1/r} \). Letting \( u_r = P_r u \), we have

\[
[H_g u, u] = \left[ \pi \, g(rz) u, u \right] = \left[ \pi - \sum_{r} \frac{g_k}{r^{k+1}} u, u \right]
\]

\[
= \sum g_{i+j} r^{-i-j} u_i u_j = \sum g_{i+j} \left( u_i r^{-i} \right) \left( u_j r^{-j} \right)
\]

\[
= [H_g P_r u, P_r u] = [P_r^* H_g P_r u, u],
\]

and hence \( H_g = P_r^* H_g P_r \), which implies \( \sigma(H_g) = \sigma(H_g) \). The invariance of the Cauchy index under these scaling operations is obvious.

By applying the previous scaling result, the proof of the Hermite-Hurwitz theorem reduces to the two cases \( q(z) = z^m \) or \( q(z) = (z^2 + 1)^m \).

To begin let \( q(z) = z^m \) and \( g = p/q \). Assume \( p(z) = p_0 + p_1 z + \cdots + p_{m-1} z^{m-1} \); then the coprimeness condition is equivalent to \( p_0 \neq 0 \). Therefore we have

\[
g(z) = p_{m-1} z^{-1} + \cdots + p_0 z^{m},
\]

which shows that

\[
I_g = I_{p/\pi^m} = \begin{cases} 0 & \text{if } m \text{ is even,} \\ \text{sign } p_0 & \text{if } m \text{ is odd.} \end{cases}
\]

On the other hand, \( \text{Ker } H_g = z^{m+1} R[z] \), and so \( \sigma(H_g) = \sigma(H_g; X, m) \). Relative to the standard basis the truncated Hankel map has the matrix representation

\[
\begin{pmatrix}
\vdots & \vdots & \vdots & \vdots & \vdots \\
p_{m-1} & \cdots & p_1 & p_0 \\
p_1 & \cdots & p_0 & 0 \\
p_0 & \cdots & 0 & p_0 \\
p_0 & \cdots & 0 & p_0 \\
p_0 & \cdots & 0 & p_0 \\
0 & \cdots & 0 & p_0 \\
0 & \cdots & 0 & p_0 \\
0 & \cdots & 0 & p_0 \\
\end{pmatrix}
\]

Now clearly the previous matrix has the same signature as the matrix

\[
\begin{pmatrix}
0 & \cdots & 0 & 0 \\
0 & \cdots & 0 & p_0 \\
0 & \cdots & 0 & p_0 \\
0 & \cdots & 0 & p_0 \\
0 & \cdots & 0 & p_0 \\
0 & \cdots & 0 & p_0 \\
\end{pmatrix}
\]
and hence

\[ \sigma(H_g) = \begin{cases} \text{sign } p_0 & \text{if } m \text{ is odd,} \\ 0 & \text{if } m \text{ is even.} \end{cases} \]

Next let \( q(z) = (z^2 + 1)^m \). Since \( q \) has no real zeros, it follows that in this case \( l_g = l_{p/q} = 0 \), and it suffices to prove that also \( \sigma(H_g) = 0 \).

Let \( g(z) = p(z)/(z^2 + 1)^m \) with \( \deg p < 2m \). Let us expand \( p \) in the form

\[ p(z) = \sum_{k=0}^{m-1} (p_k + q_k z)(z^2 + 1)^k \]

with the \( p_k \) and \( q_k \) uniquely determined. The coprimeness condition is equivalent to \( p_0 \) and \( q_0 \) not being zero together. The transfer function \( g \) has therefore the following representation:

\[ g(z) = \sum_{k=0}^{m-1} \frac{p_k + q_k z}{(z^2 + 1)^{m-k}}. \]

In much the same way every polynomial \( u \) of degree less than \( 2m \) can be written in a unique way as

\[ u(z) = \sum_{i=0}^{m-1} (u_i + v_i z)(z^2 + 1)^i. \]

Now

\[
\left[ H_gu, u \right] = \left[ \pi - \sum_{k=0}^{m-1} \left( \frac{p_k + q_k z}{(z^2 + 1)^{m-k}} \right) \sum_{i=0}^{m-1} (u_i + v_i z)(z^2 + 1)^i, \right.

\times \sum_{j=0}^{m-1} (u_j + v_j z)(z^2 + 1)^j]

- \sum_{k=0}^{m-1} \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} \left[ (p_k + q_k z)(u_i + v_i z)(u_j + v_j z), \right.

\times \frac{1}{(z^2 + 1)^{m-k-i-j}} \right].
\]
Now clearly if \( m - k - i - j = 1, 2 \), the contribution of the corresponding terms is zero. So either \( k = m - i - j - 1 \) or \( k = m - i - j - 2 \). In the first case we get

\[
\left[ (p_k + q_k z)(u_i + v_i z)(u_j + v_j z), (z^2 + 1)^{-1} \right] = q_k (u_i u_j - v_i v_j) + p_k (u_i v_j + v_i u_j),
\]

whereas in the second case we get \( q_k v_i v_j \). Thus altogether we have

\[
[H_g u, u] = \sum \sum q_{m-i-j-1} (u_i u_j - v_i v_j) + p_{m-i-j-1} (u_i v_j + u_j v_i)
\]

\[+ \sum \sum q_{m-i-j-2} v_i v_j.\]

The matrix \( M \) of this quadratic form is given, in term of the basis \( (1, z, (z^2 + 1), z(z^2 + 1), \ldots) \), by

\[
\begin{pmatrix}
q_{m-1} & p_{m-1} & q_1 & p_1 & q_0 & p_0 \\
p_{m-1} & q_{m-2} - q_{m-1} & p_1 & q_0 - q_1 & p_0 & -q_0 \\
q_0 & p_0 & q_0 & p_0 & -q_0 \\
p_0 & -q_0 & p_0 & q_0 & -q_0
\end{pmatrix}
\]

Now by our assumption on coprimeness, the matrix

\[
\begin{pmatrix}
q_0 & p_0 \\
p_0 & -q_0
\end{pmatrix}
\]

is nonsingular and has signature zero, which implies that also the signature of \( M \) is zero.

With this the proof of the Hermite-Hurwitz theorem is complete. \( \Box \)

We end this section by giving a much simplified proof of a theorem of Frobenius [20] (see also Gantmacher [31]) on the computation of the signature of a Hankel matrix induced by a rational function \( g \), based on the continued fraction representation of \( g \) or alternatively on the Euclidean algorithm. The motivation for this is the remarks following Theorem 4.1 in Kalman [45].
To this end we need a minor generalization of the Sylvester inertia theorem.

**Theorem 5.3.** Let $A$ be an $m \times m$ real symmetric matrix, $X$ an $m \times n$ matrix of full row rank, and $B$ the $n \times n$ symmetric matrix defined by

$$B = \tilde{X}AX. \tag{5.7}$$

Then the rank and signature of $A$ and $B$ coincide.

**Proof.** The statement concerning ranks is trivial. Next we note that the direct sum decomposition is

$$R^n = \text{Ker} X \oplus \text{Range} \tilde{X}. \tag{5.8}$$

From (5.7) we have the two inclusions

$$\text{Ker} B \supset \text{Ker} X \tag{5.9}$$

and

$$\text{Range} B \subset \text{Range} \tilde{X}. \tag{5.10}$$

Let us choose a basis for $R^n$ compatible with the direct sum decomposition (5.8). In that basis we have $X = (0 \ X_1)$ with $X_1$ invertible and

$$R = \begin{pmatrix} 0 & 0 \\ 0 & B_1 \end{pmatrix} = \begin{pmatrix} 0 \\ \tilde{X}_1 \end{pmatrix} A \begin{pmatrix} 0 & X_1 \end{pmatrix}$$

or $B_1 = \tilde{X}_1 A X_1$. Thus $\sigma(B) = \sigma(B_1) = \sigma(A)$, which proves the theorem. \hfill \Box

The previous theorem is applied to prove the following easy lemma on Bezoutians.

**Lemma 5.4.** Let $q$ and $r$ be two coprime polynomials, and $p$ an arbitrary nonzero polynomial. Then

$$\sigma(B(qp, rp)) = \sigma(B(q, r)).$$

**Proof.** The Bezoutian $B(qp, rp)$ is determined by the polynomial expansion of

$$p(z)q(z)r(w) - q(w)r(z) \overline{z - w} p(w),$$
which implies \( R(qp, rp) = \tilde{X}B(q, r)X \) with

\[
X = \begin{pmatrix}
p_0 & \cdots & p_m \\
\vdots & \ddots & \vdots \\
p_0 & \cdots & p_m
\end{pmatrix}.
\]

The result follows now by an application of Theorem 5.3.

Following Kalman's notation, we write the Euclidean algorithm in the form

\[
q = \xi_1 \frac{p}{v_1} - q_2, \quad \deg q_2 < \deg p,
\]

\[
p = \xi_2 \frac{q_2}{v_2} - q_3, \quad \vdots
\]

\[
\xi_{r-1} = \frac{q_{r-1}}{v_r}, \quad \deg q_r = 0.
\]

Here \( \xi_i \) are monic polynomials with \( v_i \neq 0 \) normalizing constants.

The above corresponds to a finite continued fraction representation of \( g \), namely

\[
g = \frac{p}{q} = \frac{v_1}{\xi_1 - \frac{v_1 v_2}{\xi_2 - \frac{v_2 \cdots}{\xi_r - \frac{v_r}{\xi_r}}}}.
\]

**Theorem 5.5.** Let \( g = p/q \) have the representation (5.12). Then

\[
\sigma(H_g) = \sum_{i=1}^{r} (\text{sign} v_i) \frac{1 + (-1)^{\deg \xi_i - 1}}{2}.
\]

**Proof.** By Theorem 4.7 we have \( \sigma(H_g) = \sigma(B(q, p)) \), where \( B(q, p) \) is the Bezoutian of the polynomials \( q \) and \( p \). Using the first of the equations
(5.11), we have

\[ B(q, p) = B(\tilde{\xi}_1 \frac{p}{\nu_1}, p) = p(\tilde{\xi}_1, 1) + B(p, q_2), \]

where we have used the alternating property of the Bezoutian. Now, since \( q \) and \( p \) are coprime,

\[ \text{rank}(B(q, p)) = \deg q = \deg \tilde{\xi}_1 + \deg p \]

\[ = \text{rank} \left( B\left( \tilde{\xi}_1 \frac{p}{\nu_1}, p \right) \right) + \text{rank}(B(p, q_2)). \]

The additivity of rank implies (see Bitmead and Anderson [3]) the additivity of signature, and hence

\[ \sigma(B(q, p)) = \sigma\left( \frac{1}{\nu_1} B(\tilde{\xi}_1, 1) \right) + \sigma(B(p, q_2)), \]

and we proceed by induction to obtain

\[ (B(q, p)) = \sum_{i=1}^{r} \sigma\left( \frac{1}{\nu_i} R(\tilde{\xi}_i, 1) \right). \quad (5.14) \]

Now given \( \tilde{\xi}(z) = z^k + c_{k-1}z^{k-1} + \cdots + c_0 \), we have

\[ B(\tilde{\xi}, 1) = \frac{\tilde{\xi}(z) - \tilde{\xi}(w)}{z - w} = \sum_{i=1}^{k} c_i \frac{z^i - w^i}{z - w} \]

with \( c_k = 1 \). Using the equality

\[ \frac{\tilde{\xi}^i - w^i}{\tilde{\xi} - w} = \sum_{j=0}^{i-1} \tilde{\xi}w^{i-1-j}, \]
we have

\[
B(q, p) = \begin{pmatrix}
c_1 & c_2 & \cdots & c_{k-1} & 1 \\
c_2 & & & 1 \\
\vdots & & & \vdots \\
c_{k-1} & & & 0 \\
1 & & & & 
\end{pmatrix},
\]

and hence

\[
\sigma\left(\frac{1}{\nu}B(\zeta, 1)\right) = \begin{cases}
\text{sign } \nu & \text{if } \deg \zeta \text{ is odd} \\
0 & \text{if } \deg \zeta \text{ is even}.
\end{cases}
\]

The equality (5.14) can now be rewritten as (5.13), and the proof is complete.

6. SIGNATURE SYMMETRIC REALIZATIONS

Much of linear system theory, and nowhere more clearly than in realization theory, is concerned with the interplay between external, or input-output, properties of the system and the internal properties of corresponding realizations.

Our concern in this section is the interplay between external symmetries of a given transfer function and the possibility of realizing that transfer function by a state-space system with corresponding internal symmetries.

