A Polynomial Approach to Hankel Norm and Balanced Approximations

P. A. Fuhrmann*

Department of Mathematics
Ben-Gurion University of the Negev
Beer Sheva, Israel

Dedicated to Roger W. Brockett, in remembrance of an opportunity missed.

Submitted by Hans Schneider

ABSTRACT

A new unified approach to problems of Hankel norm and balanced approximations is presented which is based on a combination of polynomial algebra and the geometry of invariant subspaces. Contrary to state space methods, where contact with external properties of systems is indirect, the approach presented yields new insights into basic properties of Hankel norm approximation and balanced realizations. Several approximation results are interpreted geometrically in terms of projections. Also duality results in this area are pointed out.

1. INTRODUCTION

In a masterful, pathbreaking series of papers, Adamjan, Arov, and Krein (1968a, 1968b, 1971, 1978) developed in great generality the theory of Hankel norm approximations. This theory, commonly referred to as AAK theory, contained the theory of optimal and suboptimal extensions of Hankel operators. Because of the close connection between Hankel operators and compressions of shifts, it also provided an alternative approach not only to Nehari’s theorem but also to Sarason’s theorem (Sarason, 1968), and its generalization to the commutant lifting theorem (Sz.-Nagy and Foias, 1970).

In Glover (1984), a greatly influential paper, the AAK theory is developed, for the rational case, in a self-contained way, using mostly linear algebraic tools. In the process Glover obtained a parametrization of all solutions and,

* The author holds the Earl Katz Family Chair in Algebraic System Theory.

LINEAR ALGEBRA AND ITS APPLICATIONS 146:133-220 (1991) 133

655 Avenue of the Americas, New York, NY 10010 0024-3795/91/$3.50
what is more important, also $L^\infty$ error bounds resulting from the truncation of balanced realizations.

Our object in this paper is to present a different approach to parts of this theory. We will also deal with the rational case, but will restrict ourselves at this stage to the scalar case. The approach we take is algebraic and coordinate free, and it reduces all computations to polynomial equations. However, the polynomial equations can be interpreted on at least three different levels. On one level we have polynomial equations, which on applying the theory of polynomial models lead to certain operator equations in finite dimensional spaces. The other level is the level of rational functions. On this level we get simple solutions to the Nehari extensions as well as to the solution of an important $H^\infty$-Bezout equation. The third level is geometric, i.e. the level of invariant subspaces of $H^2$ spaces and operator equations in these spaces. Some of Glover's results are not only rederived, but extended, inasmuch as also the Schmidt pairs are explicitly calculated and a geometric interpretation in terms of invariant subspaces and projections is given.

In the process of proving some of the results we highlight some intricate symmetries in the study of Hankel operators.

The paper is structured as follows. In Section 2 we collect some basic information about polynomial and rational models. In particular we outline the polynomial model approach to realization theory. This material will be used in the rest of the paper. It will be instrumental in the construction of balanced realizations.

In Section 3 we study Hankel operators with rational symbols, their kernel and image spaces, and their relation to coprime factorizations. We study the singular value and singular vector equation and derive them in polynomial form. This fundamental polynomial equation is not new. It appears already in the work of Kung (1980) and Harshavardhana, Jonckheere, and Silverman (1984). However, it does not seem to have been realized before how much information could be squeezed out of it. In the process a very simple proof of Nehari's theorem in the rational case is obtained. In this section we also give brief discussions of material related to the computation of singular values, Schmidt pairs, and best approximants. These topics include the determination of signed eigenvalues of a related self-adjoint Hankel operator, Nevanlinna-Pick interpolation, and some connection with geometric control theory.

Section 4 is devoted to the analysis of inversion of Hankel operators, or rather the restriction of Hankel operators to the map from the cokernel to the image. This is related to a spectral mapping theorem derived in Fuhrmann (1968a,b). In the rational context in which we work, the result is recovered from the fundamental polynomial equation (FPE) and the Schmidt pairs associated with the smallest singular value.

In Section 5 we recover a result of Glover concerning singular values of
the optimal Hankel norm approximant corresponding to the last singular value. Certain polynomial identities derived in Sections 4 and 5 are generalized. This is done directly, through polynomial algebra, starting from the FPE. These polynomial relations can be transformed into relations between rational functions. In turn these have an interpretation as orthogonality relations in $L^2$ or $H^2$ spaces (Section 6). On this level the geometry of the situation is highlighted.

Section 7 is concerned with duality in Hankel norm approximation problems. Next, in Sections 8 and 9, we pass to an analysis of balanced realizations, introduced by Moore (1981). This analysis is based on polynomial model realization theory and utilizes the information and insights about Schmidt pairs of Hankel operators that has been obtained in previous sections. We do this for two extreme cases. The first case, treated in Section 8, is the generic case, namely the case of Hankel operators with distinct singular values. The second case, treated in Section 9, is that of Hankel operators with identical singular values. This is equivalent to the analysis of realization of antistable all-pass functions. This we do through the use of continued fractions and the realizations associated to them. While this is not done in the present paper, the content of these two sections can be put together to analyse the general case. This is related to constructing a global basis made up basically of sets of orthogonal polynomials.

Finally, in Section 10, all this is put together to obtain error bounds for model reduction through the truncation of balanced realizations.

2. POLYNOMIAL MODELS

Polynomial and rational models provide the main tool in the whole paper. We proceed to give the basic definitions. Necessarily the exposition is brief, and it is suggested that the interested reader consult such other papers as Fuhrmann (1976, 1977, 1981a,b, 1983, 1984) and Helmke and Fuhrmann (1989).

Throughout the paper we will denote by $F$ an arbitrary commutative field. It might be identified later with the real number field $R$. By $F[z]$ we denote the ring of polynomials over $F$; by $F((z^{-1}))$, the set of truncated Laurent series in $z^{-1}$, i.e. the set of all formal series of the form $\sum_{n \in \mathbb{N}} a_n z^{-n}$, $n \in \mathbb{Z}$. $F((z^{-1}))$ is a vector space over $F$ as well as a field. It contains the field of $F(z)$ of rational functions as a subfield. By $F[[z^{-1}]]$ and $z^{-1}F[[z^{-1}]]$ we denote 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. Also, $z^{-1}F[[z^{-1}]]$ is isomorphic to $F((z^{-1}))/F[z]$, which is an $F[z]$-module with the module action given by

$$z \cdot h = S_{-} h - \pi_{-} z h.$$  \hfill (1)

Similarly we define

$$S_{+} f = zf \quad \text{for} \quad f \in F[z].$$  \hfill (2)

Given a monic polynomial $q$ of degree $n$, we define a projection $\pi_{q}$ in $F[z]$ by

$$\pi_{q} f = q \pi_{-} q^{-1} f \quad \text{for} \quad f \in F[z].$$  \hfill (3)

We define the polynomial model associated with $q$ to be the space

$$X_{q} = \text{Im} \pi_{q}.$$  \hfill (4)

endowed with the module structure, over the ring $F[z]$, induced by the shift map defined through

$$z \cdot f = S_{q} f = \pi_{q} S_{+} f \quad \text{for} \quad f \in X_{q}.$$  \hfill (5)

To get an understanding of the projection $\pi_{q}$, note that it simply takes the remainder of $f$ modulo the polynomial $q$. In fact, given any $f$ in $F[z]$, we can write $f = aq + r$, and this representation is unique if $\text{deg} \ r < \text{deg} \ q$. To isolate the remainder $r$ we proceed by dividing the previous equality by $q$, i.e., $fq^{-1} = a + rq^{-1}$. Now $a$ is in $F[z]$, whereas $rq^{-1}$, since $\text{deg} \ r < \text{deg} \ q$, is in $z^{-1}F[[z^{-1}]]$, so applying the projection $\pi_{-}$, we have $\pi_{-} q^{-1} f = rq^{-1}$ and $r = q \pi_{-} q^{-1} f$. The advantage of this circuitous route to the remainder $r$ is in the ease with which it generalizes to the multivariable case.

Analogously we define the rational model to be the space

$$X^{q} = \text{Im} \pi^{q}.$$  \hfill (6)

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

$$\pi^{q} h = \pi_{-} q^{-1} \pi_{+} q h \quad \text{for} \quad h \in z^{-1}F[[z^{-1}]].$$  \hfill (7)

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

$$S^{q} h = S_{-} h \quad \text{for} \quad h \in X^{q}.$$  \hfill (8)
We illustrate the last definition through a simple example. Let \( q(z) = (z - \alpha)(z - \beta) \) and let \( h(z) = 1/z \). Then

\[
\pi^q h = \pi_- (z - \alpha)^{-1}(z - \beta)^{-1} \pi_+ \left( \frac{(z - \alpha)(z - \beta)}{z} \right)
\]

\[
= \pi_- (z - \alpha)^{-1}(z - \beta)^{-1} \left[ z - (\alpha + \beta) \right]
\]

\[
= \frac{\alpha}{\alpha - \beta} \frac{1}{z - \alpha} - \frac{\beta}{\alpha - \beta} \frac{1}{z - \beta}.
\]

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

\[
\rho_q h = q h \quad \text{for} \quad h \in X^q,
\]

i.e., we have \( \rho_q S^q = S_q \rho_q \).

A map \( Z \) in \( X_q \) commutes with \( S_q \) if and only if \( Z = p(S_q) \) for some polynomial \( p \in \mathbb{F}[z] \) and \( p(S_q) \) is invertible if and only if \( p \) and \( q \) are coprime.

We define a pairing of elements of \( \mathbb{F}((z^{-1})) \) as follows: for

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

and

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

let

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

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

\[
M^\perp = \left\{ g \in \mathbb{F}((z^{-1})) \mid [f, g] = 0 \text{ for all } f \in M \right\},
\]
then in particular $F[z]^\perp = F[z]$ and $(qF[z])^\perp = X^q$. Since in general $(X/M)^* = M^{\perp}$, we have

$$X^*_q = (F[z]/qF[z])^* \approx [qF[z]]^\perp = X^q. \quad (12)$$

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

$$\langle f, g \rangle = [q^{-1}f, g]. \quad (13)$$

Relative to this bilinear form we have the important relation

$$S^*_q = S_q, \quad (14)$$

so that $S_q$ is self-adjoint.

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

$$\langle e_i, f_j \rangle = \delta_{ij}, \quad 1 \leq i, j \leq n. \quad (15)$$

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

$$f_i(z) = z^{i-1}, \quad i = 1, \ldots, n, \quad (16)$$

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

Given the polynomial $q$ as above, we define

$$e_i(z) = \pi_+ z^{i-1}q = q_i + q_{i+1}z + \cdots + z^n, \quad i = 1, \ldots, n, \quad (17)$$

and call the set $B_{co} = \{e_1, \ldots, e_n\}$ the control basis of $X_q$.

The important fact about this pair of bases is that relative to the bilinear form $\langle , \rangle$ of Equation (13) the standard and control bases are dual to each other. In particular, since $S_q^* = S_q$, we have $p(S_q)^* = p(S_q^*) = p(S_q)$, and so $p(S_q)$ is a self-adjoint operator in the indefinite metric $\langle , \rangle$. Thus the matrix representation of $p(S_q)$ relative to any dual pair of bases is symmetric.
The following theorem summarizes the most important properties of linear maps that commute with $S_q$.

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

(i) The map $S_q$ is cyclic.

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

$$Z S_q = S_q Z.$$  \hspace{1cm} (18)

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

$$Z = p(S_q).$$ \hspace{1cm} (19)

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

$$\text{Ker} \, p(S_q) = p_1 X_r$$ \hspace{1cm} (20)

and

$$\text{Im} \, p(S_q) = r X_{q_1}.$$ \hspace{1cm} (21)

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

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

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

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

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

$$p(S_q) a(S_q) + B(S_q) q(S_q) = I,$$ \hspace{1cm} (23)

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

$$p(S_q) a(S_q) = I.$$ \hspace{1cm} (24)
Note that the polynomials $a$ and $b$ solving Equation (22) can be found using the Euclidean algorithm.

We note, for later use, that

$$\tilde{C}_q = \left[ S_q \right]_{st} = \begin{pmatrix} 0 & \cdots & 0 & -q_0 \\ 1 & \cdots & \cdots & \cdots \\ \vdots & \cdots & \cdots & \cdots \\ 1 & -q_{n-1} \end{pmatrix} \quad (25)$$

and

$$C_q = \left[ S_q \right]_{co} = \begin{pmatrix} 0 & 1 & \cdots & \cdots & \cdots & \cdots \\ \vdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & -q_0 & -q_1 & \cdots & \cdots & -q_{n-1} \end{pmatrix}, \quad (26)$$

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

Given a proper rational function $\phi = n/d$, the associated realization is constructed as follows. We choose $X_d$ as the state space and define $(A, B, C, D)$ through

$$A = S_d,$$

$$B \xi = n \xi \quad \text{for} \quad \xi \in \mathbb{R},$$

$$C f = \left( d^{-1} f \right)_{-1} \quad \text{for} \quad f \in X_d,$$

$$D = \pi_+ \phi. \quad (27)$$

The realization of $\phi$ is minimal, by the coprimeness of $n$ and $d$.

We can rewrite this realization using the rational model rather than the polynomial one. Thus the state space is chosen as $X^d$, and $(A, B, C, D)$ is defined through

$$A = S^d,$$

$$B \xi = \frac{n}{d} \xi, \quad \text{for} \quad \xi \in \mathbb{R},$$

$$C f = \left( f \right)_{-1} = (zf)(\infty) \quad \text{for} \quad f \in X^d,$$

$$D = \phi(\infty). \quad (28)$$

It is this realization that we use as a basis for obtaining a balanced realization.

We will find the following representation of use.
**Lemma 2.1.** If $d$ is a nonsingular polynomial matrix, then there exists an observable pair, unique up to isomorphism, such that

$$X_d = \left\{ d(z)C(zI - A)^{-1} \xi \mid \xi \in F^n \right\}.$$ 

3. **Hankel Operators**

We will study Hankel operators defined on half plane Hardy spaces, rather than on those of the unit disc as was done by Adamjan, Arov, and Krein (1971). In this we follow the choice of Glover (1984). It seems to be a very convenient one in that all results on duality simplify significantly, due to the greater symmetry between the two half planes than between the unit disc and its exterior.

The setting is the Hardy spaces. Thus $H^2_+$ is the Hilbert space of all analytic functions in the open right half plane with

$$\|f\|^2 = \sup_{x > 0} \frac{1}{\pi} \int_{-\infty}^{\infty} |f(x + iy)|^2 \, dy.$$ 

The space $H^2_-$ is similarly defined in the open left half plane. A theorem of Fatou guarantees the existence of boundary values of $H^2_+$-functions on the imaginary axis. Thus $H^2_\pm$ can be considered as closed subspaces of $L^2(i\mathbb{R})$, the space of Lebesgue square integrable functions on the imaginary axis. It follows from the Fourier-Plancherel and Paley-Wiener theorems that

$$L^2(i\mathbb{R}) = H^2_+ \oplus H^2_-,$$

with $H^2_+$ and $H^2_-$ the Fourier-Plancherel transforms of $L^2(0, \infty)$ and $L^2(-\infty, 0)$ respectively. Also $H^\infty_+$ and $H^\infty_-$ will denote the spaces of bounded analytic functions on the open right and left half planes respectively. We will define $f^*(s) = f(-s)^*$. These spaces can be considered as subspaces of $L^\infty(i\mathbb{R})$, the space of Lebesgue measurable and essentially bounded functions on the imaginary axis. An extensive discussion of these spaces can be found in Hoffman (1962), Duren (1970), and Garnett (1981).

We proceed to define Hankel operators, and we do this directly in the frequency domain. Readers interested in the time domain definition and the details of the transformation into the frequency domain are referred to Fuhrmann (1981) and Glover (1984).

In the algebraic theory of Hankel operators the kernel and image of a
Hankel operator are directly related to the coprime factorization of the symbol over the ring of polynomials. The details can be found for example in Fuhrmann (1983). In the same way the kernel and image of a large class of Hankel operators are related to a coprime factorization over $H^\infty$. This theme, originating in the work of Douglas, Shapiro, and Shields (1971) and that of D. N. Clark (see Helton, 1974), is developed extensively in Fuhrmann (1981). Of course, if the symbol of the Hankel operator is rational and in $H^\infty$, these two coprime factorizations are easily related.

Thus assume $\phi = n/d \in H^\infty$ and $n \land d = 1$. So our assumption is that $d$ is antistable. In spite of the slight ambiguity, we will write $n = \deg d$. It will always be clear from the context what $n$ means. This leads to

$$\phi = \frac{n}{d} = \frac{n}{d^*} \frac{d^*}{d}.$$ 

Thus

$$\phi = m^* \eta$$

with

$$\eta = \frac{n}{d^*}, \quad m = \frac{d}{d^*}$$

is a coprime factorization in $H^\infty$.

We proceed to define Hankel operators.

**Definition 3.1.** Given a function $\phi \in L^\infty(i\mathbb{R})$, the Hankel operator $H_\phi : H^2_+ \to H^2_+$ is defined by

$$H_\phi f = P_-(\phi f) \quad \text{for} \quad f \in H^2_+. \quad (29)$$

The adjoint operator $(H_\phi)^* : H^2_+ \to H^2_+$ is given by

$$(H_\phi)^* f = P_+(\phi^* f) \quad \text{for} \quad f \in H^2_+. \quad (30)$$

Here $\phi^*(z) = \overline{\phi(-\overline{z})}$.

**Theorem 3.1.**

1. For every $\psi \in H^\infty_+$ the Hankel operator $H_\phi$ satisfies the functional equation

$$P_- \psi H_\phi f = H_\phi \psi f, \quad f \in H^2_+. \quad (31)$$
2. Ker $H_\phi$ is an invariant subspace, i.e., for $f \in \text{Ker } H_\phi$ and $\psi \in H_+^\infty$ we have $\psi f \in \text{Ker } H_\phi$.

It follows from a theorem of Beurling (1949) that Ker $H_\phi = mH_+^2$ for some inner function $m \in H_+^\infty$. Since we are dealing with the rational case, the next theorem can make this more specific; it characterizes the kernel and image of a Hankel operator and also clarifies the connection between them and polynomial and rational models. A closely related derivation can be found in Young (1983) and Lindquist and Picci (1985).

**Theorem 3.2.** Let $Q = n/d \in H_+^\infty$ and $n \land d = 1$. Then

1. $\text{Ker } H_\phi = \frac{d}{d^*} H_+^2$.
2. $\{\text{Ker } H_\phi\}^\perp = \left\{ \frac{d}{d^*} H_+^2 \right\}^\perp = X^{d^*}$.
3. $\text{Im } H_\phi = H_-^2 \oplus \frac{d}{d^*} H_+^2 = X^d$.

**Proof.** $\{\text{Ker } H_\phi\}^\perp$ contains only rational functions. Let $f = p/q \in \{ (d/d^*) H_+^2 \}^\perp$; then $d^*p/dq \in H_-^2$. So $q \mid d^*p$. But, as $p \land q = 1$, it follows that $q \mid d^*$, i.e. $d^* = qr$. Hence $f = rp/d^* \in X^{d^*}$.

Conversely, let $p/d^* \in X^{d^*}$. Then

$$\frac{p}{d^*} = \frac{p}{d} \frac{d}{d^*}, \quad \text{or } \frac{d^*}{d} \frac{p}{d^*} \in H_-^2.$$ 

So we have $p/d^* \in \{ (d/d^*) H_+^2 \}^\perp$. 

The previous theorem, though elementary, is central to all further development, as it provides the direct link between an infinite dimensional object, namely the Hankel operator, and the well developed theory of polynomial and rational models. This link will be continually exploited.

It is quite well known (see Gohberg and Krein, 1969) that singular values of operators are closely related to the problem of best approximation by operators of finite rank. That this basic method could be applied to the approximation of Hankel operators by Hankel operators of lower ranks through the detailed analysis of singular values and the corresponding Schmidt pairs is a fundamental contribution of Adamjan, Arov, and Krein.

We recall that, given a bounded operator $A$ on a Hilbert space, $\sigma$ is a **singular value of** $A$ if there exists a nonzero vector $f$ such that

$$A^* Af = \sigma^2 f.$$
Rather than solve the previous equation, we let \( g = (1/\sigma) A f \) and go over to the equivalent system

\[
A f = \sigma g, \\
A^* g = \sigma f,
\]

i.e., \( \sigma \) is a singular value of both \( A \) and \( A^* \).

The analysis of Schmidt pairs of Hankel operators goes back to Adamjan, Arov, and Krein (1971). Here, for the rational case, we present an algebraic derivation of some of their results.

We proceed to compute the singular vectors of the Hankel operator \( H_\phi \). In view of the preceding remarks, we have to solve

\[
H_\phi f = \sigma g, \\
H_\phi^* g = \sigma f,
\]

or

\[
P_+ \frac{n^p}{d} \frac{p}{d^*} = \sigma \frac{\hat{p}}{d}, \\
P_- \frac{n^*}{d^*} \frac{\hat{p}}{d} = \sigma \frac{p}{d^*}.
\]

This means there exist polynomials \( \pi \) and \( \xi \) such that

\[
\frac{n^p}{d} \frac{p}{d^*} = \sigma \frac{\hat{p}}{d} + \frac{\pi}{d^*}, \\
\frac{n^*}{d^*} \frac{\hat{p}}{d} = \sigma \frac{p}{d^*} + \frac{\xi}{d}.
\]

These equations can be rewritten as polynomial equations

\[
n p = \sigma d^* \hat{p} + d \pi, \quad (32) \\
n^* \hat{p} = \sigma dp + d^* \xi. \quad (33)
\]

**Remark 3.1.** Equation (32), considered as an equation modulo the polynomial \( d \), is not an eigenvalue equation, as there are too many unknowns.
More specifically, we have to find the coefficients of both $p$ and $\hat{p}$. To overcome this difficulty we study in more detail the structure of Schmidt pairs of Hankel operators.

**Lemma 3.1.** Let $\{p/d^*, \hat{p}/d\}$ and $\{q/d^*, \hat{q}/d\}$ be two Schmidt pairs of the Hankel operator $H_{n/d}$, corresponding to the same singular value $\sigma$. Then

\[
\frac{p}{\hat{p}} = \frac{q}{\hat{q}},
\]

i.e., this ratio is independent of the Schmidt pair.

