Holomorphic Spaces
MSRI Publications
Volume **33**, 1998

A Basic Interpolation Problem

HARRY DYM

Abstract. A basic interpolation problem, which includes bitangential matrix versions of a number of classical interpolation problems, is formulated and solved. Particular attention is placed on the development of the problem in a natural way and upon the fundamental role played by a special class of reproducing kernel Hilbert spaces of vector-valued meromorphic functions that originate in the work of L. de Branges. Necessary and sufficient conditions for the existence of a solution to this problem, and a parametrization of the set of all solutions to this problem when these conditions are met, are presented. Some comparisons with the methods of Katsnelson, Kheifets, and Yuditskii are made. The presentation is largely self-contained and expository.

1. Introduction

This paper presents a largely self-contained expository introduction to a number of problems in interpolation theory for matrix-valued functions, including the classical problems of Schur, Nevanlinna–Pick (NP), and Carathéodory–Fejér (CF) as special cases. The development will use little more than the elementary properties of vector-valued Hardy spaces of exponent 2.

To illustrate the scope of the paper we shall begin with sample problems, all of which are formulated in the class $\mathcal{S}^{p\times q}(\Omega_+)$ of $p\times q$ matrix-valued functions (mvf) $S(\lambda)$ that are analytic and contractive in a given region Ω_+ in the complex plane $\mathbb{C}$.

Example 1.1 (The left tangential NP problem). The data for this problem is a set of points $\omega_1,\dots,\omega_n$ in Ω_+ and two sets of vectors: $\xi_1,\dots\xi_n$ in $\mathbb{C}^p$ and $\eta_1,\dots\eta_n$ in $\mathbb{C}^q$. An mvf $S(\lambda)\in\mathcal{S}^{p\times q}(\Omega_+)$ is said to be a solution of this problem if

$$\xi_j^* S(\omega_j)=\eta_j^* \quad \text{for } j=1,\dots,n. \tag{1.1}$$

The author thanks Renee and Jay Weiss for endowing the chair that supports his research.

Example 1.2 (The right tangential NP problem). The data for this problem is exactly the same as for the preceding example, but now the interpolation conditions are imposed on the right: An mvf $S(\lambda) \in \mathcal{S}^{p\times q}(\Omega_+)$ is said to be a solution of this problem if

$$S(\omega_j)\eta_j = \xi_j \quad \text{for } j = 1, \dots, n. \tag{1.2}$$

Example 1.3 (The bitangential NP problem). The data for this problem is exactly the same as for the preceding two examples, but now the first μ interpolation constraints are imposed on the left and the last $\nu = n - \mu$ constraints are imposed on the right. An mvf $S(\lambda) \in \mathcal{S}^{p\times q}(\Omega_+)$ is said to be a solution of this problem if

$$\xi_j^* S(\omega_j) = \eta_j^* \quad \text{for } j = 1, \dots, \mu, \tag{1.3}$$

and

$$S(\omega_j)\eta_j = \xi_j \quad \text{for } j = \mu{+}1, \dots, n. \tag{1.4}$$

Example 1.4 (The left tangential CF problem). The data for this problem is a point $\omega \in \Omega_+$, a vector $\xi \in \mathbb{C}^p$ and a set of vectors $\eta_1, \dots, \eta_n$ in $\mathbb{C}^q$. An mvf $S(\lambda) \in \mathcal{S}^{p\times q}(\Omega_+)$ is said to be a solution of this problem if

$$\frac{\xi^* S^{j-1}(\omega)}{(j-1)!} = \eta_j^* \quad \text{for } j = 1, \dots, n. \tag{1.5}$$

Example 1.5 (A mixed problem). The data for this problem is a pair of points ω_1, ω_2 in Ω_+, a set of vectors ξ_1 and $\xi_{21}, \dots, \xi_{2\nu}$ in $\mathbb{C}^p$ and a set of vectors $\eta_{11}, \dots, \eta_{1\mu}$ and η_2 in $\mathbb{C}^q$. An mvf $S(\lambda) \in \mathcal{S}^{p\times q}(\Omega_+)$ is said to be a solution of this problem if

$$\frac{\xi_1^* S^{j-1}(\omega_1)}{(j-1)!} = \eta_{1j} \quad \text{for } j = 1, \dots, \mu, \tag{1.6}$$

and

$$\frac{S^{j-1}(\omega_2)\eta_2}{(j-1)!} = \xi_{2j} \quad \text{for } j = 1, \dots, \nu. \tag{1.7}$$

The basic objective for all these problems is twofold:

(1) To formulate necessary and sufficient conditions in terms of the given data for the existence of at least one mvf $S(\lambda) \in \mathcal{S}^{p\times q}(\Omega_+)$ that meets the interpolation constraints.
(2) To describe the set of all mvf $S(\lambda) \in \mathcal{S}^{p\times q}(\Omega_+)$ that meet the interpolation constraints when the conditions for existence are met.

It turns out that when the region Ω_+ is chosen to be either the open unit disc

$$\mathbb{D} = \{\lambda \in \mathbb{C} : |\lambda| < 1\},$$

or the open upper half-plane

$$\mathbb{C}_+ = \{\lambda \in \mathbb{C} : \operatorname{Im}\lambda > 0\},$$

Ω_+	$\mathbb{D}$	$\mathbb{C}_+$	Π_+
$a(\lambda)$	1	$\sqrt{\pi}(1-i\lambda)$	$\sqrt{\pi}(1+\lambda)$
$b(\lambda)$	λ	$\sqrt{\pi}(1+i\lambda)$	$\sqrt{\pi}(1-\lambda)$
$\rho_\omega(\lambda)$	$1-\lambda\omega^*$	$-2\pi i(\lambda-\omega^*)$	$2\pi(\lambda+\omega^*)$
Ω_0	$\mathbb{T}$	$\mathbb{R}$	$i\mathbb{R}$
$\langle f,g\rangle$	$\frac{1}{2\pi}\int_0^{2\pi} g(e^{i\theta})^* f(e^{i\theta})d\theta$	$\int_{-\infty}^{\infty} g(x)^* f(x)dx$	$\int_{-\infty}^{\infty} g(iy)^* f(iy)dy$
λ°	$1/\lambda^*$ if $\lambda \neq 0$	λ^*	$-\lambda^*$
$f^\#(\lambda)$	$f(\lambda^\circ)^*$	$f(\lambda^\circ)^*$	$f(\lambda^\circ)^*$
$\delta_\omega(\lambda)$	$\lambda-\omega$	$2\pi i(\lambda-\omega)$	$-2\pi(\lambda-\omega)$
$ab'-ba'$	1	$2\pi i$	-2π
$\varphi_{j,\omega}(\lambda)$	$\lambda^j/(1-\lambda\omega^*)^{j+1}$	$-1/2\pi i(\lambda-\omega^*)^{j+1}$	$(-1)^j/2\pi(\lambda+\omega^*)^{j+1}$
$(R_\alpha\rho_\omega^{-1})(\lambda)$	$\omega^*/\rho_\omega(\alpha)\rho_\omega(\lambda)$	$2\pi i/\rho_\omega(\alpha)\rho_\omega(\lambda)$	$-2\pi/\rho_\omega(\alpha)\rho_\omega(\lambda)$

Table 1. Bringing the classical regions into a unified framework.

or the open right half-plane

$$\Pi_+ = \{\lambda \in \mathbb{C} : \operatorname{Re}\lambda > 0\},$$

all these problems (as well as more complicated problems in more complicated regions) can be incorporated into a more general problem, which we call the Basic Interpolation Problem (BIP). Moreover, by exercising a little care in the choice of notation, most of the analysis for all three of the classical choices of Ω_+ mentioned above can be carried out in one stroke. Table 1 serves as a dictionary for the meaning of the symbol that is appropriate for the region Ω_+ in use.

In order to describe the BIP we need to introduce some notation.

Let $H_2^k(\Omega_+)$ denote the set of $(k\times 1)$-vector-valued functions with entries in the Hardy space $H_2(\Omega_+)$. This space is identified as a closed subspace of the Hilbert space $L_2^k(\Omega_0)$ of $(k\times 1)$-vector-valued functions that are measurable and square integrable (that is, $\langle f,f\rangle$, as defined in Table 1, is finite) on the boundary Ω_0 of Ω_+ in the usual way. The symbol $H_2^k(\Omega_+)^\perp$ designates the orthogonal complement of $H_2^k(\Omega_+)$ in $L_2^k(\Omega_0)$ with respect to the inner product indicated in Table 1,

$$\underline{\underline{p}} \text{ denotes the orthogonal projection of } L_2^k(\Omega_0) \text{ onto } H_2^k(\Omega_+),$$

and

$$\underline{\underline{q}}' = I - \underline{\underline{p}} \text{ denotes the orthogonal projection of } L_2^k(\Omega_0) \text{ onto } H_2^k(\Omega_+)^\perp.$$

The dependence of the projections on the height k of the column vectors is suppressed in order to keep the notation simple.

For each of the three listed choices of the kernel function $\rho_\omega(\lambda)$, we have

$$\Omega_+ = \{\omega \in \mathbb{C} : \rho_\omega(\omega) > 0\}$$

and

$$\Omega_0 = \{\omega \in \mathbb{C} : \rho_\omega(\omega) = 0\}.$$

We shall take

$$\Omega_- = \{\omega \in \mathbb{C} : \rho_\omega(\omega) < 0\}.$$

The use of such a flexible notation to cover problems in both $\mathbb{D}$ and $\mathbb{C}_+$ more or less simultaneously was promoted in [Alpay and Dym 1984] and [Dym 1989a]. The observation that the kernels $\rho_\omega(\lambda)$ that intervene in these problems can be expressed in terms of a pair of polynomials $a(\lambda)$ and $b(\lambda)$ as

$$\rho_\omega(\lambda) = a(\lambda)a(\omega)^* - b(\lambda)b(\omega)^* \tag{1.8}$$

was first made by Lev-Ari and Kailath [1986]. They noticed that certain fast algorithms in which the term $\rho_\omega(\lambda)$ intervenes will work if and only if $\rho_\omega(\lambda)$ can be expressed in the form (1.8). A general theory of reproducing kernels with denominators of this form and their applications was developed in [Alpay and Dym 1992; 1993a; 1993b; 1996]; for related developments see [Nudelman 1993].

The rest of the notation is fairly standard: The symbol A^* denotes the adjoint of an operator A on a Hilbert space, with respect to the inner product of the space. If A is a finite matrix, the adjoint will always be computed with respect to the standard inner product, so that in this case A^* will be the Hermitian transpose, or just the complex conjugate if A is a number. The symbol $\sigma(A)$ denotes the spectrum of a matrix A and J stands for the $m \times m$ signature matrix

$$J = \begin{bmatrix} I_p & 0 \\ 0 & -I_q \end{bmatrix}$$

with $p \geq 1$, $q \geq 1$ and $p + q = m$.

The following evaluations, which depend basically on Cauchy's formula for $H_2(\Omega_+)$, will prove useful. For details, see [Dym 1994b, Section 2.2].

LEMMA 1.6. *If $\omega \in \Omega_+$, $u \in \mathbb{C}^k$ and $S(\lambda) \in \mathcal{S}^{p\times q}(\Omega_+)$, then:*

(1) $\dfrac{u}{\rho_\omega} \in H_2^k(\Omega_+)$ *and* $\dfrac{u}{\delta_\omega} \in H_2^k(\Omega_+)^\perp$.

(2) $\underline{\underline{p}} S^* \dfrac{u}{\rho_\omega} = S(\omega)^* \dfrac{u}{\rho_\omega}$ *(when $k = p$).*

(3) $\underline{\underline{q}}' S \dfrac{u}{\delta_\omega} = S(\omega) \dfrac{u}{\delta_\omega}$ *(when $k = q$).*

(4) $\big(\underline{\underline{p}} S^* \varphi_{j,\omega} u\big)(\lambda) = \sum_{i=0}^{j} S^{(j-i)}(\omega)^* \varphi_{i,\omega}(\lambda) u$ *(when $k = p$).*

(5) $\big(\underline{\underline{q}}' S(\cdot - \omega)^{-j} u\big)(\lambda) = \sum_{i=0}^{j-1} (\lambda - \omega)^{i-j} \dfrac{S^{(i)}(\omega)}{i!} u$ *(when $k = q$).*

2. The Basic Interpolation Problem

First formulation. The data for the Basic Interpolation Problem consists of a set of $2n$ vector-valued functions $g_1, \dots, g_n, h_1, \dots, h_n$, where

$$g_1, \dots, g_\mu \in H_2^p(\Omega_+), \qquad g_{\mu+1}, \dots, g_n \in H_2^p(\Omega_+)^\perp,$$
$$h_1, \dots, h_\mu \in H_2^q(\Omega_+), \qquad h_{\mu+1}, \dots, h_n \in H_2^q(\Omega_+)^\perp.$$

An mvf $S(\lambda) \in \mathcal{S}^{p\times q}(\Omega_+)$ is said to be a solution of this problem if

$$\underline{\underline{p}} S^* g_j = h_j \quad \text{for } j = 1, \dots, \mu$$

and

$$\underline{\underline{q}}' S h_j = g_j \quad \text{for } j = \mu{+}1, \dots, n.$$

EXAMPLE 2.1. Let there be given points $\omega_1, \dots, \omega_n$ in Ω_+, vectors $\xi_1, \dots, \xi_n$ in $\mathbb{C}^p$ and a second set of vectors $\eta_1, \dots, \eta_n$ in $\mathbb{C}^q$, just as in Examples 1.1–1.3. Let

$$g_j(\lambda) = \begin{cases} \dfrac{\xi_j}{\rho_{\omega_j}(\lambda)} & \text{for } j = 1, \dots, \mu, \\[2ex] \dfrac{\xi_j}{\delta_{\omega_j}(\lambda)} & \text{for } j = \mu{+}1, \dots, n, \end{cases}$$

and

$$h_j(\lambda) = \begin{cases} \dfrac{\eta_j}{\rho_{\omega_j}(\lambda)} & \text{for } j = 1, \dots, \mu, \\[2ex] \dfrac{\eta_j}{\delta_{\omega_j}(\lambda)} & \text{for } j = \mu{+}1, \dots, n. \end{cases}$$

Then

$$g_j \in H_2^p(\Omega_+) \quad \text{and} \quad h_j \in H_2^q(\Omega_+) \quad \text{for } j = 1, \dots, \mu,$$

whereas

$$g_j \in H_2^p(\Omega_+)^\perp \quad \text{and} \quad h_j \in H_2^q(\Omega_+)^\perp \quad \text{for } j = \mu{+}1, \dots, n.$$

Therefore, by evaluations (2) and (3) in Lemma 1.6, it is readily seen that $S(\lambda)$ is a solution of the BIP based on this choice of $g_1, \dots, g_n$ and $h_1, \dots, h_n$ if and only if $S(\lambda)$ is a solution of the bitangential NP set forth in Example 1.3. Examples 1.1 and 1.2 correspond to choosing $\mu = n$ and $\nu = n$, respectively.