The simplest type of external symmetry a transfer function may possess is

\[
\tilde{G}(z) = G(z).
\]

A matrix \( J \), over the real field, will be called a signature matrix if

\[
\tilde{J} = J = J^{-1}.
\]

This means that in some basis \( J \) has a matrix representation of the form

\[
J = \begin{pmatrix}
I_p & 0 \\
0 & -I_q
\end{pmatrix}.
\]
This is equivalent to \( J \) having a matrix representation of the form

\[
\Sigma = \text{diag}(\varepsilon_1 \Sigma_1, \ldots, \varepsilon_k \Sigma_k)
\]

with \( \Sigma_i \) the \( n_i \times n_i \) matrix of the form

\[
\begin{pmatrix}
0 & \cdots & \cdots & 0 & 1 \\
\vdots & \ddots & \ddots & \vdots & \vdots \\
\vdots & \ddots & \ddots & \vdots & \vdots \\
0 & \cdots & \cdots & 1 & 0 \\
1 & 0 & \cdots & \cdots & 0
\end{pmatrix}
\]

and \( \varepsilon_i = \pm 1 \). In this case

\[
\text{rank} \, \Sigma = \sum_{i=1}^{k} n_i = p + q \quad \text{and} \quad \sigma(\Sigma) = \sum_{i=1}^{k} \varepsilon_i \left( \frac{1 + (-1)^{n_i}}{2} \right) = p - q.
\]

A realization \((A, B, C)\) is called signature symmetric if for some signature matrix \( J \) the following diagram is commutative:

\[
\begin{array}{ccc}
R^m & \xrightarrow{J} & \tilde{R}^m \\
\downarrow{B} & & \downarrow{\tilde{B}} \\
X & \xrightarrow{J} & X \\
\downarrow{A} & & \downarrow{\tilde{A}} \\
X & \xrightarrow{J} & X \\
\end{array}
\]

\tag{6.4}

The next theorem, due to Youla and Tissi [62] and Brockett and Skoog [12], establishes a relation between the two concepts. Due to the importance of this result as a model for other theorems of this type, we will give two independent proofs, one stressing the state space aspects and the other the transfer function or frequency domain aspects of the problem. While the state space approach in this instance cannot be surpassed in elegance, the transfer function approach has the advantage of providing explicit formulas for the intertwining maps. Moreover, the same calculations, at least in the scalar case, that were used to prove the Hermite-Hurwitz theorem lead also to the construction of a canonical signature symmetric realization.

**Theorem 6.1.** Let \( G \) be an \( m \times m \) real transfer function. Then \( \tilde{G}(z) = G(z) \) if and only if \( G \) admits a signature symmetric realization.
Proof. Assume \((A, B, C)\) is a signature symmetric realization, with signature matrix \(J\). Then

\[ JA = \tilde{A} J \quad \text{and} \quad JB = \tilde{C}. \]  

(6.5)

Hence \(G(z) = C(zI - A)^{-1}B\) implies

\[ \tilde{G}(z) = \tilde{B}(zI - \tilde{A})^{-1}\tilde{C} = \tilde{B}(zI - \tilde{A})^{-1}JB \]

\[ = \tilde{B}J(zI - A)^{-1}B = C(zI - A)^{-1}B = G(z). \]

To prove the converse, let \(G(z) = G(z)\), and let \((A, B, C)\) be any minimal realization of \(G\). Hence as \(G(z) = G(z)\), \((A, B, C)\) is also a minimal realization of \(G\). By the state space isomorphism theorem there exists a nonsingular map \(Z\) such that the diagram

\[ \begin{array}{ccc}
X & \xrightarrow{A} & X \\
\downarrow Z & & \downarrow A \\
X & \xrightarrow{C} & R^m \\
\downarrow \tilde{B} & & \downarrow \tilde{B} \\
\end{array} \]

is commutative. By duality we have also the commutativity of the diagram

\[ \begin{array}{ccc}
X & \xrightarrow{A} & X \\
\downarrow Z^* & & \downarrow A \\
X & \xrightarrow{C} & R^m \\
\downarrow \tilde{B} & & \downarrow \tilde{B} \\
\end{array} \]

By the uniqueness of the map intertwining two canonical realizations of the same transfer functions it follows that necessarily \(Z^* = Z\).

Now any real symmetric matrix \(Z\) can be factored as \(Z = QJ\tilde{Q}\) with \(J\) a signature matrix. For such a factorization it is easily checked that \((QAZ^{-1}, QB, CQ^{-1})\) is a signature symmetric realization.

Next we give the polynomial proof of this result. Thus let \(\tilde{G} = G\), and let \(G = Q^{-1}P\) be a left coprime matrix fraction representation. By the symmetry...
of $G$ we have

$$Q^{-1}P = \tilde{P}\tilde{Q}^{-1} \quad (6.6)$$

and hence

$$P(z)\tilde{Q}(z) = Q(z)\tilde{P}(z). \quad (6.7)$$

This relation is of course connected to module homomorphisms via Theorem 2.1. In fact if $Z : X_{\tilde{Q}} \to X_{\tilde{Q}}$ is defined by

$$Zf = \pi_Q Pf \quad \text{for} \quad f \in X_{\tilde{Q}}, \quad (6.8)$$

then $Z = Z^*$ and $ZS_{\tilde{Q}} = S_{\tilde{Q}}Z$. Applying the realization construction and Theorem 2.2, we see that the realizations associated with the matrix fractions $Q^{-1}P$ and $\tilde{P}\tilde{Q}^{-1}$ are connected by duality. Choosing any basis $B$ in $X_{\tilde{Q}}$ and its dual basis $B^*$ in $X_{\tilde{Q}}$, the matrix representation $[Z]_{B^*}$ is symmetric, and we conclude the proof as before.

We proceed now to get a polynomial representation of $g$ which exhibits more clearly the corresponding signature symmetric realization.

Assume as before that $g = p/q$ and that $q$ is a monic polynomial. Let $q = q_1 \cdots q_s$ be a factorization of $q$ into monic factors which are powers of irreducible mutually coprime polynomials. Since the field is $R$, then either

$$q_i(z) = (z - a_i)^{m_i} \quad \text{with} \quad a_i \text{ real}$$

or

$$q_i(z) = [(z - a_i)^2 + b_i^2]^{m_i}.$$ 

Assume also that

$$g = \frac{p}{q} = \sum_{i=1}^{s} \frac{p_i}{q_i}, \quad (6.9)$$

with $\deg p_i < \deg q_i$, is the partial fraction decomposition of $g$.

**Lemma 6.2.** Let $g = p/q$ with $\deg p < \deg q$ and $p(z) = p_0 + \cdots + p - 1^{m-1}$. 

(a) If \( q(z) = z^m \), then there exists a polynomial \( r \), with \( \deg r < \deg q \), such that with \( \varepsilon = \text{sign} p_0 \),
\[
p = \varepsilon r^2 \mod q. \tag{6.10}
\]

(b) If \( q(z) = (z^2 + 1)^m \), there exists a polynomial \( r \), with \( \deg r < \deg q \), such that
\[
p = r^2 \mod q. \tag{6.11}
\]

Remark. It follows from (a) that
\[
g = (r\varepsilon)q^{-1}r + \pi \tag{6.12}
\]
in the first case, and
\[
g = rq^{-1}r + \pi \tag{6.13}
\]
in the second case. Here \( \pi \) denotes a polynomial term. These representations of \( g \) in the Rosenbrock [53] style are the key to the construction of signature symmetric realizations.

Proof. (a): By the coprimeness of \( p \) and \( z^m \) we have \( p_0 \neq 0 \). Without loss of generality we assume \( p_0 > 0 \); otherwise we consider \( \varepsilon p \) with \( \varepsilon = \text{sign} p_0 \). Let \( r(z) = r_0 + \cdots + r_{m-1}z^{m-1} \), then
\[
r(z)^2 = \sum_k \left( \sum_{i+j=k} r_ir_j \right)z^k.
\]
For \( p = r^2 \mod z^m \) to hold we must have
\[
\sum_{i+j=k} r_ir_j = p_k, \quad k = 0, \ldots, m-1. \tag{6.14}
\]
For \( k = 0 \) this is the nonlinear equation \( r_0^2 = p_0 \), which is solvable by our assumption \( p_0 > 0 \). The other equations can be solved recursively, as the \( k \)th equation, given \( r_0, \ldots, r_{k-1} \), is linear in \( r_k \).

(b): In this case we expand \( p \) in powers of \( z^2 + 1 \) as
\[
p(z) = (p_0 + q_0z) + (p_1 + q_1z)(z^2 + 1)
+ \cdots + (p_{m-1} + q_{m-1}z)(z^2 + 1)^{m-1}. \tag{6.15}
\]
Now

\[ r^2 \mod (z^2 + 1)^m = \sum_{k=0}^{m} \left\{ \sum_{i+j=k} (r_i + s_i z)(r_j + s_j z) \right\} (z^2 + 1)^k \mod (z^2 + 1)^m \]

\[ = \sum_{k=0}^{m} \left\{ \sum_{i+j=k} \left[ (r_i r_j - s_i s_j) + (r_i s_j + r_j s_i) z \right] \right\} (z^2 + 1)^k \]

\[ + \sum_{k=0}^{m} \left\{ \sum_{i+j=k} s_i s_j \right\} (z^2 + 1)^{k+1} \]

\[ = \sum_{k=0}^{m} \left\{ \sum_{i+j=k} \left[ (r_i r_j - s_i s_j) + (r_i s_j + r_j s_i) z \right] \right\} (z^2 + 1)^k \]

\[ + \sum_{k=1}^{m} \left[ \sum_{i+j=k-1} s_i s_j \right] (z^2 + 1)^k. \]

Hence the equations for \( r_i, s_i \) become

\[ r_0^2 + s_0^2 + 2r_0 s_0 z = p_0 + q_0 z \] \hspace{1cm} (6.16)

for \( k = 0 \), and

\[ \sum_{i+j=k} \left[ (r_i r_j - s_i s_j) + (r_i s_j + r_j s_i) z \right] + \sum_{i+j=k-1} s_i s_j = p_k + q_k z \] \hspace{1cm} (6.17)

for \( k > 0 \).

Equating coefficients in (6.16) yields the nonlinear system

\[ r_0^2 - s_0^2 = p_0 , \]

\[ 2r_0 s_0 = q_0 . \] \hspace{1cm} (6.18)

This system, as \( p_0^2 + q_0^2 > 0 \), always has (two) nontrivial solutions, as is clear from the geometric interpretation, representing the intersection of two hyperbolas.

Equating coefficients in (6.17) yields a system which is linear in \( r_k \) and \( s_k \) with the nonsingular coefficient matrix

\[ \begin{pmatrix} r_0 & -s_0 \\ s_0 & r_0 \end{pmatrix} . \] \hspace{1cm} (6.19)
Hence these equations can also be solved recursively. Note that the nonsingularity of this matrix is equivalent to the coprimeness of \( r \) and \( (z^2 + 1)^m \).

Of course, by scaling, the previous lemma holds for any linear and quadratic factor. We state this without the obvious proof.

**Lemma 6.3.** Let \( g = p/q \) be rational with \( \deg p < \deg q \), \( p \) and \( q \) coprime.

(a) If \( q(z) = (z - a)^m \), then there exists a polynomial \( r \), with \( \deg r < \deg q \), such that, with \( \varepsilon = \text{sign} \, p(a) \)

\[
p = \varepsilon r^2 \mod q.
\] (6.20)

(b) If \( q(z) = ((z - a)^2 + b^2)^m \), \( b \neq 0 \), then there exists a polynomial \( r \) such that

\[
p = r^2 \mod q.
\] (6.21)

Let now \( q = q_1 \cdots q_s \) be a factorization of \( q \) into powers of coprime irreducible monic polynomials. We define \( d_i \) by

\[
d_i = q/q_i.
\] (6.22)

From the partial fraction decomposition of \( g \), (6.9), it follows that

\[
p = \sum_{i=1}^{s} p_i \left( \frac{q}{q_i} \right) = \sum_{i=1}^{s} p_i d_i.
\] (6.23)

Now we saw that the equations

\[
p_i = \varepsilon_i r_i^2 \mod q_i
\] (6.24)

are solvable, where \( \varepsilon_i = -1 \) if \( q_i(z) = (z - a_i)^m \), and \( p_i(a_i) = -1 \), and \( \varepsilon_i = 1 \) otherwise. Note that an equivalent way of writing (6.24) is

\[
\frac{p_i}{q_i} = r_i \left( \varepsilon_i q_i^{-1} \right) r_i + \pi_i
\] (6.25)

with \( \pi_i \) a polynomial term. We can state now the following.
THEOREM 6.4. With the previous notation let

\[ e(z) = \sum_{i=1}^{s} \epsilon_i d_i(z). \]  

(6.26)

Then there exists a polynomial \( r \), coprime with \( q \), such that

\[ p = er^2 \mod q \]  

(6.27)
or equivalently

\[ p/q = r(eq^{-1})r + \pi \]

\[ = \sum_{i=1}^{s} r_i(e_iq_i^{-1})r_i + \pi. \]  

(6.28)

REMARK. The last form is more in line with the multivariable version of this theorem.