**Proof.** The polynomials $p, \hat{p}$ correspond to one Schmidt pair; let the polynomials $q, \hat{q}$ correspond to another Schmidt pair, i.e.

\[
nq = \sigma d^*\hat{q} + d\rho,
\]

\[
n^*\hat{q} = \sigma dq + d^*\eta.
\]

Now, from Equations (32) and (35) we get

\[
0 = \sigma d(\hat{p}q - q\hat{p}) + d^*(\xi\hat{q} - \eta\hat{p}).
\]

Since $d$ and $d^*$ are coprime, it follows that $d^* \mid \hat{p}q - q\hat{p}$. On the other hand, from Equations (32) and (34), we get

\[
0 = \sigma d^*(\hat{p}q - \hat{q}p) + d(\pi q - \rho p),
\]

and hence that $d \mid \hat{p}q - \hat{q}p$. Now both $d$ and $d^*$ divide $\hat{p}q - \hat{q}p$, and, as $\deg(\hat{p}q - \hat{q}p) < \deg d + \deg d^*$, it follows that

\[
\hat{p}q - \hat{q}p = 0.
\]

Equivalently,

\[
\frac{p}{\hat{p}} = \frac{q}{\hat{q}},
\]

i.e. $p/\hat{p}$ is independent of the particular Schmidt pair associated with the singular value $\sigma$. 


Lemma 3.2. Let \( \{ p/d^*, \hat{p}/d \} \) be a Schmidt pair associated with the singular value \( \sigma \). Then \( p/\hat{p} \) is unimodular or all pass.

Proof. Going back to Equation (33) and the dual of (32), we have

\[
\begin{align*}
n^* \hat{p} &= \sigma dp + d^* \xi, \\
n^* p^* &= \sigma d(\hat{p})^* + d^* \pi^*.
\end{align*}
\]

It follows that

\[
0 = \sigma d [pp^* - \sigma (\hat{p})^*] + d^*(\xi p^* - \pi^* \hat{p}),
\]

and hence \( d^* \mid pp^* - \sigma (\hat{p})^* \). By symmetry also \( d \mid pp^* - \sigma (\hat{p})^* \), and so necessarily

\[
pp^* - \sigma (\hat{p})^* = 0.
\]

This can be rewritten as

\[
\frac{p}{\hat{p}} = 1,
\]

i.e., \( p/\hat{p} \) is all pass.

We will say that a pair of polynomials \( (p, \hat{p}) \), with \( \deg p, \deg \hat{p} < \deg d \), is a solution pair if there exist polynomials \( \pi \) and \( \xi \) such that Equations (32) and (33) are satisfied.

The next lemma characterizes all solution pairs.

Lemma 3.3. Let \( \sigma \) be a singular value of the Hankel operator \( H_{n/d} \). Then there exists a unique (up to a constant factor) solution pair \( (p, \hat{p}) \) of minimal degree. The set of all solution pairs is given by \( \{(q, \hat{q}) \mid q = pa, \hat{q} = \hat{p}a, \deg a < \deg q - \deg p\} \).

Proof. Clearly, if \( \sigma \) is a singular value of the Hankel operator, then a nonzero solution pair \( (p, \hat{p}) \) of minimal degree exists. Let \( (q, \hat{q}) \) be any other solution pair with \( \deg q, \deg \hat{q} < \deg d \). By the division rule for polynomials, \( q = ap + r \) with \( \deg r < \deg p \). Similarly, \( \hat{q} = \hat{a} \hat{p} + \hat{r} \) with \( \deg \hat{r} < \deg \hat{p} \).

From Equation (32) we get

\[
n(ap) = \sigma d^*(a\hat{p}) + d(a \pi),
\]

(36)
whereas Equation (34) yields
\[ n(ap + r) = \sigma d*(\hat{\alpha} + \hat{\beta}) + d(\tau). \]  
(37)

By subtraction we obtain
\[ nr - \sigma d*[(\hat{\alpha} - a)\hat{\beta} + \hat{\gamma}] + d(\tau - a\pi). \]  
(38)

Similarly, from Equation (33) we get
\[ n^*(\hat{\alpha} + \hat{\gamma}) = \sigma d(ap + r) + d^*\xi, \]  
(39)

whereas Equation (33) yields
\[ n^*(a\hat{\beta}) = \sigma d(ap) + d^*(a\xi). \]  
(40)

Subtracting the two gives
\[ n^*[(\hat{\alpha} - a)\hat{\beta} + \hat{\gamma}] = \sigma dr + d^*(\eta - a\xi). \]  
(41)

Equations (38) and (41) imply that \( \{r/d^*, [(\hat{\alpha} - a)\hat{\beta} + \hat{\gamma}]/d\} \) is a \( \sigma \)-Schmidt pair. Since necessarily \( \text{deg } r = \text{deg}[(\hat{\alpha} - a)\hat{\beta} + \hat{\gamma}] \), we get \( \hat{\alpha} = a \). Finally, since we assumed \( (p, \hat{p}) \) to be of minimal degree, we must have \( r = \hat{\gamma} = 0 \).

Conversely, if \( a \) is any polynomial with \( \text{deg } a < \text{deg } d - \text{deg } p \), then from Equations (32) and (33) it follows by multiplication that \( (pa, \hat{p}a) \) is also a solution pair.

**Lemma 3.4.** Let \( p, q \) be coprime polynomials with real coefficients such that \( p/q \) is all pass. Then \( q = \pm p^* \).

**Proof.** Since \( p/q \) is all pass, it follows that
\[ \frac{p}{q} \frac{p^*}{q^*} = 1, \]

or \( pp^* = qq^* \). As \( p \) and \( q \) are coprime, it follows that \( p \mid q^* \) and hence \( q^* = \pm p \).

In the general case we have the following.
Lemma 3.5. Let $p, q$ be polynomials with real coefficients such that $p \wedge q = 1$ and $p/q$ is all pass. Then, with $r = p \wedge \hat{p}$, we have:

$$p = rs,$$

$$\hat{p} = \pm rs^*.$$

Proof. Write $p = rs$, $\hat{p} = r\hat{s}$. Then $s \wedge \hat{s} = 1$, and $s/\hat{s}$ is all pass. The result follows by applying the previous lemma.

The next theorem is of central importance in that it reduces the analysis to one polynomial. Thus we get an equation which is easily reduced to an eigenvalue problem.

Theorem 3.3. Let $\sigma$ be a singular value of $H_\phi$, and let $(p, \hat{p})$ be a nonzero, minimal degree solution pair of Equations (32) and (33). Then $p$ is a solution of

$$np = \lambda d^* p^* + d\pi,$$  \hfill (42)

with $\lambda$ real and $|\lambda| = \sigma$.

Proof. Let $(p, \hat{p})$ be a nonzero, minimal degree solution pair of Equations (32) and (33). By taking their adjoints we can easily see that $(\hat{p}^*, p^*)$ is also a nonzero, minimal degree solution pair. By the uniqueness of such a solution, i.e. by Lemma 3.3, we have

$$\hat{p}^* = \epsilon p.$$  \hfill (43)

Since $\hat{p}/p$ is all pass and both polynomials are real, we have $\epsilon = \pm 1$. Let us put $\lambda = \epsilon \sigma$; then (43) can be rewritten as

$$\hat{p} = \epsilon p^*,$$

and so (42) follows from (32).

Remark 3.2. We will refer to Equation (42) as the fundamental polynomial equation. It will be the source of all future derivations.

Corollary 3.1. Let $\sigma_i$ be a singular value of $H_\phi$, and let $p_i$ be the minimal degree solution of the fundamental polynomial equation, i.e.

$$np_i = \lambda_i d^* p_i^* + d\pi_i.$$
Then:

1. We have

\[ \deg p_i = \deg p_i^* = \deg \pi_i. \]

2. Putting \( p_i(z) = \sum_{j=0}^{n-1} p_{i,j} z^j \) and \( \pi_i(z) = \sum_{j=0}^{n-1} \pi_{i,j} z^j \), we have the equality

\[ \pi_{i,n-1} = \lambda_i p_{i,n-1}. \quad (44) \]

**Corollary 3.2.** Let \( p \) be a minimal degree solution of Equation (42).

Then:

1. The set of all singular vectors of the Hankel operator \( H_{n/d} \) corresponding to the singular value \( \sigma \) is given by

\[ \text{Ker}(H_{n/d}^* - \sigma^2 I) = \left\{ \frac{pa}{d^*} \mid a \in R[z], \deg a < \deg d - \deg p \right\}. \]

2. The multiplicity of \( \sigma = \| H_\phi \| \) as a singular value of \( H_\phi \) is equal to \( m = \deg d - \deg p \), where \( p \) is the minimum degree solution of (42).

3. There exists a constant \( c \) such that \( c + n/d \) is a constant multiple of an antistable all-pass function if and only if \( \sigma_1 = \cdots = \sigma_n \).

**Proof.** We will prove part 3 only. Assume all singular values are equal to \( \sigma \). Thus the multiplicity of \( \sigma \) is \( \deg d \). Hence the minimal degree solution \( p \) of (42) is a constant and so is \( \pi \). Putting \( c = -\pi/p \), then (42) can be rewritten as

\[ \frac{n}{d} + c = \lambda \frac{d^* p^*}{dp}, \]

and this is a multiple of an antistable all-pass function.

Conversely assume, without loss of generality, that \( n/d + c \) is antistable all pass. Then the induced Hankel operator is isometric, and all its singular values are equal to 1.

Part 3 of the corollary is due to Glover (1984).

The fundamental polynomial equation is easily reduced to either a generalized eigenvalue equation or a regular eigenvalue equation. There are several reductions of this kind in the literature—e.g. Kung (1980); Harshavardhana,
Jonckheere, and Silverman (1984). The one proposed here is simple and uses polynomial models.

Starting from (42), we apply the standard functional calculus and the fact that \( d(S_d) = 0 \), i.e. the Cayley-Hamilton theorem, to obtain

\[
n(S_d)p_i = \lambda_i d^*(S_d)p_i^*.
\]

Now \( d, d^* \) are coprime, as \( d \) is antistable and \( d^* \) is stable. Thus, by Theorem 2.1, \( d^*(S_d) \) is invertible. In fact the inverse of \( d^*(S_d) \) is easily computed through the solution of the Bezout equation

\[
a(z)d(z) + b(z)d^*(z) = 1.
\]

with \( \deg a, \deg b < \deg d \). In this case the polynomials \( a \) and \( b \) are uniquely determined, which by virtue of symmetry forces the equality \( a = b^* \). Hence

\[
b^*(z)d(z) + b(z)d^*(z) = 1. \tag{46}
\]

From this we get \( b(S_d)d^*(S_d) = 1 \), or

\[
d^*(S_d)^{-1} = b(S_d).
\]

Because of the symmetry in the Bezout equation (46), we expect that some reduction in the computational complexity should be possible. This indeed turns out to be the case.

Given an arbitrary polynomial \( f \), we let

\[
\begin{align*}
f_+(z^2) &= \frac{f(z) + f^*(z)}{2}, \\
f_-(z^2) &= \frac{f(z) - f^*(z)}{2z}.
\end{align*} \tag{47}
\]

The Bezout equation can be rewritten as

\[
\left[ b_+(z^2) - zb_-(z^2) \right] \left[ d_+(z^2) + zd_-(z^2) \right] \\
+ \left[ b_+(z^2) + zb_-(z^2) \right] \left[ d_+(z^2) - zd_-(z^2) \right] = 1,
\]

or

\[
2\left[ b_+(z^2)d_+(z^2) - z^2b_-(z^2)d_-(z^2) \right] = 1.
\]
We can of course solve the lower degree Bezout equation
\[ 2 \left[ b_+(z) d_+(z) - z b_-(z) d_-(z) \right] = 1. \]
This is possible because, by the assumption that \( d \) is antistable, \( d_+ \) and \( z d_- \) are coprime. Putting \( b(z) = b_+(z^2) + z b_-(z^2) \), we get a solution to the Bezout equation (46).

Going back to Equation (45), we have
\[ n(S_d) b(S_d) p_i = \lambda_i p_i^*. \tag{48} \]
To simplify, we let \( r = \pi_d(b n) = bn \mod d \). Then (48) is equivalent to
\[ r(S_d) p_i = \lambda_i p_i^*. \tag{49} \]
If \( K : X_d \rightarrow X_d \) is given by \( K p = p^* \), then (49) is equivalent to the generalized eigenvalue equation
\[ r(S_d) p_i = \lambda_i K p_i. \]
Since \( K \) is obviously invertible and \( K^{-1} = K \), the last equation transforms into the regular eigenvalue equation
\[ K r(S_d) p_i = \lambda_i p_i. \]
To get a matrix equation one can take the matrix representation with respect to any choice of basis in \( X_d \).

**Corollary 3.3.** If \( p/d^* \) is a strictly outer \( \sigma \)-Schmidt vector, i.e. with \( p \) stable, then it is unique up to a sign.

**Proof.** Suppose \( p/d^*, q/d^* \) are strictly outer \( \sigma \)-Schmidt vectors. This means \( p \) and \( q \) are both stable. Since
\[ \frac{p}{p^*} = \frac{q}{q^*}, \]
it follows that \( pq^* = qp^* \) and hence \( p \mid q \mid p \), which implies \( p = \pm q \).

For the proof of Nehari’s theorem we need the following.

**Lemma 3.6.** Let \( \sigma = \| H_\sigma \| \). Then there exists an outer \( \sigma \)-singular vector.
Proof. Let $f$ be a $\sigma$-singular vector, i.e. $\|H_\phi f\| = \sigma \|f\|$. Let $f = \Theta F$ be an inner-outer factorization of $f$ (see Hoffman, 1962). Then, using Theorem 3.1,

$$
\|H_\phi\| \cdot \|F\| - \|H_\phi\| \cdot \|f\| = \|H_\phi f\| = \|H_\phi \Theta F\| = \|P_\Theta P_\phi F\|
$$

Therefore we must have equality throughout. This means that also $F = p/d^*$ is a $\sigma$-singular vector.

In the rational case this result can be strengthened.

**Lemma 3.7.** Let $\phi = n/d \in H_\infty$ with $d$ antistable, $d$ and $n$ coprime. Let $\sigma = \|H_\phi\|$. Then there exists a $\sigma$-Schmidt function $p/d^*$ with $p$ stable, i.e., this outer function has no zeros on the imaginary axis.

Proof. Let $p/d^*$ be any outer $\sigma$-Schmidt function. Thus

$$
np = \sigma d^* \hat{p} + d\pi.
$$

Let $r = p \wedge \hat{p}$. Thus $r$ has all its zeros on the imaginary axis. From the previous equation we get, with $p = rt$, $\hat{p} = \pm rt^*$ and putting $\lambda = \epsilon \sigma$,

$$
nrt = \lambda d^* rt^* + d\pi.
$$

This implies $r | \pi$. Therefore we get

$$
nt = \lambda d^* t^* + d\pi',
$$

with $t$ stable.

We are ready to give now a simple proof of Nehari's theorem in our rational context.

**Theorem 3.4 (Nehari).** Given a rational function $\phi = n/d \in H_\infty$ and $n \wedge d = 1$. Then

$$
\sigma_1 = \|H_\phi\| = \inf \|\phi - q\|_{\infty}, \quad q \in H_\infty^*,
$$

and this infimum is attained on a unique function $q = \phi - \sigma_1 g/f$, where $\{f, g\}$ is an arbitrary $\sigma_1$-Schmidt pair of $H_\phi$. 
Proof. Let \( \sigma_1 = \| H_\phi \| \). It follows from Equation (29) that the fact that for \( q \in H_+^\infty \) we have \( H_q = 0 \) that
\[
\sigma_1 = \| H_\phi \| = \| H_\phi - H_q \| = \| H_{\phi - q} \| \leq \| \phi - q \|_\infty,
\]
and so \( \sigma_1 \leq \inf_{q \in H_+^\infty} \| \phi - q \|_\infty \).

To complete the proof we will show there exists a \( q \in H_+^\infty \) for which equality holds.

We saw, in Lemma 3.7, that for \( \sigma_1 = \| H_\phi \| \) there exists a stable solution \( p_1 \) of
\[
n p_1 = \lambda d^* p_1^* + d \pi_1.
\]

Dividing this equation by \( dp_1 \), we get
\[
\frac{n}{d} - \frac{\pi_1}{p_1} = \lambda \frac{d^* p_1^*}{dp_1}.
\]

So, with \( q = \pi_1 / p_1 = n / d - \lambda_1 d^* p_1^* / dp_1 \in H_+^\infty \), we get
\[
\| \phi - q \|_\infty = \sigma_1 = \| H_\phi \|.
\]

For simplicity of exposition we will make now, for most of the rest of the paper, the following genericity assumption.

Assumption 3.1. For \( \phi = n / d \in H_+^\infty \) the singular values of \( H_{n/d} \) are all simple.

We conclude this section with a brief study of some problems related to Hankel operators and the polynomial methods we have developed.

3.1. Signed Singular Values

Given our definition of Hankel operators, it is impossible to use self-adjointness, as we deal with a map between two different Hilbert spaces. There is however an easy way to relate a Hankel operator to a self-adjoint operator, and we will exploit this.

To this end we define the map \( J : L^2(i\mathbb{R}) \to L^2(i\mathbb{R}) \) by
\[
Jf(z) = f^*(z) = \overline{f(-\bar{z})}.
\]
Clearly this is a unitary map in $L^2(i\mathbb{R})$, and it satisfies $JH^2_\pm = H^2_\pm$. This map is related to Hankel operators through

**Lemma 3.8.** Let $H_\phi$ be a Hankel operator and $J$ as above. Then:

1. We have
   \[ JP_\pm = P_\pm J. \quad (51) \]

2. We have
   \[ JH_\phi = H_\phi J | H^2_+. \]

3. If \( \{f, g\} \) is a Schmidt pair of $H_\phi$, then \( \{Jf, Jg\} \) is a Schmidt pair of $H_\phi^*$. 

4. Let $\sigma > 0$.
   (a) The map $\hat{U} : H^2_+ \to H^2_+$ defined by
   \[ \hat{U} = \frac{1}{\sigma} H^* J f \]
   is a bounded linear operator in $H^2_+$.
   (b) $\text{Ker}(H^* H - \sigma^2 I)$ is an invariant subspace for $\hat{U}$.
   (c) The map $U : \text{Ker}(H^* H - \sigma^2 I) \to \text{Ker}(H^* H - \sigma^2 I)$ defined by
   \[ U = \hat{U} | \text{Ker}(H^* H - \sigma^2 I) \]
   satisfies
   \[ U = U^* = U^{-1}. \]
   (d) Defining
   \[ K_+ = \frac{I + U}{2}, \quad K_- = \frac{I - U}{2}, \]
   we have
   \[ K_\pm = K_\pm^* = K^2_\pm. \]
   So $K_\pm$ are orthogonal projections and
   \[ \text{Ker}(I - U) = \text{Im} K_+, \]
   \[ \text{Ker}(I + U) = \text{Im} K_- , \]
and

\[ \text{Ker}(H^*H - \sigma^2 I) = \text{Im } K_+ \oplus \text{Im } K_- . \]

Also

\[ U = K_+ - K_- , \]

which is the spectral decomposition of \( U \), i.e., \( U \) is a signature operator.

**Proof.**

1: Obvious.

2: We compute

\[ JHf = JP_+ \phi f = P_+ J(\phi f) = P_+ \phi^* Jf = H_\phi^* (Jf) . \]

3: Let \( \{ f, g \} \) be a Schmidt pair of \( H_\phi \). From

\[ H_\phi f = \sigma g , \]

\[ H_\phi^* g = \sigma f \]

we get

\[ \sigma Jg = JH_\phi f = H_\phi^* Jf . \]

The other equation is proved analogously.

4(a): Obvious.

(b): Let \( f \in \text{Ker}(H^*H - \sigma^2 I) \). Then

\[ (H^*H - \sigma^2 I) \frac{1}{\sigma} JHf = \frac{1}{\sigma} H^* H JHf - \sigma JHf \]

\[ = \frac{1}{\sigma} JHH^* Hf - \sigma JHf = \frac{1}{\sigma} JH \sigma^2 f - \sigma JHf = 0 . \]

(c): To see that \( U \) is unitary let \( f \in \text{Ker}(H^*H - \sigma^2 I) \). This means that

\[ \| Hf \| = \sigma \| f \| . \]

Then

\[ \| Uf \| = \frac{1}{\sigma} \| H^* Jf \| = \frac{1}{\sigma} \| JHf \| = \frac{1}{\sigma} \| Hf \| = \| f \| . \]
So $U$ is isometric, i.e., $U^*U = I$. Clearly $U^2 = I$, for, given $f \in \ker(H^*H - \sigma^2 I)$, we have

$$U^2 f = \frac{1}{\sigma} H^* J \left( \frac{1}{\sigma} H^* J f \right) = \frac{1}{\sigma^2} H^* J H^* J f$$

$$= \frac{1}{\sigma^2} H^* J^2 H f = \frac{1}{\sigma^2} H^* H f = f.$$  

The following result is due to Ober (1989), who used canonical forms for balanced realizations. Similar reasoning to the following proof has been used in Fuhrmann (1975).