More elaborate examples involving derivatives can be constructed in much the same way by taking advantage of the evaluations (4) and (5) in Lemma 1.6.

A reformulation of the Basic Interpolation Problem. It is readily checked that the mvf $S(\lambda) \in \mathcal{S}^{p\times q}(\Omega_+)$ is a solution of the BIP if and only if

$$\underline{\underline{p}}(S^* g_j - h_j) = 0 \quad \text{for } j = 1, \ldots, \mu$$

and

$$\underline{\underline{q}}'(S h_j - g_j) = 0 \quad \text{for } j = \mu{+}1, \ldots, n.$$

In fact, since the first condition is automatically satisfied for $j = \mu{+}1, \ldots, n$ and the second condition is automatically satisfied for $j = 1, \ldots, \mu$, the indices may be allowed to run from 1 to n in both cases. In other words, an mvf $S(\lambda) \in \mathcal{S}^{p\times q}(\Omega_+)$ is a solution of the BIP if and only if

$$g_j - S h_j \in H_2^p(\Omega_+) \quad \text{for } j = 1, \ldots, n$$

and

$$-S^* g_j + h_j \in H_2^q(\Omega_+)^\perp \quad \text{for } j = 1, \ldots, n.$$

This last pair of constraints can be stacked, yielding our final formulation:

Final formulation of the BIP. As before, the data for this problem is two sets of vector-valued functions $g_1, \ldots, g_n$ and $h_1, \ldots, h_n$, where

$$g_1, \ldots, g_\mu \in H_2^p(\Omega_+), \qquad g_{\mu+1}, \ldots, g_n \in H_2^p(\Omega_+)^\perp,$$
$$h_1, \ldots, h_\mu \in H_2^q(\Omega_+), \qquad h_{\mu+1}, \ldots, h_n \in H_2^q(\Omega_+)^\perp.$$

Then $S(\lambda) \in \mathcal{S}^{p\times q}(\Omega_+)$ is said to be a solution of the BIP based on this given set of data if

$$\begin{bmatrix} I_p & -S \\ -S^* & I_q \end{bmatrix} \begin{bmatrix} g_j \\ h_j \end{bmatrix} \in H_2^p(\Omega_+) \oplus H_2^q(\Omega_+)^\perp$$

for $j = 1, \ldots, n$.

This formulation is meaningful even if the g_j and h_j are not themselves in $L_2^p(\Omega_0)$ and $L_2^q(\Omega_0)$, because it is the difference that comes into play. This is significant in the study of interpolation problems with constraints on the boundary Ω_0. Thus for example, if $\Omega_+ = \mathbb{D}$ and $\omega_1 \in \mathbb{T}$, then

$$g_1(\lambda) = \frac{\xi_1}{\rho_{\omega_1}(\lambda)} = \frac{\xi_1}{1 - \lambda\omega_1^*}$$

does not belong to $H_2^p(\mathbb{D})$ and

$$h_1(\lambda) = \frac{\eta_1}{\rho_{\omega_1}(\lambda)} = \frac{\eta_1}{1 - \lambda\omega_1^*}$$

does not belong to $H_2^q(\mathbb{D})$. Nevertheless it is meaningful to investigate the set of mvfs $S(\lambda) \in \mathcal{S}^{p\times q}(\mathbb{D})$ for which

$$g_1 - S h_1 \in H_2^p(\mathbb{D}).$$

We shall not pursue this question here.

3. Necessary Conditions for the Existence of a Solution to the Basic Interpolation Problem

THEOREM 3.1. *If $S(\lambda) \in \mathcal{S}^{p\times q}(\Omega_+)$ is a solution of the BIP based on the given data $g_1, \ldots, g_n$ and $h_1, \ldots, h_n$, then the $n \times n$ Hermitian matrix Q with entries*

$$q_{ij} = \begin{cases} \langle g_j, g_i\rangle - \langle h_j, h_i\rangle & \text{for } i, j = 1, \ldots, \mu, \\ -\langle Sh_j, g_i\rangle & \text{for } i = 1, \ldots, \mu \text{ and } j = \mu{+}1, \ldots, n, \\ -\langle g_j, g_i\rangle + \langle h_j, h_i\rangle & \text{for } i, j = \mu{+}1, \ldots, n, \end{cases} \tag{3.1}$$

is positive semidefinite.

PROOF. Define

$$q_{ij} = \left\langle \begin{bmatrix} I_p & -S \\ -S^* & I_q \end{bmatrix} \begin{bmatrix} g_j \\ h_j \end{bmatrix}, \begin{bmatrix} g_i \\ h_i \end{bmatrix} \right\rangle \tag{3.2}$$

for $i, j = 1, \ldots, n$. Then

$$Q = [q_{ij}] \geq 0,$$

since

$$\begin{bmatrix} I_p & -S(\lambda) \\ -S^*(\lambda) & I_q \end{bmatrix} \geq 0$$

for a.e. point $\lambda \in \Omega_0$. The rest of the proof amounts to evaluating (3.2). This is where the assumption that $S(\lambda)$ is a solution of the BIP comes into play. There are three basic cases, but it is convenient to begin the analysis of the first two together.

Suppose first that $1 \leq i \leq \mu$. Then

$$\begin{aligned} q_{ij} &= \langle g_j - Sh_j,\, g_i\rangle + \langle -S^* g_j + h_j,\, h_i\rangle \\ &= \langle g_j - Sh_j,\, g_i\rangle = \langle g_j, g_i\rangle - \langle Sh_j, g_i\rangle. \end{aligned}$$

Now, if also $1 \leq j \leq \mu$, then

$$\langle Sh_j, g_i\rangle = \langle h_j, \underline{\underline{p}} S^* g_i\rangle = \langle h_j, h_i\rangle,$$

whereas, if $\mu{+}1 \leq j \leq n$, then $\langle Sh_j, g_i\rangle$ cannot be reduced but $\langle g_j, g_i\rangle = 0$. These evaluations lead easily to the first two sets of formulas for q_{ij}. To verify the last set, assume $\mu{+}1 \leq i, j \leq n$. Then

$$\begin{aligned} q_{ij} &= \langle g_j - Sh_j, g_i\rangle + \langle -S^* g_j + h_j, h_i\rangle = \langle -S^* g_j + h_j, h_i\rangle \\ &= -\langle g_j, \underline{\underline{q}}' Sh_i\rangle + \langle h_j, h_i\rangle = -\langle g_j, g_i\rangle + \langle h_j, h_i\rangle, \end{aligned}$$

as claimed. The proof is complete. □

If $g_1, \ldots, g_n$ and $h_1, \ldots, h_n$ are chosen as in Example 2.1, it is readily checked (with the aid of Cauchy's formula) that

$$q_{ij} = \left\langle \frac{\xi_j}{\rho_{\omega_j}}, \frac{\xi_i}{\rho_{\omega_i}} \right\rangle - \left\langle \frac{\eta_j}{\rho_{\omega_j}}, \frac{\eta_i}{\rho_{\omega_i}} \right\rangle = \frac{\xi_i^* \xi_j - \eta_i^* \eta_j}{\rho_{\omega_j}(\omega_i)} \quad \text{for } i, j = 1, \ldots, \mu, \tag{3.3}$$

whereas

$$q_{ij} = -\left\langle \frac{\xi_j}{\delta_{\omega_j}}, \frac{\xi_i}{\delta_{\omega_i}} \right\rangle + \left\langle \frac{\eta_j}{\delta_{\omega_j}}, \frac{\eta_i}{\delta_{\omega_i}} \right\rangle = \frac{-\xi_i^*\xi_j + \eta_i^*\eta_j}{\rho_{\omega_i}(\omega_j)} \quad \text{for } i, j = \mu{+}1, \ldots, n. \tag{3.4}$$

These formulas are expressed totally in terms of the data of the problem. Thus a necessary condition for the existence of a solution to the bitangential NP problem of Example 1.3 is that the matrices exhibited in formulas (3.3) and (3.4) are positive semidefinite.

Theorem 3.1 gives a necessary condition for the existence of a solution to the BIP. In general this condition will not be sufficient unless additional structure is imposed on the data $g_1, \ldots, g_n$ and $h_1, \ldots, h_n$ of the problem. In particular we shall assume that the space

$$\operatorname{span}\left\{ \begin{bmatrix} g_j \\ h_j \end{bmatrix} : j = 1, \ldots, \mu \right\}$$

is a μ-dimensional R_α-invariant subspace of $H_2^m(\Omega_+)$ for some point $\alpha \in \Omega_+$ and that

$$\operatorname{span}\left\{ \begin{bmatrix} g_j \\ h_j \end{bmatrix} : j = \mu{+}1, \ldots, n \right\}$$

is a ν-dimensional R_α-invariant subspace of $H_2^m(\Omega_+)^\perp$ for some point $\alpha \in \Omega_-$, where R_α denotes the operator defined by the rule

$$(R_\alpha f)(\lambda) = \frac{f(\lambda) - f(\alpha)}{\lambda - \alpha}, \tag{3.5}$$

wherever it is meaningful.

4. R_α Invariance

In this section we study finite-dimensional spaces of vector-valued functions that are invariant under the action of R_α for at least one appropriately chosen point $\alpha \in \mathbb{C}$. The contents are taken largely from [Dym 1994b, Section 3].

THEOREM 4.1. *Let $\mathcal{M}$ be an n-dimensional vector space of $(m \times 1)$-vector-valued functions that are meromorphic in some open nonempty set $\Delta \subset \mathbb{C}$ and suppose further that $\mathcal{M}$ is R_α-invariant for some point $\alpha \in \Delta$ in the domain of analyticity of $\mathcal{M}$. Then $\mathcal{M}$ is spanned by the columns of a rational $m \times n$ matrix-valued function of the form*

$$F(\lambda) = V\{M - \lambda N\big)^{-1}, \tag{4.1}$$

where $V \in \mathbb{C}^{m\times n}$, $M, N \in \mathbb{C}^{n\times n}$,

$$MN = NM \quad \text{and} \quad M - \alpha N = I_n. \tag{4.2}$$

Moreover, $\lambda \in \Delta$ is a point of analyticity of F if and only if the $n \times n$ matrix $M - \lambda N$ is invertible.

Proof. Let $f_1, \dots, f_n$ be a basis for $\mathcal{M}$ and let

$$F(\lambda) = \begin{bmatrix} f_1(\lambda) & \cdots & f_n(\lambda) \end{bmatrix}$$

be the $m \times n$ matrix-valued function with columns $f_1(\lambda), \dots, f_n(\lambda)$. Then, because of the presumed R_α-invariance of the columns of F,

$$R_\alpha F(\lambda) = \frac{F(\lambda) - F(\alpha)}{\lambda - \alpha} = F(\lambda)E_\alpha$$

for some $n \times n$ matrix E_α independent of λ. Thus

$$F(\lambda)\big(I_n - (\lambda - \alpha)E_\alpha\big) = F(\alpha),$$

and hence, since $\det\big(I_n - (\lambda - \alpha)E_\alpha\big) \not\equiv 0$,

$$F(\lambda) = F(\alpha)\big(I_n + \alpha E_\alpha - \lambda E_\alpha\big)^{-1},$$

which is of the form (4.1) with $V = F(\alpha)$, $M = I_n + \alpha E_\alpha$ and $N = E_\alpha$.

Suppose next that F is analytic at a point $\omega \in \Delta$ and that $u \in \ker(M - \omega N)$. Then

$$F(\lambda)(M - \lambda N)u = Vu = 0,$$

first for $\lambda = \omega$, and then for every $\lambda \in \Delta$ in the domain of analyticity of F. Thus, for all such λ,

$$(\omega - \lambda)F(\lambda)Nu = F(\lambda)\big(M - \lambda N - (M - \omega N)\big)u = 0.$$

Since the columns of $F(\lambda)$ are linearly independent functions of λ, we get $Nu = 0$. But this in conjunction with the prevailing assumption $(M - \omega N)u = 0$ implies that

$$u \in \ker M \cap \ker N \implies u = 0 \implies M - \omega N \text{ is invertible.}$$

Thus we have shown that if F is analytic at ω, then $M - \omega N$ is invertible. Since the opposite implication is easy, this serves to complete the proof. $\square$

Corollary 4.2. *If* $\det(M - \lambda N) \not\equiv 0$ *and* $F(\lambda) = V(M - \lambda N)^{-1}$ *is a rational* $m \times n$ *matrix-valued function with* n *linearly independent columns, then*

(1) M *is invertible if and only if* F *is analytic at zero, and*
(2) N *is invertible if and only if* F *is analytic at infinity and* $F(\infty) = 0$.

Moreover, if M *is invertible* F *can be expressed in the form*

$$F(\lambda) = C(I_n - \lambda A)^{-1}, \tag{4.3}$$

whereas if N *is invertible* F *can be expressed in the form*

$$F(\lambda) = C(A - \lambda I_n)^{-1}. \tag{4.4}$$

PROOF. The first assertion is contained in the theorem; the second is obtained in much the same way. More precisely, if $\lim_{\lambda\to\infty} F(\lambda) = 0$ and $u \in \ker N$, then

$$F(\lambda)Mu = F(\lambda)(M - \lambda N)u = Vu.$$

Upon letting λ approach ∞, it follows that

$$Vu = 0 \implies F(\lambda)Mu = 0 \implies Mu = 0 \implies u \in \ker M \cap \ker N \implies u = 0.$$

Thus N is invertible. The other direction is easy, as are formulas (4.3) and (4.4). Just take $C = VM^{-1}$ and $A = NM^{-1}$ in the first case, and $C = VN^{-1}$ and $A = MN^{-1}$ in the second. □

COROLLARY 4.3. *Let f be an $(m\times 1)$-vector-valued function that is meromorphic in some open nonempty set $\Delta \subset \mathbb{C}$ and let $\alpha \in \Delta$ be a point of analyticity of f. Then f is an eigenfunction of R_α if and only if it can be expressed in the form*

$$f(\lambda) = \frac{v}{\rho_\omega(\lambda)}$$

for one or more choices of $\rho_\omega(\lambda)$ in Table 1 *with $\rho_\omega(\alpha) \neq 0$ and some nonzero constant vector $v \in \mathbb{C}^m$.*

Linear independence. It seems worthwhile to emphasize that herein the n columns of an $m \times n$ matrix-valued function $F(\lambda)$ are said to be linearly independent if they are linearly independent in the vector space of continuous $(m \times 1)$-vector-valued functions on the domain of analyticity of F. If

$$F(\lambda) = C(I_n - \lambda A)^{-1} \quad \text{or} \quad F(\lambda) = C(A - \lambda I_n)^{-1},$$

this is easily seen to be equivalent to

$$\bigcap_{j=0}^{n-1} \ker CA^j = 0;$$

that is, to the pair (C, A) being observable. Such a realization for F is minimal in the sense of Kalman because (in the usual terminology; see [Kailath 1980], for example) the pair (A, B) is automatically controllable:

$$\bigcap_{j=0}^{n-1} \ker B^* A^{*j} = \{0\}$$

(or, equivalently, $\operatorname{rank}\,[B \;\; AB \;\; \cdots \;\; A^{n-1}B] = n$), since $B = I_n$.