Proof. By the Chinese remainder theorem there exists a unique polynomial \( r \) such that \( \deg r < \deg q \) and

\[ r_i = r \mod q_i. \]  

(6.29)

This clearly implies that

\[ r_i^2 = r^2 \mod q_i. \]  

(6.30)

Hence

\[ er^2 \mod q_j = (\sum \epsilon_k d_k)r^2 \mod q_j \]

\[ = \sum \epsilon_k d_k r^2 \mod q_j \]

\[ = \sum \epsilon_k d_k r_i^2 \mod q_j \]

\[ - \sum c_j d_j r_i^2 \mod q_j, \]

since, for \( k \neq j \), \( q_j \) divides \( d_k \). But \( \epsilon_i r_i^2 = p_i \mod q_i \) and so

\[ er^2 \mod q_j = pjd_j \mod q_j \]

\[ = \sum p_k d_k \mod q_j = p \mod q_j, \]
i.e., \( p - er^2 = 0 \mod q \). Since the \( q_i \) are mutually coprime, (6.27) follows and so also (6.28).

Using the realization procedure outlined in Section 2, based on the Rosenbrock type representations of rational functions, we have two realizations of \( g \) based on the two representations of \( g \):

\[
g = (re)q^{-1}r + \pi = rq^{-1}(er) = \pi. \tag{6.31}
\]

Both realizations have \( X_q \) as state space and are given by \((A, B, C)\) and \((A_1, B_1, C_1)\) respectively, where

\[
A = Sq,
\]

\[
B_1 = r, \tag{6.32}
\]

\[
C_f = \left( rq^{-1}f \right)_{-1} \quad \text{for } f \in X_q,
\]

and

\[
A_1 = S_q,
\]

\[
B_1 = \pi q er = er \mod q, \tag{6.33}
\]

\[
C_1 f = \left( rq^{-1}f \right)_{-1} \quad \text{for } f \in X_q.
\]

By Theorem 2.12 of Fuhrmann [27] we have, corresponding to the bilinear form in \( X_q \) defined by

\[
\langle f, g \rangle = \left[ q^{-1}f, g \right], \tag{6.34}
\]

that

\[
A^* = A = A_1 = A_1^*, \quad C_1 = B^*, \quad B_1 = C^*.
\]

Letting

\[
Z = e(S_q), \tag{6.35}
\]

we clearly have

\[
ZA = A^*Z = AZ \tag{6.36}
\]

and

\[
B^*Z = C \quad \text{and} \quad C^* = ZB. \tag{6.37}
\]

as well as

\[
Z^* = Z. \tag{6.38}
\]
We will see that for a suitable choice of a pair of dual bases in \( X_q \) the corresponding matrix representation of \( Z \), i.e. \( [Z]_{B^*} \), is a signature matrix.

**Theorem 6.5.** Let \( B \) be the basis of \( X_q \) defined by (3.29), and \( B^* \) its dual basis given by (3.30) with \( Z \) defined by (6.35). Then \( [Z]_{B^*} \) has a matrix representation of the form

\[
[Z]_{B^*} = \text{diag}(\varepsilon_1 \Sigma_1, \ldots, \varepsilon_s \Sigma_s) \tag{6.39}
\]

with

\[
\Sigma_i = \begin{pmatrix}
0 & \cdots & 0 & 1 \\
. & \ddots & . & . \\
. & . & \ddots & . \\
0 & \cdots & \cdots & \cdots \\
1 & 0
\end{pmatrix} \tag{6.40}
\]

**Proof.** Since \( Z = e(S_q) \), then \( Z \) leaves invariant all \( S_q \)-invariant subspaces. In particular \( Zd_j X_{q_i} \subset d_j X_{q_i} \), and hence the orthogonality of these subspaces is preserved under \( Z \). Therefore it remains to calculate

\[
\langle Zd_k(d_k(S_{q_k})^{-1})f, d_k(d_k(S_{q_k})^{-1})g \rangle_X
\]

\[
= \left[q^{-1} e(S_q)d_k(d_k(S_{q_k})^{-1})f, d_k(d_k(S_{q_k})^{-1})g \right]
\]

\[
= \left[q^{-1}(\sum \varepsilon_j d_j(S_q)) d_k d_k(S_{q_k})^{-1} f, d_k d_k(S_{q_k})^{-1} g \right]
\]

\[
= q_k^{-1} \left(\sum \varepsilon_j d_j(S_q) d_k d_k(S_{q_k})^{-1} f, d_k d_k(S_{q_k})^{-1} g \right]
\]

\[
= \varepsilon_k \left[q_k^{-1} q_k \pi_k q_k^{-1} d_k d_k(S_{q_k})^{-1} f, d_k d_k(S_{q_k})^{-1} g \right]
\]

\[
= \varepsilon_k \langle \pi_{q_k} d_k d_k(S_{q_k})^{-1} f, d_k d_k(S_{q_k})^{-1} g \rangle_{X_{q_k}}
\]

\[
= \varepsilon_k \langle \pi_{q_k}^2 d_k d_k(S_{q_k})^{-1} f, d_k d_k(S_{q_k})^{-1} g \rangle_{X_{q_k}}
\]

\[
= \varepsilon_k \langle \pi_{q_k} d_k d_k(S_{q_k})^{-1} f, \pi_{q_k} d_k d_k(S_{q_k})^{-1} g \rangle_{X_{q_k}}
\]

\[
= \varepsilon_k \langle f, g \rangle_{X_{q_k}}.
\]
In particular we have

\[ \langle Zd_k \left( d_k(\sigma_{q_k})^{-1} \right) v_i^{(k)}, d_k \left( d_k(\sigma_{q_k})^{-1} \right) g \rangle = \epsilon_k v_i^{(k)}, \]

\[ v_j^{(k)} = \delta_{k, i+j-m_k-1}, \quad (6.41) \]

i.e.,

\[ Zd_k \left( d_k(\sigma_{q_k})^{-1} \right) v_i^{(k)} = \epsilon_k d_k v_i^{(k)}. \quad (6.42) \]

Naturally this can be verified directly by observing that

\[ Zd_k \left( d_k(\sigma_{q_k})^{-1} \right) f - \sum_j \epsilon_j d_j(\sigma_q) d_k d_k(\sigma_{q_k})^{-1} f \]

\[ = \sum_k \epsilon_j \pi q_j d_k d_k(\sigma_{q_k})^{-1} f \]

\[ = \epsilon_k d_k q_k \pi q_k^{-1} d_k^{-1} d_k(\sigma_{q_k})^{-1} f \]

\[ = \epsilon_k d_k d_k(\sigma_{q_k})^{-1} f \]

\[ = \epsilon_k d_k d_k(\sigma_{q_k}) d_k(\sigma_{q_k})^{-1} f \]

\[ = \epsilon_k d_k f, \]

i.e.,

\[ Zd_k \left( d_k(\sigma_{q_k})^{-1} \right) f = \epsilon_k d_k f. \quad (6.43) \]

The theorem follows now from (6.43).

With this we have now at our disposal all the machinery for exhibiting an explicit signature symmetric realization of g in matrix terms.

**Theorem 6.6.** Let \( g = p/q \) be a strictly proper real rational function. Then g admits a signature symmetric realization. One such realization is
given, in the notation of this section, by

\[ A = \begin{bmatrix} S_q \end{bmatrix}, \]
\[ B_1 = \begin{bmatrix} r \end{bmatrix}, \quad \text{(6.44)} \]
\[ C = \tilde{B}\Sigma, \]

where

\[ \Sigma = \text{diag}(\varepsilon_1\Sigma_1, \ldots, \varepsilon_s, \Sigma_s). \quad \text{(6.45)} \]

The signature of \( \Sigma \), namely \( \sigma(\Sigma) \), is given by

\[ \sigma(\Sigma) = \sum_{j=1}^{s} \varepsilon_j \left( \frac{1 + (-1)^{m_j-1}}{2} \right). \quad \text{(6.46)} \]

is uniquely determined by \( g \), and is equal to \( \sigma(H_g) \) or equivalently to \( I_g \), the Cauchy index of \( g \).

7. A MULTIVARIABLE CHINESE REMAINDER THEOREM

We saw in Section 6 the application of the scalar Chinese remainder theorem in the piecing together, or interpolating, of local canonical forms to a global one. With the intention of putting it to the same use, we derive now a multivariable version of Theorem 3. A result similar to the existence part of Theorem 7.2 has been proved by Gohberg, Kaashoek, Lerer, and Rodman [32]. For the classic result one can consult Lang [50] and Newman [52].

**Definition 7.1.** Given nonsingular polynomial matrices \( Q \in F^{m \times m}[z] \), \( i = 1, \ldots, s \), we will say that the \( Q_i \) are **mutually left coprime** if for each \( i \), \( Q_i \) is left coprime with the (unique up to a right unimodular factor) least common right multiple of all \( Q_j, j \neq i \). Mutual right coprimeness is analogously defined.

Note that this is a stronger condition than pairwise coprimeness. An alternative statement for mutual left coprimeness is

\[ \det Q = \prod_{i=1}^{s} \det Q_i, \quad \text{(7.1)} \]

where

\[ QF^m[z] = \bigcap_{j=1}^{s} Q_jF^m[z]. \quad \text{(7.2)} \]
Yet another equivalent statement is that for $Q$ defined by (7.2) we have the
direct sum decomposition

$$X^{\tilde{Q}} = X^{\tilde{Q}_1} \oplus \cdots \oplus X^{\tilde{Q}_s}. \quad (7.3)$$

**Theorem 7.2.**

(a) Let $Q_i \in F^{m \times m}[[z]]$ be nonsingular and mutually left coprime, and let
$Q$ be defined by (7.2), i.e., $Q$ is the least common right multiple of all $Q_i$. Then
given $A_i \in F^{m \times n}[[z]]$, $i = 1, \ldots, s$, such that $Q_i^{-1}A_i$ is strictly proper,
there exists a unique $A \in F^{m \times n}[[z]]$ such that

$$A = A_i + Q_iB_i \quad (7.4)$$

and $Q^{-1}A$ is strictly proper.

(b) Let $Q_i \in F^{m \times m}[[z]]$ be nonsingular and mutually right coprime, and
let $Q$ be the least common left multiple of all $Q_i$, $j = 1, \ldots, s$. Then
given $A_i \in F^{n \times m}[[z]]$ such that $A_iQ_i^{-1}$ is strictly proper, there exists a unique
$A \in F^{n \times m}[[z]]$ such that

$$A = A_i + B_iQ_i \quad (7.5)$$

and $AQ^{-1}$ is strictly proper.

**Proof.** (a): Clearly it suffices to prove the theorem in the case $n = 1$. We
will use the duality pairing between $X_Q$ and $X^{\tilde{Q}}$. Since (7.2) implies the
inclusion $QF^m[z] \subset Q_iF^m[z]$, there exist a nonsingular polynomial matrix $D_i$
such that

$$Q = Q_iD_i. \quad (7.6)$$

Moreover $Q_iX_{D_i}$ is a submodule of $X_Q$ or equivalently an $S_Q$-invariant
subspace, and

$$(Q_iX_{D_i})^\perp = \tilde{D}_iX_{\tilde{Q}_i}X^{\tilde{Q}}. \quad (7.7)$$

Hence

$$\left( \sum_{j \neq i} \tilde{D}_jX_{\tilde{Q}_j} \right)^\perp = \bigcap_{j \neq i} Q_jX_{D_j} \quad (7.8)$$

is an invariant subspace of $X_Q$ and so of the form $D'_iX_{Q'_i}$, i.e., we have the factorizations

$$Q = Q_iD_i = D'_iQ'_i, \quad i = 1, \ldots, s. \quad (7.9)$$
The equality
\[ \bigcap_{j \neq i} Q_j X_{D_j} = D'_i X_{Q'_i}, \quad (7.10) \]
together with (7.9), clearly implies the equality
\[ \bigcap_{j \neq i} F^m[z] = D'_i F^m[z], \]
i.e., \( D'_i \) is the l.c.r.m. of all \( Q_j, j \neq i \). In particular there exist \( R_{ji} \) such that
\[ D'_i = Q_j R_{ji}, \quad i \neq j. \quad (7.11) \]
By our assumption of the mutual left coprimeness of the \( Q_j \) it follows that \( Q_i \) and \( D'_i \) are left coprime. From (7.8) one obtains
\[ \sum_{j \neq i} \tilde{D}_j X_{\tilde{Q}_j} = \tilde{Q}_i X_{\tilde{D}_i}, \]
which in turn implies
\[ \sum_{j \neq i} \tilde{D}_j F^m[z] = \tilde{Q}_i F^m[z], \]
i.e., that \( \tilde{Q}_i \) is the g.c.l.d. of all \( \tilde{D}_j, j \neq i \). Since \( X_{\tilde{Q}} = Q_i X_{D_i} + D'_i X_{Q'_i} \) is a direct sum decomposition, so is \( X_{\tilde{Q}} = Q_i X_{\tilde{D}_i} + \tilde{D}_i X_{\tilde{Q}_i} \). Hence \( \tilde{Q}_i' \) and \( \tilde{D}_i \) are left coprime, or equivalently \( D_i \) and \( Q_i' \) are right coprime. Using slightly different terminology, \( \tilde{Q}_i \) and \( D_i \) are skew coprime, and the same holds for \( D'_i \) and \( Q'_i \).