**Theorem 3.5.** Let $\sigma$ be a singular value of the Hankel operator $H_{\phi}$, let $J$ be defined by (50), and let $p$ be the minimal degree solution of (42). Assume $\deg d = n$ and $\deg p = m$. If $\epsilon = \lambda / \sigma$, then

$$\dim \ker(H_{\phi} - \lambda J) = \left\lceil \frac{n - m}{2} \right\rceil,$$

$$\dim \ker(H_{\phi} + \lambda J) = \left\lfloor \frac{n - m}{2} \right\rfloor.$$  

Also

$$\dim \ker(H_{\phi}^* H_{\phi} - \sigma^2 I) = n - m,$$

$$\dim \ker(H_{\phi} - \lambda J) - \dim \ker(H_{\phi} + \lambda J) = \begin{cases} 0, & n - m \text{ even,} \\ 1, & n - m \text{ odd.} \end{cases}$$

**Proof.** We put $k = n - m$. Then $\{(p / d^*) z^i \mid i = 1, \ldots, k\}$ is a basis for $\ker(H_{\phi}^* H_{\phi} - \sigma^2 I)$. From (42) we get

$$n(z^i)p = (-1)^i \lambda d^*(z^i p)^* + d(z^i \pi).$$

It follows that

$$\frac{z^i p}{d^*} \in \begin{cases} \ker(H_{\phi} - \lambda J) & \text{for } i \text{ even,} \\ \ker(H_{\phi} + \lambda J) & \text{for } i \text{ odd.} \end{cases}$$  

\(\blacksquare\)
3.2. Hamiltonian Structure

We turn now to a short study of some connections between singular value computations and Hamiltonian structure. Starting from Equation (42) and its dual, i.e. from

\[ np_i = \lambda_i d^* p_i^* + d \pi_i, \]

\[ n^* p_i^* = \lambda_i d p_i + d^* \pi_i^*, \]

we get

\[ np_i = \lambda_i d^* \frac{\lambda_i d p_i + d^* \pi_i^*}{n^*} + d \pi_i, \]

which leads to

\[ \left( \lambda_i^2 d^* d - n n^* \right) p_i = - \left[ \lambda_i (d^*)^2 \pi_i^* + n^* d \pi_i \right], \]

or

\[ p_i = - \frac{\lambda_i (d^*)^2 \pi_i^* + n^* d \pi_i}{\lambda_i^2 d^* d - n n^*} = - \frac{\lambda_i d^* \pi_i^* + n^*}{\lambda_i^2 - \frac{n n^*}{d^*}}, \]

and

\[ \frac{p_i}{d^*} = - \frac{\left( \frac{d}{d^*} \right) n^* \pi_i + \lambda_i d^* \pi_i^*}{\lambda_i^2 d^* d - n n^*}. \]  

(56)

**Theorem 3.6.** Let \( \phi = n / d \in H^\infty \) and \( n \wedge d = 1 \). Then \( \sigma \) is a singular value of \( H_\phi \) if and only if there exists a polynomial \( \pi \), of degree \( \leq n - 1 \), such that

\[ \frac{d}{d^*} n^* \pi_i + \lambda_i d^* \pi_i^* = 0 \mod \sigma_i^2 d^* d - n n^*. \]  

(57)

If the condition (57) is satisfied, the corresponding singular vectors are given by (56).

The polynomial \( \Omega = \sigma_i^2 d^* d - n n^* \) is Hamiltonian symmetric polynomial, i.e., it satisfies \( \Omega = \Omega^* \). Similarly, \( \sigma_i^2 - n n^* / d^* \) is a Hamiltonian symmetric
transfer function. This observation leads to an interesting appearance of Hamiltonian maps in this problem. This connection is summed up in the following.

**Theorem 3.7.** Let $\phi(z) = C(zI - A)^{-1}B$ be in $R^\infty(H)$, with $(A, B, C)$ a minimal realization of $\phi$. Let $F_\sigma$ be the Hamiltonian matrix

$$F_\sigma = \begin{pmatrix} A & \frac{1}{\sigma}B\tilde{B} \\ \frac{1}{\sigma}C\tilde{C} & -\tilde{A} \end{pmatrix}.$$  

Then $\sigma$ is a singular value of $H_\phi$ if and only if

$$\det [m(F_\sigma)]_{11} = 0,$$

where $m = d/d^*$.  

The connection was first pointed and proved in a special case by Zhou and Khargonekar (1987), and in the general case by Lypchuk, Smith, and Tannenbaum (1988), who showed that the conjectured matrix of Zhou and Khargonekar [1987] is equivalent to that given in Foias and Tannenbaum (1987). Another proof is given in Smith [1989]. For a treatment in the spirit of this paper see Fuhrmann [1990].

3.3. **Application to Nevanlinna-Pick Interpolation**

We discuss now briefly the connection between Nehari's theorem in the rational case and the finite Nevanlinna-Pick interpolation problem.

To begin we describe the interpolation problem.

**Definition 3.2.** Given points $\lambda_1, \ldots, \lambda_n$ in the open right half plane and complex numbers $c_1, \ldots, c_n$, then $\psi \in H^\infty_+$ is a Nevanlinna-Pick interpolant if it is a function of minimum $H^\infty_+$ norm that satisfies

$$\psi(\lambda_i) = c_i, \quad i = 1, \ldots, n. \quad (58)$$

We define the polynomial $d$ by

$$d(z) = \prod_{i=1}^n (z - \lambda_i).$$
Clearly \( d \) is antistable and \( d^* \) stable. We construct now one \( H^\infty_+ \) interpolant.

Let \( n \) be the unique polynomial, with \( \deg n < \deg d \), that satisfies the following interpolation constraints:

\[
n(\lambda_i) = d^*(\lambda_i)c_i, \quad i = 1, \ldots, n. \tag{59}
\]

This interpolant can be easily constructed by Lagrange interpolation or any other equivalent method. We note that as \( d^* \) is stable \( d^*(h) \neq 0 \) for \( i = 1, \ldots, n \) and \( n/d^* \in H^\infty_+ \). Moreover Equation (59) implies

\[
\frac{n(\lambda_i)}{d^*(\lambda_i)} = c_i, \quad i = 1, \ldots, n, \tag{60}
\]

i.e., \( n/d^* \) is an \( H^\infty_+ \) interpolant.

Any other interpolant is of the form \( n/d^* - (d/d^*)\theta \) for some \( \theta \in H^\infty_+ \).

To find \( \inf_{\theta \in H^\infty_+} \| n/d^* - (d/d^*)\theta \|_\infty \) is equivalent, \( d/d^* \) being inner, to finding \( \inf_{\theta \in H^\infty_+} \| n/d - \theta \|_\infty \). However this is just the content of Nehari's theorem and

\[
\inf_{\theta \in H^\infty_+} \| n/d - \theta \|_\infty = \sigma_1. \tag{61}
\]

Moreover the minimizing function is

\[
\theta = \frac{\pi_1}{p_1} = \frac{n}{d} - \lambda_1 \frac{d^*p^*_1}{dp_1}.
\]

Going back to the interpolation problem, we get for the Nevanlinna-Pick interpolant \( \psi \)

\[
\psi = \frac{n}{d^*} - \frac{d}{d^*} \theta = \frac{d}{d^*} \lambda_1 \frac{d^*p^*_1}{dp_1} = \frac{\lambda_1 p^*_1}{p_1}.
\]

Thus we can state:

Theorem 3.8. Given the Nevanlinna-Pick interpolation problem of Definition 3.2, let \( d(z) = \Pi_{i=1}^n (z - \lambda_i) \), and let \( n \) be the minimal degree polynomial satisfying the interpolation constraints

\[
n(\lambda_i) = d^*(\lambda_i)c_i.
\]
Let \( p_1 \) be the minimal degree solution of

\[
n p_1 = \lambda_1 d^* p_1^* + d \pi_1
\]

corresponding to the largest singular value \( \sigma_1 \). Then the Nevanlinna-Pick interpolant is given by

\[
\psi = \frac{p_1^*}{p_1}.
\]

That the optimal Nevanlinna-Pick interpolant is a multiple of a Blaschke product is classical and probably goes back to the pioneering and extraordinary work of Schur (1917, 1918). In this connection see Nikolskii (1985) and Delsarte, Genin, and Kamp (1981). The interesting, probably already known insight in this version of the result is, in view of Lemma 3.3, the connection between the (McMillan) degree of the Nevanlinna-Pick interpolant and the multiplicity of \( \sigma_1 \) as a singular value.

3.4. Connection with Geometric Control Theory

Controlled and conditioned invariant subspaces, introduced and studied by Basile and Marro (1969) and by Wonham and Morse (1970), have played a prominent role in the development of system theory and the solution of control design problems. In this connection see Wonham (1974) and Schumacher (1982).

The following lemmas relate spaces of singular vectors to geometric control objects. In view of the work of Ball and Helton (1989) this possibly opens up another way of studying fractional linear representations. However, this theme will not be pursued further in this paper.

Recall that, given the pair \((C, A)\), a subspace \( V \) of the state space \( X \) is called a conditioned invariant subspace if there exists a linear transformation \( H \) such that \((A + HC)V \subseteq V\), or equivalently that

\[
A(V \cap \text{Ker } C) \subseteq V. \tag{62}
\]

**Lemma 3.9.** Let \( \phi = n/d \in H^\infty \) and \( n \wedge d = 1 \), i.e., \( d \) is antistable. Let \( \sigma_i \) be a singular value of \( H_{n/d} \). Let \((A, B, C)\) be the shift realization of \( n/d \) constructed in Section 2. Then \( \gamma_i = \text{Ker}(H_{n/d}^* H_{n/d} - \sigma_i^2 I) \) is a conditioned invariant subspace of \( X^d \).

**Proof.** This follows from Theorem 3.6 in Fuhrmann [1981a]. However, it
can be easily proved directly. Thus we have to show that

$$A(\mathcal{V}_i \cap \text{Ker } \mathcal{C}) \subset \mathcal{V}_i.$$  

Now, for \( f \in X^d \), \( \mathcal{C}f = (d^{-1}f)_- \), so if \( f = p/d^* \in \mathcal{V}_i \), we have \( f = p_i a/d \) for some \( a \). Therefore \( f \in \text{Ker } \mathcal{C} \) if and only if \( \deg a < \deg d - \deg p_i - 1 \). For such an \( f \) we have \( S^d f = zf = p_i(zu)/d \in \mathcal{V}_i \).

\[ \text{Lemma 3.10.} \quad \text{Let } \phi = n/d \in H^\infty \text{ and } n \wedge d = 1, \text{ i.e., } d \text{ is antistable. Assume the generic case, i.e., the singular values of } H_{n/d} \text{ are distinct. Let } a_i \text{ be a singular value of } H_{n/d}. \text{ Let } (A, B, C) \text{ be the shift realization of } n/d \text{ constructed in Section 2. Then } \mathcal{W}_i = \sum_{j \neq i} \text{Ker}(H_{n/d}^* H_{n/d} - a_j^2 I) \text{ is a controlled invariant subspace of } X^d. \]

\[ \text{Proof.} \quad \text{Let } f \in \mathcal{W}_i \text{ so } f = \sum_{j \neq i} c_j p_j^*/d. \text{ Since } \{ p_j^* \mid j = 1, \ldots, n \} \text{ is a basis for } X^d \text{ and since the coordinates of } n \text{ in this basis are all nonzero (this will be proved in Section 8), clearly there exists a constant } \xi_f \text{ such that } \]

$$S_d p - \xi_f n = \sum_{j \neq i} c_j p_j^*.$$  

As \( \xi_f \) depends linearly on \( f \), there exists a map \( K : X^d \to \mathbb{R} \) such that \( \xi_f = Kf \). Thus

$$S_d p - \xi_f n = (A - BK)p.$$  

\[ \text{4. INVERSION} \]

In this section we discuss the relation between Hankel operators and compressions of multiplication operators. We discuss the invertibility properties of compression operators and derive invertibility results for Hankel operators. A detailed account can be found in Fuhrmann [1981b].

To begin let \( m, \Theta \in H^\infty_+ \) with \( m \) an inner function. Let \( H(m) = \{ mH^2 \}_+ \), and let \( T_\Theta : H(m) \to H(m) \) be defined by

$$T_\Theta f = P_{H(m)} \Theta f \quad \text{for } f \in H(m).$$  

Clearly, if \( \Theta \in H^\infty_+ \), we have \( \| T_\Theta \| \leq \| \Theta \|_{\infty} \).
Remark 4.1. This class of operators is of extreme importance and plays a crucial role in the commutant lifting theorem. In this connection we refer to Sarason [1967], Sz.-Nagy and Foias [1970], and Fuhrmann [1981b].

The next theorem sums up duality properties of operators commuting with shifts.

Theorem 4.1. Let $\Theta, m \in H^\infty_+$ with $m$ an inner function, and let $T_\Theta$ be defined by

$$T_\Theta f := P_{H(m)} \Theta f \quad \text{for } f \in H(m).$$

Then:

1. Its adjoint $T_\Theta^*$ is given by

$$T_\Theta^* f = P_+ \Theta^* f \quad \text{for } f \in H(m).$$

2. The operator $\tau_m : H(m) \to H(m)$ defined by

$$\tau_m f := mf^*$$

is unitary.

3. The operators $T_{\Theta^*}$ and $T_\Theta^*$ are unitarily equivalent. More specifically we have

$$T_\Theta \tau_m = \tau_m T_\Theta^*.$$ 

Proof. 1: Let $f, g \in H(m)$. Then

$$\langle T_\Theta f, g \rangle = \langle P_{H(m)} \Theta f, g \rangle = \langle m P_\Theta f, m^* g \rangle = \langle P_\Theta^* f, m^* g \rangle = \langle \Theta f, m^* g \rangle = \langle f, \Theta^* g \rangle = \langle P_+ f, \Theta^* g \rangle = \langle f, P_+ \Theta^* g \rangle = \langle f, T_\Theta^* g \rangle.$$ 

Here we have used the fact that $g \in H(m)$ if and only if $m^* g \in H^\infty_-$.

2: Clearly the map $\tau_m$, as a map in $L^2$, is unitary. From the orthogonal direct sum decomposition

$$L^2 = H^2_+ \oplus H(m) \oplus mH^2_+$$
it follows, by conjugation, that

\[ L^2 = m^*H^2_+ \oplus \{ H^2 \oplus m^*H^2_- \} \oplus H^2_+ . \]

Hence \( m(H^2 \oplus m^*H^2_-) = H(m) \).

3: We compute

\[ T_\Theta m f = T_\Theta mf^* = P_{H(m)} \Theta mf^* = mP_- m^* \Theta mf^* = mP_- \Theta f^* . \]

Now

\[ \tau_m T_\Theta^* = \tau_m (P_+ \Theta^* f) = m(P_+ \Theta^* f)^* = mP_- \Theta f^* . \]

The following spectral mapping theorem has been proved in Fuhrmann (1968a). A vectorial generalization is given in Fuhrmann (1968b). This will be instrumental in the analysis of Hankel operators restricted to their cokernels.

**Theorem 4.2** (Fuhrmann). Let \( \Theta, m \in H^\infty_+ \) with an inner function. The following statements are equivalent:

1. The operator \( T_\Theta \) defined in (63) is invertible.
2. There exists a \( \delta > 0 \) such that
   \[ |\Theta(s)| + |m(s)| \geq \delta \quad \text{for all } s \text{ with } \text{Re } s > 0 . \]  \hspace{1cm} (64)
3. There exist \( \xi, \eta \in H^\infty_+ \) that solve the Bezout equation
   \[ \xi \Theta + \eta m = 1 . \]  \hspace{1cm} (65)

In this case we have

\[ T_\Theta^{-1} = T_\xi . \]

**Proof.** We will not give a proof, which can be found in Fuhrmann (1969, 1981b). We remark only that by the Carleson corona theorem (Carleson, 1962), the strong coprimeness condition of (64) is equivalent to the solvability of the Bezout equation (65) over \( H^\infty_+ \).

**Theorem 4.3** (Adamjan, Arov, Krein). Let \( \phi = n/d \in H^\infty_+ \). Let the singular values of \( H_\phi \) satisfy \( \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{n-m} \geq \sigma_{n-m+1} = \cdots = \sigma_n \), i.e., \( \sigma_n \)
is a singular value of multiplicity $m$. Let $p$ be the minimum degree solution of (42) corresponding to $\sigma_n$. Then $p$ is antistable and of degree $n - m$.

We will say that two Hilbert space operators $T : H_1 \to H_2$ and $T' : H_3 \to H_4$ are equivalent if there exist unitary operators $U : H_1 \to H_3$ and $V : H_2 \to H_4$ such that

$$VT = T'U.$$

**Lemma 4.1.** Let $T : H_1 \to H_2$ and $T' : H_3 \to H_4$ be equivalent. Then $T$ and $T'$ have the same singular values.

**Proof.** Let $T^*Tx = \sigma^2 x$. Since $VT = T'U$, it follows that

$$U^*T'^*T'Ux = T^*V^*VTx = T^*Tx = \sigma^2 x,$$

or

$$T'^*T'(Ux) = \sigma^2(Ux).$$

**Remark 4.2.** If $x$ is a singular vector of the operator $T$ corresponding to the singular value $\sigma$, i.e. $T^*Tx = \sigma^2 x$, then

$$T^{-1}(T^{-1})^*x = \sigma^{-2}x,$$

i.e., $x$ is also a singular vector for $(T^{-1})^*$ corresponding to the singular value $\sigma^{-1}$.

In view of this remark, it is of interest to compute $[(H_\phi \mid H(m))^{-1}]^*$. Before proceeding with this, we compute the inverse of a related operator. This is a special case of Theorem 4.2 for the rationals. Note that, since $\|T_\Theta^{-1}\| = \sigma_n^{-1}$, there exists, by Sarason's theorem, a $\xi \in H_\phi^\infty$ such that $T_\Theta^{-1} = T_\xi$ and $\|\xi\|_\infty = \sigma_n^{-1}$. The next theorem provides this $\xi$. For an algebraic analogue of the next two theorems we refer to Helmke and Fuhrmann (1989).

**Theorem 4.4.** Let $\phi = n / d \in H_\phi^\infty$. Let $\theta = n / d^* \in H_\theta^\infty$. The operator $T_\Theta$ defined by Equation (63) is invertible, and its inverse given by $T_{(1/\lambda_n)p_n/p_n^*}$, where $\lambda_n$ is the last signed singular value of $H_\phi$ and $p_n$ is the minimal degree solution of

$$np_n = \lambda_n d^*p_n^* + d\pi_n.$$
Proof. From the previous equation we obtain the Bezout equation

$$\frac{n}{d^*} \left( \frac{1}{\lambda_n} \frac{p_n}{p_n^*} \right) - \frac{d}{d^*} \left( \frac{\pi_n}{\lambda_n p_n^*} \right) = 1. \quad (66)$$

By Theorem 4.3 the polynomial $p_n$ is antistable, so $p_n/p_n^* \in H_+$. This, by Theorem 4.2, implies the result.

It is well known that stabilizing controllers are related to solutions of Bezout equations over $H^\infty$. Thus we expect Equation (66) to lead to a stabilizing controller. The next corollary is a result of this type.

**Corollary 4.1.** Let $\phi = n/d \in H^\infty$. The controller $k = p_n/p_n^*$ stabilizes $\phi$. If the multiplicity of $\sigma_n$ is $m$, there exists a stabilizing controller of degree $n - m$.

**Proof.** Since $p_n$ is antistable, we get from (42) that $np_n - d\pi_n = \lambda_n d^* p_n^*$ is stable. We compute

$$\frac{\phi}{1 - k\phi} = \frac{d}{1 - \frac{p_n}{\pi_n} \frac{n}{d}} = \frac{-n\pi_n}{np_n - d\pi_n}. \quad \square$$

This corollary is related to questions of robust control. For more on this see Glover (1986).

**Theorem 4.5.** Let $\phi = n/d \in H^\infty$. Let $H : X^{d^*} \to X^d$ be defined by $H = H_\phi \mid X^{d^*}$. Then

1. $H_\phi^{-1} : X^d \to X^{d^*}$ is given by

$$H_\phi^{-1} h = \frac{1}{\lambda_n} \frac{d}{d^*} p - \frac{p_n h}{p_n^* h}. \quad (67)$$

2. $(H_\phi^{-1})^* : X^{d^*} \to X^d$ is given by

$$\left( H_\phi^{-1} \right)^* f = \frac{1}{\lambda_n} \frac{d^*}{d} p^* \frac{p_n^* f}{p_n f}. \quad (68)$$

**Proof.** 1: Let $m = d/d^*$, and let $T$ be the map given by $T = mH_{n/d}$. 


Thus we have the following commutative diagram:

\[
\begin{array}{c}
X^{d^*} \xrightarrow{H_{\theta}} X^d \\
\downarrow m \quad \downarrow m \\
X^{d^*}
\end{array}
\]

Now

\[
Tf = \frac{d}{d^*} P - \frac{n}{d^*} f = \frac{d}{d^*} P - \frac{d^*}{d} \frac{n}{d^*} f
\]

\[
= p_{H(d/d^*)} \frac{n}{d^*} f = P_{X^d} \frac{n}{d^*} f,
\]

i.e., \( T = T_\theta \), where \( \theta = n/d^* \). Now, from \( T_\theta = mH_{n/d} \) we have, by Theorem 4.4,

\[
T^{-1}_\theta = T_{(1/\lambda_n)p_n/p^*_n}.
\]

So, for \( h \in X^d \),

\[
H^{-1}_{n/d} h = \frac{1}{\lambda_n} p_{H(d/d^*)} \frac{p_n}{p^*_n} \frac{d}{d^*} h
\]

\[
= \frac{1}{\lambda_n} \frac{d}{d^*} p_{n} \frac{d}{d^*} \frac{p_n}{p^*_n} h
\]

\[
= \frac{1}{\lambda_n} \frac{d}{d^*} \frac{p_n}{p^*_n} h. \tag{69}
\]

2: Equation (69) can be written also as

\[
H^{-1}_{n/d} h = T_{(1/\lambda_n)p_n/p^*_n} mh.
\]

Therefore, using Theorem 4.1, we have, for \( f \in X^{d^*} \),

\[
(H^{-1}_{\theta}) f = m^*(T_{(1/\lambda_n)p_n/p^*_n})^* = \frac{d^*}{d} P + \frac{1}{\lambda_n} \frac{p_n}{p^*_n} f = \frac{1}{\lambda_n} \frac{d^*}{d} P + \frac{p_n}{p^*_n} f.
\]

\[
\text{\textsf{Corollary 4.2.}} \quad \text{There exist polynomials } \alpha_i, \text{ of degree } \leq n - 2, \text{ such that}
\]

\[
\lambda_i p^*_n p_i - \lambda_n p_n p^*_i = \lambda_i d^* \alpha_i, \quad i = 1, \ldots, n - 1.
\]

This holds also formally for \( i = n \) with \( \alpha_n = 0 \).

\[
\text{Proof.} \quad \text{Since}
\]

\[
H_{n/d} \frac{p_i}{d^*} = \lambda_i \frac{p^*_i}{d},
\]
it follows that

\[
\left( H_{n/d}^{-1} \right)^* \frac{p_i}{d^*} = \lambda_i^{-1} \frac{p_i^*}{d}.
\]

So, using Equation (68), we have

\[
\frac{1}{\lambda_n} \frac{d^*}{d} P + \frac{p_n^*}{p_n} \frac{p_i}{d^*} = \frac{1}{\lambda_i} \frac{p_i^*}{d},
\]

i.e.

\[
\frac{\lambda_n}{\lambda_i} \frac{p_i^*}{d^*} = P + \frac{p_n^*}{p_n} \frac{p_i}{d^*}.
\]

This implies, by partial fraction decomposition, the existence of polynomials \( \alpha_i, \ i = 1, \ldots, n \), such that \( \deg \alpha_i < \deg p_n = n - 1 \), and

\[
\frac{p_n^*}{p_n} \frac{p_i}{d^*} = \frac{\lambda_n}{\lambda_i} \frac{p_i^*}{d^*} + \frac{\alpha_i}{p_n},
\]

i.e.

\[
\lambda_i p_n^* p_i - \lambda_n p_n p_i^* = \lambda_i d^* \alpha_i. \tag{70}
\]

We saw, in Theorem 4.5, that for the Hankel operator \( H_{\phi} \) the map \( (H_{\phi}^{-1})^* \) is not a Hankel map. However, there is an equivalent Hankel map. We sum this up in the following.