R_α-invariant subspaces of $H_2^m(\Omega_+)$ and $H_2^m(\Omega_+)^\perp$. Let

$$F_1(\lambda) = \begin{bmatrix} g_1(\lambda) & \cdots & g_\mu(\lambda) \\ h_1(\lambda) & \cdots & h_\mu(\lambda) \end{bmatrix} \tag{4.5}$$

be an $m \times \mu$ mvf with columns

$$f_j(\lambda) = \begin{bmatrix} g_j(\lambda) \\ h_j(\lambda) \end{bmatrix} \in H_2^p(\Omega_+) \oplus H_2^q(\Omega_+),$$

and suppose that $\{F_1(\lambda)u : u \in \mathbb{C}^\mu\}$ is a μ-dimensional R_α-invariant subspace of $H_2^m(\Omega_+)$ for some point $\alpha \in \Omega_+$. This is meaningful because $F_1(\lambda)$ is analytic in Ω_+. Then it follows from Theorem 3.1 that $F_1(\lambda)$ admits a representation of the form

$$F_1(\lambda) = V(M - \lambda N)^{-1}.$$

Hence, since $F_1(\lambda)$ is analytic at the point $\lambda = 0$ if $\Omega_+ = \mathbb{D}$ and at the point $\lambda = \infty$ with $F_1(\infty) = 0$ if $\Omega_+ = \mathbb{C}_+$ or $\Omega_+ = \Pi_+$, it follows also that

$$F_1(\lambda) = \begin{cases} \begin{bmatrix} C_{11} \\ C_{21} \end{bmatrix} (I_\mu - \lambda A_1)^{-1} \text{ with } \sigma(A_1) \subset \mathbb{D} & \text{if } \Omega_+ = \mathbb{D}, \\ \begin{bmatrix} C_{11} \\ C_{21} \end{bmatrix} (\lambda I_\mu - A_1)^{-1} \text{ with } \sigma(A_1) \subset \Omega_- & \text{if } \Omega_+ = \mathbb{C}_+ \text{ or } \Omega_+ = \Pi_+, \end{cases} \tag{4.6}$$

where $C_{11} \in \mathbb{C}^{p\times\mu}$, $C_{21} \in \mathbb{C}^{q\times\mu}$ and of course $A_1 \in \mathbb{C}^{\mu\times\mu}$.

Next let

$$F_2(\lambda) = \begin{bmatrix} g_{\mu+1}(\lambda) & \cdots & g_n(\lambda) \\ h_{\mu+1}(\lambda) & \cdots & h_n(\lambda) \end{bmatrix} \tag{4.7}$$

be an $m \times \nu$ mvf with columns

$$f_j(\lambda) = \begin{bmatrix} g_j(\lambda) \\ h_j(\lambda) \end{bmatrix} \in H_2^p(\Omega_+)^\perp \oplus H_2^q(\Omega_+)^\perp$$

for $j = \mu+1, \dots, n$ and suppose that $\{F_2(\lambda)u : u \in \mathbb{C}^\nu\}$ is a ν-dimensional R_α-invariant subspace of $H_2^p(\Omega_+)^\perp$ for some choice of $\alpha \in \Omega_-$. This too is meaningful because $F_2(\lambda)$ is analytic in Ω_-. Then, by another application of Theorem 3.1, $F_2(\lambda)$ admits a representation of the form

$$F_2(\lambda) = \begin{bmatrix} C_{12} \\ C_{22} \end{bmatrix} (\lambda I_\nu - A_2)^{-1} \quad \text{with } \sigma(A_2) \subset \Omega_+, \tag{4.8}$$

where $C_{12} \in \mathbb{C}^{p\times\nu}$, $C_{22} \in \mathbb{C}^{q\times\nu}$ and $A_2 \in \mathbb{C}^{\nu\times\nu}$.

More explicit formulas, based on the Jordan decomposition of A_1 and A_2, may be found in [Dym 1994b, Section 3.3; Dym 1989b].

5. Back to the Basic Interpolation Problem

From now on we assume that the data of the BIP is chosen so that span$\{f_j(\lambda) : j = 1,\dots,\mu\}$ is a μ-dimensional R_α-invariant subspace of $H_2^m(\Omega_+)$ for some point $\alpha \in \Omega_+$ and span$\{f_j(\lambda) : j = \mu+1,\dots,n\}$ is a ν-dimensional R_α-invariant subspace of $H_2^m(\Omega_+)^\perp$ for some choice of $\alpha \in \Omega_-$. In view of the analysis in the previous section, this means that the data of the BIP is of the form

$$F(\lambda) = C(M - \lambda N)^{-1}, \tag{5.1}$$

where

$$C = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} = \begin{array}{c} \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \begin{array}{l} \}p \\ \}q \end{array} \\ \begin{array}{cc} \underbrace{\qquad}_{\mu} & \underbrace{\qquad}_{\nu} \end{array} \end{array}, \tag{5.2}$$

$$M - \lambda N = \begin{cases} \begin{bmatrix} I_\mu - \lambda A_1 & 0 \\ 0 & \lambda I_\nu - A_2 \end{bmatrix} & \text{if } \Omega_+ = \mathbb{D}\ , \\ \begin{bmatrix} \lambda I_\mu - A_1 & 0 \\ 0 & \lambda I_\nu - A_2 \end{bmatrix} & \text{if } \Omega_+ = \mathbb{C}_+ \text{ or } \Pi_+, \end{cases} \tag{5.3}$$

$\mu + \nu = n$,

$$\sigma(A_1) \subset \mathbb{D} \text{ if } \Omega_+ = \mathbb{D}, \quad \sigma(A_1) \subset \Omega_- \text{ if } \Omega_+ = \mathbb{C}_+ \text{ or } \Pi_+, \tag{5.4}$$

and

$$\sigma(A_2) \subset \Omega_+ \text{ for all three of the classical choices of } \Omega_+. \tag{5.5}$$

This means that the BIP is now fully specified in terms of the matrices C, M and N, or equivalently, in terms of their block decompositions C_{11}, C_{12}, C_{21}, C_{22}, A_1 and A_2.

THEOREM 5.1. *If the BIP that is specified in terms of C, M and N admits a solution, then there exists an $n \times n$ matrix $P \geq 0$ which solves the Lyapunov–Stein equation*

$$M^*PM - N^*PN = C^*JC \qquad \textit{if } \Omega_+ = \mathbb{D}, \tag{5.6a}$$

$$M^*PN - N^*PM = 2\pi i C^*JC \qquad \textit{if } \Omega_+ = \mathbb{C}_+, \tag{5.6b}$$

$$M^*PN + N^*PM = -2\pi C^*JC \qquad \textit{if } \Omega_+ = \Pi_+. \tag{5.6c}$$

PROOF. Let $S(\lambda) \in \mathcal{S}^{p\times q}(\Omega_+)$ be a solution of the BIP and define P by the rule

$$v^*Pu = \left\langle \begin{bmatrix} I_p & -S(\lambda) \\ -S(\lambda)^* & I_q \end{bmatrix} F(\lambda)u,\, F(\lambda)v \right\rangle, \tag{5.7}$$

where $F(\lambda)$ is given by (5.1) and u and v are any vectors in $\mathbb{C}^n$. Then clearly $P \geq 0$. The rest of the proof depends largely upon Theorem 3.1 and the special form of the data. We shall present details only for the case $\Omega_+ = \mathbb{D}$. Details for the other two classical choices of Ω_+ may be found in [Dym 1994b]. Because of the special block form of M and n, it is convenient to write P in the block form

$$P = \begin{array}{cc} \left[\begin{array}{cc} P_{11} & P_{12} \\ P_{21} & P_{22} \end{array}\right] & \begin{array}{l} \}\mu \\ \}\nu \end{array} \\ \begin{array}{cc} \underbrace{\hphantom{P_{11}}}_{\mu} & \underbrace{\hphantom{P_{12}}}_{\nu} \end{array} & \end{array}.$$

Then P is a solution of (5.6a) if and only if the following three equations are satisfied:

$$P_{11} - A_1^*P_{11}A_1 = C_{11}^*C_{11} - C_{21}^*C_{21}, \tag{5.8}$$

$$A_1^*P_{12} - P_{12}A_2 = C_{11}^*C_{12} - C_{21}^*C_{22}, \tag{5.9}$$

$$A_2^*P_{22}A_2 - P_{22} = C_{12}^*C_{12} - C_{22}^*C_{22}. \tag{5.10}$$

By Theorem 3.1,

$$y^*P_{11}x = \langle JF_1(\lambda)x, F_1(\lambda)y\rangle$$

for every choice of x and y in $\mathbb{C}^\mu$. Thus

$$\begin{aligned} y^*(P_{11} - A_1^*P_{11}A_1)x \\ &= \langle JF_1(\lambda)x, F_1(\lambda)y\rangle - \langle JF_1(\lambda)A_1x, F_1(\lambda)A_1y\rangle \\ &= \langle JF_1(\lambda)x,\, F_1(\lambda)(I_\mu - \lambda A_1)y\rangle + \langle JF_1(\lambda)(I_\mu - \lambda A_1)x,\, F_1(\lambda)\lambda A_1y\rangle \\ &= \left\langle JF_1(\lambda)x,\, \begin{bmatrix} C_{11} \\ C_{21} \end{bmatrix} y \right\rangle + \left\langle J \begin{bmatrix} C_{11} \\ C_{21} \end{bmatrix},\, F_1(\lambda)\lambda A_1 y \right\rangle \\ &= y^*[C_{11}^*\ C_{21}^*]JF_1(0)x + O \\ &= y^*(C_{11}^*C_{11} - C_{21}^*C_{21})x, \end{aligned}$$

which serves to prove (5.8).

Next, by the middle formula of (3.1),

$$\begin{aligned} y^*(A_1^*P_{12} - P_{12}A_2)x &= -\langle S(\lambda)C_{22}(\lambda I_\nu - A_2)^{-1}x,\, C_{11}(I_\mu - \lambda A_1)^{-1}A_1y\rangle \\ &\quad + \langle S(\lambda)C_{22}(\lambda I_\nu - A_2)^{-1}A_2x,\, C_{11}(I_\mu - \lambda A_1)^{-1}y\rangle \\ &= -\langle S(\lambda)C_{22}(\lambda I_\nu - A_2)^{-1}\lambda x,\, C_{11}(I_\mu - \lambda A_1)^{-1}\lambda A_1y\rangle \\ &\quad + \langle S(\lambda)C_{22}(\lambda I_\nu - A_2)^{-1}A_2x,\, C_{11}(I_\mu - \lambda A_1)^{-1}y\rangle \end{aligned}$$

for every choice of $x \in \mathbb{C}^\nu$ and $y \in \mathbb{C}^\mu$. But, by adding the term

$$\big\langle S(\lambda)C_{22}(\lambda I_\nu - A_2)^{-1}\lambda x,\, C_{11}(I_\mu - \lambda A_1)^{-1}y\big\rangle$$

to the first inner product and subtracting it from the second, the last expression can be rewritten as ① − ②, where

$$
\begin{aligned}
① &= \big\langle \lambda S(\lambda) C_{22}(\lambda I_\nu - A_2)^{-1}x,\, C_{11}(I_\mu - \lambda A_1)^{-1}(I_\mu - \lambda A_1)y \big\rangle \\
&= \big\langle \lambda \underline{\underline{q}}'\big(S(\lambda) C_{22}(\lambda I_\nu - A_2)^{-1}x\big),\, C_{11}y \big\rangle \\
&= \big\langle \lambda C_{12}(\lambda I_\nu - A_2)^{-1}x,\, C_{11}y \big\rangle = y^* C_{11}^* C_{12} x
\end{aligned}
$$

and

$$
\begin{aligned}
② &= \big\langle S(\lambda) C_{22}(\lambda I_\nu - A_2)^{-1}(\lambda I_\nu - A_2)x,\, C_{11}(I_\mu - \lambda A_1)^{-1}y \big\rangle \\
&= \big\langle S(\lambda) C_{22}x,\, C_{11}(I_\mu - \lambda A_1)^{-1}y \big\rangle \\
&= \big\langle C_{22}x,\, \underline{\underline{p}}\big(S(\lambda)^* C_{11}(I_\mu - \lambda A_1)^{-1}y\big) \big\rangle \\
&= \big\langle C_{22}x,\, C_{21}(I_\mu - \lambda A_1)^{-1}y \big\rangle = y^* C_{21}^* C_{22} x.
\end{aligned}
$$

Thus

$$
① - ② = y^*(C_{11}^* C_{12} - C_{21}^* C_{22})x,
$$

which serves to verify (5.9).

Finally, (5.10) is obtained in much the same way from the formula

$$
y^* P_{22} x = -\big\langle J F_2(\lambda)x, F_2(\lambda)y \big\rangle,
$$

which is valid for every choice of x and y in $\mathbb{C}^q$ and is itself obtained from Theorem 3.1. □

Theorem 5.1 admits a converse: If the Lyapunov–Stein equation (5.6) admits a nonnegative solution P, then the BIP is solvable. However, this is only part of the story because if both $\mu \ge 1$ and $\nu = n - \mu \ge 1$, then (5.6) can have many solutions. To be more precise, under the spectral conditions that were imposed on A_1 and A_2 in (5.4) and (5.5), the P_{11} and P_{22} blocks of every solution P of (5.6) are the same, however, the P_{12} block is not unique unless $\sigma(A_1^*) \cap \sigma(A_2) = \varnothing$. This extra freedom can be used to impose more interpolation conditions and leads to a more refined problem which we shall refer to as the augmented BIP. For an instructive example, see [Dym 1989b, Section 10].