The equality (7.9), together with the coprimeness conditions, implies the invertibility of the maps \( Z_i : X_{Q'_i} \to X_{Q_i} \) defined by
\[ Z_i f = \eta_{Q_i} D'_i f, \quad f \in X_{Q'_i}. \quad (7.12) \]
Assume now \( a \in X_{Q}, \) i.e., \( Q^{-1}a \) is strictly proper. Using the direct sum decomposition
\[ X_{Q} = D'_1 X_{Q'_1} \oplus \cdots \oplus D'_s X_{Q'_s}, \quad (7.13) \]
we can write, in a unique way,
\[ a = \sum_{j=1}^{s} D'_j g_j \quad (7.14) \]
with $g_j X_Q^j$. We compute now

$$\pi_{Q_i} a = \pi_{Q_i} \sum_{j=1}^{s} D_j g_j = \sum_{j=1}^{s} \pi_{Q_i} D_j g_j.$$ 

By (7.11), $\pi_{Q_i} D_j g_j = 0$ for $j \neq i$, and so

$$\pi_{Q_i} a = \pi_{Q_i} D_i g_i = Z_i g_i = a_i,$$

and this last equation is solvable for $g_i$, as the map $Z_i$ is invertible. In summary,

$$a = \sum_{j=1}^{s} D_j (Z_j^{-1} a_j)$$

(7.15)

is a solution to the system of equations $\pi_{Q_i} a = a_i$, $i = 1, \ldots, s$, and $a \in X_Q$.

To prove uniqueness assume $a$ and $a'$ are in $X_Q$ and satisfy $\pi_{Q_i} a = \pi_{Q_i} a'$, or $\pi_{Q_i}(a - a') = 0$, $i = 1, \ldots, s$. This implies $a - a' \in Q_i F^m[z]$ for $i = 1, \ldots, s$, and so $a - a' \in Q F^m[z]$. Since $Q^{-1}(a - a')$ is strictly proper, this implies $a = a'$ and uniqueness is proved.

Part (b) follows by duality.

The next theorem relates local coprimeness to a global one.

**Theorem 7.3.** Let $Q_1, \ldots, Q_s$ be nonsingular mutually left coprime polynomial matrices, and let $Q$ be the l.c.r.m. of all $Q_j$, $j = 1, \ldots, s$. If $U = U_j + Q_j L_j$, $j = 1, \ldots, s$, then $U$ and $Q$ are left coprime if and only if $U_j$ and $Q_j$ are left coprime for all $j = 1, \ldots, s$. Similarly for right coprimeness.

**Proof.** From our assumption follows the factorization $Q = Q_j D_j$. Assume $U$ and $Q$ are left coprime. Then there exist polynomial matrices $X$ and $Y$ such that

$$UX + QY = I,$$  (7.16)

which in turn implies $(U_j + Q_j L_j)X + Q_j D_j Y = I$, or

$$U_j X + Q_j (L_j X + D_j Y) = I,$$  (7.17)

i.e., that $U_j$ and $Q_j$ are left coprimes for all $j = 1, \ldots, s$. 

Conversely, assume $U_j$ and $Q_j$ are left coprime for all $j = 1, \ldots, s$. Therefore there exist polynomial matrices $X_j$ and $Y_j$ such that

$$U_j X_j + Q_j Y_j = I. \quad (7.18)$$

Using the equality $U = U_j + Q_j L_j$, we have

$$(U - Q_j L_j) X_j + Q_j Y_j = UX_j + Q_j (L_j X_j + Y_j) = I,$$

i.e., $U$ and $Q_j$ are left coprime for all $j = 1, \ldots, s$. We will show that this implies the left coprimeness of $U$ and $Q$. Indeed, if $U$ and $Q$ are not left coprime, then $UF^m(z) + QF^m(z)$ is a proper full submodule of $F^m[z]$ and hence of the form $EF^m[z]$ for some nonsingular polynomial matrix $E$. Moreover, since $QF^m(z) \subseteq EF^m[z]$, we have the inclusion $X^E \subset X^0 = X^0 \oplus \cdots \oplus X^0$. Let $h \in X^E$, and write $h = h_1 + \cdots + h_s$ with $h_i \in X^0$. Assume for some index $i$, $h_i \neq 0$. Then $h_i \perp UF^m(z) + QF^m(z)$, which contradicts the proven left coprimeness of $U$ and $Q_i$. This completes the proof.

8. ON PARTIAL FRACTION DECOMPOSITIONS

The partial fraction decomposition of a scalar rational function with real coefficients proved to be a central tool in the study of the relation between the Cauchy index and the signature of the Hankel matrix, as well as in the explicit construction of a signature symmetric realization. One expects to be able to follow the same line in the study of the multivariable case. To this end we study matrix partial fraction decompositions in somewhat more detail.

Let us consider an extremely simple situation. Let $g = p/q$ with $p, q$ polynomials and $\deg p < \deg q$. Moreover let $q(z) = \prod_{j=1}^s (z - z_j)$ with the $z_j$ distinct. Clearly in this case $g(z) = \prod_{j=1}^s g_j/(z - z_j)$ is a partial fraction decomposition. Furthermore $X_q$ is the space of all polynomials of degree $n$, and $(q(z)/(z - z_i) : i = 1, \ldots, s)$ is a basis of eigenfunctions of $S_q$. Since $p(S_q)S_q = S_q p(S_q)$, then $p(S_q)$ commutes with $S_q$, i.e., it is an intertwining map.

Now an intertwining map leaves the spectral decomposition invariant. So if $ZS_q = S_q Z$, then $Z$ induces a map $Z_i$ in $V_i = \text{span}(q(z)/(z - z_i))$, and $Z_i$ commutes with $S_q : V_i$. As $Z$ is the direct sum of the $Z_i$, we have rank $Z = \sum \text{rank } Z_i$ as well as $\sigma(Z) = \sum \sigma(Z_i)$.

All of this extends to the multivariable case, and we will study this in somewhat more detail. Since our interest is mainly in signatures, we will focus on symmetric transfer functions.
Let $G(z)$ be an $m \times m$ strictly proper rational matrix function, and assume $G$ is symmetric, i.e. $\tilde{G}(z) = G(z)$. Moreover let $G(z) = Q(z)^{-1}P(z)$ be a left matrix fraction representation of $G$. So by symmetry

$$G = Q^{-1}P = \tilde{P}\tilde{Q}^{-1}, \quad (8.1)$$

and consequently

$$P\tilde{Q} = Q\tilde{P}. \quad (8.2)$$

Let $q$ be the minimal polynomial of $\tilde{S}_Q$ (we could just as well use $\det Q$, the characteristic polynomial of $S_Q$), and let

$$q = q_1 \cdots q_s \quad (8.3)$$

be a factorization of $q$ into mutually coprime factors. Writing $g = d_i q_i$ with $d_i = (\prod_{j \neq i} q_j)$, this induces factorizations of $\tilde{Q}$ of the form (for details see Fuhrmann and Willems [30, Theorem 2.16])

$$\tilde{Q} = D_i Q_i \quad (8.4)$$

with $q_i$ the minimal polynomial of $\tilde{S}_{Q_i}$, $q_i q_j$ the minimal polynomial of $\tilde{S}_{Q_j}$, and $q_i q_j$ the minimal polynomial of $\tilde{S}_{D_i}$. Such a factorization yields $\tilde{S}_Q$-invariant subspaces of $X_Q$ given by $D_i X_{Q_i}$ and a direct sum decomposition of $X_Q$ of the form

$$X_Q = \sum_i D_i X_{Q_i}. \quad (8.5)$$

The subspace $\sum_i D_i X_{Q_i}$ is itself invariant under $\tilde{S}_Q$ and so of the form $Q_i X_{D_i}$ with $Q = Q_i D_i$. In particular we have

$$X_Q = D_i X_{Q_i} + Q_i X_{D_i}. \quad (8.6)$$

Similarly for $X_{\tilde{Q}}$ we have

$$X_{\tilde{Q}} = \tilde{D}_i X_{\tilde{Q}_i} + \tilde{Q}_i X_{\tilde{D}_i}, \quad (8.7)$$

and we have (see Fuhrmann [27])

$$\left(D_i X_{Q_i}\right)^\perp = \tilde{Q}_i X_{\tilde{D}_i}. \quad (8.8)$$

But since we also have

$$X_{\tilde{Q}} = \sum_i \tilde{D}_i X_{\tilde{Q}_i} = \tilde{D}_i X_{\tilde{Q}_i} + \sum_{j \neq i} \tilde{D}_j X_{\tilde{Q}_j}, \quad (8.9)$$
it follows that
\[ \tilde{Q}_i X \tilde{D}_i = \sum_{j \neq i} D'_j X \tilde{Q}'_j. \]  
(8.10)

In particular we have the orthogonality relations
\[ \langle D_i X Q'_i, \tilde{D}'_j X \tilde{Q}'_j \rangle = 0 \quad \text{if} \quad j \neq i, \]  
(8.11)
or equivalently
\[ \left[ Q^{-1} D_i f, \tilde{D}'_j g \right] = 0 \]  
(8.12)
for all \( f \in X Q_i \) and \( g \in X \tilde{Q}_j \).

But the equality (8.12) certainly holds when \( f \in Q_i F^m[z] \) or \( g \in \tilde{Q}_j F^m[z] \), so it holds for all polynomials \( f \) and \( g \) in \( F^m[z] \). Thus
\[ \left[ D'_j Q^{-1} D_i f, g \right] = 0 \]  
(8.13)
for all \( f, g \in F^m[z] \), which implies the following.

**Lemma 8.1.** Let \( Q \) be a nonsingular polynomial matrix, and let \( D_i \) and \( D'_j \) be defined as above. Then \( D'_j Q^{-1} D_i \) is a polynomial matrix for \( j \neq i \).

**Lemma 8.2.** Let \( G = Q^{-1} P = \tilde{P} Q^{-1} \in F^{m \times m}[z] \) be strictly proper, and \( Q_j \) be defined as before. Then there exist uniquely determined \( P_j \in F^{m \times m}[z] \) such that
\[ P = \sum_j D_j P_j \]  
(8.14)
and \( Q^{-1}_j P_j \) is strictly proper.

**Proof.** The existence of \( P_j \) such that (8.14) is satisfied follows from the left coprimeness of the \( D_j \). If \( Q_j P_j^{-1} \) is not strictly proper, we can write \( P_j = P'_j + Q_j H_j \) with \( Q_j^{-1} P'_j \) strictly proper. But then
\[ P = \sum_j D_j P_j = \sum_j D_j (P'_j + Q_j H_j) = \sum_j D_j P'_j + Q \sum_j H_j, \]  
and since \( Q^{-1} P \) is strictly proper, it follows that \( \sum H_j = 0 \).
\[ \blacksquare \]
THEOREM 8.3. Let \( G = Q^{-1}P = \hat{Q}Q^{-1} \), and let \( Z: X_{\hat{Q}} \to X_Q \) be defined by
\[
Zf = \pi_Q Pf \quad \text{for} \quad f \in X_{\hat{Q}}.
\] (8.15)
Then
\[
Z(\hat{D}_i X_{\hat{Q}_i}) = D_i X_Q. \] (8.16)

Proof. \( Z\hat{D}_i f = \pi_Q \hat{P}\hat{D}_i f = Q\pi_\hat{Q}^{-1}P\hat{D}_i f = Q\pi_\hat{Q}^{-1}\hat{Q}\hat{D}_i f. \) Now by the left coprimeness of the \( D_j \) there exist \( P_j \) such that
\[
P = \sum D_j P_j
\] (8.17)
and
\[
\hat{P} = \sum \hat{P}_j \hat{D}_j,
\] (8.18)
and so \( Z\hat{D}_i f = Q\pi_\hat{Q}^{-1} \sum \hat{P}_j \hat{D}_j \hat{Q}^{-1}\hat{D}_i f. \) Now for \( j \neq i \) it follows from Lemma 8.1, and the fact that
\[
\hat{D}_j \hat{Q}^{-1}\hat{D}_i = (D_j Q^{-1}D_j)^{-1},
\]
that \( \hat{D}_j \hat{Q}^{-1}\hat{D}_i \) is a polynomial matrix. So
\[
Z\hat{D}_i f = Q\pi_\hat{Q}^{-1} \hat{P}_i \hat{D}_i \hat{Q}^{-1}\hat{D}_i f = D_i Q_i \pi_\hat{Q}^{-1} \hat{P}_i \hat{D}_i \hat{Q}^{-1}\hat{D}_i f = D_i \pi_\hat{Q}_i P_i \hat{D}_i \hat{Q}^{-1}\hat{D}_i f = D_i \pi_\hat{Q}_i P_i \hat{D}_i f \in D_i X_Q,
\]
as \( Q = D_i Q_i. \)