**Theorem 4.6.** Let \( \phi = n/d \in H_\infty \). Let \( H : X^{d^*} \to X^d \) be defined by \( H = H_{\phi} \big| X^{d^*} \). Then:

1. The operator \( (H_{\phi}^{-1})^* \) is equivalent to the Hankel operator \( H_{(1/\lambda_n)d^*p_n/dp_n^*} \).
2. The Hankel operator \( H_{(1/\lambda_n)d^*p_n/dp_n^*} \) has singular values \( \sigma_i^{-1} < \cdots < \sigma_n^{-1} \).
3. The Schmidt pairs of \( H_{(1/\lambda_n)d^*p_n/dp_n^*} \) are \( \{ p_i^*/d^*, p_i/d \} \).

**Proof.** We saw that

\[
\left( H_{n/d}^{-1} \right)^* = \frac{d^*}{d} T_{(1/\lambda_n)p_n/dp_n^*}.
\]

Since multiplication by \( d^*/d \) is a unitary map of \( X^{d^*} \) onto \( X^d \), the operator
(H_{n/d})^* has, by Lemma 4.1, the same singular values as \( T_{(1/\lambda_n)p_n/p_n^*} \). These are the same as those of the adjoint operator \( T_{(1/\lambda_n)p_n/p_n^*}^* \). However, the last operator is equivalent to the Hankel operator \( H_{(1/\lambda_n)d^*p_n/dp_n^*} \). Indeed,

\[
\frac{d^*}{d} T_{(1/\lambda_n)p_n/p_n^*} f = \frac{d^*}{d} \left( T_{H(d/d^*)} - \frac{1}{\lambda_n} p_n^* f \right) = \frac{d^*}{d} p_n^* f - \frac{d^*}{d} \frac{1}{\lambda_n} p_n^* f - H_{(1/\lambda_n)d^*p_n/dp_n^*} f.
\]

This Hankel operator has singular values \( \sigma_1^{-1} < \cdots < \sigma_n^{-1} \), and its Schmidt pairs are \( \{ p_i^*/d^*, p_i/d \} \). Indeed,

\[
H_{d^*p_n/dp_n^*} p_i^* = \left( \frac{d^*}{d} p_n^* \right)^* p_i = - \left( \frac{d^*}{d} p_n^* \right) \left( \frac{dp_n^*}{d^*} \right) = - \frac{p_n p_i^*}{dp_n^*}.
\]

Now, from Equation (70) we get

\[
p_n p_i^* = \frac{\lambda_i}{\lambda_n} p_n^* p_i - \frac{\lambda_i}{\lambda_n} d^* \alpha_i,
\]

or taking the dual of that equation,

\[
p_n p_i^* = \frac{\lambda_n}{\lambda_i} p_n^* p_i + d \alpha_i^*.
\]

So

\[
\frac{p_n p_i^*}{dp_n^*} = \frac{\lambda_i}{\lambda_n} p_n^* p_i + \frac{d \alpha_i^*}{dp_n^*} = \frac{\lambda_n}{\lambda_i} p_i + \frac{\alpha_i^*}{p_n^*}.
\]

Hence

\[
P - \frac{p_n p_i^*}{dp_n^*} = \frac{\lambda_n}{\lambda_i} p_i.
\]

Therefore

\[
\frac{1}{\lambda_n} H_{d^*p_n/dp_n^*} \frac{p_i^*}{d^*} = \frac{1}{\lambda_i} \frac{p_i}{d}.
\]

5. HANKEL APPROXIMANT SINGULAR VALUES

The following theorem extends the analysis (see Glover, 1984) of singular values of the Hankel approximant corresponding to the smallest singular value.
Theorem 5.1. Let $\phi = n/d \in H_\infty^2$, and let $p_i$ be the minimal degree solutions of

$$np_i = \lambda_id^*p_i^* + dp_i.$$ 

Consider $p_n/p_n - n/d - \lambda_n d^* p_n^*/dp_n$. Then

1. $\pi_n/p_n \in H_\infty^2$, and $H_{\pi_n/p_n}$ has the singular values $\sigma_i = |\lambda_i|$, $i = 1, \ldots, n - 1$.

2. The $\sigma$-Schmidt pairs of $H_{\pi_n/p_n}$ are given by $\{\alpha_i/p_n, \alpha_i^*/p_n\}$, where the $\alpha_i$ are given by

$$\lambda_i p_n^*p_i - \lambda_n p_n p_i^* = \lambda_i d^* \alpha_i.$$  \hspace{1cm} (71)

3. Moreover, we have

$$\frac{\alpha_i}{p_n^*} = P_X p_i d^*,$$

i.e., the singular vectors of $H_{\pi_n/p_n}$ are projections of the singular vectors of $H_{n/d}$ onto $X^{n^2}$, the orthogonal complement of $\text{Ker} \ H_{\pi_n/p_n} = (p_n/p_n^*)H_\infty^2$.

Proof. Rewrite Equation (70) as

$$\lambda_i \frac{p_i}{d^*} - \lambda_n \frac{p_n p_i^*}{d^* p_n^*} = \lambda_i \frac{\alpha_i}{p_n^*}.$$  \hspace{1cm} (72)

So

$$\lambda_i \frac{\pi_n}{p_n} \frac{p_i}{d^*} - \lambda_n \frac{\pi_n}{p_n} \frac{p_n p_i^*}{d^* p_n^*} = \lambda_i \frac{\pi_n}{p_n} \frac{\alpha_i}{p_n^*}.$$ 

Projecting on $H_2^2$, and recalling that $p_n$ is antistable, we get

$$P - \frac{\pi_n}{p_n} \frac{p_i}{d^*} = P - \frac{\pi_n}{p_n} \frac{\alpha_i}{p_n^*}.$$ 

So

$$p_i/d^* - \alpha_i/p_n^* \in \text{Ker} \ H_{\pi_n/p_n}.$$ 

This is also clear from

$$\frac{p_i}{d^*} - \frac{\alpha_i}{p_n^*} = \frac{\lambda_n}{\lambda_i} \frac{p_n}{p_n^*} \frac{p_i}{d^*} \frac{p_n}{p_n^*} H_\infty^2 = \text{Ker} \ H_{\pi_n/p_n}.$$ 

Now

$$\frac{\pi_n}{p_n} = \frac{n}{d} - \frac{\lambda_n d^* p_n^*}{dp_n},$$
So finally we get

\[ H_{\pi_n/p_n} \frac{p_i}{d^*} = H_{n/d - \lambda_n d^{*} p_n^*/d p_n} \frac{p_i}{d^*} = H_{n/d} \frac{p_i}{d^*} - \lambda_n H_{d^* p_n^*/d p_n} \frac{p_i}{d^*} = \lambda_i \frac{p_i}{d} - \lambda_n p_i \frac{d^{*} p_n^*}{d p_n} = \lambda_i \frac{p_i}{d} - \frac{\lambda_n p_i}{d} \frac{d^{*} p_n^*}{d p_n} = \lambda_i \frac{d \alpha_i^*}{d p_n} = \frac{\alpha_i^*}{p_n}. \]

So finally we get

\[ H_{\pi_n/p_n} \frac{\alpha_i}{p_n^*} = \lambda_i \frac{\alpha_i^*}{p_n}. \]

Note that Equation (72) can be written as

\[ \frac{p_i}{d^*} = \frac{\alpha_i}{p_n^*} + \frac{\lambda_n}{\lambda_i} \frac{p_n}{p_n^*} \frac{p_i}{d^*}. \]

Since \( p_i^{*}/d^* \in H_{\pi_n}^2 \), this yields, projecting on \( X \pi_n^* = \{(p_n/p_n^*)H_{\pi_n}^2\} \perp \)

\[ P_{X\pi_n^*} \frac{p_i}{d^*} = \frac{\alpha_i}{p_n^*}. \]

**Corollary 5.1.** There exist polynomials \( \xi_i \) of degree \( \leq n - 3 \) such that

\[ \pi_n \alpha_i - \lambda_i p_n^* \xi_i = p_n \xi_i, \quad i = 1, \ldots, n - 1. \]

**6. Orthogonality Relations**

We devote this section to the derivation of some polynomial identities out of the singular-value–singular-vector equations. We proceed to interpret these relations as orthogonality relations between singular vectors associated with different singular values. Furthermore the same equations provide useful orthogonal decompositions of singular vectors.
Equation (71), in fact more general relations, could be derived directly. This we proceed to do. Starting from the singular value equations, i.e.

\[ np_i = \lambda_i d^* p_i^* + d \pi_i, \]
\[ np_j = \lambda_j d^* p_j^* + d \pi_j, \]

we get

\[ 0 = d^* \left( \lambda_i p_i^* p_j - \lambda_j p_i p_j^* \right) + d \left( \pi_i p_j - \pi_j p_i \right). \] (73)

Since \( d \) and \( d^* \) are coprime, there exist polynomials \( \alpha_{ij} \), of degree \( \leq n - 2 \), for which

\[ \lambda_i p_i^* p_j - \lambda_j p_i p_j^* = d \alpha_{ij} \] (74)

and

\[ \pi_i p_j - \pi_j p_i = -d^* \alpha_{ij}. \]

**Remark 6.1.** We know that for a self-adjoint operator, eigenvectors corresponding to different eigenvalues are orthogonal. Thus, under the assumption \( \sigma_i \neq \sigma_j \), we must have

\[ \left( \frac{p_i}{d^*}, \frac{p_j}{d^*} \right)_{H_*} = 0. \] (75)

This orthogonality relation could be derived from the polynomial equations by contour integration in the complex plane. Indeed, Equation (74) could be rewritten as

\[ \frac{p_i^* p_j}{d d^*} - \frac{\lambda_i p_i p_j^*}{d d^*} = \frac{\alpha_{ij}}{d^*}. \]

This equation can be integrated over the boundary of a half disc of radius \( R \), centered at the origin, which lies in the right half plane. Since \( d^* \) is stable, the integral on the right hand side is zero. A standard estimate, using the fact that \( \deg p_i^* p_j \leq 2n - 2 \), leads, in the limit as \( R \to \infty \), to

\[ \left( \frac{p_j}{d^*}, \frac{p_i}{d^*} \right)_{H_*} = \frac{\lambda_j}{\lambda_i} \left( \frac{p_i}{d^*}, \frac{p_j}{d^*} \right)_{H_*}. \]

This implies the orthogonality relation (75).
Equation (74) can be rewritten as

$$ p_i^* p_j = \frac{\lambda_j}{\lambda_i} p_i p_j^* + \frac{1}{\lambda_i} d\alpha_{ij}, $$

However if we rewrite (74) as

$$ \lambda_j p_i p_j^* = \lambda_i p_i^* p_j - d\alpha_{ij}, $$

then after conjugation we get

$$ p_i^* p_j = \frac{\lambda_j}{\lambda_i} p_i p_j^* - \frac{1}{\lambda_j} d^*\alpha_{ij}^*. $$

Equating the two expressions leads to

$$ \begin{pmatrix} \frac{\lambda_i}{\lambda_j} - \frac{\lambda_j}{\lambda_i} \\ \frac{\lambda_j}{\lambda_i} \end{pmatrix} p_i p_j^* = \frac{1}{\lambda_i} d\alpha_{ij} + \frac{1}{\lambda_j} d^*\alpha_{ij}^*. $$

Putting $j = i$, we get

$$ \alpha_{ii} = 0. $$

Otherwise we have

$$ p_i p_j^* = \frac{1}{\lambda_i^2 - \lambda_j^2} \left( \lambda_j d\alpha_{ij} - \lambda_i d^*\alpha_{ij}^* \right). \quad (76) $$

Conjugating this last equation and interchanging indices leads to

$$ p_i p_j^* = \frac{1}{\lambda_j^2 - \lambda_i^2} \left( \lambda_j d\alpha_{ji} - \lambda_i d^*\alpha_{ji}^* \right). \quad (77) $$

Comparing the two expressions leads to

$$ \alpha_{ji} = -\alpha_{ij}. \quad (78) $$

We end by studying two special cases. For the case $j = n$ we put
\( \alpha_i = \frac{\alpha_i^*}{\lambda_i} \) to obtain

\[
\lambda_i p_n^* p_i - \lambda_n p_n p_i^* = \lambda_i d_i^* \alpha_i^*, \quad i = 1, \ldots, n - 1, \tag{79}
\]

or

\[
\lambda_i p_i^* p_n - \lambda_n p_i p_n^* = \lambda_i d_i^* \alpha_i^*, \quad i = 1, \ldots, n - 1. \tag{80}
\]

From Equation (73) it follows that

\[
\pi_n p_i - \pi_i p_n = \lambda_i d_i^* \alpha_i^*,
\]

or, equivalently,

\[
\pi_i p_i^* - \pi_i^* p_i = \lambda_i d_i \alpha_i.
\]

If we specialize now to the case \( i = 1 \), we obtain

\[
\pi_n p_1 - \pi_1 p_n = \lambda_1 d_1^* \alpha_1^*,
\]

which, after dividing through by \( p_1, p_n \) and conjugating, yields

\[
\frac{\pi_n^*}{p_n^*} - \frac{\pi_i^*}{p_i^*} = \lambda_1 \frac{d_1}{p_i^* p_n^*}. \tag{81}
\]

Similarly, starting from Equation (72) and putting \( i = 1 \), we get

\[
\lambda_1 p_i^* p_1 - \lambda_1 p_1 p_i^* = d_1 \alpha_1;
\]

putting also \( \beta_i = \lambda_i^{-1} \alpha_1 \) we get

\[
\lambda_1 p_i^* p_1 - \lambda_1 p_1 p_i^* = \lambda_i d_1 \beta_i, \tag{83}
\]

and of course \( \beta_1 = 0 \). This can be rewritten as

\[
p_i^* p_1 - \frac{\lambda_1}{\lambda_i} p_1 p_i^* = d_1 \beta_i, \tag{84}
\]

or

\[
p_1 p_i^* - \frac{\lambda_i}{\lambda_1} p_i^* p_i = d_i^* \beta_i^*. \tag{85}
\]
This is equivalent to

\[ \frac{p_i^*}{d^*} = \frac{\beta_i^*}{p_1} + \frac{\lambda_i}{\lambda_1} \frac{p_i^*}{p_1} \frac{p_i}{d^*}. \] (86)

We note that Equation (86) is nothing else but the orthogonal decomposition of \( p_i^*/d^* \) relative to \( H_+^{2} = X p_i \oplus (p_i^*/p_i) H_+^{2} \). Therefore we have

\[ \left\| \frac{\beta_i^*}{p_1} \right\|^2 = \left\| \frac{p_i^*}{d^*} \right\|^2 - \left( \frac{\lambda_i}{\lambda_1} \right)^2 \left\| \frac{p_i^*}{d^*} \right\|^2 = \left\| \frac{p_i}{d^*} \right\|^2 \left( 1 - \left( \frac{\lambda_i}{\lambda_1} \right)^2 \right). \] (87)

Notice that if we specialize Equation (80) to the case \( i = 1 \) and Equation (83) to the case \( i = n \), we obtain the relation

\[ \beta_n = \alpha_n^*. \] (88)

7. DUALITY IN HANKEL NORM APPROXIMATION

In the present section we will shed some light on intrinsic duality properties of problems of Hankel norm approximation and extensions. Results strongly suggesting an underlying duality have appeared before. In fact a comparison of Lemmas 9.1 and 9.4 in Glover's paper suggests that one should be derivable from the other by some duality considerations. This in fact turns out to be the case, though the duality analysis is far from being obvious. In the process we will prove a result dual to Theorem 5.1. This extends in a sense Lemma 9.4 in Glover (1984). While this analysis, after leading to the form of the Schmidt pairs in Theorem 7.2, is not necessary for the proof, it is felt that it is of independent interest and its omission would leave all intuition out of the exposition.

The analysis of duality can be summed up in the scheme shown in Figure 1, which exhibits the relevant Hankel operators, their singular values, and the corresponding Schmidt pairs.

We would like to analyse the truncation that corresponds to the largest singular value. To this end we invert (I) the Hankel operator \( H_{n/d} \) and conjugate (C) it, i.e. take its adjoint, as in Theorem 4.5. This operation preserves Schmidt pairs and inverts singular values; however, the operator so obtained is not a Hankel operator. This we correct by replacing it with an equivalent Hankel operator (E). This preserves singular values but changes the Schmidt pairs. Thus ICE in Figure 1 stands for a sequence of these three
operations. To the Hankel operator so obtained, i.e. to $H_{(1/\lambda_n) d^* \alpha^* p / p^*_n}$, we apply Theorem 5.1, which leads to the Hankel operator $H_{(1/\lambda_n) d^* \alpha^* p / p^*_n}$. This is done in Theorem 7.1. To this Hankel operator we apply again the sequence of three operations ICE, and this leads to Theorem 7.2.

We proceed to study this Hankel map.

**Theorem 7.1.** For the Hankel operator $H_{(1/\lambda_n) d^* \alpha^* p / p^*_n}$ the Hankel norm approximant corresponding to the least singular value, i.e. to $\sigma_1^{-1}$, is $(1/\lambda_n) d^* \alpha^* / p^* p^*_n$.

For the Hankel operator $H_{(1/\lambda_n) d^* \alpha^* p / p^*_n}$ we have:

1. $\text{Ker} H_{(1/\lambda_n) d^* \alpha^* p / p^*_n} = \frac{p^*_1}{p_1} H_{+}^2$.
2. $X_{p_1} = \left\{ \frac{p^*_1}{p_1} H_{+}^2 \right\}^\perp$.
3. The singular values of $H_{(1/\lambda_n) d^* \alpha^* p / p^*_n}$ are $\sigma_1^{-1} < \cdots < \sigma_n^{-1}$.
4. The Schmidt pairs of $H_{(1/\lambda_n) d^* \alpha^* p / p^*_n}$ are $\{ \beta^*_i / p_1, \beta_i / p^*_1 \}$, where

$$\frac{\beta^*_i}{p_1} = P_{X_{p_1}} \left( \frac{p^*_i}{d^*} \right)$$
For \( \varphi_1 \), the Schmidt pairs for \( H_{(1/\lambda_n) \varphi_1} \) are \( \{ p_1^*/d^*, p_1/d \} \). Therefore the best Hankel norm approximant associated with \( \varphi_1 \) is, using also Equation (79),

\[
\begin{align*}
\frac{1}{\lambda_n} \frac{d^* p_n}{dp^n_*} - \frac{1}{\lambda_1} \frac{p_1}{d} = \frac{1}{\lambda_n} \frac{d^* p_n}{dp^n_*} - \frac{1}{\lambda_1} \frac{d^* p_1}{dp^*_1} \\
= \frac{1}{\lambda_1 \lambda_n} \frac{d^* (\lambda_1 p_n p^*_1 - \lambda_n p_1 p^*_n)}{dp^*_1 p^n_*} \\
= \frac{1}{\lambda_1 \lambda_n} \frac{d^* (\lambda_1 d^* \alpha_1^*)}{dp^*_1 p^n_*} \\
= \frac{1}{\lambda_n} \frac{d^* \alpha_1^*}{p^*_1 p^n_*}.
\end{align*}
\]

1: Let \( f \in (p^*_1/p_1) H_s^2 \), i.e., \( f = (p^*_1/p_1) g \) for some \( g \in H_s^2 \). Then

\[
P - \frac{1}{\lambda_n} \frac{d^* \alpha_1^*}{p^n_* p^*_1} g = \frac{1}{\lambda_n} P - \frac{d^* \alpha_1^*}{p^*_1 p^n_*} g = 0,
\]

as \( d^* \alpha_1^*/p_1 p^n_* \in H_s^\infty \) and \( g \in H_s^2 \). Conversely, let \( f \in \text{Ker } H_{(1/\lambda_n) \varphi_1} \), i.e.

\[
P - \frac{d^* \alpha_1^*}{p^*_1 p^n_*} f = 0.
\]

This implies \( p^*_1 \| d^* \alpha_1 \). Now \( p_1 \) and \( d \) are coprime, as the first polynomial is stable whereas the second antistable. This implies naturally the coprimeness of \( p^*_1 \) and \( d^* \). Also we have

\[
\lambda_1 p^*_1 p_n - \lambda_n p_1 p^*_n = \lambda_1 d^* \alpha_1^*.
\]

If \( p^*_1 \) and \( \alpha_1^* \) are not coprime, then, by the previous equation, \( p^*_1 \) has a common factor with \( p_1 p^n_* \). However \( p_1 \) and \( p_n \) are coprime, as the first is
stable and the second antistable. So are $p_1$ and $p_1^*$, and for the same reason. Therefore we must have that $f/p_1^*$ is analytic in the right half plane. So $(p_1/p_1^*) f \in H_+^2$, i.e. $f \in (p_1^*/p_1) H_+^2$.

2: Follows from the previous part.

3: This is a consequence of Theorem 5.1.