6. The Augmented Basic Interpolation Problem

The augmented BIP is formulated in terms of the data C,

$$
M = \begin{array}{c} \begin{bmatrix} M_{11} & 0 \\ 0 & M_{22} \end{bmatrix} \begin{matrix} \}\mu \\ \}\nu \end{matrix} \\ \begin{matrix} \underbrace{\quad}_{\mu} & \underbrace{\quad}_{\nu} \end{matrix} \end{array}, \qquad
N = \begin{array}{c} \begin{bmatrix} N_{11} & 0 \\ 0 & N_{22} \end{bmatrix} \begin{matrix} \}\mu \\ \}\nu \end{matrix} \\ \begin{matrix} \underbrace{\quad}_{\mu} & \underbrace{\quad}_{\nu} \end{matrix} \end{array},
$$

and a solution $P \ge 0$ of the Lyapunov–Stein equation (5.6): more precisely, $S(\lambda) \in \mathcal{S}^{p\times q}(\Omega_+)$ is said to be a solution of the augmented BIP if

(1) $S(\lambda)$ is a solution of the BIP, and

(2) $-\langle S(\lambda)C_{22}(M_{22}-\lambda N_{22})^{-1}y,\, C_{11}(M_{11}-\lambda N_{11})^{-1}x\rangle = x^*P_{12}y$ for every choice of $x \in \mathbb{C}^\mu$ and $y \in \mathbb{C}^\nu$.

LEMMA 6.1. *If $S(\lambda)$ is a solution of the BIP based on C, M and N (subject to the spectral conditions* (5.4) *and* (5.5)) *and if P is a nonnegative solution of the Lyapunov–Stein equation* (5.6), *the following conditions are equivalent*:

$$-\langle S(\lambda)C_{22}(M_{22} - \lambda N_{22})^{-1}y,\, C_{11}(M_{11} - \lambda N_{11})^{-1}x\rangle = x^*P_{12}y \tag{1}$$

for every choice of $x \in \mathbb{C}^\mu$ and $y \in \mathbb{C}^\nu$.

$$\left\langle \begin{bmatrix} I_p & -S(\lambda) \\ -S(\lambda)^* & I_q \end{bmatrix} F(\lambda)v,\, F(\lambda)u \right\rangle = u^*Pv \tag{2}$$

for every choice of u and v in $\mathbb{C}^n$.

$$\left\langle \begin{bmatrix} I_p & -S(\lambda) \\ -S(\lambda)^* & I_q \end{bmatrix} F(\lambda)u,\, F(\lambda)u \right\rangle = u^*Pu \tag{3}$$

for every choice of $u \in \mathbb{C}^n$.

$$\left\langle \begin{bmatrix} I_p & -S(\lambda) \\ -S(\lambda)^* & I_q \end{bmatrix} F(\lambda)u,\, F(\lambda)u \right\rangle \le u^*Pu \tag{4}$$

for every choice of $u \in \mathbb{C}^n$.

PROOF. Suppose first that (4) holds for some given solution $S(\lambda)$ of the BIP and some given solution $P \geq 0$ of the Lyapunov–Stein equation (5.6). Let u^*Qv denote the inner product on the left-hand side of the equality in (2) for this choice of $S(\lambda)$. By (4), the matrix

$$X = P - Q = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}$$

is nonnegative. Therefore, since X_{11} is the $\mu \times \mu$ zero matrix and X_{22} is the $\nu \times \nu$ zero matrix, the matrix

$$\begin{bmatrix} 0 & X_{12} \\ X_{12}^* & 0 \end{bmatrix}$$

is nonnegative. But this is only possible if $X_{12} = 0$. This serves to establish the nontrivial half of the equivalence of (4) and (3). The equivalence of (3) and (2) is a standard argument; the equivalence of (1) and (2) rests heavily on the proof of Theorem 3.1 and the fact that $S(\lambda)$ is assumed to be a solution of the BIP. The details are left to the reader. $\square$

7. Reproducing Kernel Hilbert Spaces

A Hilbert space $\mathcal{H}$ of $(m \times 1)$-vector-valued functions defined on some subset Δ of $\mathbb{C}$ is said to be a reproducing kernel Hilbert space (RKHS) if there exists an $m \times m$ mvf $K_\omega(\lambda)$ on $\Delta \times \Delta$ such that, for every choice of $\omega \in \Delta$, $u \in \mathbb{C}^m$, and $f \in \mathcal{H}$, we have $K_\omega u \in \mathcal{H}$ (as a function of λ), and

$$\langle f, K_\omega u\rangle_{\mathcal{H}} = u^* f(\omega). \tag{7.1}$$

The main facts are these:

- The RK (reproducing kernel) is unique; that is, if $K_\omega(\lambda)$ and $L_\omega(\lambda)$ are both RK's for the same RKHS, then $K_\omega(\lambda) = L_\omega(\lambda)$ for every choice of ω and λ in Δ.
- $K_\alpha(\beta)^* = K_\beta(\alpha)$. (7.2)
- For every choice of $\omega_1, \dots, \omega_n$ in Δ and $u_1, \dots, u_n$ in $\mathbb{C}^m$, we have

$$\sum_{i,j=1}^{n} u_j^* K_i(\omega_j) u_i \geq 0. \tag{7.3}$$

EXAMPLE 7.1. $H_2^m(\Omega_+)$ is an RKHS with RK

$$K_\omega(\lambda) = I_m / \rho_\omega(\lambda)$$

for each of the classical choices of Ω_+, where $\rho_\omega(\lambda)$ and the corresponding inner product are specified in Table 1. Basically this is just Cauchy's theorem for $H_2(\Omega_+)$.

EXAMPLE 7.2 ($\mathcal{H}(S)$ SPACES). For each choice of $S(\lambda) \in \mathcal{S}^{p\times q}(\Omega_+)$, the kernel

$$L_\omega(\lambda) = \frac{I_p - S(\lambda)S(\omega)^*}{\rho_\omega(\lambda)} \tag{7.4}$$

is positive in the sense exhibited in inequality (7.3). Perhaps the easiest way to see this is to observe that

$$\sum_{i,j=1}^{n} \xi_i^* \Lambda_{\alpha_j}(\alpha_i) \xi_j = \langle g, g\rangle - \big\langle \underline{\underline{p}} S^* g, \underline{\underline{p}} S^* g\big\rangle$$

with $g = \sum_{j=1}^{n} \xi_j / \rho_{\alpha_j}$; see Lemma 1.6 for help with the evaluation, if need be.

Because of this positivity, it follows on general grounds (see [Aronszajn 1950], for example) that $L_\omega(\lambda)$ is the reproducing kernel of exactly one RKHS, which we shall designate by $\mathcal{H}(S)$. The following beautiful characterization of $\mathcal{H}(S)$ is due to de Branges and Rovnyak [1966].

THEOREM 7.3. *Let $S \in \mathcal{S}^{p\times q}(\Omega_+)$, and for $f \in H_2^p(\Omega_+)$ let*

$$\kappa(f) = \sup\big\{\|f + Sg\|^2 - \|g\|^2 : g \in H_2^q \Omega_+\big\}. \tag{7.5}$$

Then

$$\mathcal{H}(S) = \big\{f \in H_2^p(\Omega_+) : \kappa(f) < \infty\big\}$$

and

$$\|f\|^2_{\mathcal{H}(S)} = \kappa(f).$$

PROOF. A proof for $\Omega_+ = \mathbb{D}$ can be found in [de Branges and Rovnyak 1966], but it goes through for the other two cases in just the same way. ☐

It is an instructive exercise to check that if $S(\lambda)$ is isometric a.e. on Ω_0, then $p \geq q$ and

$$\mathcal{H}(S) = H_2^p \ominus SH_2^q.$$

A number of useful properties of the space $\mathcal{H}(S)$ as well as references to more extensive lists are provided in [Dym 1994b, Section 6].

EXAMPLE 7.4. Let $\mathcal{M} = \{F(\lambda)u : u \in \mathbb{C}^n\}$, where $F(\lambda)$ is an $m \times n$ mvf with linearly independent columns $f_1(\lambda), \ldots, f_n(\lambda)$, and let P be any $n \times n$ positive definite matrix (that is, $P > 0$). Then the space $\mathcal{M}$, endowed with the inner product

$$\big\langle F(\lambda)u, F(\lambda)v \big\rangle_{\mathcal{M}} = v^* P u \tag{7.6}$$

for every choice of u and v in $\mathbb{C}^n$, is an RKHS with RK

$$K_\omega(\lambda) = F(\lambda)P^{-1}F(\omega)^*. \tag{7.7}$$

The verification is by direct computation.

8. A Special Class of Reproducing Kernel Hilbert Spaces

We shall be particularly interested in RKHS's of $(m \times 1)$-vector-valued meromorphic functions in $\mathbb{C}$ with RK's of a special form that will be described below in Theorem 8.1. The theorem is an elaboration of a fundamental result from [de Branges 1963]. It is formulated in terms of the polynomials $a(\lambda)$ and $b(\lambda)$ given in Table 1 in order to obtain a statement that is applicable to each of the three classical choices of Ω_+.

A set Δ is said to be symmetric with respect to Ω_0 (or $\rho_\omega(\lambda)$) if for every $\lambda \in \Delta$ (except 0 for $\Omega_0 = \mathbb{T}$) the point λ^o belongs to Δ; note that $\rho_\omega(\omega^o) = 0$. Recall that $\rho_\omega(\lambda) = a(\lambda)a(\omega)^* - b(\lambda)b(\omega)^*$.

THEOREM 8.1. *Let $\mathcal{H}$ be an RKHS of $(m \times 1)$-vector-valued functions that are analytic in an open nonempty subset Δ of $\mathbb{C}$ symmetric with respect to Ω_0. Then the reproducing kernel $K_\omega(\lambda)$ can be expressed in the form*

$$K_\omega(\lambda) = \frac{J - \Theta(\lambda)J\Theta(\omega)^*}{\rho_\omega(\lambda)}, \tag{8.1}$$

for some choice of $m \times m$ matrix-valued function $\Theta(\lambda)$ analytic in Δ and some signature matrix J, if and only if the following two conditions hold:

(1) *$\mathcal{H}$ is R_α-invariant for every $\alpha \in \Delta$.*

(2) *The structural identity*

$$\big\langle R_\alpha(bf), R_\beta(bg)\big\rangle_{\mathcal{H}} - \big\langle R_\alpha(af), R_\beta(ag)\big\rangle_{\mathcal{H}} = |ab' - ba'|^2 g(\beta)^* J f(\alpha) \tag{8.2}$$

holds for every choice of α, β *in* Δ *and* f, g *in* $\mathcal{H}$.

Moreover, in this case, the function $\Theta(\lambda)$ *that appears in* (8.1) *is unique up to a* J *unitary constant factor on the right. If there exists a point* $\gamma \in \Delta \cap \Omega_0$, *it can be taken equal to*

$$\Theta(\lambda) = I_m - \rho_\gamma(\lambda) K_\gamma(\lambda) J. \tag{8.3}$$

This formulation is adapted from [Alpay and Dym 1993b]; see especially Theorems 4.1, 4.3, and 4.4 of that reference. The restriction to the three choices of $a(\lambda)$ and $b(\lambda)$ specified earlier permits some simplification in the presentation, because the terms $r(a,b;\alpha)f$ and $r(b,a;\alpha)f$ that intervene there are constant multiples of $R_\alpha(af)$ and $R_\alpha(bf)$, respectively.

For the three cases of interest, the structural identity (8.2) can be reexpressed as

$$\big\langle (I + \alpha R_\alpha) f,\, (I + \beta R_\beta) g\big\rangle_{\mathcal{H}} - \langle R_\alpha f, R_\beta g\rangle_{\mathcal{H}} = g(\beta)^* J f(\alpha). \tag{8.4}$$

if $\Omega_+ = \mathbb{D}$,

$$\langle R_\alpha f, g\rangle_{\mathcal{H}} - \langle f, R_\beta g\rangle_{\mathcal{H}} - (\alpha - \beta^*)\langle R_\alpha f, R_\beta g\rangle_{\mathcal{H}} = 2\pi i g(\beta)^* J f(\alpha) \tag{8.5}$$

if $\Omega_+ = \mathbb{C}_+$, and

$$\langle R_\alpha f, g\rangle_{\mathcal{H}} + \langle f, R_\beta g\rangle_{\mathcal{H}} + (\alpha + \beta^*)\langle R_\alpha f, R_\beta g\rangle_{\mathcal{H}} = -2\pi g(\beta)^* J f(\alpha) \tag{8.6}$$

if $\Omega_+ = \Pi_+$.

Formula (8.5) appears in [de Branges 1963]; formula (8.4) is equivalent to a formula of Ball [1975], who adapted de Branges' work to the disc, including an important technical improvement from [Rovnyak 1968].

The role of the two conditions in Theorem 8.1 becomes particularly transparent when $\mathcal{H}$ is finite-dimensional. Indeed, if the n-dimensional space $\mathcal{M}$ considered in Example 7.4 is R_α invariant for some point α in the domain of analyticity of $F(\lambda)$, then, by Theorem 4.1, $F(\lambda)$ can be expressed in the form

$$F(\lambda) = V(M - \lambda N)^{-1} \tag{8.7}$$

with M and N satisfying (4.2). Thus R_α-invariance forces the elements of $\mathcal{M}$ to be rational of the indicated form. Since

$$(R_\beta F)(\lambda) = F(\lambda) N (M - \beta N)^{-1}$$

for every point β at which the matrix $M - \beta N$ is invertible, that is, for every $\beta \in \mathcal{A}_F$, the domain of analyticity of F, it is readily checked that

$$\begin{aligned}\langle R_\alpha F u, F v\rangle_{\mathcal{M}} &= \langle F N (M - \alpha N)^{-1} u, F v\rangle_{\mathcal{M}} \\ &= v^* P N (M - \alpha N)^{-1} u,\end{aligned} \tag{8.8}$$

and similarly that

$$\langle Fu, R_\beta Fv\rangle_{\mathcal{M}} = v^*(M^* - \beta^* N^*)^{-1} N^* P u \tag{8.9}$$

and

$$\langle R_\alpha Fu, R_\beta v\rangle_{\mathcal{M}} = v^*(M^* - \beta^* N^*)^{-1} N^* P N (M - \alpha N)^{-1} u \tag{8.10}$$

for every choice of α, β in $\mathcal{A}_F$ and u, v in $\mathbb{C}^n$. For each of the three special choices of Ω_+ under consideration, it is now readily checked that the structural identity (8.2) reduces to a matrix equation for P by working out (8.4)–(8.6) with the aid of (8.8)–(8.10). In other words, in a finite-dimensional R_α-invariant space $\mathcal{M}$ with Gram matrix P, the de Branges structural identity is equivalent to a Lyapunov–Stein equation for P. This was first established explicitly in [Dym 1989b] by a considerably lengthier calculation. If F is analytic at zero, we may presume that $M = I_n$ in (8.7) and take $\alpha = \beta = 0$ in the structural identity (8.2).