THEOREM 8.4. Let \( G = Q^{-1}P = PQ^{-1} \) be strictly proper. Then with the previous notation and \( G_j = Q_j^{-1}P_j \) we have
\[
G = \sum_{j=1}^{s} G_j,
\] (8.19)
\[
H_G = \sum_{j=1}^{s} H_{G_j},
\] (8.20)
\begin{equation}
\text{rank } H_G = \sum_{j=1}^{s} \text{rank } H_{G_j},
\tag{8.21}
\end{equation}

or equivalently

\begin{equation}
\delta(G) = \sum_{j=1}^{s} \delta(G_j),
\tag{8.22}
\end{equation}

\begin{equation}
\sigma(H_G) = \sum_{j=1}^{s} \sigma(H_{G_j}),
\tag{8.23}
\end{equation}

and if \( Z: X_\tilde{C} \to X_Q \) is defined by (8.14), then

\begin{equation}
\sigma(Z) = \sum_{i=1}^{s} \sigma(Z_i),
\tag{8.24}
\end{equation}

where \( Z_i: X_{\tilde{C}_i} \to X_{Q_i} \) is defined by

\begin{equation}
Z_i u = \pi_{Q_i} P D_i' u_i = \pi_{Q_i} R_i u_i
\tag{8.25}
\end{equation}

with \( PD_i' = Q_i' L_i + R_i, (Q_i')^{-1} R_i \) strictly proper, and \( R_i \tilde{Q}_i = Q_i' \tilde{R}_i \). Equivalently

\begin{equation}
\sigma(\Gamma(Q, \tilde{P}, P, \tilde{Q})) = \sum \sigma(\Gamma(Q_i', \tilde{R}_i, R_i, \tilde{Q}_i')).
\tag{8.26}
\end{equation}

**Proof.** Since \( P = \sum_{j=1}^{s} D_j P_j \) and \( Q = D_j Q_j \), it follows (assuming without loss of generality \( Q_j^{-1} P_j \) to be strictly proper) that

\begin{equation}
G = Q^{-1} P = \sum_{j=1}^{s} Q^{-1} D_j P_j = \sum_{j=1}^{s} Q_j^{-1} P_j = \sum_{j=1}^{s} \tilde{P}_j \tilde{Q}_j^{-1} = \tilde{G}.
\end{equation}

Hence (8.22) follows, which of course implies (8.23). Now \( G \) is the transfer function of the parallel connection of systems with transfer functions \( G_j \), so by Fuhrmann [23] it follows from the mutual right coprimeness of the \( Q_j \) and consequent left coprimeness of the \( \tilde{Q}_j \) that (8.21) holds.

Finally we have \( X_\tilde{C} = \sum_{i=1}^{s} \tilde{D}_i X_{\tilde{C}_i} \), so if \( f \in X_\tilde{C} \), then \( f = \sum_{i=1}^{s} \tilde{D}_i g_i \) with \( g_i \in X_{\tilde{C}_i} \). Now, using the orthogonality relation (8.11) and the previous
theorem, we have

\[ \langle Zf, f \rangle = \sum_{i=1}^{n} \langle Z\tilde{D}_i g_i, \tilde{D}_i g_i \rangle \]

\[ = \sum_{i=1}^{n} \langle \pi_{Q_i} P\tilde{D}_i g_i, \tilde{D}_i g_i \rangle \]

\[ = \sum_i [Q^{-1}R_{Q_i} D_i Q^{-1}P\tilde{D}_i g_i, \tilde{D}_i g_i] \]

\[ = \sum_i [(Q_i')^{-1}Q_i'D_i Q^{-1}P\tilde{D}_i g_i, g_i] \]

\[ = \sum_i [\pi_{-} (Q_i')^{-1}P\tilde{D}_i g_i, g_i] \]

\[ = \sum_i \langle \pi_{Q_i} P\tilde{D}_i g_i, g_i \rangle \]

\[ = \sum_i \langle \pi_{Q_i} R_i g_i, g_i \rangle \]

\[ = \sum_i \langle Z_i g_i, g_i \rangle, \]

and since we have a direct sum decomposition, (8.24) follows.

Since the maps $Z_i$ are all self-dual, we use now Theorem 4.7 to obtain (8.26).

One should note that no assumptions on the left coprimeness of $Q$ and $P$ have been made. The coprimeness conditions were those on the $Q_P$ which are part of the construction.

9. SIGNATURE SYMMETRIC REALIZATIONS: 
THE MULTIVARIABLE CASE

In the previous sections we saw, given a real symmetric transfer function, the relation between the signatures of the Hankel matrix and that of the associated intertwining map and Bezoutian. In trying to suitably generalize
the Hermite-Hurwitz theorem one would like also to have an appropriate
generalization of the Cauchy index. Though this section is heavily indebted to
the fundamental paper of Bitmead and Anderson [3], their definition of the
matrix Cauchy index is not adopted. However, using basically the calculations
of Bitmead and Anderson one can put their results in terms of local canonical
forms. This enables us to define the notion of local sign characteristic as well
as a local Cauchy index and use it to pass to a global definition. The stress will
be on the connection with signature symmetric realizations. For a geometric
approach to the matrix Cauchy index the reader is referred to Byrnes and
Duncan [13] and Byrnes [14].

Given two real symmetric rational functions \( G_1 \) and \( G_2 \) with coprime
factorizations
\[
G_i = P_i Q_i^{-1},
\]
we say that \( G_1 \) is congruent equivalent to \( G_2 \), and write \( G_2 \simeq G_1 \), if there
exists a polynomial matrix \( U \) such that \( U \) and \( Q_2 \) are left coprime and
\[
G_2 = \tilde{U} G_1 U + \Pi,
\]
where \( \Pi \) denotes a symmetric polynomial matrix. We note that (9.2) implies
that the McMillan degree of \( G_2 \) is less than or equal to that of \( G_1 \). The
coprimeness condition is not as asymmetric as it seems. Indeed, by the
symmetry of \( G_i \) we have \( G_i = \tilde{Q}_i^{-1} \tilde{P}_i \), and the left coprimeness of \( U \) and \( Q_i \) is
the same as the right coprimeness of \( U \) and \( Q_i \).

**Lemma 9.1.** The previously defined relation is a bona fide equivalence
relation, i.e., it is reflexive, symmetric, and transitive.

*Proof.* That \( G \simeq G \) is trivial. Assume \( G_2 \simeq G_1 \) in the previous sense.
Since \( U \) and \( Q_1 \) are left coprime, there exist polynomial matrices \( X \) and \( Y \)
such that
\[
UX + Q_1 Y = I.
\]
Hence
\[
\tilde{X} G_2 X = \tilde{X} \tilde{U} G_1 U X + \tilde{X} \Pi X
\]
\[
= (I - \tilde{Y} \tilde{Q}_1) G_1 (I - Q_1 Y) + \tilde{X} \Pi X
\]
\[
= G_1 - \tilde{Y} \tilde{Q}_1 G_1 - G_1 Q_1 Y + \tilde{Y} \tilde{Q}_1 G_1 Q_1 Y + \tilde{X} \Pi X.
\]
Since $G_1 = P_1 Q_1^{-1} = \tilde{Q}_1^{-1} \tilde{P}_1$, it is clear that $G_1 Q_1 Y$, $\tilde{Y}_1 \tilde{Q}_1$, and $\tilde{Y}_1 \tilde{Q}_1 G_1 Q_1 Y$ are polynomial matrices. Letting

$$\Pi_1 = \tilde{Y} \tilde{Q}_1 G_1 + G_1 Q_1 Y - \tilde{Y} \tilde{Q}_1 G_1 Q_1 Y - \tilde{X} \Pi X,$$

we have

$$G_1 = \tilde{X} G_2 X + \Pi_1. \quad (9.4)$$

It remains to check that $X, Q_2$ are left coprime. Indeed, if $X$ and $Q_2$ were not left coprime, we would have the inequality $\delta(G_2) < \delta(G_1)$ between the McMillan degrees of $G_1$ and $G_2$, in contradiction to the inverse inequality noted before. Thus in particular congruence equivalence implies equality of McMillan degrees. This proves symmetry.

To prove transitivity let $G_2 \approx G_1$ and $G_3 \approx G_2$. So there exist polynomial matrices $U_i, i = 1, 2$, such that $U_i, Q_1$ are left coprime and the relations

$$G_2 = \tilde{U}_1 G_1 U_1 + \Pi_1 \quad (9.5)$$

and

$$G_3 = \tilde{U}_2 G_2 U_2 + \Pi_2 \quad (9.6)$$

hold. This implies

$$G_3 = \tilde{U}_2 G_2 U_2 + \Pi_2$$
$$= \tilde{U}_2 (\tilde{U}_1 G_1 U_1 + \Pi_1) U_2 + \Pi_2$$
$$= (U_1 U_2)^{-1} G_1 (U_1 U_2) + (\tilde{U}_2 \Pi_1 U_2 + \Pi_2)$$
$$= (U_1 U_2)^{-1} G_1 (U_1 U_2) + I_{3},$$

from which it is also clear that $\delta(G_3) < \delta(G_1)$. Now, by the left coprimeness assumptions there exist polynomial matrices $X_i$ and $Y_i, i = 1, 2$, such that

$$U_i X_i + Q_1 Y_i = I. \quad (9.7)$$
This implies

\[
\tilde{X}_1 \tilde{X}_2 G_3 X_1 X_1 = \tilde{X}_1 \tilde{X}_2 \tilde{U}_1 \tilde{U}_3 G_4 U_1 U_2 X_1 X_1 + \tilde{X}_1 \tilde{X}_2 \Pi_3 X_2 X_1 \\
= \tilde{X}_1 (I - \tilde{Y}_2 \tilde{Q}_2) \tilde{U}_1 G_1 U_1 (I - Q_2 Y_2) X_1 \\
= \tilde{X}_1 \tilde{U}_1 G_1 U_1 X_1 - \tilde{X}_1 \tilde{Y}_2 \tilde{Q}_2 \tilde{U}_1 G_1 U_1 X_1 \\
- \tilde{X}_1 \tilde{U}_1 G_1 U_1 X_1 + \tilde{X}_1 \tilde{Y}_2 \tilde{Q}_2 \tilde{U}_1 G_1 U_1 Q_2 Y_2 X_1 \\
= (U_1 X_1)^{-1} G_1 (U_1 X_1) + \Pi_4 \\
= (I - \tilde{Y}_2 \tilde{Q}_2) G_1 (I - Q_2 Y_2) + \Pi_4 \\
- G_1 + \Pi_4 - \tilde{Y}_2 \tilde{Q}_2 G_1 - C_1 Q_1 Y_1 + (Q_1 Y_1) G_1 (Q_3 Y_1) \\
= G_1 + \Pi_5.
\]

From this it is necessary that \( X_2 X_1 \) be left coprime with \( Q_3 \); otherwise we would have \( \delta(G_1) \leq \delta(G_3) \), in contradiction to the inequality \( \delta(G_3) \leq \delta(G_1) \) noted before. Thus transitivity has been proved. 

The next theorem is an extension of Lemma 3.1 in Bitmead and Anderson [1977].

**Theorem 9.2.** Let \( G_1, G_2 \) be two real symmetric rational functions with coprime factorizations \( G_i = P_i Q_i^{-1} \). Then \( G_2 = G_1 \) implies

\( \begin{align*}
(i) \quad & \text{rank}(H_{G_2}) = \text{rank}(H_{G_1}). \\
(ii) \quad & \sigma(H_{G_2}) = \sigma(H_{G_1}).
\end{align*} \)

**Proof.** Since \( \text{rank} H_G \) is equal to the McMillan degree of \( G \), part (i) follows from the proof of the previous theorem.

To prove (ii) note that \( \text{Ker} H_G = Q F^m[z] \) and \( \text{Range} H_G = X \tilde{Q} \), so \( \sigma(H_G) = \sigma(\tilde{H}_G) \), where \( \tilde{H}_G : X_Q \to X_0 \) is defined by

\[
\tilde{H}_G = H_G \mid X_Q.
\]

Since \( G_2 = G_1 \), it follows there exist \( U \) and \( \Pi \), with \( U \) and \( Q_1 \) left coprime, such that

\[
G_2 = \tilde{U} G_1 U + \Pi.
\]
Now
\[ [H_{G_2} f, f] = [H_{\tilde{G}_1} U + \pi f, f] = [\pi_-(\tilde{G}_1 U + \pi) f, f] \]
\[ = [\pi_-=\tilde{G}_1 U f, f] = [G_1 U f, U f] \]
\[ = [H_{G_1} U f, U f] = [H_{\pi G_1} U f, \pi G_1 U f]. \]

But the left coprimeness of $U$ and $Q_1$ implies
\[ U F^m(z) + Q_1 F^m(z) = F^m(z), \quad \text{(9.10)} \]
and so
\[ \pi_{Q_1} U F^m(z) = X_{Q_1} \quad \text{(9.11)} \]
and the equality of signatures follows.