4: Follows also from Theorem 5.1, as the singular vectors of $H_{(1/\lambda_\alpha)} \pi_{\alpha}/p_1 p_1^*$ are given by

$$P_{X_{\alpha_{i1}}} \frac{p_1^*}{d^*}.$$ 

This can be computed. Indeed, starting from Equation (83) we have

$$\beta_i = \frac{\lambda_i p_1^* p_i - \lambda_i p_1 p_i^*}{\lambda_i d}.$$ 

We compute

$$\lambda_i \beta_n p_i^* - \lambda_n \beta_i p_n^* = \lambda_i \frac{\lambda_1 p_1^* p_n - \lambda_n p_1 p_n^*}{\lambda_1 d} p_i^* - \lambda_n \frac{\lambda_1 p_1^* p_i - \lambda_i p_1 p_i^*}{\lambda_1 d} p_n^* = \frac{\lambda_1 p_1^*}{\lambda_1 d} \left( \lambda_i p_1^* p_n - \lambda_n p_1 p_n^* \right) = \frac{p_i^*}{d} \left( \lambda_i d \alpha_i^* \right)\left( \lambda_i p_1^* p_n - \lambda_n p_1 p_n^* \right) = \lambda_i p_1^* \alpha_i^*.$$ 

Recalling that

$$\frac{\beta_i^*}{p_1} = \frac{p_i^*}{d^*} - \frac{\lambda_i p_i^*}{\lambda_1 p_1^*}$$ 

and $\beta_n = \alpha_i^*$, we have

$$H_{(1/\lambda_\alpha)} \pi_{\alpha}/p_1 p_1^* - \frac{1}{\lambda_i} \frac{\beta_i^*}{p_1} = P_{\beta^*} \left( 1 \frac{\beta_i}{\lambda_i p_i^*} p_1^* \right) \left( \frac{1}{\lambda_n} d^* \beta_n \beta_i^* - \frac{1}{\lambda_i} p_1 p_i^* \beta_i \right).$$ 

Thus it suffices to show that the last term is zero. Now, from Equation (85),

$$d^* \beta_i^* = p_1 p_i^* - \frac{\lambda_i}{\lambda_1} p_1^* p_1.$$
Hence

\[
P \frac{1}{p_1^* p_n^* p_1} \left( \frac{1}{\lambda_n} d^* \beta_n \beta_1^* - \frac{1}{\lambda_1} p_1 p_n^* \beta_1 \right) = P \frac{1}{p_1^* p_n^* p_1} \left( \frac{1}{\lambda_n} p_1 p_1^* \beta_n - \lambda_1 \frac{1}{\lambda_1} p_1 p_1^* \beta_1 \right)
\]

\[
= P \frac{1}{p_1^* p_n^* p_1} \left( \frac{1}{\lambda_n} p_1 p_1^* \beta_n - \lambda_1 \frac{1}{\lambda_1} p_1 p_1^* \beta_1 \right)
\]

\[
= P \frac{1}{p_1^* p_n^* p_1} \left( \frac{1}{\lambda_n} \left( \lambda_1 p_1^* \beta_n - \lambda_n p_n^* \beta_1 \right) - \lambda_1 \frac{1}{\lambda_1} p_1 p_1^* \beta_n \right)
\]

\[
= P \frac{1}{p_1^* p_n^* p_1} \left( \frac{1}{\lambda_n} \left( \lambda_1 p_1^* \beta_n - \lambda_n p_n^* \beta_1 \right) \right)
\]

\[
= P \frac{1}{p_1^* p_n^* p_1} \left( \frac{1}{\lambda_n} \frac{1}{\lambda_1} \left( \lambda_1 p_1^* \beta_n - \lambda_n p_n^* \beta_1 \right) \right) = 0,
\]

as \( p_1 p_n^* \) is stable.

The main result of the following theorem has been proved in Glover (1984), and a trace class extension is given in Glover, Curtain, and Partington (1989). However, it seems to have been first derived by Hruscev and Peller; see Theorem 1.6 and Corollary 1.8 in Appendix 5 of Nikolskii (1985). In this connection the work of Ober (1987b,c) is also relevant.

**Theorem 7.2.** Let \( \phi = n / d \in \mathbb{H}^\infty \) and \( n \wedge d = 1 \), i.e., \( d \) is antistable. Let \( \pi_1 / p_1 \) be the optimal causal approximant to \( n / d \). Then:

1. \( n / d - \pi_1 / p_1 \) is all pass.
2. The singular values of the Hankel operator \( H_{\pi_1 / p_1} \) are \( \sigma_2 > \cdots > \sigma_n \), and the corresponding Schmidt pairs of \( H_{\pi_1 / p_1} \) are \( \{ \beta_1 / p_1, \beta_1^* / p_1^* \} \), where the \( \beta_i \) are defined by

\[
\beta_i = \frac{\lambda_1 p_1^* p_i - \lambda_i p_1 p_i^*}{\lambda_1 d_1}.
\]

**Proof.**

1. Since

\[
\frac{n}{d} - \frac{\pi_1}{p_1} = \lambda_1 \frac{d^* p_1^*}{d p_1}.
\]

2: We saw, in Equation (81), that

\[
\frac{\pi_1^*}{p_1^*} = \frac{\pi_n^*}{p_n^*} - \lambda_1 \frac{d \alpha_1}{p_1^* p_n^*}.
\]
Since \( \tau_n^* / p_n^* \in H_+^\infty \), the associated Hankel operator is zero. Hence

\[
H_{x^f/p_1^f} = H_{-\lambda_n d\alpha_i/p_1 p_i^*}.
\]

Thus we have to show that

\[
H_{-\lambda_n d\alpha_i/p_1 p_i^*} \frac{\beta_i}{p_1} = \frac{\beta_i^*}{p_1^*}.
\]

To this end we start from Equation (84), which, multiplied by \( \alpha_1 \), yields

\[
d\alpha_1 \beta_i = p_1^* p_1 \alpha_1 - \frac{\lambda_i}{\lambda_1} p_1 p_i^* \alpha_1.
\]  

(90)

This in turn implies

\[
\frac{d\alpha_1}{p_1^* p_n^*} \frac{\beta_i}{p_1} = \frac{p_1 \alpha_1}{p_1 p_n^*} - \frac{\lambda_i}{\lambda_1} \frac{\alpha_1 p_i^*}{p_1 p_n^*}.
\]  

(91)

Since \( p_1 \alpha_1 / p_1 p_n^* \in H_+^2 \), we have

\[
d(\beta_n p_n - d\beta_i p_i) = \frac{\lambda_i}{\lambda_1} p_1 p_i^* p_n.
\]  

(92)

All we have to do is to obtain a partial fraction decomposition of the last term. To this end we go back to Equation (84), from which we get

\[
d\beta_i p_n = p_1^* p_1 p_n - \frac{\lambda_i}{\lambda_1} p_1 p_i^* p_n,
\]  

(93)

\[
d\beta_n p_i = p_1^* p_n p_i - \frac{\lambda_i}{\lambda_1} p_1 p_n^* p_i.
\]

Hence

\[
d(\beta_n p_i - d\beta_i p_n) = \frac{p_1}{\lambda_1} (\lambda_i p_1^* p_n - \lambda_n p_n^* p_i) = \frac{p_1}{\lambda_1} \lambda_i d\alpha_i^*
\]  

(94)

and

\[
\beta_n p_i - d\beta_i p_n = \frac{\lambda_i}{\lambda_1} p_1^* \alpha_i^*.
\]  

(95)

Now

\[
\beta_n^* p_i^* - \beta_i^* p_n^* = \alpha_1 p_i^* - \beta_i^* p_n^* = \frac{\lambda_i}{\lambda_1} p_i^* \alpha_i.
\]
Dividing through by $p^*_1 p^*_n$, we get
\[
\frac{\alpha_1 p^*_1}{p^*_1 p^*_n} = \frac{\beta^*_1}{p^*_1} + \frac{\lambda_i \alpha_i}{\lambda_1 p^*_n},
\]
and from this it follows that
\[
P - \frac{\alpha_1 p^*_1}{p^*_1 p^*_n} = \frac{\beta^*_1}{p^*_1},
\]
Using Equation (92), we have
\[
-\lambda_i P - \frac{d \alpha_1}{p^*_p p^*_n} \frac{\beta_i}{p^*_1} = \frac{\beta^*_1}{p^*_1},
\]
and this completes the proof.

As an immediately corollary we obtain the dual of Corollary 5.1.

**Corollary 7.1.** There exist polynomials $\omega_i$ of degree $\leq n - 2$ such that
\[
\pi^*_i \beta_i - \lambda_i \beta^*_i p^*_1 = p^*_i \omega_i, \quad i = 2, \ldots, n.
\]

There is another way of looking at duality, and this is summed up in the diagram in Figure 2. We will not go into the details except for the following.

**Theorem 7.3.** The Hankel operator $(1/\lambda_{n-1})H_{p_n^* \sigma_{n-1}/p_n \sigma_{n-1}^*} : X^{p^*_1} \to X^{p_n}$ has singular values $\sigma_1^{-1} < \cdots < \sigma_{n-1}^{-1}$ and Schmidt pairs \{\(\alpha^*_i / p^*_i, \alpha_i / p_n\)\}.
Proof. Starting from
\[ \pi_n \alpha_i = \lambda_i p_n^* \alpha_i^* + p_n \zeta_i, \]
\[ \pi_n \alpha_j = \lambda_j p_n^* \alpha_j^* + p_n \zeta_j, \]
we get
\[ 0 = p_n^* (\lambda_i \alpha_j^* - \lambda_j \alpha_i^*) + p_n (\alpha_j \zeta_i - \alpha_i \zeta_j). \]

For \( j = n - 1 \) we can write
\[ \lambda_i \alpha_{n-1} \alpha_i^* - \lambda_{n-1} \alpha_i \alpha_{n-1}^* = \lambda_i p_n \kappa_i, \]
\[ \alpha_{n-1} \alpha_i^* = p_n \kappa_i + \frac{\lambda_{n-1}}{\lambda_i} \alpha_i \alpha_{n-1}^*, \]
i.e.
\[ \frac{\alpha_i^*}{p_n} = \frac{\kappa_i}{\alpha_{n-1}} + \frac{\lambda_{n-1}}{\lambda_i} \frac{\alpha_i}{\alpha_{n-1} \ p_n}. \]

Thus we have
\[ \frac{1}{\lambda_{n-1}} H_{p_n^* \alpha_n^* \ p_n^{\ alpha_n^*}} \frac{\alpha_i^*}{p_n^*} = \frac{1}{\lambda_{n-1}} p_n \frac{\lambda_{n-1}}{\lambda_i} \left( p_n \kappa_i + \frac{\lambda_{n-1}}{\lambda_i} \alpha_i \alpha_{n-1}^* \right) \]
\[ = \frac{1}{\lambda_{n-1}} p_n \frac{\kappa_i}{\alpha_{n-1}^*} + \frac{1}{\lambda_i} p_n \frac{\alpha_i}{p_n} \]
\[ = \frac{1}{\lambda_i} \frac{\alpha_i}{p_n}. \]

8. BALANCED REALIZATIONS

A fundamental problem in system theory is that of model reduction, namely the replacement of high order state space models by reduced order models, where however one would like to keep the resulting errors small in some given
measure. The measure of smallness is generally taken as some standard norm; the most frequently encountered ones are the $L^\infty$, $L^2$, and Hankel norms.

Now, in the case of state space representations, there are standard ways of reducing dimension by elimination of nonobservable and nonreachable states. With this in mind one would go further and try to eliminate the less controllable and less observable states. However, a state which is not important for observation may be crucial for the control of the system, and vice versa. To get around this difficulty, Moore (1981) introduced the notion of balanced realizations for asymptotically stable transfer functions. In such a realization the reachability and observability Gramians are equal and diagonal. This opens up the possibility of approximating the original system by a lower order one through the elimination of states that correspond to simultaneous low controllability and observability properties. The existence of balanced realizations for asymptotically stable transfer functions has been established by Moore. A simple reduction to a balanced system is given in Laub (1980) and Glover (1984).

What distinguishes balanced model reduction is its remarkable behavior with respect to truncations. Truncation preserves asymptotic stability; see Pernebo and Silverman (1982). An extension of their result is given by Ober (1989), who also obtained canonical forms for balanced realizations for a wide class of transfer functions. Other related papers are Young (1985), Gregson and Young (1988), and Yeh, Yang, and Tsai [1988].

However, the most significant progress in the application of balanced realizations is no doubt the work of Glover (1984). In that paper Glover obtains $L^\infty$ estimates on the approximation error using the truncation of balanced realizations. Some infinite dimensional generalizations are given in Young (1986), Ober (1987a–d), and Glover, Curtain, and Partington (1989).

In this paper we take a different approach to balanced realizations. We take advantage of the extensive analysis of the Hankel operators and their Schmidt pairs that has been carried out in the previous sections. Using this we will show that the polynomial model realization method, outlined in Section 2, can be used to obtain, for an asymptotically stable transfer function, a balanced realization. All one has to do is to compute, in the generic case (i.e. the case of distinct singular values), a matrix representation with respect to a basis constructed of the suitably normalized Hankel singular vectors. While this is certainly a roundabout way of approaching the balancing problem, it has the great advantage in the geometric insight it provides into the problem of truncation. Theorem 10.2 is a case in point.

From this point of view, balancing means (assuming the polynomial model realization theory) a correct choice of an orthogonal basis in a (left) invariant subspace of $H^2$. This can be reduced to the algebra of orthogonal polynomials on the line. In particular, for the case of asymptotically stable all-pass transfer
functions we identify these orthogonal polynomials and relate them to a continued fraction expansion. This recovers the canonical form obtained previously by Ober (1987a–d). For the general case, as Ober (1989) correctly observed, the canonical form for the all-pass case turns out to be the main building block for the construction of the balanced canonical form in the general case. This is reflected in the functional representation by piecing together a suitable orthogonal basis made up, basically, of polynomials orthogonal with respect to measures related to Hankel singular vectors. We will not go into the details of this in this paper.

To begin, we recall the definition.

**Definition 8.1.** A minimal asymptotically stable system $(A, B, C, D)$ is called balanced if there exists a diagonal matrix $\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_n)$ such that

\[
A\Sigma + \Sigma A = -B\tilde{B}, \\
\tilde{A}\Sigma + \Sigma \tilde{A} = -\tilde{C}C.
\]

(104)

The matrix $\Sigma$ is called the Gramian of the system $(A, B, C, D)$, and its diagonal entries are called the singular values of the system.

Our definition of Hankel operators is based on antistable transfer functions, rather than stable ones as in Glover (1984). As a consequence, in the derivation of balanced realizations it is convenient to extend the notion of balancing to (asymptotically) antistable transfer functions.

We will say that $(A, B, C, D)$ is a balanced realization of an (asymptotically) antistable transfer functions $\phi$ if $(-\tilde{A}, \tilde{B}, -\tilde{C}, \tilde{D})$ is a balanced realization of the (asymptotically) stable transfer functions $\phi^*$—equivalently, if for a diagonally matrix $\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_n)$ we have

\[
-\tilde{A}\Sigma + \Sigma(-A) = -\tilde{C}C, \\
-A\Sigma + \Sigma(-\tilde{A}) = -B\tilde{B},
\]

or

\[
\Lambda\Sigma + \Sigma\tilde{A} = B\tilde{B}, \\
\tilde{A}\Sigma + \Sigma A = \tilde{C}C.
\]

(105)

Balanced realizations have another symmetry property. With this in mind we recall the following definition.
DEFINITION 8.2. A realization \((A, B, C)\) is called signature symmetric if for some signature matrix \(J\), i.e. a matrix of the form \(J = \text{diag}(\pm 1, \ldots, \pm 1)\), the diagram in Figure 3 is commutative.


We proceed now to the establishment of the existence of balanced realizations. We will use the polynomial model realization theory outlined in Section 2. Our procedure is to compute a matrix representation of that realization of \(\phi\) with respect to the basis \(\{p_i^*/d, i = 1, \cdots, n\}\) of \(X^d\), with a suitable normalization.

We begin by studying the generic case, namely the case where all singular values are distinct.

\[ \text{THEOREM 8.1.} \] Let \(\phi = n/d \in H^\infty\). Let \(p_i\) be the minimal degree solutions of the singular value equation

\[
n p_i = \lambda_i d^* p_i^* + d \pi_i,
\]

normalized so that

\[
\frac{\|p_i\|_{d^*}^2}{\|d^*\|} = \sigma_i.
\] (106)

Then:

1. The function \(\phi\) has a balanced realization of the form

\[
A = \begin{pmatrix} \epsilon_i b_i b_j \\ \lambda_i + \lambda_j \end{pmatrix},
\]

\[
B = (b_1, \ldots, b_n),
\]

\[
C = (\epsilon_1 b_1, \ldots, \epsilon_n b_n),
\]

\[
D = \phi(\infty).
\]
with

\[ b_i = (-1)^n \epsilon_i p_{i,n-1}, \]
\[ c_i = (-1)^{n-1} p_{i,n-1} = -\epsilon_i b_i. \]

2. The balanced realization is sign symmetric. Specifically, with \( \epsilon_i = \lambda_i / \sigma_i \) and \( J = \text{diag}(\epsilon_1, \ldots, \epsilon_n) \), we have

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

3. Relative to a conformal block decomposition

\[ \Sigma = \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{pmatrix}, \quad A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \]

we have

\[ B\tilde{B} = \begin{pmatrix} A_{11} \Sigma_1 + \Sigma_1 \tilde{A}_{11} & A_{12} \Sigma_2 + \Sigma_1 \tilde{A}_{21} \\ A_{21} \Sigma_1 + \Sigma_2 \tilde{A}_{12} & A_{22} \Sigma_2 + \Sigma_2 \tilde{A}_{22} \end{pmatrix}. \]

4. With respect to the constructed balanced realization, we have the following representation:

\[ \frac{p_1^*(z)}{d(z)} = C(zI - A)^{-1}e_i. \]

Proof. 1: Setting \( D = \phi(\infty) \), we may assume without loss of generality that \( n/d \in H^\infty \) is strictly proper. In that case, obviously, \( n/d \in X^d \). We begin by computing the input map \( B = (b_1, \ldots, b_n)^- \) of the realization. Since \( \{ p_i^*/d, i = 1, \ldots, n \} \) is a basis for \( X^d \), we can write

\[ \frac{n}{d} = \sum_{i=1}^{n} b_i \frac{p_i^*}{d}, \]

or equivalently

\[ n = \sum_{i=1}^{n} b_i p_i^*. \]

We want to compute the \( b_i \). By orthogonality of the singular vectors of the
Hankel operator $H_{n/d}$ we have

$$b_i = \frac{(n/d, p_i^*/d)_{H^2}}{(p_i^*/d, p_i^*/d)_{H^2}}. \quad (107)$$

Now

$$np_i = \lambda_i d^* p_i^* + d\pi_i,$$

so

$$\frac{n}{d} \frac{p_i}{d^*} = \frac{\lambda_i}{d} \frac{p_i^*}{d^*} + \frac{\pi_i}{d^*}.$$

Integrating this equality over the imaginary axis, we have

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{n}{d} \frac{p_i}{d^*} \, dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \lambda_i \frac{p_i^*}{d} \, dt + \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{\pi_i}{d^*} \, dt.$$

Now

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{n}{d} \frac{p_i}{d^*} \, dt = \frac{1}{2\pi i} \int_{-\infty}^{+\infty} \frac{n}{d} \frac{p_i}{d^*} \, dz = \lim_{R \to \infty} \frac{1}{2\pi i} \int_{\gamma} \frac{n}{d} \frac{p_i}{d^*} \, dz$$

$$= \lim_{R \to \infty} \frac{1}{2\pi i} \int_{\hat{\gamma}} \frac{n}{d} \frac{p_i}{d^*} \, dz.$$

Here $\gamma$ and $\hat{\gamma}$ are the semicircular contours shown in Figure 4. Note that $\gamma$ is positively oriented whereas $\hat{\gamma}$ is negatively oriented. So, using the fact that $d$
HANKEL NORM AND BALANCED APPROXIMATIONS

is antistable, and taking $R$ large enough,

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{p_i}{d} \, dt = \frac{1}{2\pi i} \oint_{\gamma} \frac{p_i^*}{d} \, dz = \frac{1}{2\pi i} \oint_{\gamma} \frac{\pi_i}{d^*} \, dz.$$ 

Expanding $p_i^*/d$ at infinity, we have

$$\frac{p_i^*}{d} = (-1)^{n-1} \frac{p_{i,n-1}}{z} + \cdots$$

and therefore

$$\frac{1}{2\pi} \int_{\gamma} \frac{p_i^*}{d} \, dz = -(1)^{n-1} p_{i,n-1}.$$ 

In the same way, by expanding $\pi_i/d^*$ at infinity we compute the other integral. Thus we have

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{p_i}{d} \, dt = \frac{\lambda_i}{2\pi i} \oint_{\gamma} \frac{p_i^*}{d} \, dz = (-1)^n \lambda_i p_{i,n-1},$$

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{p_i}{d} \, dt = \frac{1}{2\pi i} \oint_{\gamma} \frac{\pi_i}{d^*} \, dz = (1)^n \pi_{i,n-1}.$$  

(108)

Of course the two results are the same: by Equation (44), we have

$$\pi_{i,n-1} = \lambda_i p_{i,n-1}.$$ 

Therefore we have, from Equation (107) and using the normalization (106),

$$b_i = \frac{(-1)^n \lambda_i p_{i,n-1}}{\sigma_i} = (-1)^n \epsilon_i p_{i,n-1}.$$  

(109)

Next we compute the output map $C$. This is simple, as

$$c_i = C \left( \frac{p_i^*}{d} \right) = \left( \frac{p_i^*}{d} \right)_{-1} = (-1)^{n-1} p_{i,n-1}.$$  

(110)
Comparing this with (109), we get
\[ c_i = -\epsilon_i b_i. \]  

To compute the generating matrix \( A \) we compute the matrix representation of \( S^d \) with respect to the basis \( \{ p_i^* / d, i = 1, \ldots, n \} \) of \( X^d \). We have
\[ S^d p_i^* \frac{d}{d} = \frac{zp_i^* - \xi_i d}{d}, \]
where \( \xi_i = (-1)^n \frac{1}{p_i, n-1} \). Let now \( S^d p_i^* / d = \sum_{j=1}^n a_{ji} p_j^* / d \). Once again using orthogonality, we get
\[ a_{ji} = \left( \frac{S^d p_i^*}{d}, \frac{p_j^*}{d} \right)_{H^2} = \left( \frac{zp_i^* - \xi_i d}{d}, \frac{p_j^*}{d} \right)_{H^2}. \]