THEOREM 8.2. *Let $F(\lambda) = V(M - \lambda N)^{-1}$ be an $m \times n$ matrix-valued function with* $\det(M - \lambda N) \not\equiv 0$ *and linearly independent columns, and let the vector space*

$$\mathcal{M} = \{F(\lambda)u : u \in \mathbb{C}^n\}$$

be endowed with the inner product

$$\langle Fu, Fv\rangle_{\mathcal{M}} = v^* P u,$$

based on an $n \times n$ positive definite matrix P. Then $\mathcal{M}$ is a finite-dimensional RKHS with RK $K_\omega(\lambda)$ given by (7.7).

The RK can be expressed in the form

$$K_\omega(\lambda) = \frac{J - \Theta(\lambda) J \Theta(\omega)^*}{\rho_\omega(\lambda)}$$

with $\rho_\omega(\lambda)$ as in Table 1 *if and only if P is a solution of the equation*

$$M^*PM - N^*PN = V^*JV \qquad \text{for } \Omega_+ = \mathbb{D}, \tag{8.11}$$

$$M^*PN - N^*PM = 2\pi i V^*JV \qquad \text{for } \Omega_+ = \mathbb{C}_+, \tag{8.12}$$

$$M^*PN + N^*PM = -2\pi V^*JV \qquad \text{for } \Omega_+ = \Pi_+. \tag{8.13}$$

Moreover, in each of these cases $\Theta(\lambda)$ is uniquely specified up to a J unitary constant multiplier on the right by the formula

$$\Theta(\lambda) = I_m - \rho_\gamma(\lambda) F(\lambda) P^{-1} F(\gamma)^* J \tag{8.14}$$

for any choice of the point $\gamma \in \Omega_0 \cap \mathcal{A}_F$.

PROOF. This is an easy consequence of Theorem 8.1 and the discussion preceding the statement of this theorem. The basic point is that, because of the special form of F, (8.4) holds if and only if P is a solution of (8.11); similarly (8.5) holds if and only if P is a solution of (8.12), and (8.6) if and only if P is a solution of (8.13). □

It is well to note that formula (8.14) is a realization formula for $\Theta(\lambda)$, and that in the usual notation of (4.3) and (4.4) it depends only upon A, C and P. It can be reexpressed in one of the standard A, B, C, D forms by elementary manipulations.

Formulas (8.3) and (8.14) for $\Theta(\lambda)$ are obtained by matching the right-hand sides of (8.1) and (8.9). This leads to the formula

$$\Theta(\lambda)J\Theta(\omega)^* = J - \rho_\omega(\lambda)F(\lambda)P^{-1}F(\omega)^*,$$

which is clearly a necessary constraint on $\Theta(\lambda)$ since $\mathcal{M}$ has only one reproducing kernel, and hence any two recipes for it must agree. The final formula emerges upon setting $\omega = \gamma \in \Omega_0 \cap \mathcal{A}_F$ and then discarding J unitary constant factors on the right such as $\Theta(\gamma)^{-1}$ and J. Thus the general theory of "structured" reproducing kernel spaces as formulated in Theorem 8.1 yields formula (8.14). However, once the formula is available, it can be used to check that

$$F(\lambda)P^{-1}F(\omega)^* = \frac{J - \Theta(\lambda)J\Theta(\omega)^*}{\rho_\omega(\lambda)} \tag{8.15}$$

for every pair of points λ, ω in $\mathcal{A}_F$ by straightforward calculation, using only the fact that P is a solution of one of the equations (8.11)–(8.13), according to the choice of Ω_+. More information and references may be found in [Dym 1994b, Section 5], which was used heavily in the preparation of this section.

9. Diversion on J-Inner Functions

An $m \times m$ mvf $\Theta(\lambda)$ that is meromorphic in Ω_+ is said to be J-inner if it meets the following two conditions, where $\mathcal{A}_\Theta$ denotes the domain of analyticity of the mvf $\Theta(\lambda)$:

$$\Theta(\lambda)^*J\Theta(\lambda) \le J \quad \text{for } \lambda \in \Omega_+ \cap \mathcal{A}_\Theta. \tag{9.1}$$

$$\Theta(\lambda)^*J\Theta(\lambda) = J \quad \text{for a.e. point} \lambda \in \Omega_0. \tag{9.2}$$

The evaluations in (9.2) are taken as nontangential boundary limits. Such limits exist, because the inequality (9.1) insures that every entry in the mvf $\Theta(\lambda)$ can be expressed as the ratio of two functions in $\mathcal{S}^{1\times 1}(\Omega_+)$; see, for example, [Dym 1989a, Theorem 1.1].

The identity (8.15) implies that

$$J - \Theta(\omega)J\Theta(\omega)^* = \rho_\omega(\omega)F(\omega)P^{-1}F(\omega)^*$$

and hence that

$$\Theta(\omega)J\Theta(\omega)^* \le J \quad \text{for } \omega \in \Omega_+ \cap \mathcal{A}_\Theta, \tag{9.3}$$

$$\Theta(\omega)J\Theta(\omega)^* = J \quad \text{for } \omega \in \Omega_0 \cap \mathcal{A}_\Theta. \tag{9.4}$$

The constraints (9.2) and (9.3) are equivalent to the assertion that $\Theta(\lambda)$ is J-inner, even though the "stars" are on the wrong side. The equivalence of (9.2) and (9.4) is self-evident. That of (9.1) and (9.3) takes a little more doing; see [Dym 1989a, pp. 16, 21], for example, for a couple of approaches. In any event, we shall be able to derive most of what we need directly from (9.3) and (9.4) without making use of the equivalence.

Upon writing

$$\Theta(\lambda) = \begin{array}{l}\left[\begin{matrix} \Theta_{11}(\lambda) & \Theta_{12}(\lambda) \\ \underbrace{\Theta_{21}(\lambda)}_{p} & \underbrace{\Theta_{22}(\lambda)}_{q} \end{matrix}\right] \begin{matrix} \}p \\ \}q \end{matrix}\end{array}$$

in the indicated block form, it is readily seen from the $(2,2)$ block of the inequality (9.3) that

$$\Theta_{22}(\lambda)\Theta_{22}(\lambda)^* \ge I_q + \Theta_{21}(\lambda)\Theta_{21}(\lambda)^*$$

for every point $\lambda \in \Omega_+ \cap \mathcal{A}_\Theta$. This implies that $\Theta_{22}(\lambda)$ is invertible for all such points and hence that the $m \times m$ mvf

$$\Sigma(\lambda) = \begin{bmatrix} I_p & -\Theta_{12}(\lambda) \\ 0 & -\Theta_{22}(\lambda) \end{bmatrix}^{-1} \begin{bmatrix} \Theta_{11}(\lambda) & 0 \\ \Theta_{21}(\lambda) & -I_q \end{bmatrix} \tag{9.5}$$

is well defined for $\lambda \in \Omega_+ \cap \mathcal{A}_\Theta$. Let $[\ ]^{-*}$ stand for $\big([\ ]^*\big)^{-1}$. The identity

$$I_m - \Sigma(\lambda)\Sigma(\omega)^* = \begin{bmatrix} I_p & -\Theta_{12}(\lambda) \\ 0 & -\Theta_{22}(\lambda) \end{bmatrix}^{-1} \big(J - \Theta(\lambda)J\Theta(\omega)^*\big) \begin{bmatrix} I_p & -\Theta_{21}(\omega) \\ 0 & -\Theta_{22}(\omega) \end{bmatrix}^{-*}, \tag{9.6}$$

which is readily checked by direct calculation, implies that $\Sigma(\lambda)$ is contractive and analytic in $\Omega_+ \cap \mathcal{A}_\Theta$ and hence in fact analytic in all of Ω_+. Thus $\Sigma(\lambda) \in \mathcal{S}^{m\times m}(\Omega_+)$. The mvf $\Sigma(\lambda)$ is termed the Potapov–Ginzburg transform of $\Theta(\lambda)$. It is also unitary a.e. on Ω_0 (when, as in the present case, $\Theta(\lambda)$ is J-inner; that is, $\Theta(\lambda)$ being J-inner implies that $\Sigma(\lambda)$ is inner).

By (9.5), the entries in the block decomposition

$$\Sigma(\lambda) = \begin{bmatrix} \Sigma_{11}(\lambda) & \Sigma_{12}(\lambda) \\ \Sigma_{21}(\lambda) & \Sigma_{22}(\lambda) \end{bmatrix}$$

are given by the formulas

$$\begin{aligned} \Sigma_{11}(\lambda) &= \Theta_{11}(\lambda) - \Theta_{12}(\lambda), \Theta_{22}(\lambda)^{-1}\Theta_{21}(\lambda), & \Sigma_{12}(\lambda) &= \Theta_{12}(\lambda)\Theta_{22}(\lambda)^{-1} \\ \Sigma_{21}(\lambda) &= -\Theta_{22}(\lambda)^{-1}\Theta_{21}(\lambda), & \Sigma_{22}(\lambda) &= \Theta_{22}(\lambda)^{-1}. \end{aligned} \tag{9.7}$$

Moreover, since $\Sigma(\lambda) \in \mathcal{S}^{m\times m}(\Omega_+)$, it follows that $\Sigma_{11}(\lambda) \in \mathcal{S}^{p\times p}(\Omega_+)$, $\Sigma_{12}(\lambda) \in \mathcal{S}^{p\times q}(\Omega_+)$, $\Sigma_{21}(\lambda) \in \mathcal{S}^{q\times p}(\Omega_+)$, and $\Sigma_{22}(\lambda) \in \mathcal{S}^{q\times q}(\Omega_+)$. Consequently,

$$\Theta_{21}(\lambda)\mathcal{E}(\lambda) + \Theta_{22}(\lambda) = \Theta_{22}(\lambda)\big(I_q - \Sigma_{21}(\lambda)\mathcal{E}(\lambda)\big)$$

is invertible in $\Omega_+ \cap \mathcal{A}_\Theta$ for every choice of $\mathcal{E}(\lambda) \in \mathcal{S}^{p\times q}(\Omega_+)$. Thus the linear fractional transformation

$$T_\Theta[\mathcal{E}] = \big(\Theta_{11}(\lambda)\mathcal{E}(\lambda) + \Theta_{12}(\lambda)\big)\big(\Theta_{21}(\lambda)\mathcal{E}(\lambda) + \Theta_{22}(\lambda)\big)^{-1} \tag{9.8}$$

is well defined for $\mathcal{E}(\lambda) \in \mathcal{S}^{p\times q}(\Omega_+)$. The following facts are readily checked:

$$T_\Theta[\mathcal{E}] \in \mathcal{S}^{p\times q}(\Omega_+);$$

$$T_\Theta[0] = \Sigma_{12}(\lambda); \tag{9.9}$$

$$T_\Theta[\mathcal{E}] = \Sigma_{12}(\lambda) + \Sigma_{11}(\lambda)\mathcal{E}(\lambda)\big(I_q - \Sigma_{21}(\lambda)\mathcal{E}(\lambda)\big)^{-1}\Sigma_{22}(\lambda). \tag{9.10}$$

It is also easy to check that, if $\Psi(\lambda)$ also is J-inner, then

$$T_{\Theta\Psi}[\mathcal{E}] = T_\Theta\big[T_\Psi[\mathcal{E}]\big]. \tag{9.11}$$

Linear fractional transformations of the form (9.10) were extensively studied in [Redheffer 1960].

THEOREM 9.1. *If $\Theta(\lambda)$ is J-inner and $\Sigma(\lambda)$ denotes the Potapov–Ginzburg transform of $\Theta(\lambda)$ (that is, if $\Sigma(\lambda)$ is given by* (9.5)), *and if*

$$f(\lambda) = \begin{bmatrix} g(\lambda) \\ h(\lambda) \end{bmatrix}$$

belongs to the RKHS $\mathcal{H}(\Theta)$, we have:

(1) $\begin{bmatrix} I_p & -\Sigma_{12}(\lambda) \\ -\Sigma_{12}(\lambda)^* & I_q \end{bmatrix}\begin{bmatrix} g(\lambda) \\ h(\lambda) \end{bmatrix} \in H_2^p(\Omega_+) \oplus H_2^q(\Omega_+)^\perp$.

(2) $\Sigma_{11}^* g \in H_2^p(\Omega_+)^\perp$.

(3) $\Sigma_{22} h \in H_2^q(\Omega_+)$ *and*

$$\|f\|^2_{\mathcal{H}(\Theta)} = \|g - \Sigma_{12}h\|^2 + \|\Sigma_{22}h\|^2. \tag{9.12}$$

PROOF. This is Theorem 2.7 of [Dym 1989a] restated in the current notation. □

COROLLARY 9.2. *In the setting of the theorem*,

$$\|f\|^2_{\mathcal{H}(\Theta)} = \left\langle \begin{bmatrix} I_p & -\Sigma_{12}(\lambda) \\ -\Sigma_{12}(\lambda)^* & I_q \end{bmatrix}\begin{bmatrix} g(\lambda) \\ h(\lambda) \end{bmatrix}, \begin{bmatrix} g(\lambda) \\ h(\lambda) \end{bmatrix} \right\rangle. \tag{9.13}$$

PROOF. The right-hand side of (9.12) is equal to

$$\langle g - \Sigma_{12}h, g - \Sigma_{12}h\rangle + \langle \Sigma_{22}h, \Sigma_{22}h\rangle$$
$$= \langle g - \Sigma_{12}h, g\rangle - \langle \Sigma_{12}^* g, h\rangle + \big\langle (\Sigma_{12}^*\Sigma_{12} + \Sigma_{22}^*\Sigma_{22})h,\, h\big\rangle.$$

But this is equal to the right-hand side of (9.13) since

$$\Sigma_{12}(\lambda)\Sigma_{12}(\lambda) + \Sigma_{22}(\lambda)^*\Sigma_{22}(\lambda) = I_q$$

for $\lambda \in \Omega_0$. ☐

10. Sufficiency When $P > 0$

To this point we know that if the BIP based on the matrices C, M and N (subject to the spectral constraints imposed in (5.4) and (5.5)) admits a solution, then there exists a solution $P \geq 0$ of the Lyapunov–Stein equation (5.6) and of the augmented BIP based on C, M, N and P. Our next objective is to establish a converse when $P > 0$.

THEOREM 10.1. *If P is a positive definite solution of the Lyapunov–Stein equation* (5.6), *and if $\Theta(\lambda)$ is the J-inner mvf defined by* (8.14), *then $T_\Theta[\mathcal{E}]$ is a solution of the augmented BIP for every choice of $\mathcal{E}(\lambda) \in \mathcal{S}^{p\times q}(\Omega_+)$.*

PROOF. Let $S(\lambda) = T_\Theta[\mathcal{E}]$ for some choice of $\mathcal{E}(\lambda) \in \mathcal{S}^{p\times q}(\Omega_+)$ and let

$$f_j(\lambda) = \begin{bmatrix} g_j(\lambda) \\ h_j(\lambda) \end{bmatrix}$$

denote the j'th element in the associated RKHS $\mathcal{H}(\Theta)$, where $f_j(\lambda) \in H_2^m(\Omega_+)$ for $j = 1, \ldots, \mu$ and $f_j(\lambda) \in H_2^m(\Omega_+)^\perp$ for $j = \mu+1, \ldots, n$. Also, let

$$f(\lambda) = \begin{bmatrix} g(\lambda) \\ h(\lambda) \end{bmatrix} = F(\lambda)u$$

denote an arbitrary element of the same RKHS. We now proceed in steps.