The converse of the previous theorem is not true. However, this naturally raises the question of finding the congruence equivalence invariants or, alternatively, of reducing the symmetric transfer functions to some canonical form. This we proceed to do, and the main result is Theorem 9. We begin by a preliminary analysis which studies the problem locally.

**Theorem 9.3.**

(a) Let $G(z) = \sum_{i=1}^r \Gamma_i/(z - a)^i$ with $\Gamma_i$ real symmetric matrices. Then $G \simeq G'$, where
\[ G'(z) = \text{diag}(D_1(z-a)^{-k_1}, \ldots, D_t(z-a)^{-k_t}) \quad \text{(9.12)} \]
and
\[ D_i = \begin{pmatrix} I_i^{(+)} & 0 \\ 0 & -I_i^{(-)} \end{pmatrix} \quad \text{(9.13)} \]
The dimensions of the identity matrices $I_i^{(+)}$ and $I_i^{(-)}$ are uniquely determined.

(b) Let $G(z) = \sum_{i=1}^r (A_i + B_i z)/[(z - a)^2 + b^2]^i$ with $A_i, B_i$ real symmetric matrices. Then $G \simeq G'$, where
\[ G'(z) = \text{diag}(I_1[(z-a)^2 + b^2]^{-k_1}, \ldots, I_t[(z-a)^2 + b^2]^{-k_t}). \quad \text{(9.14)} \]
The dimensions of the identity matrices $I_i$ are uniquely determined.
Before proceeding with the proof of Theorem 9.3 we prove some lemmas which deal with the special case of nonsingularity of the highest coefficient in the local partial fraction expansion.

**Lemma 9.4.** Let $G(z) = \sum_{j=1}^{\infty} A_j z^{-j}$ with $A_j$ $n \times n$ real symmetric and $A_j$ nonsingular. Then there exists a signature symmetric matrix $\Sigma$ and polynomial matrices $X$ and $\Pi$ such that $X$ and $z^t \Pi$ are left coprime and

$$G(z) = \tilde{X}(z) \Sigma z^{-s} X(z) + \Pi(z). \quad (9.15)$$

The signature matrix $\Sigma$ is uniquely determined by $G$.

**Proof.** Let $X(z) = \sum_{j=0}^{s-1} X_j z^j$. Then (9.15) is equivalent to the system

$$\tilde{X}_0 \Sigma X_0 = A_s,$$

$$\tilde{X}_0 \Sigma X_1 + \tilde{X}_1 \Sigma X_0 = A_{s-1},$$

$$\vdots$$

$$\tilde{X}_0 \Sigma X_{s-1} + \tilde{X}_1 \Sigma X_{s-2} + \cdots + \tilde{X}_{s-1} \Sigma X_0 = A_0. \quad (9.16)$$

Now the first equation is solvable by reducing the symmetric matrix $A_s$ to congruence canonical form. Clearly $X_0$ is nonsingular. The other equations can be solved recursively, and all are of the form

$$\tilde{X}_0 J X_k + \tilde{X}_k J X_0 = C_k \quad (9.17)$$

with $C_k$ symmetric. By defining $W_k = \tilde{X}_0 J X_k$, Equation (9.17) reduces to $\tilde{W}_k + W_k = C_k$, which has as a solution $W_k = C_k / 2$. So

$$X_k = C_k X_0^{-1} J / 2 \quad (9.18)$$

is a solution. 

**Lemma 9.5.** Let $A, B$ be real symmetric $n \times n$ matrices. Then $A \pm iB$ are nonsingular polynomial matrices if and only if

$$\begin{pmatrix} A & -B \\ B & A \end{pmatrix}$$

is nonsingular.
**Proof.** \( \begin{pmatrix} x \\ y \end{pmatrix} \) is a null vector of \( \begin{pmatrix} A & -B \\ B & A \end{pmatrix} \) if and only if

\[
Ax - Bu = 0, \\
By + Ax = 0.
\]  

(9.19)

But these are also the equations resulting from separating the equation

\[
(A + Bi)(x + iy) = 0
\]  

(9.20)

into its real and imaginary parts.

**Lemma 9.6.** Let \( A, B \) be real symmetric matrices, and assume \( A + iB \) are invertible. Then there exist real matrices \( X, Y \) such that

\[
XX - YY = A, \\
XY + YX = B.
\]  

(9.21)

**Proof.** Our assumption is equivalent to the nonsingularity of the symmetric matrix

\[
\begin{pmatrix} A & B \\ B & -A \end{pmatrix}.
\]  

(9.22)

Let \( \lambda \) be an eigenvalue of this matrix, with a corresponding eigenvector \( \begin{pmatrix} x \\ y \end{pmatrix} \). Then it is easy to check that \( \begin{pmatrix} -y \\ x \end{pmatrix} \) is an eigenvector corresponding to the eigenvalue \( -\lambda \). This implies the existence of an orthogonal matrix of the form

\[
\begin{pmatrix} X & -Y \\ Y & X \end{pmatrix}
\]

and a real diagonal matrix \( L \) such that

\[
\begin{pmatrix} A & B \\ B & -A \end{pmatrix}\begin{pmatrix} X & -Y \\ Y & X \end{pmatrix} = \begin{pmatrix} X & -Y \\ Y & X \end{pmatrix}\begin{pmatrix} L & 0 \\ 0 & -L \end{pmatrix}.
\]

Factoring \( L \) as \( \hat{K}K \) and redefining \( X \) and \( Y \), we obtain the equality

\[
\begin{pmatrix} A & B \\ B & -A \end{pmatrix} = \begin{pmatrix} \hat{X} & -\hat{Y} \\ \hat{Y} & \hat{X} \end{pmatrix}\begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix}\begin{pmatrix} X & Y \\ Y & X \end{pmatrix},
\]  

(9.23)

which is equivalent to (9.21).
Lemma 9.7. Let

\[ G(z) = \sum_{j=1}^{s} (A_j \cdot B_j z)(z^2 + 1)^{-j} \]

with \( A_j, B_j \) real symmetric matrices. Furthermore, assume \( A_{j_1} \pm iB_{j_1} \) are nonsingular, which is equivalent to the coprimeness of \( \sum_{j=1}^{s}(A_j + B_j z)(z^2 + 1)^{s-j} \) and \( (z^2 + 1)^s I \). Then

\[ G(z) = \frac{1}{(z^2 + 1)^s}. \]  

(9.24)

Proof. We will show there exists a polynomial matrix

\[ X(z) = \sum_{i=0}^{s-1} (X_i + Y_i z)(z^2 + 1)^i \]  

(9.25)

with real matrix coefficients for which

\[ G(z) = \tilde{X}(z)(z^2 + 1)^{-s}X(z) + \Pi(z) \]  

(9.26)

holds. Equation (9.26) is equivalent to the system of equations

\[ A_z + B_z z = (\tilde{X}_0 X_0 - \tilde{Y}_0 Y_0) + (\tilde{X}_0 Y_0 + \tilde{Y}_0 X_0) z, \]

\[ A_{s-1} + B_{s-1} z = (\tilde{X}_1 X_1 - \tilde{Y}_1 Y_1 + \tilde{X}_1 X_0 - \tilde{Y}_1 Y_0 + \tilde{Y}_1 X_0) + (\tilde{X}_0 Y_1 + \tilde{Y}_0 X_1 + \tilde{X}_1 Y_0 + \tilde{Y}_1 X_0) z, \]  

(9.27)

By Lemma 9.6 the first equation has a solution with

\[ \begin{pmatrix} X_0 & Y_0 \\ -Y_0 & X_0 \end{pmatrix} \]

nonsingular. But this nonsingularity condition is equivalent to the coprimeness of \( X \) and \( (z^2 + 1)^s I \).
The $k$th equation can be put in the form
\[ \dot{X}_0X_k + \dot{X}_kX_0 + \dot{Y}_0Y_k - \dot{Y}_kY_0 = A'_k \]
\[ \dot{X}_0Y_k + \dot{Y}_kX_0 + \dot{Y}_0X_k + \dot{X}_kY_0 = B'_k \]
with $A'_k, B'_k$ real symmetric. This can be rewritten in the matrix form
\[
\begin{pmatrix}
\dot{X}_0 & \dot{Y}_0 \\
\dot{Y}_0 & -\dot{X}_0
\end{pmatrix}
\begin{pmatrix}
X_k & Y_k \\
-Y_k & X_k
\end{pmatrix} +
\begin{pmatrix}
\dot{X}_k & \dot{Y}_k \\
\dot{Y}_k & \dot{X}_k
\end{pmatrix}
\begin{pmatrix}
X_0 & Y_0 \\
0 & -X_0
\end{pmatrix} =
\begin{pmatrix}
A'_k & B'_k \\
B'_k & -A'_k
\end{pmatrix}.
\]  
(9.29)

This can be solved immediately by
\[
\begin{pmatrix}
X_k & Y_k \\
-Y_k & X_k
\end{pmatrix} = \frac{1}{2}
\begin{pmatrix}
\dot{X}_0 & \dot{Y}_0 \\
\dot{Y}_0 & -\dot{X}_0
\end{pmatrix}^{-1}
\begin{pmatrix}
A'_k & B'_k \\
B'_k & -A'_k
\end{pmatrix}.
\]
(9.30)

Proof of Theorem 9.3. (a): By using scaling generations we may assume without loss of generality that $a = 0$. Thus $G(z) = \sum_{i=1}^{s} C_i z^{-i}$. If $G_s$ is nonsingular, we apply Lemma 9.4. Otherwise let $T$ be a nonsingular constant matrix such that
\[
TG_sT = \begin{pmatrix}
G_s^{(11)} & 0 \\
0 & 0
\end{pmatrix}
\]
with $G_s^{(11)}$ symmetric and nonsingular. Hence $G = G'$, where
\[
G(z) = \begin{pmatrix}
G^{(11)}(z) & G^{(12)}(z) \\
G^{(21)}(z) & G^{(22)}(z)
\end{pmatrix}
\]
with $G^{(11)}$ satisfying the conditions of Lemma 9.4. Since obviously
\[
\text{Range } H_{G^{(12)}} \subset \text{Range } H_{G^{(11)}},
\]
(9.30)
there exists a polynomial matrix $C$ such that

$$G^{(12)} = \pi_+ G^{(11)} C. \quad (9.31)$$

For a proof of this implication see Fuhrmann [28]. It follows now, using also the symmetry of $G'$, that

$$
\begin{pmatrix}
I & 0 \\
-C & I
\end{pmatrix}
G'(I 
0 \\
I 
0 \\
C 
I 
H 
0 
0 
0

\end{pmatrix} + \Pi
$$

with $\Pi$ a polynomial matrix and $H(z) = \sum_{i=-1}^{1} H_i z^{-i}$. To $G^{(11)}$ we apply Lemma 9.4 and repeat the same process with $H$. The proof is completed by induction.

(b): Let us consider, without loss of generality,

$$G(z) = \sum_{j=1}^{s} A_j + B_j z
$$

If $A_s \pm i B_s$ are nonsingular, then Lemma 9.7 is applicable. Otherwise there exists a nonsingular matrix

$$
\begin{pmatrix}
X_0 & Y_0 \\
-Y_0 & X_0
\end{pmatrix}
$$

such that

$$
\begin{pmatrix}
A_s & B_s \\
B_s & -A_s
\end{pmatrix} = \begin{pmatrix}
\tilde{X}_0 & -\tilde{Y}_0 \\
\tilde{Y}_0 & \tilde{X}_0
\end{pmatrix}
\begin{pmatrix}
P & 0 \\
0 & P
\end{pmatrix}
\begin{pmatrix}
X_0 & Y_0 \\
-Y_0 & X_0
\end{pmatrix}
$$

with $P$ an orthogonal projection of the form $\begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}$. Relative to this representation of $P$ and with $X = X_0 + X_0 z$, we have

$$\pi \tilde{X}G X = \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix},$$

with $G_{11}$ satisfying the conditions of Lemma 9.7 and Range $H_{C,12} \subset$ Range $H_{G,11}$. We finish the proof basically as in part (a).
To prove uniqueness in case (a) assume $G$ is congruent equivalent to both
\[ \text{diag}(J_1z^{-k_1}, \ldots, J_s z^{-k_s}) \] (9.32)
and
\[ \text{diag}(J'_1z^{-l_1}, \ldots, J'_s z^{-l_s}). \] (9.33)

We will prove $s = t$, $k_i = l_i$, and $J_i = J'_i$ for $i = 1, \ldots, s$. By realization theory and the state space isomorphism theorem, the invariant factors of the generators in the two minimal realizations based on (9.32) and (9.33) are $z^{k_j}$ counted \( \dim J_j \) times, \( j = 1, \ldots, s \), and $z^{l_j}$ counted \( \dim J'_j \) times, \( j = 1, \ldots, t \), respectively. Thus necessarily $s = t$, $k_i = l_i$ for $i = 1, \ldots, s$ and $\dim J_i = \dim J'_i$. It remains to prove that $\sigma(J_i) = \sigma(J'_i)$ for $i = 1, \ldots, s$. Clearly if $G = G'$, then also $\pi_- z^i G = \pi_- z^i G'$. Hence, by the transitivity of congruence equivalence and Theorem 9.2,

\[ \pi_- z^i(J_1z^{-k_1}, \ldots, J_s z^{-k_s}) \]
and
\[ \pi_- z^i(J'_1z^{-l_1}, \ldots, J'_s z^{-l_s}) \]
induce Hankel maps with the same signatures. By choosing $j = k_i$, $i = 1, \ldots, s - 1$, we obtain
\[ \pi_- z^{k_i-1}(J_1z^{-k_1}, \ldots, J_s z^{-k_s}) = (0, \ldots, 0, J_i z^{-(k_i-k_{i-1})}, \ldots, J_s z^{-(k_s-k_{s-1})}) \]
and
\[ \pi_- z^{k_i}(J'_1z^{-l_1}, \ldots, J'_s z^{-l_s}) = (0, \ldots, 0, J'_i z^{-(k_i-k_{i-1})}, \ldots, J'_s z^{-(k_s-k_{s-1})}) \].