Now, using the previous definition of the contours \( \gamma, \tilde{\gamma} \),
\[ \left( S^d p_i^*, \frac{p_j^*}{d} \right)_{H^2} = \left( \frac{zp_i^* - \xi_i d}{d}, \frac{p_j^*}{d} \right) = \frac{1}{2\pi \int_{-\infty}^{\infty} \frac{z p_i^* - \xi_i d}{d} \cdot \frac{p_j}{d^*} dt} = \lim_{n \to \infty} \frac{1}{2\pi i} \int_{\gamma} \frac{zp_i^* - \xi_i d}{d} \cdot \frac{p_j}{d^*} dz = \frac{1}{2\pi i} \int_{\tilde{\gamma}} \left\{ \frac{zp_i^* p_j}{dd^*} - \frac{\xi_i d}{d^*} \right\} dz \]
\[ = \frac{1}{2\pi i} \int_{\tilde{\gamma}} \frac{zp_i^* p_j}{dd^*} dz. \]
Now, using Equation (76), we have
\[ \frac{p_i p_j^*}{dd^*} = \frac{1}{\lambda_i - \lambda_j^*} \left( \lambda_j \frac{\alpha_{ij}^*}{d} + \lambda_i \frac{\alpha_{ij}}{d} \right). \]  

So
\[ \frac{zp_i p_j^*}{dd^*} = \frac{1}{\lambda_i - \lambda_j^*} \left( \lambda_j \frac{z \alpha_{ij}^*}{d} + \lambda_i \frac{z \alpha_{ij}}{d} \right). \]
Integrating this expression over the contour $\gamma$, we get

$$\int_\gamma \frac{zp_i p_j^*}{dd^*} \, dz = \frac{\lambda_i}{\lambda_i^2 - \lambda_j^2} \int_\gamma \frac{z\alpha_{ij}^*}{d} \, dz. \quad (114)$$

By increasing $R$ and expanding $z\alpha_{ij}^*/d$ at $\infty$, noting that $\deg \alpha_{ij} \leq n - 2$, we obtain

$$\lim_{R \to \infty} \int_\gamma \frac{zp_i p_j^*}{dd^*} = -\frac{\lambda_i}{\lambda_i^2 - \lambda_j^2} \alpha_{ij,n-2} (-1)^{n-2} = (-1)^{n-1} \frac{\lambda_i}{\lambda_i^2 - \lambda_j^2} \alpha_{ij,n-2}. \quad (115)$$

However, integrating over $\gamma$,

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{zp_i p_j^* - \xi_i d}{d} \cdot \frac{p_j}{d^*} \, dt$$

$$= \frac{1}{2\pi i} \int_\gamma \frac{zp_i^* - \xi_i d}{d} \cdot \frac{p_j}{d^*} \, dz$$

$$= \frac{1}{2\pi i} \int_\gamma \frac{zp_i^* p_j}{dd^*} \, dz - \frac{\xi_i}{2\pi i} \int_\gamma \frac{p_j}{d^*} \, dz$$

$$= \frac{\lambda_i}{\lambda_i^2 - \lambda_j^2} \int_\gamma \frac{z\alpha_{ij}}{d^*} \, dz - \xi_i (-1)^n p_{i,n-1}$$

$$= \frac{\lambda_i}{\lambda_i^2 - \lambda_j^2} (-1)^n \alpha_{ij,n-2} - (-1)^{n-1} p_{i,n-1} (-1)^n p_{j,n-1}$$

$$= \frac{\lambda_i}{\lambda_i^2 - \lambda_j^2} (-1)^n \alpha_{ij,n-2} + p_{i,n-1} p_{j,n-1}.$$}

Equating the two expressions, we get

$$(-1)^{n-1} \frac{\lambda_j}{\lambda_i^2 - \lambda_j^2} \alpha_{ij,n-2} = \frac{\lambda_i}{\lambda_i^2 - \lambda_j^2} (-1)^n \alpha_{ij,n-2} + p_{i,n-1} p_{j,n-1}. \quad (116)$$

So

$$(-1)^{n-1} \alpha_{ij,n-2} \left( \frac{\lambda_j}{\lambda_i^2 - \lambda_j^2} + \frac{\lambda_i}{\lambda_i^2 - \lambda_j^2} \right) = (-1)^{n-1} \frac{\alpha_{ij,n-2}}{\lambda_i - \lambda_j} = p_{i,n-1} p_{j,n-1},$$
or

\[ \alpha_{ij, n-2} = (-1)^{n-1} (\lambda_i - \lambda_j) p_{i, n-1} p_{j, n-1}. \]

Finally

\[
\frac{1}{2 \pi i} \int_{\gamma} \frac{zp_i^* p_j}{d d^*} dz = (-1)^{n-1} \frac{\lambda_j}{\lambda_i^2 - \lambda_j^2} \alpha_{ij, n-2}
\]

\[
= ( -1 )^{n-1} \frac{\lambda_j}{\lambda_i^2 - \lambda_j^2} ( -1 )^{n-1} ( \lambda_i - \lambda_j ) p_{i, n-1} p_{j, n-1}
\]

\[
= \frac{\lambda_j}{\lambda_i + \lambda_j} p_{i, n-1} p_{j, n-1}.
\]

Thus we have

\[
\left( \frac{S^d p_i^*}{d}, \frac{p_j^*}{d} \right)_{H^2} = \frac{\lambda_j}{\lambda_i + \lambda_j} p_{i, n-1} p_{j, n-1}.
\]

We use now the normalization (106) as well as the equalities (109) and (110) to get

\[
\alpha_{ij} = \frac{\left( S^d p_j^*/d, p_i^*/d \right)_{H^2}}{\left( p_j^*/d, p_j^*/d \right)_{H^2}} = \frac{1}{a_j} \frac{\lambda_j}{\lambda_i + \lambda_j} p_{i, n-1} p_{j, n-1}
\]

\[
= \frac{\epsilon_j}{\lambda_i + \lambda_j} (-1)^n \epsilon_i b_i (-1)^{n-1} c_j = \frac{\epsilon_j}{\lambda_i + \lambda_j} (-1)^n \epsilon_i b_i (-1)^{n-1} (-\epsilon_j b_j)
\]

\[
= \frac{\epsilon_i b_i b_j}{\lambda_i + \lambda_j}.
\]

(116)

Interchanging indices, we get

\[ a_{ji} = \frac{\epsilon_j b_j b_i}{\lambda_i + \lambda_j}, \]

and hence we get the relation

\[ \epsilon_i a_{ji} = a_{ij} \epsilon_j. \]

This means simply, with \( J = \text{diag}(\epsilon_1, \ldots, \epsilon_n) \), that \( JA = \tilde{A} J \). That \( C = \tilde{B} J \) is
clear from (111). Now, from (116), it follows that

\[ b_i b_j = \epsilon_j (\lambda_i + \lambda_j) a_{ij} = \lambda_i \epsilon_j a_{ij} + a_{ij} (\epsilon_j \lambda_j) \]

\[ = \lambda_i (\epsilon_i a_{ji}) + a_{ij} \sigma_j = \sigma_i a_{ji} + a_{ij} \sigma_j. \]

However, this is just the first (anti)-Liapunov equation in (105). The other Liapunov equation is proved analogously or can be derived from this one by signature symmetry. Thus we have proved that our realization is balanced.

2: The proof is a simple computation.

3: By Lemma 2.1 there exist vectors \( e'_i \) such that

\[ \frac{p_i^*(z)}{d(z)} = C(zI - A)^{-1} e'_i. \]

Now

\[ \frac{p_i^*(z)}{d(z)} \cdot \frac{p_i(z)}{d^*(z)} = \tilde{e}_i (-zI - \tilde{A})^{-1} \tilde{C} C(zI - A)^{-1} e'_i \]

\[ = \tilde{e}_i (-zI - \tilde{A})^{-1} [\Sigma (zI - A) + (-zI - \tilde{A}) \Sigma] (zI - A)^{-1} e'_j \]

\[ = \tilde{e}_i (-zI - \tilde{A})^{-1} \Sigma e'_j + \tilde{e}_i \Sigma (zI - A)^{-1} e'_j. \]

Integrating this identity, applying the reasoning used already several times, leads to

\[ \frac{1}{2 \pi i} \int_{-\infty}^{+\infty} \frac{p_i^*(z)}{d} \cdot \frac{p_i(z)}{d^*} \, dt = \tilde{e}_i \Sigma e'_j. \]

However, by the orthogonality relation (75) and the chosen normalization \( \sigma_i = (p_i^*/d, p_i^*/d)_{H^2} \), we have

\[ \frac{1}{2 \pi i} \int_{-\infty}^{+\infty} \frac{p_i^*(z)}{d} \cdot \frac{p_i(z)}{d^*} \, dt = \sigma_i \delta_{ij}. \]

Thus

\[ \sigma_i \delta_{ij} = \tilde{e}_i \Sigma e'_j, \]

and so necessarily \( e'_i = e_i, i = 1, \ldots, n \)
We can apply the same method to obtain a balanced realization for the Nehari extension of a strictly proper rational function \( n/d \in H^\infty \).

**Theorem 8.2.** Let \( \phi = n/d \in H^\infty \), and let \( \pi_1/p_1 \) of Theorem 3.4 be the Nehari extension of \( n/d \). Let \((A, B, C)\) be the balanced realization of \( \phi \) given by Theorem 8.1. Then \( \pi_1/p_1 \) admits a balanced realization \((A_N, B_N, C_N)\) with

\[
A_N = -\left\{ \frac{\epsilon_j \mu_j b_j}{\lambda_i + \lambda_j} \right\}, \\
B_N = (\mu_2 b_2, \ldots, \mu_nb_n), \\
C_N = (\mu_2 \epsilon_2 b_2, \ldots, \mu_n \epsilon_n b_n), \\
D_N = \lambda_1,
\]

where

\[
\mu_i = \sqrt{\frac{\lambda_i - \lambda_i}{\lambda_i + \lambda_i}} \quad \text{for} \quad i = 2, \ldots, n.
\]

**Proof.** We will use the shift realization of Section 2 and compute a matrix representation with respect to a suitable normalization of the basis \( \{\beta_i/p_1 | i = 2, \ldots, n\} \) of \( X_{p_1} \) consisting of the singular vectors of \( H_{\pi_1/p_1}^* \).

Thus we put \( f_i = \nu_i \beta_i/p_1, \quad i = 2, \ldots, n \). As in Theorem 8.1, we will use the normalization \( \|f_i\|^2 = \sigma_i, \quad i = 2, \ldots, n \). Now, from Equation (87) we have

\[
\frac{\|\beta_i\|^2}{\|p_1\|^2} = \frac{p_i d^*}{\|d^*\|^2} \left( 1 - \frac{\lambda_i^2}{\lambda_1^2} \right) = \sigma_i \left( 1 - \frac{\lambda_i^2}{\lambda_1^2} \right) = \sigma_i \left( \frac{\|\lambda_i 2 - \lambda_i 2\|}{\|\lambda_i 2\|} \right),
\]

as \( \|p_i/d^*\|^2 = \sigma_i \). Hence

\[
\|f_i\|^2 = \sigma_i \nu_i^2 \left( \frac{\lambda_i 2 - \lambda_i 2}{\lambda_1 2} \right),
\]

or

\[
\frac{\|\beta_i\|^2}{\|p_1\|^2} = \frac{\sigma_i}{\nu_i^2} \quad (118)
\]
and

\[ v_i = \sqrt{\frac{|\lambda_i|^2}{|\lambda_i|^2 - |\lambda_i|^2}}. \] (119)

We proceed to compute the realization of \( \pi_1 / p_1 \).

For the constant term we have, by Equation (44),

\[ D_N = \frac{\pi_1}{p_1}(\infty) = \frac{\pi_{1,n-1}}{p_{1,n-1}} = \lambda_1. \]

Next we compute the output map \( C_N \). Now \( C_N = (C_{N,2}, \ldots, C_{N,n}) \) with

\[ C_{N,i} = C_{N,f_i} = (f_i)_{-1} = \left( \frac{v_i}{p_i} \right)_{-1} = \frac{v_i}{p_{1,n-1}}, \quad i = 2, \ldots, n. \] (120)

Now, by Equation (83), we have, computing the highest degree coefficient,

\[ \lambda_1(-1)^{n-1} p_{1,n-1} p_{1,n-1} - \lambda_i p_{1,n-1}(-1)^{n-1} p_{1,n-1} = \lambda_1 \beta_{i,n-2}, \]

or

\[ \frac{\beta_{i,n-2}}{p_{1,n-1}} = (-1)^{n-1} \frac{\lambda_1 - \lambda_i}{\lambda_1} p_{i,n-1}. \] (121)

Hence

\[ C_{N,i} = v_i \frac{\beta_{i,n-2}}{p_{1,n-1}} = \sqrt{\frac{|\lambda_1|^2}{|\lambda_i|^2 - |\lambda_i|^2}} (-1)^{n-1} \frac{\lambda_1 - \lambda_i}{\lambda_1} p_{i,n-1} \]

\[ = \frac{|\lambda_1|}{\lambda_1} \sqrt{\frac{\lambda_1 - \lambda_i}{(\lambda_1 - \lambda_i)(\lambda_1 + \lambda_i)}} (-1)^{n-1} p_{i,n-1} \]

\[ = \varepsilon_1 \frac{\lambda_1 - \lambda_i}{\lambda_1 + \lambda_i} (-1)^{n-1} p_{i,n-1} \]

\[ = \varepsilon_1 \mu_i (-1)^{n-1} p_{i,n-1} = \varepsilon_1 \mu_i c_i. \]
To compute the input map $B$ we note that, for $\xi \in \mathbb{R}$, and using Equation (44),

$$B\xi = \pi \frac{\pi_1}{p_1} \xi = \left( \frac{\pi_1}{p_1} - \frac{\pi_1}{p_1}(\infty) \right) \xi$$

$$= \left( \frac{\pi_1}{p_1} - \frac{\pi_1, n-1}{p_1, n-1} \right) \xi = \left( \frac{\pi_1}{p_1} - \lambda_1 \right) \xi = \left( \frac{\pi_1 - \lambda_1 p_1}{p_1} \right) \xi.$$ 

So we have to compute the coefficients $b_{N,i}, i = 2, \ldots, n$, in the expansion

$$\frac{\pi_1 - \lambda_1 p_1}{p_1} = \sum b_{N,i} f_i = \sum b_{N,i} \mu_i \frac{\beta_i}{p_1}.$$ 

Using the orthogonality, in $H_+^2$, of the functions $\beta_i/p_1$, we have

$$b_{N,i} = \frac{\left( \frac{\pi_1 - \lambda_1 p_1}{p_1}, \frac{\nu_i \beta_i}{p_1} \right)_{H_+^2}}{\left( \frac{\nu_i \beta_i}{p_1}, \frac{\nu_i \beta_i}{p_1} \right)_{H_+^2}} = \frac{1}{\nu_i} \frac{\left( \frac{\pi_1 - \lambda_1 p_1}{p_1}, \frac{\beta_i}{p_1} \right)_{H_+^2}}{\left( \frac{\beta_i}{p_1}, \frac{\beta_i}{p_1} \right)_{H_+^2}}.$$ 

Using once again Equation (118), we have

$$b_{N,i} = \frac{\nu_i}{\sigma_i} \left( \frac{\pi_1 - \lambda_1 p_1}{p_1}, \frac{\beta_i}{p_1} \right)_{H_+^2} = \frac{1}{\sigma_i} \left| \frac{\lambda_1}{\sqrt{\lambda_1^2 - \lambda_1^2}} \left( \frac{\pi_1 - \lambda_1 p_1}{p_1}, \frac{\beta_i}{p_1} \right)_{H_+^2} \right|.$$ 

To compute the inner product we transform the integral over the imaginary axis to a contour integral, over the contour $\gamma$, using the fact that $\deg(\pi_1 - \lambda_1 p_1) \beta_i^* \leq \deg p_1 p_1^* - 2$, to get

$$\left( \frac{\pi_1 - \lambda_1 p_1}{p_1}, \frac{\beta_i}{p_1} \right)_{H_+^2} = \frac{1}{2 \pi} \int_{-\infty}^{\infty} \frac{\pi_1 - \lambda_1 p_1}{p_1} \frac{\beta_i^*}{p_1^*} dt$$

$$= \frac{1}{2 \pi i} \int_{\gamma} \frac{\pi_1 - \lambda_1 p_1}{p_1} \frac{\beta_i^*}{p_1^*} dz$$

$$= \frac{1}{2 \pi i} \int_{\gamma} \frac{\pi_1}{p_1} \frac{\beta_i^*}{p_1^*} dz - \frac{\lambda_1}{2 \pi i} \int_{\gamma} \frac{\beta_i^*}{p_1^*} dz$$

$$= \frac{1}{2 \pi i} \int_{\gamma} \frac{\pi_1}{p_1} \frac{\beta_i^*}{p_1^*} dz.$$
for the other integral vanishes by Cauchy's theorem and the fact that $\beta_i^*/p_i^* \in H^2$.

In order to evaluate the last integral we use Equation (97) rewritten as

$$\frac{\pi^+ \beta_i}{p_i^*} = \lambda_i \frac{\beta_i^*}{p_i^*} + \frac{\omega_i}{p_1}, \quad (122)$$

or dually as

$$\frac{\pi_1 \beta_i^*}{p_1^*} = \lambda_i \frac{\beta_i}{p_1} + \frac{\omega_i^*}{p_1^*}. \quad (123)$$

Now $\omega^*/p_i^* \in H^2$, and so, integrating the previous equality over the contour $\gamma$ and expanding at $\infty$, we are led to

$$\frac{1}{2\pi i} \int_{\gamma} \frac{\pi_1 \beta_i^*}{p_1^*} \, dz = \frac{\lambda_i}{2\pi i} \int_{\gamma} \frac{\beta_i}{p_1} \, dz = \lambda_i \frac{\beta_{i,n-2}}{p_{1,n-1}}. \quad (124)$$

To evaluate the last term we use Equation (86), i.e.

$$\frac{p_i^*}{d^*} = \frac{\beta_i^*}{p_1} + \frac{\lambda_i}{p_1} \frac{p_i}{d^*}. \quad (124)$$

Expansion at $\infty$ yields

$$\frac{(-1)^{n-1}}{(-1)^n} p_{i,n-1} = \frac{(-1)^{n-2} \beta_{i,n-2}}{p_{1,n-1}} + \frac{\lambda_i (-1)^{n-1} p_{i,n-1}}{(-1)^n},$$

or

$$- \left( 1 - \frac{\lambda_i}{\lambda_1} \right) p_{i,n-1} = \frac{(-1)^{n-2} \beta_{i,n-2}}{p_{1,n-1}},$$

and so

$$\frac{\beta_{i,n-2}}{p_{1,n-1}} = (-1)^{n-1} \left( 1 - \frac{\lambda_i}{\lambda_1} \right) p_{i,n-1}.$$ 

Note that $1 - \lambda_i/\lambda_1 = (\lambda_1 - \lambda_i)/\lambda_1 > 0$. 
Putting all this together, we have

\[
\begin{align*}
    b_{N,i} &= \frac{1}{\sigma_i} \left| \lambda_i \right| \pi_1 \left( \frac{\beta_i}{p_1}, \frac{\nu_i \beta_i}{p_1} \right)_{H^*_i} = \frac{1}{\sigma_i} \left| \lambda_i \right| \frac{\beta_i, n-2}{p_{1, n-1}} \\
    &= \frac{1}{\sigma_i} \left| \lambda_i \right| \lambda_i (-1)^{n-1} \left( \frac{\lambda_1 - \lambda_i}{\lambda_1} \right) p_{i, n-1} \\
    &= \varepsilon_i (-1)^{n-1} \sqrt{\frac{\lambda_1 - \lambda_i}{\lambda_1 + \lambda_i}} p_{i, n-1} \\
    &= -\mu_i b_i.
\end{align*}
\]

Our last step is the derivation of the generator matrix \( A_N \). Actually the result follows from Theorem 8.1, but we prefer to give also a direct computational derivation.