Step 1. $\big(g(\lambda) - S(\lambda)h(\lambda)\big) \in H_2^p(\Omega_+)$.

Proof of Step 1. Theorem 9.1 guarantees that

$$\big(g(\lambda) - \Sigma_{12}(\lambda)h(\lambda)\big) \in H_2^p(\Omega_+).$$

Therefore, in view of formula (9.10), it remains only to show that

$$\Sigma_{11}(\lambda)\mathcal{E}(\lambda)\big(I_q - \Sigma_{21}\mathcal{E}(\lambda)\big)^{-1}\Sigma_{22}(\lambda)h(\lambda) \in H_2^p(\Omega_+) \tag{10.1}$$

and hence, since

$$\Sigma_{11}(\lambda)\mathcal{E}(\lambda)\big(I_q - \Sigma_{21}(\lambda)\mathcal{E}(\lambda)\big)^{-1} \in H_\infty^{p\times q}(\Omega_+) \tag{10.2}$$

in the present setting, this follows from part (3) of Theorem 9.1.

Step 2. $\big(-S(\lambda)^*g(\lambda) + h(\lambda)\big) \in H_2^q(\Omega_+)^\perp$.

Proof of Step 2. The proof is similar to that of Step 1: Theorem 9.1 guarantees that

$$\big(-\Sigma_{21}(\lambda)^* g(\lambda) + h(\lambda)\big) \in H_2^q(\Omega_+)^\perp$$

and therefore, by (9.10), it remains only to show that

$$\Sigma_{22}(\lambda)^*\big(I_q - \Sigma_{21}(\lambda)\mathcal{E}(\lambda)\big)^{-*}\mathcal{E}(\lambda)^*\Sigma_{11}(\lambda)^* g(\lambda) \in H_2^q(\Omega_+)^\perp.$$

But this goes through much as before except that now we invoke assertion (2) instead of (3) of Theorem 9.1.

Step 3. $-\big\langle \Sigma_{12}(\lambda)h_j(\lambda), g_i(\lambda)\big\rangle = p_{ij}$ for $i = 1, \dots, \mu$ and $j = \mu{+}1, \dots, n$, where p_{st} denotes the (s,t) entry of P for $s, t = 1, \dots, n$.

Proof of Step 3. By the corollary to Theorem 9.1,

$$\begin{aligned} p_{ij} &= \big\langle f_j(\lambda), f_i(\lambda)\big\rangle_{\mathcal{H}(\Theta)} \\ &= \left\langle \begin{bmatrix} I_p & -\Sigma_{12}(\lambda) \\ -\Sigma_{12}(\lambda)^* & I_q \end{bmatrix} \begin{bmatrix} g_j(\lambda) \\ h_j(\lambda) \end{bmatrix}, \begin{bmatrix} g_i(\lambda) \\ h_i(\lambda) \end{bmatrix} \right\rangle \\ &= \big\langle g_j(\lambda) - \Sigma_{12}(\lambda)h_j(\lambda),\, g_i(\lambda)\big\rangle + \big\langle -\Sigma_{12}(\lambda)^* g_j(\lambda) + h_j(\lambda),\, h_i(\lambda)\big\rangle \end{aligned}$$

for every choice of $i, j = 1, \dots, n$. But now, if $i = 1, \dots, \mu$ and $j = \mu{+}1, \dots, n$, this is easily seen to reduce to the asserted identity with the help of Step 2.

Step 4. $-\big\langle S(\lambda)h_j(\lambda), g_i(\lambda)\big\rangle = p_{ij}$ for $i = 1, \dots, \mu$ and $j = \mu{+}1, \dots, n$.

Proof of Step 4. In view of Step 3 and formula (9.10), it remains to show that

$$-\big\langle \Sigma_{11}(\lambda)\mathcal{E}(\lambda)\big(I_q - \Sigma_{21}(\lambda)\mathcal{E}(\lambda)\big)^{-1}\Sigma_{22}(\lambda)h_j(\lambda),\, g_i(\lambda)\big\rangle = 0$$

for $i = 1, \dots, \mu$ and $j = \mu{+}1, \dots, n$. But this is easily checked since Theorem 9.1 guarantees that

$$\begin{aligned} \Sigma_{11}(\lambda)^* g_i(\lambda) &\in H_2^p(\Omega_+)^\perp \quad &&(\text{even for } i = 1, \dots, n), \\ \Sigma_{22}(\lambda)h_j(\lambda) &\in H_2^q(\Omega_+) \quad &&(\text{even for } j = 1, \dots, n), \end{aligned}$$

and

$$\big(I_q - \Sigma_{21}(\lambda)\mathcal{E}(\lambda)\big)^{-1} \in H_\infty^{q\times q}(\Omega_+).$$

This completes the proof of the step and the theorem. □

We remark that most of this analysis goes through in one form or another for problems with infinitely many interpolation constraints. Indeed Theorem 9.1 is applicable to infinite-dimensional spaces and the only point in the proof of Theorem 10.1 that depends critically on the specified form of the BIP (including the assumptions on $\sigma(A_1)$ and $\sigma(A_2)$) is the assertion that

$$\big(I_q - \Sigma_{21}(\lambda)\mathcal{E}(\lambda)\big)^{-1} \in H_\infty^{q\times q}(\Omega_+), \tag{10.3}$$

which was used implicitly in Steps 1 and 2 and explicitly in Step 4. In the present setting, (10.3) can be justified by direct estimates based on the specific formula for $\Sigma_{21}(\lambda)$ provided in (11.25). In more general settings, the most that can be said is that $\big(I_q - \Sigma_{21}(\lambda)\mathcal{E}(\lambda)\big)^{-1}$ is an outer function in the Smirnov class $\mathcal{N}_+^{q\times q}(\Omega_+)$. The needed estimates to justify the proof of Theorem 10.1 are then obtained by applications of the maximum principle in the Smirnov class. (Thus for example, in the proof of Step 1, (10.1) is established by showing first that $\Sigma_{11}\mathcal{E}\big(I_q - \Sigma_{21}\mathcal{E}\big)^{-1} \in \mathcal{N}_+^{p\times q}(\Omega_+)$. Then, since $\Sigma_{22}h \in H_2^q(\Omega_+)$, the product of these two terms belongs to $\mathcal{N}_+^p(\Omega_+) \cap L_2^p(\Omega_+)$ and hence by the maximum principle to $H_2^p(\Omega_+)$.) The author first learned the power of estimates in matrix Smirnov classes from [Arov 1973]; much useful information may also be found in [Katsnelson and Kirstein 1997].

The next theorem states that all solutions to the augmented BIP are obtained by the parametrization furnished in Theorem 10.1. In order to keep the length of the discussion under control, we shall rely a little more on outside references than has been our practice to this point.

THEOREM 10.2. *If P is a positive definite solution of the Lyapunov–Stein equation* (5.6), *then every solution $S(\lambda)$ of the augmented BIP based on this choice of P can be expressed as a linear fractional transformation of the form $S(\lambda) = T_\Theta[\mathcal{E}]$ in terms of the associated $\Theta(\lambda)$ for some choice of $\mathcal{E}(\lambda) \in \mathcal{S}^{p\times q}(\Omega_+)$.*

PROOF. Let $S(\lambda) \in \mathcal{S}^{p\times q}(\Omega_+)$ be a solution of the augmented BIP based on P and let $X(\lambda) = [I_p \;\; -S(\lambda)]$. Then, as follows from [Dym 1994b, pp. 205–206, 210], $X(\lambda)F(\lambda) \in \mathcal{H}(S)$ and the corresponding Gram matrix Q defined by the rule

$$u^*Qv = \big\langle X(\lambda)F(\lambda)v,\, X(\lambda)F(\lambda)u\big\rangle_{\mathcal{H}(S)}$$

for every choice of u and v in $\mathbb{C}^n$ is majorized by P, that is,

$$Q \le P. \tag{10.4}$$

Now, in terms of the notation introduced in Example 7.2, the kernel

$$\begin{aligned}\frac{X(\lambda)\Theta(\lambda)J\Theta(\omega)^*X(\omega)^*}{\rho_\omega(\lambda)} &= \frac{X(\lambda)JX(\omega)^*}{\rho_\omega(\lambda)} - X(\lambda)\Big(\frac{J-\Theta(\lambda)J\Theta(\omega)^*}{\rho_\omega(\lambda)}\Big)X(\omega)^* \\ &= \Lambda_\omega(\lambda) - X(\lambda)F(\lambda)P^{-1}F(\omega)^*X(\omega)^*.\end{aligned}$$

Our next objective is to show that this kernel is positive, that is, that

$$\sum_{i,j=1}^{k} y_i^*\big(\Lambda_{\alpha_j}(\alpha_i) - (XF)(\alpha_i)P^{-1}(XF)(\alpha_j)^*\big)y_j \ge 0$$

for every choice of vectors $y_1,\dots,y_k$ in $\mathbb{C}^p$ and points $\alpha_1,\dots,\alpha_k$ in Ω_+. In view of (10.4), it is enough to show that this inequality holds with Q^{-1} in place of P^{-1}. But now as

$$F(\lambda) = \big[f_1(\lambda) \;\; \cdots \;\; f_n(\lambda)\big],$$

the second term in the sum with P replaced by Q can be reexpressed as

$$y_i^*(XF)(\alpha_i)Q^{-1}(XF)(\alpha_j)^*y_j = \sum_{u,v=1}^{n} y_i^*(Xf_u)(\alpha_i)(Q^{-1})_{uv}(Xf_v)(\alpha_j)^*y_j$$

$$= \sum_{u,v=1}^{n} \langle Xf_u, \Lambda_{\alpha_i}y_i\rangle_{\mathcal{H}(S)}(Q^{-1})_{uv}\langle \Lambda_{\alpha_j}y_j, Xf_v\rangle_{\mathcal{H}(S)}.$$

Thus, upon setting

$$f_0 = \sum_{ju=1}^{k} \Lambda_{\alpha_j}y_j,$$

it now follows readily that

$$\sum_{i,j=1}^{k} y_i^*(XF)(\alpha_i)P^{-1}(XF)(\alpha_j)^*y_j \le \sum_{u,v=1}^{n} \langle Xf_u, f_0\rangle_{\mathcal{H}(S)}(Q^{-1})_{uv}\langle f_0, Xf_v\rangle_{\mathcal{H}(S)}$$

$$= \left\|\Pi f_0(\lambda)\right\|^2_{\mathcal{H}(S)},$$

where Π denotes the orthogonal projection of $\mathcal{H}(S)$ onto the subspace spanned by the elements $X(\lambda)f_u(\lambda)$, for $u = 1, \dots, n$. The asserted inequality is now clear since

$$\sum_{i,j=1}^{k} y_i^*\Lambda_{\alpha_j}(\alpha_i)y_j = \left\|f_0(\lambda)\right\|^2_{\mathcal{H}(S)}.$$

The desired conclusion now follows directly from [Dym 1989a, Theorem 3.8], since the "admissibility" needed to invoke that theorem amounts to the kernel dealt with above being positive. □

11. Explicit Formulas

In this section we shall provide explicit formulas for the mvfs $\Theta(\lambda)$ and $\Sigma(\lambda)$ in terms of the data C, M and N of the BIP and the solution P of the Lyapunov–Stein equation when $P > 0$. We shall do this in a more general setting than is needed for the three classical choices of Ω_+ that were considered earlier because it enhances the usefulness of the formulas and does not involve any extra work, just a little extra notation.

To this end, let $a(\lambda)$ and $b(\lambda)$ denote a pair of functions that are defined and analytic in an open nonempty connected set $\Omega \subset \mathbb{C}$ and assume that the subsets

$$\Omega_+ = \left\{\lambda \in \Omega : |a(\lambda)|^2 - |b(\lambda)|^2 > 0\right\}$$

and

$$\Omega_- = \left\{\lambda \in \Omega : |a(\lambda)|^2 - |b(\lambda)|^2 < 0\right\}$$

are both nonempty. This implies that

$$\Omega_0 = \left\{\lambda \in \Omega : |a(\lambda)|^2 - |b(\lambda)|^2 = 0\right\}$$

is nonempty and that in fact, as is shown in [Alpay and Dym 1996], Ω_0 contains an open arc. These definitions of $\Omega_\pm$ and Ω_0 are consistent with the earlier ones and the choices $\Omega = \mathbb{C}$ with $a(\lambda)$ and $b(\lambda)$ as in Table 1.

Let $\mathfrak{A} \in \mathbb{C}^{n\times n}$, $\mathfrak{B} \in \mathbb{C}^{n\times n}$, $C_1 \in \mathbb{C}^{p\times n}$, $C_2 \in \mathbb{C}^{q\times n}$ and $P \in \mathbb{C}^{n\times n}$ be fixed matrices such that $P \geq 0$,

$$\mathfrak{A}^*P\mathfrak{A} - \mathfrak{B}^*P\mathfrak{B} = C_1^*C_1 - C_2^*C_2, \tag{11.1}$$

and the determinant of the mvf

$$G(\lambda) = a(\lambda)\mathfrak{A} - b(\lambda)\mathfrak{B} \tag{11.2}$$

does not vanish identically: $\det G(\lambda) \not\equiv 0$ in Ω.

We shall also make frequent use of the $n \times n$ mvf

$$H(\lambda) = b(\lambda)\mathfrak{A}^* - a(\lambda)\mathfrak{B}^* \tag{11.3}$$

and the $m \times n$ mvf

$$F(\lambda) = CG(\lambda)^{-1}, \tag{11.4}$$

where

$$C = \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} \quad \text{and} \quad m = p + q.$$

The functions $\rho_\omega(\lambda)$ and $\delta_\omega(\lambda)$ are expressed in terms of $a(\lambda)$ and $b(\lambda)$ just as before:

$$\rho_\omega(\lambda) = a(\lambda)a(\omega)^* - b(\lambda)b(\omega)^* \tag{11.5}$$

and

$$\delta_\omega(\lambda) = a(\omega)b(\lambda) - b(\omega)a(\lambda) \tag{11.6}$$

for every choice of λ and ω in Ω.