The equality of signatures implies the following set of equalities:
\[ \sum_{j=i} k_j \sigma(J_j) = \sum_{j=i} k_j \sigma(J'_j), \quad i = 1, \ldots, s. \] (9.34)

This implies
\[ \sigma(J_i) = \sigma(J'_i), \quad i = 1, \ldots, s, \] (9.35)
and we are done.
In case (b) the proof is even simpler, for assume $G$ is congruent equivalent to both

$$\text{diag} \left( I_1(z^2 + 1)^{-k}, \ldots, I_s(z^2 + 1)^{-k} \right) \quad (9.36)$$

and

$$\text{diag} \left( I_1(z^2 + 1)^{-l}, \ldots, I_s(z^2 + 1)^{-l} \right). \quad (9.37)$$

Then again by an application of realization theory and the state space isomorphism theorem, the invariant factors of the generators in the canonical realizations based on (9.36) and (9.37) respectively are $(z^2 + 1)^k$ counted $\dim I_j$ times and $(z^2 + 1)^l$ counted $\dim I_j'$ times. This implies $t = s$, $k_j = l_j$ and $I_j = I_j'$ for $j = 1, \ldots, s$, and we are done.

We state now the main result of this section. In the next theorem we state it in polynomial terms, whereas in Theorem 9.12 we give a matrix representation of the result.

**Theorem 9.8.** Let $G(z)$ be an $m \times m$ real symmetric strictly proper rational matrix. Then $G$ has the representation

$$G(z) = \sum_{i=1}^{s} \tilde{X}_i D_i^{-1} X_i + \Pi \quad (9.38)$$

with $X_i, D_i$ left coprime and either

$$D_i(z) = \text{diag}(\Sigma_1^{(i)}(z - a_i)^{k_1}, \ldots, \Sigma_m^{(i)}(z - a_i)^{k_m}) \quad (9.39)$$

or

$$D_i(z) = \text{diag}(I_1^{(i)} \left[ (z - a_i)^2 + b_i^2 \right]^{k_1}, \ldots, I_m^{(i)} \left[ (z - a_i)^2 + b_i^2 \right]^{k_m}). \quad (9.40)$$

Here $\Sigma_i$ are signature matrices. The signature matrices $\Sigma_i$ and the numbers $a_i, b_i$ and $k_j^{(i)}, j = 1, \ldots, m_i, i = 1, \ldots, s$, are uniquely determined.

**Remark.** The set of polynomials $(z - a_i)^k$ and $[(z - a_i)^2 + b_i^2]^k$ are the elementary divisors of $A$ in any canonical realization $(A, B, C)$ of $G$. In
particularly we have that
\[ \sum_{j=1}^{m} k_j^{(i)} \dim \Sigma_j^{(i)}, \]  
(9.41)

which is the local multiplicity of \( A \), is less than or equal to \( m \).

Proof. Let \( G = \sum G_j \) be the partial fraction decomposition of \( G \). Thus the minimal polynomial of \( G_j \) is a power of an irreducible polynomial. To each of the \( G_j \) we apply Theorem 9.3 and the result follows.

Following Gohberg et al. [36], we will call (9.39) and (9.40) the signature characteristic of \( G \). It is clear that the signature characteristic carries all the spectral and signature information of \( G \). The signature of \( H_G \), the Hankel matrix induced by \( G \), is determined by the signature characteristic in the following way.

**Corollary 9.9.** Let \( G \) be as in Theorem 9.8. Then
\[ \sigma(H_G) = \sum \sigma\left( \Sigma_j^{(i)} \right), \]  
(9.42)

where the summation extends over all elementary divisors which are odd powers of a linear polynomial.

The result of Theorem 9.8 lends itself to further reduction. Let us unify notation in (9.39) and (9.40) and write now
\[ D_i(z) = (\varepsilon_1^{(i)} q_1^{(i)}, \ldots, \varepsilon_m^{(i)} q_m^{(i)}) \]  
(9.43)

with the convention that \( q_j^{(i)} q_k^{(i)} \) for \( j > k \), \( \varepsilon_j^{(i)} = \pm 1 \), and for each multiple elementary divisor the + signs precede the − signs. Thus \( D_i \) is uniquely determined by (9.39) and (9.40). Define now
\[ q_j = \prod_{i=1}^{\infty} q_j^{(i)} \]  
(9.44)

and
\[ d_j^{(i)} = \prod_{k \neq i} q_k^{(k)} = \frac{q_j}{q_j^{(i)}} \]  
(9.45)
and

\[ e_j(z) = \sum_i \epsilon_j^{(i)} d_j^{(i)}. \quad (9.46) \]

Next we define

\[ E(z) = \text{diag}(e_1(z), \ldots, e_m(z)) \quad (9.47) \]

and

\[ D(z) = \text{diag}(q_1(z), \ldots, q_m(z)). \quad (9.48) \]

We obviously have the relation

\[ ED = DE, \quad (9.49) \]

and so we can state

**Corollary 9.10.** Let \( G(z) \) be as in Theorem 9.8. Then \( G \) is congruent equivalent to

\[ G'(z) = E(z)D(z)^{-1} = D(z)^{-1}E(z) = \text{diag}(e_1q_1^{-1}, \ldots, e_mq_m^{-1}). \quad (9.50) \]

We will call (9.50) the *congruence McMillan* form of \( G \). Note that it completely determines both the invariant factors, namely the polynomials \( q_1, \ldots, q_m \), as well as all the signature information.

By a suitable choice of basis we can obtain a convenient matrix representation for the canonical form of congruence equivalence.

**Theorem 9.12.** Let \( G \) be an \( m \times m \) real symmetric strictly proper rational matrix. Then \( G \) admits a canonical signature symmetric realization \((A, B, C)\) with signature matrix

\[ A = \text{diag}(J_1, \ldots, J_k), \quad (9.51) \]

\[ \Sigma = \text{diag}(\epsilon_1 \Sigma_1, \ldots, \epsilon_k \Sigma_k), \quad (9.52) \]
and either

\[ J_i = \begin{pmatrix} a_i & 1 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 0 & \cdots & a_i \end{pmatrix} \quad (9.53) \]

or

\[ J_i = \begin{pmatrix} a_i - b_i^2 & 1 \\ 1 & a_i \\ 0 & 1 & a_i - b_i^2 \\ 0 & 0 & 1 & a_i \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & a_i - b_i^2 \\ 0 & 0 & 1 & a_i \end{pmatrix} \quad (9.54) \]

and

\[ \Sigma_i = \begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 1 \end{pmatrix} \quad (9.55) \]

where the \( \varepsilon_i = \pm 1 \), \( J_i \), and \( \Sigma_i \) are uniquely determined.

Proof: Since, by Corollary 9.10, \( G \) is congruent equivalent to \( G' \) of (9.50), by the symmetry property of congruence equivalence we have the representation

\[ G(z) = \tilde{Y}(z)E(z)D(z)^{-1}Y(z) + \Pi(z) \]

for some polynomial matrices \( Y \), \( \Pi \) for which \( Y \) and \( D \) are left coprime.
We apply now the realization procedure of Theorem 2.3. By a choice of basis in

\[ X_D = X_{q_1} \oplus \cdots \oplus X_{q_n}, \]

which is constructed by choosing basis elements in the direct summands as in Section 6, the result follows.

10. ON A THEOREM OF FROBENIUS

As an application of polynomial models we give a simple proof of a theorem of Frobenius [21] and a generalization of it. The underlying field is arbitrary.

**Theorem 10.1.** Every square matrix over an arbitrary field \( F \) is the product of two symmetric ones.

Before proving the theorem we show that there is an equivalent way of stating it. This is due to Taussky and Zassenhaus [55].

**Theorem 10.2.** Let \( A \) be a square matrix over an arbitrary field \( F \). There exists a symmetric nonsingular matrix \( S \) intertwining \( A \) and \( \tilde{A} \) if and only if \( A \) is the product of two symmetric matrices one of which is nonsingular.

**Proof.** Assume \( S \) is symmetric nonsingular and satisfies \( AS = \tilde{S}A \). Then

\[ A = S\tilde{A}S^{-1} = ST, \tag{10.1} \]

where \( T \) is defined by \( T = \tilde{A}S^{-1} \). Since

\[ \tilde{T} = S^{-T}A = S^{-1}A = \tilde{A}S^{-1} = T, \]

\( T \) is symmetric.

Conversely, assume \( A = ST \), with \( \tilde{S} = S \) and \( \tilde{T} = T \), and assume \( S \) is nonsingular. By the symmetry of \( S \) and \( T \) we have \( A = TS \), and therefore

\[ AS = STS = S\tilde{A}, \tag{10.2} \]

or \( S \) intertwines \( A \) and \( \tilde{A} \).
Proof of Theorem 10.1. Since $zI - A$ and $zI - \tilde{A}$ have the same Smith form $D(z)$, there exist unimodular polynomial matrices $U$ and $V$ such that

$$D(z) = U(z)(zI - A)V(z) = \tilde{V}(z)(zI - \tilde{A})\tilde{U}(z),$$

and hence

$$(V^{-1}U)(zI - A) = (zI - \tilde{A})(\tilde{V}V^{-1}),$$

so that, with $W$ the unimodular matrix defined by

$$W(z) = \tilde{U}(z)V(z)^{-1},$$

we have

$$\tilde{W}(z)(zI - A) = (zI - \tilde{A})W(z).$$

Now the previous equality implies that the map $Z: X_{zI - A} \rightarrow X_{zI - \tilde{A}}$ defined by

$$Zf = \pi_{zI - \tilde{A}}\tilde{W}f \quad (10.3)$$

is invertible by Theorem 2.1, and self-dual by Theorem 2.9 in Fuhrmann [27]. Thus we have

$$ZS_{zI - A} = S_{zI - \tilde{A}}Z. \quad (10.4)$$

By taking matrix representations and noting that the standard basis in $F^n$ is a self-dual basis, the result follows.

Actually we can prove a bit more.

**Theorem 10.3.** Given a reachable pair of matrices $(A, B)$, there exists a matrix $C$ such that $(C, A)$ is an observable pair and $C(zI - A)^{-1}B$ is symmetric.

**Proof.** Let $HD^{-1}$ be a right coprime factorization of the i/s (input-state) transfer function $(zI - A)^{-1}B$. Since $D$ and $\tilde{D}$ have the same Smith canonical form, there exists a unimodular matrix $U$ such that $UD = \tilde{D}\tilde{U}$. Let the
polynomial matrix $N$ be defined by

$$N = \left( \pi_-(UD^{-1}) \right) D. \quad (10.5)$$

Then $U = N + MD$, and $N$ and $D$ are right coprime. In fact

$$I = U^{-1}U = U^{-1}N + U^{-1}MD = XN + YD$$

shows the right coprimeness. Also it follows from the definition of $N$ that $ND^{-1}$ is strictly proper. We can apply now Theorem 2.3 to infer the existence of a constant matrix $C$ for which $N = CH$. This in turn implies

$$G(z) = N(z)D(z)^{-1} = CH(z)D(z)^{-1} = C(zI - A)^{-1}B.$$

The observability of the pair $(C, A)$ is equivalent, by Theorem 2.2, to the right coprimeness of $N$ and $D$. Now the equality $UD = \hat{D}\hat{U}$ implies $ND = \hat{D}\hat{N}$ and so

$$\hat{N}D^{-1} = \hat{D}^{-1}N$$

which shows that $G$ is symmetric. Moreover the map $Z: X_D \rightarrow X_{\hat{D}}$ given by

$$Zf = \pi_{\hat{D}}Nf \quad \text{for} \quad f \in X_D \quad (10.6)$$

is a self-adjoint map. 

As a special case consider $B = I$. This certainly makes the pair $(A, B)$ reachable. By the previous theorem there exists a $C$ such that

$$G(z) = C(zI - A)^{-1} = (zI - \hat{A})^{-1}\hat{C} = \hat{C}(z).$$

In this case $V - X_{zI - A}$ and so

$$Zx = \pi_{zI - A}Cx = Cx$$

is invertible, i.e., $C$ is invertible.