Since \( \pi_1 / p_1 \) is realized in the state space \( X^{p_1} \), we compute a matrix representation of \( S^{p_1} \) with respect to the basis \( \{ f_i = \nu_i \beta_i / p_1 \mid i = 2, \ldots, n \} \). We have

\[
S^{p_1} \frac{\beta_i}{p_1} = \pi \frac{z \beta_i}{p_1} = \frac{z \beta_i - \eta_i p_1}{p_1},
\]

where

\[
\eta_i = \frac{\beta_i, n-2}{p_{1, n-1}}.
\]

Let now \( S^{p_1} f_i = \sum_{j=2}^n a_{N, ji} f_j \). Using \( H^2_+ \) orthogonality as well as the normalization \( \| f_j \|^2 = \sigma_j \), we have

\[
\begin{align*}
    a_{N, ji} &= \left( S^{p_1} f_i, f_j \right)_{H^*_i} = \frac{1}{\sigma_j} \left( S^{p_1} f_i, f_j \right)_{H^*_i} = \frac{\nu_i \nu_j}{\sigma_j} \left( S^{p_1} \frac{\beta_i}{p_1}, \frac{\beta_j}{p_1} \right)_{H^*_i} \\
    &= \frac{\nu_i \nu_j}{\sigma_j} \left( \frac{z \beta_i - \eta_i p_1}{p_1}, \frac{\beta_j}{p_1} \right)_{H^*_i} \\
    &= \frac{\nu_i \nu_j}{\sigma_j} \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{z \beta_i - \eta_i p_1}{p_1} \frac{\beta_j^*}{p_1^*} dt \\
    &= \frac{\nu_i \nu_j}{\sigma_j} \lim_{R \to \infty} \frac{1}{2\pi i} \int_{\gamma} \frac{z \beta_i - \eta_i p_1}{p_1} \frac{\beta_j^*}{p_1^*} dz \\
    &= \frac{\nu_i \nu_j}{\sigma_j} \lim_{R \to \infty} \frac{1}{2\pi i} \int_{\gamma} \frac{z \beta_i - \eta_i p_1}{p_1} \frac{\beta_j^*}{p_1^*} dz.
\end{align*}
\]
Now, using the fact that $p_1$ is a stable polynomial, we have

$$\lim_{R \to \infty} \frac{1}{2\pi i} \int_{\gamma} \frac{z\beta_i - \eta_i p_1}{p_1} \frac{\beta_i^*}{\beta_i^*} \, dz = \frac{1}{2\pi i} \int_{\gamma} \frac{z\beta_i \beta_i^*}{p_1} \, dz - \frac{\eta_i}{2\pi i} \int_{\gamma} \frac{\beta_i^*}{\beta_i^*} \, dz,$$

whereas

$$\lim_{R \to \infty} \frac{1}{2\pi i} \int_{\gamma} \frac{z\beta_i - \eta_i p_1}{p_1} \frac{\beta_i^*}{p_i^*} \, dz = \frac{1}{2\pi i} \int_{\gamma} \frac{z\beta_i \beta_i^*}{p_1} \, dz - \frac{\eta_i}{2\pi i} \int_{\gamma} \frac{\beta_i^*}{p_i^*} \, dz.$$

In order to evaluate the necessary integrals we start from Equation (97), i.e. from

$$\pi_1 \beta_i^* = \lambda_i p_i^* \beta_i^* + p_1 \omega_i^*.$$

$$\pi_1 \beta_j^* = \lambda_j p_j^* \beta_j^* + p_1 \omega_j^*.$$  \hspace{1cm} (126)

We are led to

$$0 = p_1^*(\lambda_i \beta_i \beta_i^* - \lambda_j \beta_j \beta_j^*) + p_1(\omega_i^* \beta_j^* - \omega_j^* \beta_i^*).$$

In particular there exist polynomials $p_{ij}$ of degree \( \leq n - 3 \) such that

$$\lambda_i \beta_i \beta_i^* - \lambda_j \beta_j \beta_j^* = p_1 p_{ij},$$

and hence

$$\beta_i \beta_i^* = \frac{\lambda_j}{\lambda_i} \beta_j \beta_j^* + \frac{1}{\lambda_i} p_1 p_{ij}.$$

Exchanging indices and taking adjoints, we get

$$\beta_i \beta_j^* = \frac{\lambda_i}{\lambda_j} \beta_j \beta_j^* + \frac{1}{\lambda_j} p_j^* p_{ij}^*.$$
Equating the last two expressions, we get

$$\left( \frac{\lambda_i \lambda_j - \lambda_j \lambda_i}{\lambda_i - \lambda_j} \right) \beta_j \beta_i^* = \frac{1}{\lambda_i} p_{ij} \rho_{ij} - \frac{1}{\lambda_j} p_{i \rho_{ij}^*}, \quad (127)$$

and, by an argument similar to that used in the derivation of Equation (78), we have

$$\rho_{ij} = -\rho_{ji}. \quad (128)$$

So, from (127), it follows that

$$\frac{\beta_j \beta_i^*}{p_{ij}^*} = \frac{\lambda_j \rho_{ij}}{\lambda_i^2 - \lambda_j^2} + \frac{\lambda_i \rho_{ij}^*}{\lambda_i^2 - \lambda_j^2} p_{ij}$$

and also

$$\frac{z \beta_j \beta_i^*}{p_{ij}^*} = \frac{\lambda_j z \rho_{ij}}{\lambda_i^2 - \lambda_j^2} + \frac{\lambda_i z \rho_{ij}^*}{\lambda_i^2 - \lambda_j^2} p_{ij}.$$

Integrating over the contour $\gamma$, we get

$$\frac{1}{2 \pi i} \int_\gamma \frac{z \beta_j \beta_i^*}{p_{ij}^*} \, dz = (-1)^{n-3} \frac{\lambda_j \rho_{ij, n-3}}{\lambda_i^2 - \lambda_j^2} p_{i, n-1}.$$

However, integrating over the contour $\hat{\gamma}$ and using Equation (125), we get

$$\frac{1}{2 \pi i} \int_{\hat{\gamma}} \frac{z \beta_j \beta_i^*}{p_{ij}^*} \, dz - \frac{\eta_i}{2 \pi i} \int_{\hat{\gamma}} \frac{\beta_i^*}{p_{ij}^*} \, dz = \frac{1}{2 \pi i} \int_{\hat{\gamma}} \frac{\lambda_j \, z \rho_{ij}}{\lambda_i^2 - \lambda_j^2} \, dz - \frac{\eta_i}{2 \pi i} \int_{\hat{\gamma}} \frac{\beta_j^*}{p_{ij}^*} \, dz$$

$$= - \frac{\lambda_j}{\lambda_i^2 - \lambda_j^2} (-1)^{n-1} \rho_{ij, n-3} \frac{p_{i, n-1}}{p_{i, n-1}} - (-1)^{n-2} \frac{\lambda_j (-1)^{n-2} \beta_{i, n-2}}{-p_{i, n-1}}$$

$$= (-1)^{n-2} \frac{\lambda_i \rho_{ij, n-3}}{\lambda_i^2 - \lambda_j^2} p_{i, n-1} - \frac{\beta_{i, n-2} \beta_{i, n-2}}{p_{i, n-1}}.$$
Equating the two expressions implies

\[ (-1)^{n-3} \rho_{ij, n-3} \left( \frac{\lambda_i}{\lambda_i^2 - \lambda_j^2} + \frac{\lambda_j}{\lambda_i^2 - \lambda_j^2} \right) = - \frac{\beta_{i, n-2} \beta_{j, n-2}}{p_{1, n-1}} \]

and so

\[ \rho_{ij, n-3} = (-1)^{n-2} \left( \lambda_i - \lambda_j \right) \frac{\beta_{i, n-2} \beta_{j, n-2}}{p_{1, n-1}}. \]

Finally, substituting back and using (120), we get

\[
a_{N, H} = \frac{\nu_i \nu_j}{\sigma_j} \frac{1}{2 \pi i} \int \frac{z \beta_i \beta_j^*}{p_1 \ p_1^*} \ dz = \frac{\nu_i \nu_j}{\sigma_j} (-1)^{n-3} \left( \frac{\lambda_i}{\lambda_i^2 - \lambda_j^2} \right) \frac{\rho_{ij, n-3}}{p_{1, n-1}}
\]

\[
= - \frac{\nu_i \nu_j}{\sigma_j} \frac{\lambda_j}{\lambda_i^2 - \lambda_j^2} \left( \lambda_i - \lambda_j \right) \frac{\beta_{i, n-2} \beta_{j, n-2}}{p_{1, n-1}}
\]

\[
= - \frac{\nu_i \nu_j}{\sigma_j} \frac{\beta_{i, n-2} \beta_{j, n-2}}{p_{1, n-1}^2}.
\]

Now, by equation (121) and using (119),

\[
- \epsilon_i \frac{\nu_i \nu_j}{\sigma_j} \frac{\beta_{i, n-2} \beta_{j, n-2}}{p_{1, n-1}^2}
\]

\[
= - \epsilon_i \frac{\nu_i \nu_j}{\sigma_j} \frac{\lambda_i}{\lambda_i + \lambda_j} \frac{\lambda_i}{\sqrt{\lambda_i^2 - \lambda_j^2}} \frac{\lambda_j}{\sqrt{\lambda_i^2 - \lambda_j^2}} \left( \lambda_i - \lambda_j \right) \frac{p_{i, n-1} \ p_{j, n-1}}{p_{i, n-1} \ p_{j, n-1}}
\]

\[
= - \epsilon_i \mu_i \mu_j \frac{\lambda_i}{\lambda_i + \lambda_j} \frac{\lambda_j}{\lambda_i + \lambda_j} \left( \lambda_i - \lambda_j \right) \frac{p_{i, n-1} \ p_{j, n-1}}{p_{i, n-1} \ p_{j, n-1}}
\]

\[
= - \frac{\epsilon_i \nu_i \nu_j}{\lambda_i + \lambda_j} \left( \lambda_i - \lambda_j \right) \frac{p_{i, n-1} \ p_{j, n-1}}{p_{i, n-1} \ p_{j, n-1}}
\]

So, interchanging indices,

\[ a_{N, ij} = - \frac{\epsilon_i \nu_i \nu_j b_i b_j}{\lambda_i + \lambda_j}. \]
9. ALL-PASS TRANSFER FUNCTIONS

We continue the study of balanced realizations by constructing such a realization for the case of an asymptotically stable all-pass function.

Thus let $g = d^*/d$ be inner in $H^\infty$. This means that $d$ is stable and the singular values satisfy $\sigma_1 = \cdots = \sigma_n = 1$. Then

$$d(z) = d_+(z^2) + zd_-(z^2). \quad (129)$$

$$d^*(z) = d_+(z^2) - zd_-(z^2). \quad (130)$$

Then

$$g = \frac{d^* - d}{d} + 1 = \frac{-2zd_-(z^2)}{d_+(z^2) + zd_-(z^2)} + 1.$$

Next we put

$$f = \frac{-zd_-(z^2)}{d_+(z^2)},$$

i.e., $f$ is constant output feedback equivalent to $g$. Hence, once we get a canonical form for $f$, the one for $g$ will easily follow. Next we introduce yet another auxiliary strictly proper rational function by defining

$$h = \frac{-zd_-(z^2)}{d_+(-z^2)}.$$

By a well-known theorem (see Gantmacher, 1959, p. 000) the Cauchy index $l_h$ of $h$ is equal to $n = \deg d$. This piece of information implies two specific representations for $h$, one additive and the other a continued fraction representation. For the additive representation we must have $n$ real simple poles for $h$ with positive residues. Thus

$$h(z) = \sum_{k=1}^{n} \frac{\gamma_k^2}{z - \alpha_k}$$

with $\alpha_1 < \cdots < \alpha_n$. The other representation, and the more interesting one
for our purposes, is a continued-fraction representation of the form

\[ h(z) = \frac{\beta_0}{a_1(z) - \frac{\beta_1}{a_2(z) - \frac{\beta_2}{a_3(z) - \cdots - \frac{\beta_{n-2}}{a_{n-1}(z) - \frac{\beta_{n-1}}{a_n(z)}}}}}, \]

where \( a_i(z), i = 1, \ldots, n \) are monic of degree one and all the \( \beta_i \) are positive.

We will distinguish between two cases.

**Case 1.** Assume that \( n = \deg d \) is even. Therefore the function \( h \) has an even denominator and an odd numerator. Thus we can be more explicit as far as the continued fraction is concerned. Indeed, put \( q_{-1}(z) = d_1(-z^2) \) and \( q_0(z) = -zd_{-1}(-z^2) \), and set up the Euclidean algorithm with the recursion

\[ q_{i+1}(z) = a_i(z)q_i(z) - q_{i+1}(z), \quad i = 1, \ldots, n. \]

Let \( a_i = 2 - \alpha_i, i = 1, \ldots, n \). We will show by induction that the sequence of polynomials is alternatingly even and odd. In fact the initialization ensures that \( q_{-1} \) is even and \( q_0 \) is odd. So assume parity alternates for all indices till \( i \). From Equation (131) it follows that

\[ \frac{1}{q_{i-1}(z)} = \frac{\beta_i}{a_i(z)q_i(z) - q_{i+1}(z)} \]

\[ \frac{q_i(z)}{q_{i-1}(z)} = \frac{\beta_i}{a_i(z) - \frac{q_{i+1}(z)}{q_i(z)}}. \]

Now \( q_{i}(z)/q_{i-1}(z) \) is odd by the induction hypothesis. So \( a_i(z) - q_{i+1}(z)/q_{i}(z) \) is also odd. This forces both \( a_i(z) \) and \( q_{i+1}(z)/q_{i}(z) \) both to be odd. Thus \( a_i(z) = z \), and \( q_{i+1} \) has opposite parity to \( q_i \). Thus \( h \) has the
representation

\[ h(z) = \frac{\alpha_0^2}{z - \frac{\alpha_1^2}{z - \frac{\alpha_2^2}{z - \cdots - \frac{\alpha_{n-2}^2}{z - \frac{\alpha_{n-1}^2}{z}}}}} \]  

(134)

By Theorem 9.4 in Helmke and Fuhrmann (1989) \( h(z) \) has a realization of the form

\[
\begin{pmatrix}
0 & \alpha_1^2 & 0 \\
1 & \ddots & \ddots \\
& \ddots & \ddots & 0 \\
& & \ddots & \alpha_{n-1}^2 \\
& & & 1 & 0
\end{pmatrix}
, \begin{pmatrix}
0 \\
\vdots \\
\alpha_0 \\
0 & \cdots & \alpha_0
\end{pmatrix}
\]

which is similar, through the matrix \( \text{diag}(\rho_0, \ldots, \rho_{n-1}) \) with \( \rho_0 = 1 \) and \( \rho_{i+1}/\rho_i = \alpha_i^2 \), to the matrix

\[
\begin{pmatrix}
0 & \alpha_1 & 0 \\
\alpha_1 & \ddots & \ddots \\
& \ddots & \ddots & 0 \\
& & \ddots & \alpha_{n-1} \\
& & & \alpha_{n-1} & 0
\end{pmatrix}
, \begin{pmatrix}
0 \\
\vdots \\
\alpha_0 \\
0 & \cdots & \alpha_0
\end{pmatrix}
\]

To go back from \( h \) to \( f \) we note that

\[ f(z) = -ih(iz) = -i(-iz) \frac{d_-(z^2)}{d_+(z^2)}. \]

So \( f \) has only simple imaginary axis zeros.
From the continued fraction expansion (134) we get

\[ f(z) = \frac{-i\alpha_0^2}{iz - \frac{\alpha_1^2}{iz - \frac{\alpha_2^2}{iz - \frac{\alpha_n-2^2}{iz - \frac{\alpha_n-1^2}{iz}}}}} = \frac{-\alpha_0^2}{iz - \frac{\alpha_1^2}{z + \frac{\alpha_2^2}{z + \frac{\alpha_n-2^2}{z + \frac{\alpha_n-1^2}{z}}}}} \]

(135)

Using again the continued fraction canonical form, we get a realization for \( f \) given by

\[
\begin{pmatrix}
0 & \alpha_1 & 0 \\
-1 & \cdots & \cdots \\
\alpha_n & \cdots & 0 \\
-1 & 0
\end{pmatrix}
\begin{pmatrix}
0 \\
\vdots \\
\alpha_0 \\
0
\end{pmatrix}
\begin{pmatrix}
0 & \cdots & \alpha_0
\end{pmatrix}
\]

and by similarity to

\[
\begin{pmatrix}
0 & \alpha_1 & 0 \\
-\alpha_1 & \cdots & \cdots \\
\alpha_n & \cdots & 0 \\
-\alpha_n & 0
\end{pmatrix}
\begin{pmatrix}
0 \\
\vdots \\
\alpha_0 \\
0
\end{pmatrix}
\begin{pmatrix}
0 & \cdots & \alpha_0
\end{pmatrix}
\]

Now \( g \) is obtained back from \( f \) by the inverse output-feedback transforma-
tion. Thus $g$ is realized by
\[
\begin{pmatrix}
0 & \alpha_1 & 0 \\
-\alpha_1 & \ddots & \ddots \\
& \ddots & \ddots & \ddots \\
& & \ddots & \ddots & 0 \\
& & & -\alpha_{n-1} & -k
\end{pmatrix},
\]
with $k = \alpha_0^2/2$.

Clearly the realization in (136) is balanced.

Case 2. Assume that $n$ is odd. This case can be treated similarly; the details will be skipped.

We note in passing that the above realization is a special case of Darlington synthesis; see Brockett (1970) and Krishnaprasad (1980).

10. BALANCED MODEL REDUCTION

In this section we study the approach to model reduction that is based on the truncation of the last mode in a balanced realization. Since the construction of balanced realizations was done utilizing a basis made out of singular vectors of the Hankel operator, it is not surprising that there is a connection between the Hankel norm based reduced and the balanced reduction. This connection was first pointed out by Glover (1984). Here the connection is highlighted from a geometric perspective.

We begin with two very useful computations. For these I am indebted to R. Ober.

**Lemma 10.1.** Let
\[
A = \begin{pmatrix}
A_{11} & A_{12} \\
A_{21} & A_{22}
\end{pmatrix}
\]
and \( A_{11} \) be square matrices. Set

\[
W(s) = sI - A_{22} - A_{21}(sI - A_{11})^{-1}A_{12}.
\]

Then:

1. We have

\[
(sI - A)^{-1} = \begin{bmatrix} I & (sI - A_{11})^{-1}A_{12} \\ 0 & I \end{bmatrix}
\times \begin{bmatrix} (sI - A_{11})^{-1} & 0 \\ 0 & W(s)^{-1} \end{bmatrix}
\times \begin{bmatrix} I & 0 \\ A_{21}(sI - A_{11})^{-1} & I \end{bmatrix}.
\]

2. If \( A \) is \( n \times n \) and \( A_{11} \) is \((n - 1) \times (n - 1)\), then

\[
W(s) = \frac{\det(sI - A)}{\det(sI - A_{11})}.
\]

Proof. Part 1 follows because

\[
sI - A = \begin{bmatrix} I & 0 \\ -A_{21}(sI - A_{11})^{-1} & I \end{bmatrix}
\times \begin{bmatrix} sI - A_{11} & 0 \\ 0 & W(s) \end{bmatrix}
\times \begin{bmatrix} I & 0 \\ 0 & -(sI - A_{11})^{-1}A_{12} \end{bmatrix}.
\]

2. Taking determinants in the above formula, we have

\[
\det(sI - A) = W(s) \det(sI - A_{11}).
\]

Lemma 10.2. Let \((A, B, C)\) be the balanced realization of \( n/d \) constructed in Theorem 8.1, let \((A, B, C)\) be conformally partitioned as

\[
A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \quad B = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, \quad C = \begin{pmatrix} C_1 & C_2 \end{pmatrix},
\]

with \( B_1 = (b_1, \ldots, b_{n-1})^T \), \( B_2 = b_n \), \( C_1 = (c_1, \ldots, c_{n-1}) \), and \( C_2 = c_n \), and let
\( \Sigma \) be conformally partitioned as

\[
\Sigma = \begin{pmatrix}
\Sigma_1 & 0 \\
0 & \Sigma_2
\end{pmatrix}
\]

with \( \Sigma_1 = \text{diag}(\sigma_1, \ldots, \sigma_{n-1}) \) and \( \Sigma_2 = \sigma_n \). Let \( \{e_1, \ldots, e_n\} \) be the standard basis in \( \mathbb{R}^n \). Let \((A_{11}, B_1, C_1)\) be the last mode truncation of the balanced system, and let

\[
ge_b(z) = \frac{n_b(z)}{d_b(z)} = C_1(zI - A_{11})^{-1}B_1. \tag{141}
\]

Then:

1. \( W(z) \) defined by Equation (137) satisfies

\[
W(z) = \frac{d(z)}{d(b(z))},
\]

or

\[
W(z)^{-1} = \frac{d_b(z)}{d(z)}.
\]

2. We have

\[
C_1(zI - A_{11})^{-1}e_n + C_2 = \frac{p_n^*(z)}{d_b(z)}. \tag{142}
\]

Proof. Part 1 follows from Equation (139).

2: From Equation (138) we obtain

\[
C(zI - A)^{-1}e_n = \begin{pmatrix}
C_1 & C_1(zI - A_{11})^{-1}A_{12} + C_2
\end{pmatrix} \begin{pmatrix}
0 \\
W(z)^{-1}
\end{pmatrix}
= \left[ C_1(zI - A_{11})^{-1}A_{12} + C_2 \right] W(z)^{-1}.
\]

However, by Theorem 8.1, \( p^*_n(z)/d(z) = C(zI - A)^{-1}e_i \) and \( W(z)^{-1} = d_b(z)/d(z) \), so (142) follows.
**Lemma 10.3.** Let $p_n/d^*$ be the $n$th Hankel singular vector. Then the following equation is satisfied:

$$
\sigma_n(d^*d_d + d^*d_b) = p_n p_n^*.
$$

**Proof.** Let $(A, B, C)$ be a balanced realization of $n/d$. We have seen that $p_n^*(z)/d(z) = C(zI - A)^{-1}e_n$. Using Lemma 10.2, we have

$$
\frac{p_n^*(z)}{d^*(z)} = W(z)^{-1} C (zI - A_{11})^{-1} A_{12} + C_e.
$$

So

$$
\frac{p_n(z) p_n^*(z)}{d(z) d^*(z)} = W(z)^{-1} W(-z)^{-1} \left( \tilde{A}_{12} (zI - \tilde{A}_{11})^{-1} 1 \right)
\times \tilde{C} C \left( (zI - A_{11})^{-1} A_{12} \right).
$$

By balancing we have

$$
\tilde{A} \Sigma + \Sigma A = -\tilde{C} C,
$$

which implies

$$
(-zI - \tilde{A}) \Sigma + \Sigma (zI - A) = \tilde{C} C.
$$

This allows the passage from a multiplicative to an additive representation. More specifically

$$
\left( \tilde{A}_{12} (zI - \tilde{A}_{11})^{-1} 1 \right) \tilde{C} C \left( (zI - A_{11})^{-1} A_{12} \right)
$$

$$
= \left( \tilde{A}_{12} (zI - \tilde{A}_{11})^{-1} 1 \right) \left[ (-zI - \tilde{A}) \Sigma + \Sigma (zI - A) \right] \left( (zI - A_{11})^{-1} A_{12} \right).
$$
Computing one of the terms,
\[
\left( \tilde{A}_{12}(zI - \tilde{A}_{11})^{-1} \ 1 \right) \Sigma(zI - A) \left( (zI - A_{11})^{-1} A_{12} \right)
\]
\[
= \left( \tilde{A}_{12}(zI - \tilde{A}_{11})^{-1} \ 1 \right) \left( \begin{array}{cc}
\Sigma_1 & 0 \\
n_0 & \sigma_n
\end{array} \right)
\times \left( \begin{array}{cc}
zI - A_{11} & -A_{12} \\
-A_{21} & zI - A_{22}
\end{array} \right) \left( (zI - A_{11})^{-1} A_{19} \right)
\]
\[
= \left( \tilde{A}_{12}(zI - \tilde{A}_{11})^{-1} \ 1 \right) \left( \begin{array}{cc}
zI - A_{11} & -A_{12} \\
-A_{21} & zI - A_{22}
\end{array} \right) \left( (zI - A_{11})^{-1} A_{19} \right)
\]
\[
= \left( \tilde{A}_{12}(zI - \tilde{A}_{11})^{-1} \ 1 \right) \left( \begin{array}{cc}
zI - A_{11} & -A_{12} \\
-A_{21} & zI - A_{22}
\end{array} \right) \left( (zI - A_{11})^{-1} A_{19} \right)
\]
\[
= \sigma_n W(z).
\]
The other term is computed similarly. Putting the two terms together, we are led to
\[
\frac{p_n(z)p_n(z)}{d(z)d^*(z)} = W(z)^{-1}W(-z)^{-1} \left( \tilde{A}_{12}(zI - \tilde{A}_{11})^{-1} \ 1 \right)
\times \tilde{C}C \left( (zI - A_{11})^{-1} A_{12} \right)
\]
\[
= W(z)^{-1}W(-z)^{-1} \left[ \sigma_n W(z) + \sigma_n W(-z) \right]
\]
\[
= \sigma_n \left( \frac{d_b(z)}{d(z)} + \frac{d_b^*(z)}{d^*(z)} \right),
\]
and this implies (143).