We shall refer to (11.1) as the General Lyapunov–Stein (GLS) equation. This usage too is easily seen to be consistent with the Lyapunov–Stein equation (5.6), which intervenes for the classical choices of Ω_+ when one defines $\mathfrak{A}$ and $\mathfrak{B}$ in terms of M and N via the formula

$$M - \lambda N = a(\lambda)\mathfrak{A} - b(\lambda)\mathfrak{B} \tag{11.7}$$

for the three sets of choices of $a(\lambda)$ and $b(\lambda)$ appearing in Table 1.

Some algebraic identities.

LEMMA 11.1. *The identity*

$$\begin{aligned}\rho_\omega(\lambda)G(\lambda)^*PG(\omega) &+ \rho_\omega(\lambda)^*H(\lambda)PH(\omega)^* \\ &= \rho_\lambda(\lambda)G(\omega)^*PG(\omega) + \rho_\omega(\omega)H(\lambda)PH(\lambda)^*\end{aligned} \tag{11.8}$$

holds for every pair of points λ and ω in Ω. It is independent of (11.1).

Proof. This is a tedious but straightforward calculation, which does not depend upon the basic identity (11.1). The proof amounts to identifying the coefficients of like terms. Thus, for example, the coefficient of $\mathfrak{A}^*P\mathfrak{A}$ on the left hand side of (11.8) is equal to

$$\rho_\omega(\lambda)a(\lambda)^*a(\omega) + \rho_\omega(\lambda)^*b(\lambda)b(\omega)^* = \rho_\lambda(\lambda)a(\omega)^*a(\omega) + \rho_\omega(\omega)b(\lambda)b(\lambda)^*,$$

which is equal to the coefficient of $\mathfrak{A}^*P\mathfrak{A}$ on the right-hand side of (11.8). The coefficients of $\mathfrak{A}^*P\mathfrak{B}$, $\mathfrak{B}^*P\mathfrak{A}$ and $\mathfrak{B}^*P\mathfrak{B}$ are identified in the same way. □

Lemma 11.2. *The identity*

$$\rho_\lambda(\lambda)G(\omega)^*PG(\omega) + \rho_\omega(\omega)H(\lambda)PH(\lambda)^* \\ = \rho_\omega(\omega)G(\lambda)^*PG(\lambda) + \rho_\lambda(\lambda)H(\omega)PH(\omega)^* \quad (11.9)$$

holds for every pair of points λ and ω in Ω. It is independent of (11.1).

Proof. Let $L_\omega(\lambda)$ denote the left-hand side of (11.8). Then since

$$L_\omega(\lambda) = L_\lambda(\omega)^*,$$

the same invariance must hold true for the right-hand side of (11.8). □

Lemma 11.3. *The identities*

$$\rho_\omega(\lambda)G(\lambda)^*PG(\omega) + \rho_\omega(\lambda)^*G(\omega)^*PG(\lambda) \\ = \rho_\lambda(\lambda)G(\omega)^*PG(\omega)+\rho_\omega(\omega)H(\lambda)PH(\lambda)^*+\left|\rho_\omega(\lambda)\right|^2(\mathfrak{A}^*P\mathfrak{A} - \mathfrak{B}^*P\mathfrak{B}) \quad (11.10)$$

and

$$\rho_\omega(\lambda)H(\omega)PH(\lambda)^* + \rho_\omega(\lambda)^*H(\lambda)PH(\omega)^* \\ = \rho_\lambda(\lambda)G(\omega)^*PG(\omega)+\rho_\omega(\omega)H(\lambda)PH(\lambda)^*-\left|\rho_\omega(\lambda)\right|^2(\mathfrak{A}^*P\mathfrak{A} - \mathfrak{B}^*P\mathfrak{B}) \quad (11.11)$$

hold for every pair of points λ and ω in Ω. They are both independent of (11.1).

This, too, can be verified by a tedious but straightforward calculation.

The mvf $\Delta_\omega(\lambda)$. Let

$$\Delta_\omega(\lambda) = G(\omega)^*PG(\lambda) + \rho_\omega(\lambda)C_2^*C_2 \quad (11.12)$$

for every pair of points λ and ω in Ω. Clearly

$$\Delta_\omega(\lambda) = \Delta_\lambda(\omega)^*. \quad (11.13)$$

Lemma 11.4. *If P is a solution of the GLS equation* (11.1), *then*

$$\Delta_\omega(\lambda) = H(\lambda)PH(\omega)^* + \rho_\omega(\lambda)C_1^*C_1 \quad (11.14)$$

for every pair of points λ and ω in Ω.

PROOF. In view of (11.1), it suffices to show that

$$G(\omega)^*PG(\lambda) - H(\lambda)PH(\omega)^* = \rho_\omega(\lambda)\big(\mathfrak{A}^*P\mathfrak{A} - \mathfrak{B}^*P\mathfrak{B}\big). \tag{11.15}$$

But this is a straightforward computation. □

THEOREM 11.5. *If $P \geq 0$ is a solution of the GLS equation* (11.1) *and if $\omega \in \Omega_+ \cup \Omega_0$, and*

$$\operatorname{rank}\begin{bmatrix} P^{1/2}G(\omega) \\ P^{1/2}H(\omega)^* \\ C \end{bmatrix} = n, \tag{11.16}$$

then the following conclusions hold:

(1) $\Delta_\omega(\lambda)$ *is invertible for every point* $\lambda \in \Omega_+$.
(2) *If* $\omega \in \Omega_+$, *then* $\Delta_\omega(\lambda)$ *is invertible for every point* $\lambda \in \Omega_+$ *at which*

$$\operatorname{rank}\begin{bmatrix} P^{1/2}G(\lambda) \\ C \end{bmatrix} = n. \tag{11.17}$$

(3) *If* $\omega \in \Omega_0$ *and* $G(\omega)$ *is invertible, then* $\Delta_\omega(\lambda)$ *is invertible for every point* $\lambda \in \Omega_0 \setminus \{\omega\}$ *at which the rank condition* (11.17) *holds.*
(4) *If* $\omega \in \Omega_0$ *and* $G(\omega)$ *is invertible, then* $\Delta_\omega(\omega)$ *is invertible if and only if* $P > 0$.

PROOF. Suppose first that

$$\Delta_\omega(\lambda)y = 0$$

for some choice of λ and ω in $\Omega_+ \cup \Omega_0$ and $y \in \mathbb{C}^n$. Then, by (11.12) and (11.14),

$$\begin{aligned} 0 &= y^*\big(\rho_\omega(\lambda)^*\Delta_\omega(\lambda) + \rho_\omega(\lambda)\Delta_\omega(\lambda)^*\big)y \\ &= y^*\big(\rho_\omega(\lambda)^*(H(\lambda)PH(\omega)^* + \rho_\omega(\lambda)C_1^*C_1) + \rho_\omega(\lambda)(G(\lambda)^*PG(\omega) + \rho_\omega(\lambda)C_2^*C_2)\big)y \\ &= y^*\big(\rho_\lambda(\lambda)G(\omega)^*PG(\omega) + \rho_\omega(\omega)H(\lambda)PH(\lambda)^* + |\rho_\omega(\lambda)|^2(C_1^*C_1 + C_2^*C_2)\big)y. \end{aligned}$$

Therefore, since each of the summands is nonnegative for λ and ω in $\Omega_+ \cup \Omega_0$, it follows from Lemma 11.2 and the last line that

$$\begin{aligned} \rho_\lambda(\lambda)G(\omega)^*PG(\omega)y &= 0, & \rho_\omega(\omega)G(\lambda)^*PG(\lambda)y &= 0, \\ \rho_\lambda(\lambda)H(\omega)PH(\omega)^*y &= 0, & \rho_\omega(\omega)H(\lambda)PH(\lambda)^*y &= 0, \\ \rho_\omega(\lambda)C_1y &= 0, & \rho_\omega(\lambda)C_2y &= 0. \end{aligned}$$

The rest of the argument proceeds in cases that amount to figuring out which of the preceding six identities really come into play.

Case 1. If $\lambda \in \Omega_+$, then $\rho_\lambda(\lambda) > 0$ and $|\rho_\omega(\lambda)| > 0$, and hence

$$\Delta_\omega(\lambda)y = 0 \implies \begin{bmatrix} P^{1/2}G(\omega) \\ P^{1/2}H(\omega)^* \\ C \end{bmatrix} y = 0 \implies y = 0,$$

in view of the rank assumption (11.16). This serves to establish assertion (1).

Case 2. If $\lambda \in \Omega_0$ but $\omega \in \Omega_+$, then $|\rho_\omega(\omega)| > 0$ and $|\rho_\omega(\lambda)| > 0$ and hence

$$\Delta_\omega(\lambda)y = 0 \implies \begin{bmatrix} P^{1/2}G(\omega) \\ C \end{bmatrix} y = 0 \implies y = 0,$$

in view of the rank assumption (11.17). This serves to establish assertion (2).

Case 3. If $\lambda \in \Omega_0$ and $\omega \in \Omega_0$ but $\lambda \neq \omega$, then $|\rho_\omega(\lambda)| > 0$. Hence $\Delta_\omega(\lambda)y = 0$ implies $Cy = 0$. This conclusion comes from the last two of the six identities established above. It now follows further from the definition of $\Delta_\omega(\lambda)$ that $G(\omega)^*PG(\lambda)y = 0$ and hence, since $G(\omega)$ is assumed to be invertible, that $PG(\lambda)y = 0$. But this in turn implies that $P^{1/2}G(\lambda)y = 0$, since $P \geq 0$. Thus we see that, in this case,

$$\Delta_\omega(\lambda)y = 0 \implies \begin{bmatrix} P^{1/2}G(\omega) \\ C \end{bmatrix} y = 0 \implies y = 0$$

in view of the rank condition (11.17). This serves to complete the third assertion and so too the proof, since the fourth assertion is obvious. □

COROLLARY 11.6. *If $P > 0$ and $G(\omega)$ is invertible for some point $\omega \in \Omega_+ \cup \Omega_0$, then*:

(1) $\Delta_\omega(\lambda)$ *is invertible for every point* $\lambda \in \Omega_+$.
(2) $\Delta_\omega(\lambda)$ *is invertible for every point* $\lambda \in \Omega_0$ *at which*

$$\operatorname{rank} \begin{bmatrix} G(\lambda) \\ C \end{bmatrix} = n. \tag{11.18}$$

We now assume that $P > 0$ is a positive definite solution of the GLS equation (11.1) and, for a fixed point $\gamma \in \Omega_0$ at which $G(\gamma)$ is invertible, define

$$\Theta(\lambda) = I_m - \rho_\gamma(\lambda)F(\lambda)P^{-1}F(\gamma)^*J$$

for every point $\lambda \in \Omega$ at which $G(\lambda)$ is invertible, just as in (8.3). (Strictly speaking, it would be better to write $\Theta_\gamma(\lambda)$ instead of $\Theta(\lambda)$ in order to indicate the dependence upon the "normalization" point γ, but this makes the formulas involving subblocks awkward.)

Thus, upon setting

$$\varphi_\omega(\lambda) = \rho_\omega(\lambda)G(\lambda)^{-1}P^{-1}G(\omega)^{-*}, \tag{11.20}$$

for those points λ and ω in Ω at which the indicated inverses exist, it is readily seen that

$$\Theta_{11}(\lambda) = I_p - C_1\varphi_\gamma(\lambda)C_1^*, \tag{11.21}$$

$$\Theta_{12}(\lambda) = C_1\varphi_\gamma(\lambda)C_2^*, \tag{11.22}$$

$$\Theta_{21}(\lambda) = -C_2\varphi_\gamma(\lambda)C_1^*, \tag{11.23}$$

$$\Theta_{22}(\lambda) = I_q + C_2\varphi_\gamma(\lambda)C_2^*. \tag{11.24}$$

Since $\Theta(\lambda)$ is J-inner with respect to Ω_+, the Potapov–Ginzburg transform $\Sigma(\lambda)$ is well defined (by formulas (9.5) and (9.7)) and is inner.

THEOREM 11.7. *If $P > 0$ is a solution of the GLS equation* (11.1) *and if $\gamma \in \Omega_0$ and $G(\gamma)$ is invertible, then*

$$\Sigma(\lambda) = I_m - \rho_\gamma(\lambda) \begin{bmatrix} C_1 \\ -C_2 \end{bmatrix} \Delta_\gamma(\lambda)^{-1} \begin{bmatrix} C_1^* & -C_2^* \end{bmatrix} \tag{11.25}$$

and

$$\Sigma(\lambda)\Sigma(\lambda)^* = I_m - \rho_\lambda(\lambda) \begin{bmatrix} C_1 \\ -C_2 \end{bmatrix} \Delta_\gamma(\lambda)^{-1} G(\gamma)^* P G(\gamma) \Delta_\gamma(\lambda)^{-*} \begin{bmatrix} C_1^* & -C_2^* \end{bmatrix} \tag{11.26}$$

for every point $\lambda \in \Omega$ at which the indicated inverses exist.

The verification of these formulas is by direct calculation, with the help of the following well known result:

LEMMA 11.8. *If $X \in \mathbb{C}^{n\times n}$, $Y \in \mathbb{C}^{p\times n}$, $Z \in \mathbb{C}^{n\times p}$ and if X and $X + ZY$ are invertible, then $I_p + YX^{-1}Z$ is invertible and*

$$(I_p + YX^{-1}Z)^{-1} = I_p - Y(X + ZY)^{-1}Z. \tag{11.27}$$

PROOF. It suffices to check that

$$\big(I_p + YX^{-1}Z\big)\big(I_p - Y(X + ZY)^{-1}Z\big) = I_p.$$

But this is a straightforward calculation. □

PROOF OF THEOREM 11.7. Suppose first that $G(\lambda)$ is invertible. Then, since

$$\Theta_{22}(\lambda) = I_q + \rho_\gamma(\lambda) C_2 \big(G(\gamma)^* P G(\lambda)\big)^{-1} C_2^*$$

and

$$G(\gamma)^* P G(\lambda) + \rho_\gamma(\lambda) C_2^* C_2 = \Delta_\gamma(\lambda)$$

is invertible for $\lambda \in \Omega_+ \cup \Omega_0$ under the present assumptions by Theorem 11.5, Lemma 11.8 guarantees that $\Theta_{22}(\lambda)$ is invertible and that

$$\Sigma_{22}(\lambda) = \Theta_{22}(\lambda)^{-1} = I_q - \rho_\gamma(\lambda) C_2 \Delta_\gamma(\lambda)^{-1} C_2^*. \tag{11.28}$$

This serves to verify the $(2, 2)$ block entry of (11.25).