Now $\hat{C}(zI - A) = (zI - \hat{A})C$ implies

$$C = \hat{C} \quad \text{and} \quad \hat{C}A = \hat{A}C \quad (10.7)$$

and Frobenius' theorem follows.
Since in the case of the complex field the duality pairing is defined differently, the previous theorems have to be modified. In this case we define

\[ G(z) = G(z)^*, \]  

(10.8)

where \( M^* \) denotes the Hermitian adjoint of \( M \).

**Theorem 10.4.** Given a reachable pair of matrices \((A, B)\) over the complex field, then there exists a matrix \( C \) such that \((C, A)\) is observable and \( G(z) = C(zI - A)^{-1}B \) satisfies (10.8) if and only if \( A \) and \( A^* \) are similar.

**Proof.** Assume such a \( C \) exists. Then by (10.8)

\[ C(zI - A)^{-1}B = B^*(zI - A^*)^{-1}C^*. \]  

(10.9)

By the state space isomorphism theorem, the systems \((A, B, C)\) and \((A^*, C^*, B^*)\) are similar and in particular so are \( A \) and \( A^* \).

Conversely, assume \( A \) and \( A^* \) are similar. Again let \( HD^{-1} \) be a right coprime factorization of \((zI - A)^{-1}B\). Since \( D \) and \( D(z) = D(\bar{z})^* \) have the same invariant factors, i.e. those of \( A \) or \( A^* \), there exists a unimodular matrix \( U \) for which \( UD = D\bar{U} \). The rest of the proof follows that of Theorem 10.3.

**Corollary 10.5.** A complex matrix \( A \) is the product of two Hermitian matrices of which one is nonsingular if and only if \( A \) and \( A^* \) are similar.

**Proof.** Assume

\[ A = ST, \]  

(10.10)

\( S \) and \( T \) Hermitian, and \( S \) nonsingular. From (10.10) it follows that \( A^* = TS \), and so

\[ AS = STS = SA^*, \]  

(10.11)

i.e., \( A \) and \( A^* \) are similar.

Conversely, we apply Theorem 10.4 to infer the existence of a complex matrix \( C \) for which \( C(zI - A)^{-1} = (zI - A^*)^{-1}C^* \), or equivalently

\[ C^*(zI - A) = (zI - A^*)C. \]  

(10.12)
In particular $C = C^*$, and $D = C^*A$ is Hermitian too. The invertibility of $C$ is proved as in Theorem 10.3, and so we get

$$A = C^{-1}D.$$  

(10.13)

11. SELF-ADJOINT OPERATORS IN INDEFINITE METRIC SPACES

In this section we present the polynomial model approach to the study of self-adjoint operators in indefinite metric spaces, and especially their reduction to canonical form under the group of orthogonal matrices in this metric. These are very closely related to the study of symmetric polynomial matrices. We will treat the real case and in the end indicate what modifications have to be made for this approach to work also for the Hermitian case.

Given the real $n$-dimensional Euclidean space $\mathbb{R}^n$ with the usual inner product $(x, y)$, then, given a symmetric nonsingular matrix $H$, we define a new, indefinite inner product on $\mathbb{R}^n$ by

$$[x, y] = (Hx, y).$$  

(11.1)

A linear transformation $A$ in $\mathbb{R}^n$ is a self-adjoint with respect to the indefinite metric if

$$[Ax, y] = [x, Ay]$$  

(11.2)

for all $x, y \in \mathbb{R}^n$. In terms of $H$ this clearly reduces to

$$HA = \tilde{A}H.$$  

(11.3)

Letting $G = HA$, we have $\tilde{G} = G$ and

$$A = H^{-1}G.$$  

(11.4)

Thus $A$ is the product of two symmetric matrices of which one is nonsingular. This shows the connection with Frobenius' theorem.

Given any real $n \times n$ matrix $A$, then by Frobenius' theorem $A = H^{-1}G$ with $H, G$ symmetric. This is equivalent to saying that $A$ is self-adjoint in the metric induced by $H$. 
Next we relate this to the study of symmetric polynomial matrices. Consider a real symmetric nonsingular polynomial matrix \( D(z) = D_0 + D_1 z + \cdots + D_s z^s \). We saw in Fuhrmann [27] that, due to the symmetry of \( D \), the shift map \( S_D \) acting in \( X_D \) is self-adjoint, namely, for all \( f, g \in X_D \)
\[
\langle S_D f, g \rangle = \langle f, S_D g \rangle. \tag{11.5}
\]

Thus \( S_D \) is self-adjoint in the indefinite metric of \( X_D \) induced by \( D \).

To make contact with some classical results we consider the special case of \( D_s = I \), i.e. \( D(z) \) being a monic polynomial matrix. In this case \( X_D \) coincides with the set of all vector polynomials of degree \( \leq s - 1 \) and
\[
X_D = X_0 \oplus \cdots \oplus X_{s-1}, \tag{11.6}
\]
where \( X_i \) is the subspace of all vector polynomials of the form \( x z^i \). Relative to this direct sum decomposition we have the block matrix representation \( C \) of \( S_D \) of the form
\[
\begin{pmatrix}
0 & \cdots & \cdots & -D_0 \\
I & \ddots & \ddots & \vdots \\
& \ddots & \ddots & \ddots \\
& & I & -D_{s-1}
\end{pmatrix} \tag{11.7}
\]

The dual direct sum decomposition of \( X_D \) is given by
\[
X_D = Y_0 \oplus \cdots \oplus Y_{s-1}, \tag{11.8}
\]
with
\[
Y_i = \{ E_{i-1}(z) x | x \in \mathbb{R}^m \},
\]
where
\[
E_i = \pi_+ z^{-i} D, \quad i = 1, \ldots, s. \tag{11.9}
\]

For details see Fuhrmann [26]. The block matrix representation of \( S_D \) relative to this direct sum decomposition is
\[
\begin{pmatrix}
0 & I \\
& \ddots & \ddots & \vdots \\
& & I \\
-D_0 & \cdots & \cdots & -D_{s-1}
\end{pmatrix}, \tag{11.10}
\]
i.e., it is $\tilde{C}$. Since the identity map in $X_D$ relative to these direct sum decompositions has the block matrix representation

$$H = \begin{pmatrix} D_1 & \cdot & \cdot & \cdot & D_{s-1} & I \\ \cdot & \cdot & \cdot & \cdot & I \\ \cdot & \cdot & \cdot & \cdot \\ D_{s-1} & \cdot & \cdot & \cdot & 0 \\ I \end{pmatrix},$$

we have

$$H\tilde{A} = \tilde{A}H,$$

i.e., $H$ is a symmetric matrix for $A$. That the companion matrix $C$ is self-adjoint in the metric induced by $H$ has been observed before by Langer [51] and by Gohberg, Lancaster, and Rodman [36].

Suppose now that instead of starting with a symmetric polynomial matrix we start with an $H$-self-adjoint matrix $A$, with $H$ naturally assumed symmetric and nonsingular. Thus $HA = \tilde{A}H$, and equivalently

$$AH^{-1} = H^{-1}\tilde{A} = (\tilde{A}H^{-1}),$$

and clearly

$$D(z) = (zI - A)H^{-1} = zH^{-1} - AH^{-1}$$

is a symmetric polynomial matrix. Clearly $X_D$ coincides with $\mathbb{R}^n$ in this case, and

$$\langle x, y \rangle = [D^{-1}x, y]$$

$$- \left[(((zI - A)H^{-1})^{-1}x, y\right] - \left[H(zI - A)^{-1}x, y\right] = (Hx, y),$$

i.e., the metric of $X_D$ coincides with the $H$-metric of $\mathbb{R}^n$.

Under a change of basis $x \rightarrow Rx$ the matrix $A$ transforms by similarity into $A_1 = R^{-1}AR$, whereas the metric given by $H$ transforms by congruence into $H_1 = RHBR$. Clearly, as

$$H_1A_1 = (RHBR)(R^{-1}AR) = RHAR$$

$$= \tilde{R}\tilde{A}HR = \tilde{R}\tilde{A}\tilde{R}^{-1}\tilde{R}HR = \tilde{A}_1H_1,$$
A_1 is self-adjoint in the \( H_1 \)-metric. This raises the question of reducing a pair \( A, H \) with \( A \) being \( H \)-selfadjoint to its simplest form in a canonical way. That this is possible and an easy consequence of realization theory will be demonstrated next.

**Theorem 11.1.** Let \( H \) be a nonsingular symmetric \( n \times n \) matrix, and let \( A \) be a real \( n \times n \) matrix for which (11.3) holds. Then there exists a nonsingular matrix \( R \) such that

\[
R^{-1}AR = \text{diag}(J_1, \ldots, J_s) \quad (11.15)
\]

and

\[
\bar{R}HR = \text{diag}(\epsilon_1 \Sigma_1, \ldots, \epsilon_s \Sigma_s), \quad (11.16)
\]

where \((J_1, \ldots, J_s)\) is the real Jordan canonical form of \( A \), \( \Sigma_i \) has the form

\[
\Sigma_i = \begin{pmatrix}
0 & \cdots & \cdots & \cdots & 0 & 1 \\
\vdots & & & & \vdots & 1 \\
\vdots & & & & \vdots & \\
0 & 1 & & & 0 \\
1 & & & &
\end{pmatrix} \quad (11.17)
\]

and \( \epsilon_i = \pm 1 \). The matrices \( J_i, \Sigma_i \) and the signs \( \epsilon_i \) are uniquely determined up to order.

**Proof.** Define \( D \) by (11.14); then \( D \) is a nonsingular symmetric polynomial matrix. Note that in this case \( X_D \) coincides with \( \mathbb{R}^n \), and since, for \( x \in \mathbb{R}^n \),

\[
S_D x = \pi_D xx = (zI - A)H^{-1} \pi_- H(zI - A)^{-1} xx
\]

\[
= (zI - A)H^{-1} \pi_- H(zI - A)^{-1}(zI - A + A)x
\]

\[
= (zI - A)H^{-1} H(zI - A)^{-1} Ax = Ax,
\]

we have

\[
S_D = A. \quad (11.18)
\]
We consider now $\pi_+ D^{-1}$, which is a symmetric transfer function of McMillan degree $n$. We apply Theorem 9.12 to infer the existence of a signature symmetric realization $(A, B, C)$ in $X_D = \mathbb{R}^n$ with $A = S_D$ such that (11.15) and (11.16) hold.

We end this paper by showing the equivalence of the study of self-adjoint operators in indefinite metric spaces with the classical problem, solved by Kronecker [48] and Weierstrass [58], of reducing simultaneously by congruence a pair of real symmetric matrices, one being nonsingular, to canonical form. For recent accounts which contain historical remarks and further references one may consult Uhlig [56, 57].

Given two real symmetric matrices $S$ and $T$, the problem is to find a canonical form for the pair $(S, T)$ under the group action

$$(S, T) \rightarrow (\tilde{R}SR, \tilde{R}TR),$$

$R$ nonsingular. We make further assumption that $S$ is nonsingular. In this case we define

$$A = S^{-1}T,$$  \hspace{1cm} (11.20)

which implies, by the symmetry of $T$, that

$$SA = \tilde{A}S.$$  \hspace{1cm} (11.21)

Of course (11.21) means that $A$ is $S$-self-adjoint and hence Theorem 11.1 can be applied.

Thus in a suitable basis $S$ has the matrix representation

$$\tilde{R}SR = \text{diag}(\epsilon_1 \Sigma_1, \ldots, \epsilon_s \Sigma_s),$$

and $A$ has the matrix representation

$$R^{-1}AR = \text{diag}(J_1, \ldots, J_s).$$

This implies that

$$\tilde{R}TR = \tilde{R}SAR = (\tilde{R}SR)(R^{-1}AR) = \text{diag}(\epsilon_1 \Sigma_1 J_1, \ldots, \epsilon_s \Sigma_s J_s).$$

Therefore we have rederived the following theorem.
THEOREM 11.2. Let $S,T$ be real symmetric matrices of which $S$ is assumed nonsingular. Let $\text{diag}(J_1, \ldots, J_s)$ be the Jordan canonical form of $A = S^{-1}T$. Then $S, T$ are simultaneously congruent to
\[
\text{diag}(\epsilon_1\Sigma_1, \ldots, \epsilon_s\Sigma_s)
\]
and
\[
\text{diag}(\epsilon_1\Sigma_1J_1, \ldots, \epsilon_s\Sigma_sJ_s)
\]
respectively.

Much of the work on this paper has been done at the Department of Mathematics of Rutgers University. For its support and the hospitality of my host Hector J. Sussmann I am deeply grateful.

REFERENCES

47 P. P. Khargonekar and E. Emre, Further results on polynomial characterization of \((F, C)\)-invariant subspaces, to appear.


Received 18 February 1982; revised 25 October 1982