Remark 10.1. Equation (143) is the polynomial equivalent of the Lia-
punov equation. For a discussion of this we refer to Willems and Fuhrmann [1989].

**Corollary 10.1 (Pernebo and Silverman).** The polynomial $d_b$ is anti-
stable.

Now we are in a position to prove the main result of this section. This is a
slightly more specific derivation of Lemma 9.5 in Glover (1984), in the context
in which we are working. The extra information is the concrete representation
of the various all-pass functions that are encountered.

**Theorem 10.1.**

1. Let $g_b(s)$ be the balanced $n - 1$ first approximant of the stable transfer
   function $g(s)$. Then the approximation error $e(s) = g(s) - g_b(s)$ satisfies
   
   $$e(z) = \epsilon_n \frac{d}{d_b} \left( \frac{p_n^*}{d} \right)^2 = \lambda_n \frac{p_n^*}{p_n} \left( \frac{d^*}{d} + \frac{d_b^*}{d_b} \right),$$

   and hence
   
   $$|e(s)| \leq 2\sigma_n.$$

2. Let $g_h(s)$ be the $n - 1$ first Hankel norm approximant of the stable
   transfer function $g(s)$, i.e. $g_h(s) = g(s) - \lambda_n d^* p_n^*/dp_n$. Then the difference
   between these two approximants is all-pass; specifically,
   
   $$g_b - g_h = -\lambda_n \frac{p_n^* d_b^*}{p_n d_b}.$$

**Proof.** 1: From Equation (138) we get

$$g(s) = C(sI - A)^{-1} B$$

$$= \begin{pmatrix} C_1 & C_1(zI - A_{11})^{-1} A_{12} + C_2 \end{pmatrix} \begin{pmatrix} (zI - A_{11})^{-1} 0 \\ 0 & W(z)^{-1} \end{pmatrix}$$

$$\times \begin{pmatrix} B_1 \\ A_{21}(zI - A_{11})^{-1} B_1 + B_2 \end{pmatrix}$$

$$= C_1(sI - A)^{-1} B_1$$

$$+ \left[ C_1(zI - A_{11})^{-1} A_{12} + C_2 \right] W(z)^{-1} \left[ A_{21}(zI - A_{11})^{-1} B_1 + B_2 \right].$$

Since, by Theorem 8.1, the balanced realization is signature symmetric, with
the signature matrix \( J = \text{diag}(\epsilon_1, \ldots, \epsilon_n) \) which we write as \( \text{diag}(J_1, \epsilon_n) \) and \( J_1 = \text{diag}(\epsilon_1, \ldots, \epsilon_{n-1}) \). we have

\[
A_{21}(zI - A_{11})^{-1} B_1 + B_2 = \epsilon_n \left[ C_1(zI - A_{11})^{-1} A_{12} + C_2 \right].
\]

Hence for the error term \( e(z) \) we have

\[
e(z) = \epsilon_n W(z)^{-1} \left[ C_1(zI - A_{11})^{-1} A_{12} + C_2 \right]^2.
\]

But, by Lemma 10.2, \( C_1(zI - A_{11})^{-1} e_n + C_2 = p_n^*/d_b \), so

\[
e(z) = \epsilon_n W(z)^{-1} \left( \frac{p_n^*}{d_b} \right)^2 = \epsilon_n \frac{d_b}{d} \left( \frac{p_n^*}{d_b} \right)^2 = \epsilon_n \left( \frac{p_n^*}{d_b} \right)^2.
\]

On the other hand

\[
e(z) = g - g_b = \frac{n}{d} - \frac{n_b}{d_b}.
\]

2: We compute, applying Lemma 10.3,

\[
g_b - g_h = g_b - \left( g - \lambda_n \frac{d^* p_n^*}{dp_n} \right) = (g_b - g) + \lambda_n \frac{d^* p_n^*}{dp_n}
\]

\[
= -\epsilon_n \frac{p_n^*}{d_b} + \lambda_n \frac{d^* p_n^*}{dp_n} = -\epsilon_n \frac{p_n^*}{d} \left( \frac{p_n^*}{d_b} - \frac{\sigma_n}{p_n} \right)
\]

\[
= -\epsilon_n \frac{p_n^*}{d} \left( \frac{p_n^* p_n - \sigma_n d^* d_b}{d_b p_n} \right)
\]

\[
= -\epsilon_n \frac{p_n^*}{d} \frac{\sigma_n d_b^*}{d_b p_n} = -\lambda_n \frac{p_n^* d_b^*}{p_n d_b}.
\]

We go back now to the error term \( e \):

\[
e = g - g_b = (g - g_h) + (g_h - g_b)
\]

\[
= \lambda_n \frac{d^* p_n^*}{dp_n} + \lambda_n \frac{d_b^* p_n^*}{d_b p_n}
\]

\[
= \lambda_n \frac{p_n^*}{p_n} \left( \frac{d^*}{d} + \frac{d_b^*}{d_b} \right).
\]
Since $p_n^*/p_n$, $d^*/d$, and $d_b^*/d_b$ are all (antistable) all-pass functions, we get the estimate
\[ \| e \|_\infty \leq 2 \sigma_n. \]

**Corollary 10.2.** $n_b$ and $d_b$ satisfy the polynomial equation
\[ d_n - nd_b = -\epsilon_n \left( p_n^* \right)^2. \]

The formula above is "hidden" in the derivation by Enns (1984) of the error bound for balanced approximations.

In the next theorem we study the Hankel operator associated with the truncation of the balanced realization.

**Theorem 10.2.** Let $(A, B, C)$ be a balanced realization of $g = n/d$ as constructed in Theorem 8.1, and let $(A_{11}, B_1, C_1)$ be the last mode truncation of the balanced system, with

\[ g_b(z) = C_1(zI - A_{11})^{-1} B_1 = \frac{n_b(z)}{d_b(z)}. \]

Let the polynomials $q_i$, $i = 1, \ldots, n - 1$, be defined by
\[ \frac{q_i(z)}{d_b(z)} = C_1(zI - A_{11})^{-1} \tilde{e}_i, \]
where $\tilde{e}_i$ stands for the $i$th unit vector in the $(n - 1)$-dimensional Euclidean space. Then:

1. We have
\[ \frac{p_i}{d^*} - \frac{q_i}{d_b^*} \in \text{Ker } H_{x^*/n_n}. \]

2. $H_{g_b}$ has singular values $\sigma_1 > \cdots > \sigma_{n-1}$ and Schmidt pairs $\{ q_i^*/d_b, q_i/d_b^* \}$.

3. We have
\[ \frac{\alpha_i}{p_n^*} = P_{X^*/n_n} \frac{q_i}{d_b^*}. \]

4. The normalization condition
\[ \left\| \frac{q_i^*}{d_b} \right\|^2 = \sigma_i, \quad i = 1, \ldots, n - 1, \]
holds.
Proof. 1: We saw, in Theorem 8.1, that for the balanced realization 
\((A, B, C)\) of \(g\) we have

\[
\frac{p_i^*(z)}{d(z)} = C(zI - A)^{-1}e_i, \quad i = 1, \ldots, n. \tag{148}
\]

By the same token we have

\[
\frac{q_i^*(z)}{d(z)} = C_1(zI - A_{11})^{-1}e_i, \quad i = 1, \ldots, n - 1. \tag{149}
\]

Clearly \(\{q_i/d_b^*, q_i^*/d_b\}\) must be the \(\sigma_i\)-Schmidt pairs of \(H_{gs}\). We will prove this directly from the polynomial data. Notice that

\[
e_i = (\hat{e}_i \quad 0), \quad i = 1, \ldots, n - 1.
\]

Therefore, using Equation (138), we have

\[
\frac{p_i^*(z)}{d(z)} = C(zI - A)^{-1}e_i = \begin{pmatrix} C_1 & C_2 \end{pmatrix} \begin{pmatrix} I & (zI - A_{11})^{-1}A_{12} \\ 0 & I \end{pmatrix} \begin{pmatrix} (zI - A_{11})^{-1} & 0 \\ 0 & W(z)^{-1} \end{pmatrix}
\]

\[
= \begin{pmatrix} C_1 & C_2(zI - A_{11})^{-1}A_{11} + C_2 \end{pmatrix} \begin{pmatrix} (zI - A_{11})^{-1} & 0 \\ 0 & W(z)^{-1} \end{pmatrix}
\]

\[
\times \begin{pmatrix} \hat{e}_i \\ A_{21}(zI - A_{11})^{-1} \hat{e}_i \end{pmatrix}
\]

\[
= C_1(zI - A_{11})^{-1} \hat{e}_i + \left[ C_1(zI - A_{11})^{-1}A_{12} + C_2 \right] \times W(z)^{-1}A_{21}(zI - A_{11})^{-1}
\]

\[
= \frac{q_i^*(z)}{d_b(z)} + \frac{p_i^*(z)}{d_b(z)} \frac{d_b(z)}{d_y^*(z)} \left[ A_{21}(zI - A_{11})^{-1} \hat{e}_i \right].
\]
In the last step we used Lemma 10.2. Putting

\[ A_{21}(zI - A_{11})^{-1} e_i = \frac{r_i(z)}{d_b(z)}, \]

with \( \deg r_i < \deg d_b \), we get

\[ \frac{p_i^*}{d} = \frac{q_i^*}{d_b} + \frac{p_n^* r_i}{dd_b} = \frac{q_i^*}{d_b} + \frac{p_n^* p_n r_i}{p_n dd_b}, \]  \hspace{1cm} (150)

and equivalently

\[ \frac{p_i}{d^*} = \frac{q_i}{d_b^*} + \frac{p_n p_n^* r_i^*}{p_n^* d^* d_b^*}, \]  \hspace{1cm} (151)

This proves that

\[ \frac{p_i}{d^*} - \frac{q_i}{d_b^*} \in \frac{p_n}{p_n^*} H_{r^*} = \text{Ker} \ H_{*,n}/\nu = \text{Ker} \ H_{g^*}. \]

2: Applying the orthogonal projection on \( X_{p^*} \) to Equation (151), and noting that, by Theorem 5.1, \( P_{X_{p^*}} \frac{p_i}{d^*} = \alpha_i / p_n^* \), we get

\[ \frac{\alpha_i}{p_n^*} = P_{X_{p^*}} \frac{p_i}{d^*} - P_{X_{p^*}} \frac{q_i}{d_b^*}, \]  \hspace{1cm} (152)

i.e., for \( i = 1, \ldots, n - 1 \), \( p_i/d^* \) and \( q_i/d_b^* \) have the same projections on \( X_{p^*} \). Note also that from (151) we get

\[ (-1)^n p_{i,n-1} = (-1)^{n-1} q_{i,n-2}. \]  \hspace{1cm} (153)

3: We will show that

\[ H_{g^*} \frac{q_i}{d_b^*} = \lambda_i \frac{q_i^*}{d_b}, \quad i = 1, \ldots, n - 1. \]  \hspace{1cm} (154)

We saw, in Equation (151), that

\[ \frac{p_i}{d^*} = \frac{q_i}{d_b^*} + \frac{p_n}{p_n^*} f^* \]  \hspace{1cm} (155)
for some $f' \in H^2_+$. On the other hand, from Equation (71) we get

$$\frac{p_i}{d^*} = \frac{\alpha_i}{p_n^*} + \frac{\lambda_n}{\lambda_i} \frac{p_n}{p_n^*} \frac{p_i^*}{d^*}. \quad (156)$$

So from these two equations we get

$$\frac{q_i}{d_b^*} = \frac{\alpha_i}{p_n^*} \frac{p_n}{p_n^*} + p_n f \quad (157)$$

for some $f \in H^2_+$. Next we compute

$$H_{g_b} \frac{q_i}{d_b^*} = (H_{g_b} + H_{g_b - g_b}) \frac{q_i}{d_b^*} = H_{g_b} \frac{q_i}{d_b^*} + H_{g_b - g_b} \frac{q_i}{d_b^*}$$

$$= H_{g_b/p_n} \left( \frac{\alpha_i}{p_n^*} + \frac{p_n}{p_n^*} f \right) + H_{g_b - g_b} \frac{q_i}{d_b^*}. \quad (158)$$

Now $(p_n/p_n^*)f \in \text{Ker } H_{g_n/p_n}$, whereas, by Theorem 5.1,

$$H_{g_n/p_n} = \lambda_i \frac{\alpha_i^*}{p_n}.$$

Moreover, by Theorem 10.1,

$$g_b - g_b = -\lambda_n \frac{p_n^* d_b}{p_n d_b}.$$

So

$$H_{g_b - g_b} \frac{q_i}{d_b^*} = -\lambda_n \frac{p_n^* d_b}{p_n d_b} \frac{q_i}{d_b^*} = -\lambda_n \frac{p_n^*}{p_n} \frac{q_i}{d_b^*}.$$

Putting this together yields

$$H_{g_b} \frac{q_i}{d_b^*} = \lambda \left( \frac{\alpha_i^*}{p_n} - \frac{\lambda_n}{\lambda_i} \frac{p_n^*}{p_n} \frac{q_i}{d_b} \right). \quad (159)$$

We will show that

$$\frac{\alpha_i^*}{p_n} - \frac{\lambda_n}{\lambda_i} \frac{p_n^*}{p_n} \frac{q_i}{d_b} = \frac{q_i^*}{d_b}. \quad (160)$$

Since

$$\frac{\alpha_i^*}{p_n} - \frac{\lambda_n}{\lambda_i} \frac{p_n^*}{p_n} \frac{q_i}{d_b} \in \text{Im } H_{g_b} = X_{d_b},$$
we must have
\[ p_n \left| \lambda_i d_b \alpha_i^* - \lambda_n p_n^* q_i \right. \]
Thus, there exist polynomials \( \xi_i \), with \( \text{deg} \ \xi_i \leq n - 2 \), such that
\[ \lambda_i d_b \alpha_i^* - \lambda_n p_n^* q_i = \lambda_n p_n \xi_i^* \quad \text{(161)} \]
or equivalently
\[ \frac{\xi_i^*}{d_b} = \frac{\alpha_i^*}{p_n^*} - \frac{\lambda_n p_n^* q_i}{\lambda_i p_n d_b} \quad \text{(162)} \]
which, after conjugation, leads to
\[ \frac{\xi_i}{d_b} = \frac{\alpha_i}{p_n^*} - \frac{\lambda_n p_n q_i^*}{\lambda_i p_n^* d_b} \quad \text{(163)} \]
In particular (159) can be rewritten as
\[ H_{gb} \frac{q_i}{d_b} = \lambda_i \frac{\xi_i^*}{d_b} \]
Next, imitating a previous computation,
\[ H_{gb} \frac{\xi_i}{d_b} = H_{gb} \left( \frac{\alpha_i}{p_n^*} - \frac{\lambda_n p_n q_i^*}{\lambda_i p_n^* d_b} \right) = \left( H_{gb} + H_{gb-gb} \right) \left( \frac{\alpha_i}{p_n^*} - \frac{\lambda_n p_n q_i^*}{\lambda_i p_n^* d_b} \right) \]
\[ = H_{gb} \left( \frac{\alpha_i}{p_n^*} - \frac{\lambda_n p_n q_i^*}{\lambda_i p_n^* d_b} \right) + H_{gb-gb} \frac{\xi_i}{d_b} \]
\[ = \lambda_i \frac{\alpha_i}{p_n} - \lambda_n p - \frac{p_n^* d_b^*}{p_n^* d_b} \frac{\xi_i}{d_b} = \lambda_i \left( \frac{\alpha_i}{p_n} - \frac{\lambda_n p_n^* \xi_i}{\lambda_i p_n d_b} \right) \]
\[ = \lambda_i \frac{\lambda_i d_b \alpha_i^* - \lambda_n p_n^* \xi_i}{\lambda_i p_n d_b} = \lambda_i \frac{\lambda_n p_n^* q_i + \lambda_i p_n \xi_i^* - \lambda_n p_n^* \xi_i}{\lambda_i p_n d_b} \]
\[ = \lambda_i \frac{\xi_i^*}{d_b} + \lambda_n \frac{p_n^* q_i - \xi_i}{d_b} \].
Now $H_{eb} \zeta_i / d_b^* \in \text{Im } H_{eb} = X^{d_b}$, and clearly $\zeta_i / d_b \in X^{d_b}$. So we must have

$$\frac{p_n^*}{p_n} \frac{q_i - \zeta_i}{d_b} \in X^{d_b}.$$ 

This implies $p_n | q_i - \zeta_i$. Since $\deg q_i - \zeta_i \leq n - 2$, we must have $\zeta_i = q_i$, $i = 1, \ldots, n - 1$. So we have

$$\frac{q_i^*}{d_b} = \frac{\alpha_i^*}{p_n} - \frac{\lambda_n}{\lambda_i} \frac{p_n^*}{p_n} \frac{q_i}{d_b}$$  \hspace{2cm} (164)$$

and

$$\frac{q_i}{d_b^*} = \frac{\alpha_i}{p_n^*} - \frac{\lambda_n}{\lambda_i} \frac{p_n^*}{p_n} \frac{q_i^*}{d_b^*}.$$  \hspace{2cm} (165)$$

This, together with

$$\frac{p_i}{d^*} = \frac{\alpha_i}{p_n^*} - \frac{\lambda_n}{\lambda_i} \frac{p_n^*}{p_n} \frac{p_i^*}{d^*},$$  \hspace{2cm} (166)$$

yields

$$\frac{p_i}{d^*} = \frac{q_i}{d_b^*} + \frac{\lambda_n}{\lambda_i} \frac{p_n}{p_n^*} \left( \frac{q_i^*}{d_b^*} - \frac{p_i^*}{d^*} \right),$$ \hspace{2cm} (167)$$

or

$$\frac{p_i}{d^*} = \frac{q_i}{d_b^*} + \frac{\lambda_n}{\lambda_i} \frac{p_n}{p_n^*} \frac{d^* q_i^* - d_b^* p_i^*}{d^* d_b^*}.$$ \hspace{2cm} (168)$$

This will lead to an $H^2$ estimate later on.

From (165) we obtain, using orthogonality,

$$\left\| \frac{q_i^*}{d_b} \right\|^2 = \left\| \frac{\alpha_i^*}{p_n} \right\|^2 + \left\| \frac{\lambda_n}{\lambda_i} \right\|^2 \left\| \frac{q_i}{d_b} \right\|^2.$$ 

Since $\left\| \frac{q_i^*}{d_b} \right\|^2 = \left\| \frac{q_i}{d_b} \right\|^2$, we get

$$\left\| \frac{\alpha_i^*}{p_n} \right\|^2 = \left( 1 - \frac{\sigma_i^2}{\sigma_i^2} \right) \left\| \frac{q_i^*}{d_b} \right\|^2.$$ 

By the same reasoning we get from Equation (166)

$$\left\| \frac{\alpha_i}{p_n} \right\|^2 = \left( 1 - \frac{\sigma_i^2}{\sigma_i^2} \right) \left\| \frac{p_i^*}{d} \right\|^2.$$
By comparison we get
\[ \left\| \frac{p_i^*}{d} \right\|^2 = \left\| \frac{q_i^*}{d_i} \right\|^2 . \]
However, \( p_i^*/d \) was normalized so that \( \| p_i^*/d \|^2 = \sigma_i \), and hence also \( \| q_i^*/d_i \|^2 = \sigma_i \).

Now we can go back to (167) and estimate the \( H^2 \) norms. Clearly
\[ \left\| \frac{p_i}{d^*} - \frac{q_i}{d_i^*} \right\|^2 \leq \frac{\sigma_i^2}{\sigma_i^2} \left( \left\| \frac{q_i^*}{d_i} \right\|^2 + \left\| \frac{p_i^*}{d^*} \right\|^2 \right) = \frac{\sigma_i^2}{\sigma_i^2} 4 \sigma_i = 4 \frac{\sigma_i^2}{\sigma_i^2} . \quad (169) \]

The research in this paper was done over an extended period. I would like first to thank P. Khargonekar and A. Tannenbaum for getting me to break a close to ten year old decision to avoid working in areas directly related to \( H^\infty \)-control theory. I would like to thank A. Lindquist of the Royal Institute of Technology, D. Hinrichsen of the Department of Mathematics of the University of Bremen, and D. Prätzel-Wolters of the Department of Mathematics in Kaiserslautern for their interest and support. In all these institutions, as in the Department of Mathematics in Groningen and the Department of Electrical Engineering in Padova, I found a readiness to listen to a strange mixture of half-finished results, conjectures, and speculations. The ensuing discussions were of great help to me. Particularly I would like to thank the Department of Mathematics in Kaiserslautern, which also made an extremely useful visit by R. Ober possible. Finally, it is a great pleasure to acknowledge R. Ober’s help in a critical phase of this work. I learned a lot from his insights into balanced realizations. Also, in the paper I have incorporated some results of his, in particular Lemma 9.1 and some of the error formulas in Theorem 9.2.

REFERENCES


Douglas


Laub


Received 30 October 1989; final manuscript accepted 3 July 1990