Next, by (11.28) and (11.23),

$$\begin{aligned} \Sigma_{21}(\lambda) &= -\Theta_{22}(\lambda)^{-1}\Theta_{21}(\lambda) \\ &= \big(I_q - \rho_\gamma(\lambda) C_2 \Delta_\gamma(\lambda)^{-1} C_2^*\big) C_2 \varphi_\gamma(\lambda) C_1^* \\ &= C_2 \big(I_q - \rho_\gamma(\lambda) \Delta_\gamma(\lambda)^{-1} C_2^* C_2\big) \varphi_\gamma(\lambda) C_1^*, \end{aligned}$$

which is readily seen to confirm the $(2, 1)$ block entry of (11.25).

The verification of the formulas for $\Sigma_{12}(\lambda)$ and $\Sigma_{11}(\lambda)$ is similar, though the latter requires a bit more work (but just a bit, if you take advantage of the fact that $-\Theta_{12}\Theta_{22}^{-1}\Theta_{21} = \Theta_{12}\Sigma_{21}$).

Finally, the verification of (11.26) is left to the reader. □

We remark that although $\Sigma(\lambda)$ has been derived from $\Theta(\lambda)$, the given formulas are meaningful at those points λ at which $\Delta_\gamma(\lambda)$ is invertible. It is evident from Theorem 11.5 that this can happen even if P and $G(\lambda)$ are not invertible. Thus, for example, Theorem 11.5 guarantees that:

THEOREM 11.9. *If $P \geq 0$ is a solution of the GLS equation* (11.1) *and if $\gamma \in \Omega_0$ and*

$$G(\gamma)^*PG(\gamma) + C_1^*C_1 + C_2^*C_2 > 0,$$

then:

(1) $\Delta_\gamma(\lambda)$ *is invertible at every point $\lambda \in \Omega_+$.*
(2) $\Delta_\gamma(\lambda)$ *is invertible at every point $\lambda \in \Omega_0$ at which*

$$G(\lambda)^*PG(\lambda) + \left|\rho_\gamma(\lambda)\right| (C_1^*C_1 + C_2^*C_2) > 0. \tag{11.29}$$

Nevertheless, we shall not pursue this level of generality here. It is instructive, however, to see how to work directly from formula (11.25). To this end, it is convenient to first summarize the key properties of $\Delta_\gamma(\lambda)$ that come into play.

THEOREM 11.10. *If $P > 0$ is a solution of the GLS equation* (11.1) *and if $G(\lambda)$ is invertible at every point $\lambda \in \Omega_0$, then*:

(1) $\Delta_\gamma(\lambda)$ *is invertible at every point $\lambda \in \Omega_+ \cup \Omega_0$.*
(2) $\Sigma(\lambda)$ *is analytic on $\Omega_+ \cup \Omega_0$ and unitary at every point $\lambda \in \Omega_0$.*
(3) $\Sigma_{11}(\lambda)$ *is invertible at every point $\lambda \in \Omega_0$ and*

$$\det \Sigma_{11}(\lambda) = \frac{\det\big(H(\lambda)PH(\gamma)^*\big)}{\det\big(\Delta_\gamma(\lambda)\big)}. \tag{11.30}$$

(4) $\Sigma_{22}(\lambda)$ *is invertible at every point $\lambda \in \Omega_0$ and*

$$\det \Sigma_{22}(\lambda) = \frac{\det\big(G(\gamma)^*PG(\lambda)\big)}{\det\big(\Delta_\gamma(\lambda)\big)}. \tag{11.31}$$

Moreover, if Ω_+ is chosen to be one of the three classical settings, then:

(5) $\Sigma_{12}(\lambda)$ *and* $\Sigma_{21}(\lambda)$ *are strictly contractive on $\Omega_+ \cup \Omega_0$; that is, there exists a positive number $\delta < 1$ such that*

$$\left\|\Sigma_{12}(\lambda)\right\| \leq \delta \quad \textit{and} \quad \|\Sigma_{21}\| \leq \delta$$

for every point $\lambda \in \Omega_+ \cup \Omega_0$.

(6) $\big(I_q - \Sigma_{21}(\lambda)\mathcal{E}(\lambda)\big)^{-1} \in H_\infty^{q\times q}(\Omega_+)$ *for every choice of $\mathcal{E}(\lambda) \in \mathcal{S}^{p\times q}(\Omega_+)$.*
(7) $\Delta_\gamma(\lambda)^{-1}u \in H_2^n(\Omega_+)$ *for every choice of $u \in \mathbb{C}^n$.*

(8) $\rho_\gamma(\lambda)\Delta_\gamma(\lambda)^{-1} \in H_\infty^{n\times n}(\Omega_+)$.

Proof. The first four assertions are immediate from Theorems 11.9 and 11.7, except perhaps for the formulas for the determinant. But these too are easy if you take advantage of the fact that for $X \in \mathbb{C}^{p\times q}$ and $Y \in \mathbb{C}^{q\times p}$,

$$\det\big(I_p - XY\big) = \det\big(I_q - YX\big).$$

Next, in view of the identities

$$\begin{aligned}\Sigma_{21}(\lambda)\Sigma_{21}(\lambda)^* &= I_q - \Sigma_{22}(\lambda)\Sigma_{22}(\lambda)^*,\\ \Sigma_{12}(\lambda)\Sigma_{12}(\lambda)^* &= I_q - \Sigma_{22}(\lambda)\Sigma_{22}(\lambda)^*,\end{aligned}$$

valid for $\lambda \in \Omega_0$, it suffices to show that

$$\|\Sigma_{22}\| \geq \delta_1 > 0$$

for every point $\lambda \in \Omega_0$. Since $\Sigma_{22}(\lambda)$ is contractive, it follows readily from the singular value decomposition of $\Sigma_{22}(\lambda)$ that

$$\big\|\Sigma_{22}(\lambda)\big\| \geq \big|\det \Sigma_{22}(\lambda)\big|.$$

Now if Ω_0 is compact, then this does the trick, since $\big|\det \Sigma_{22}(\lambda)\big| \neq 0$ for any point $\lambda \in \Omega_0$ by formula (11.31) the assumptions on $G(\lambda)$ and the first conclusion. The final four assertions are easily established when $\Omega_+ = \mathbb{D}$. The analysis for the other two classical choices of Ω_+ is more subtle, but may be completed by a finer investigation of $\Delta_\gamma(\lambda)$ as in [Dym 1996, pp. 206–208]. The details are left to the reader. □

With the aid of Theorem 11.10, it is now not too difficult to show directly that (under the hypotheses of that theorem) if

$$S(\lambda) = \Sigma_{12}(\lambda) + \Sigma_{11}(\lambda)\mathcal{E}(\lambda)\big(I_q - \Sigma_{21}(\lambda)\mathcal{E}(\lambda)\big)^{-1}\Sigma_{22}(\lambda), \tag{11.35}$$

then $\big(C_1 - S(\lambda)C_2\big)G(\lambda)^{-1}u \in H_2^p(\Omega_+)$,

$$\big(-S(\lambda)^*C_1 + C_2\big)G(\lambda)^{-1}u \in H_2^q(\Omega_+)^\perp,$$

and

$$-\left\langle S(\lambda)C_2G(\lambda)^{-1}\begin{bmatrix}0\\ y\end{bmatrix},\ C_1G(\lambda)^{-1}\begin{bmatrix}x\\ 0\end{bmatrix}\right\rangle = \big[x^*\ 0\big]\,P\begin{bmatrix}0\\ y\end{bmatrix}$$

for every choice of $\mathcal{E}(\lambda) \in \mathcal{S}^{p\times q}(\Omega_+)$, $u \in \mathbb{C}^n$, $x \in \mathbb{C}^\mu$ and $y \in \mathbb{C}^\nu$. This yields an independent check of Theorem 10.1, and exhibits a strategy that can be imitated even when $P \geq 0$ is singular.

12. $P \geq 0$ and Other Remarks

We begin by looking backwards.

About the first ten sections. The story begins with the observation that matrix versions of a number of classical interpolation problems are all special cases of a single general problem described in Section 2: the BIP. This problem is formulated in terms of the columns $f_j(\lambda)$ of an $m \times n$ mvf

$$F(\lambda) = [f_1(\lambda) \ \cdots \ f_n(\lambda)]$$

whose first μ columns span a μ-dimensional subspace $\mathcal{M}_1$ of $H_2^m(\Omega_+)$ and whose last $\nu = n - \mu$ columns span a ν-dimensional subspace $\mathcal{M}_2$ of $H_2^m(\Omega_+)^\perp$.

In Section 3 it is shown that if the BIP admits at least one solution $S(\lambda) \in \mathcal{S}^{p\times q}(\Omega_+)$, then the $n\times n$ matrix Q defined by (3.1) must be positive semidefinite; that is, the condition $Q \geq 0$ is a necessary condition for the existence of a solution to the BIP. The remarkable fact is that if $\mathcal{M}_1$ is an R_α-invariant subspace of $H_2^m(\Omega_+)$ for some $\alpha \in \Omega_+$ and $\mathcal{M}_2$ is an R_α-invariant subspace of $H_2^m(\Omega_+)^\perp$ for some $\alpha \in \Omega_-$, then this condition is also sufficient. The point is that the R_α-invariance assumptions force $F(\lambda)$ to be of the special form

$$F(\lambda) = C(M - \lambda N)^{-1}, \tag{12.1}$$

where the matrices M and N are subject to certain spectral assumptions, as is explained in (5.3)–(5.5). It then follows further that Q is a solution of the Lyapunov–Stein equation (5.6). For one sided problems (that is, when $\mu = n$ or $\nu = n$) this is the whole story because under the special constraints (5.4) and (5.5), the Lyapunov–Stein equation has only one solution. For two sided problems this is not true unless $\sigma(A_1^*) \cap \sigma(A_2) = \varnothing$. This extra freedom is used to define the augmented BIP in Section 6. Now it turns out that if $P > 0$ is any positive definite solution of the Lyapunov–Stein equation (5.6), then the space

$$\mathcal{M} = \left\{F(\lambda)u : u \in \mathbb{C}^n\right\} \tag{12.2}$$

(with $F(\lambda)$ as in (12.1)) endowed with the inner product

$$\left\langle F(\lambda)v, F(\lambda)u\right\rangle_{\mathcal{M}} = u^*Pv \tag{12.3}$$

is an RKHS of the special kind considered in Section 8. In particular its RK is given in terms of a rational J-inner mvf $\Theta(\lambda)$ (of McMillan degree n) that is uniquely determined by the space (that is, by the elements and the inner product) up to a J unitary constant factor on the right. Moreover, as is explained in Section 10,

$$\left\{T_\Theta[\mathcal{E}] : \mathcal{E}(\lambda) \in \mathcal{S}^{p\times q}(\Omega_+)\right\}$$

is a complete description of the set of all solutions to the augmented BIP based on $F(\lambda)$ and P (that is, on the elements of $\mathcal{M}$ and the inner product imposed on $\mathcal{M}$).

What if $P \geq 0$ is singular? If the rank of the $n \times n$ matrix P is equal to k, where $k < n$, then the space $\mathcal{M}$ endowed with the inner product P is no longer a Hilbert space. However, it turns out that there exists a k-dimensional subspace $\widetilde{\mathcal{M}}$ of $\mathcal{M}$ that is R_α-invariant for every point $\alpha \in \mathcal{A}_F$ which is an RKHS with an RK of the special form (8.1) based on a rational J-inner mvf $\widetilde{\Theta}(\lambda)$ (of McMillan degree k). Moreover, in this instance,

$$\left\{T_{\widetilde{\Theta}}[\mathcal{E}] : \mathcal{E}(\lambda) \in \mathcal{S}^{p\times q}(\Omega_+)\right\}$$

fulfill k of the n interpolation conditions. The remaining $n-k$ conditions (which are not all independent) are then met by imposing extra constraints on $\mathcal{E}(\lambda)$. Some special cases illustrating this are given in [Dym 1989a, Chapter 7]. For other approaches see [Ball and Helton 1986; Bruinsma 1991; Dubovoj 1984].

Existence and representation formulas for the case of singular P may also be obtained by the methods introduced by Katsnelson, Kheifets, and Yuditskii [Katsnelson et al. 1987]; see also [Kheifets and Yuditskii 1994; Kheifets 1988a; 1988b; 1990; 1998], the last of which appears in this volume. Later in this section (after the next paragraph) we will discuss applications of this method to the augmented BIP problem.

About Section 11. In this section explicit formulas are presented for $\Theta(\lambda)$ and its Potapov–Ginzburg transform $\Sigma(\lambda)$ in terms of the data C, M, N and P when $P > 0$. It is then indicated how to verify that

$$\Sigma_{12}(\lambda) + \Sigma_{11}(\lambda)\mathcal{E}(\lambda)\big(I_q - \Sigma_{21}(\lambda)\mathcal{E}(\lambda)\big)^{-1}\Sigma_{22}(\lambda) \tag{12.4}$$

is a solution of the augmented BIP for every choice of $\mathcal{E}(\lambda) \in \mathcal{S}^{p\times q}(\Omega_+)$ directly from these formulas. This gives a second independent proof of this fact in this more general setting. In fact every solution of the augmented BIP can be expressed in this form. Moreover, although we do not pursue this here, large parts of this analysis in terms of $\Sigma(\lambda)$ is valid for matrices $P \geq 0$ that are singular.

The Abstract Interpolation Problem of Katsnelson, Kheifets, and Yuditskii. The problem of establishing the existence and parametrization of solutions to the augmented BIP in the disc fits naturally into the domain of problems that can be resolved within the framework of the Abstract Interpolation Problem of [Katsnelson et al. 1987] that was referred to just above. The starting point is the assumption that there exists an $n \times n$ positive semidefinite solution P of the GLS equation (5.6a). The idea is to rewrite (5.6a) as

$$M^*PM + C_2^*C_2 = N^*PN + C_1^*C_1 \tag{12.5}$$

and then to define an isometric colligation $V : \mathcal{D}_V \to \mathcal{R}_V$, where

$$\mathcal{D}_V = \left\{ \begin{bmatrix} P^{1/2}M \\ C_2 \end{bmatrix} x : x \in \mathbb{C}^n \right\} \subset \mathbb{C}^n \oplus \mathbb{C}^q,$$
$$\mathcal{R}_V = \left\{ \begin{bmatrix} P^{1/2}N \\ C_1 \end{bmatrix} x : x \in \mathbb{C}^n \right\} \subset \mathbb{C}^n \oplus \mathbb{C}^p.$$

The following facts then follow from the general analysis in [Katsnelson et al. 1987]:

(1) The set of all solutions to the augmented BIP is equal to the set of characteristic functions of those unitary colligations U that extend V (that is, $U|_{\mathcal{D}_V} = V$) and have the same "input" space $\mathbb{C}^q$ and the same "output" space $\mathbb{C}^p$ as V.
(2) The set of all such characteristic functions (and hence the set of all solutions to the augmented BIP in the disc) is equal to the set of all $p \times q$ mvf's of the general form (12.4) except that here $\Omega_+ = \mathbb{D}$ and $\mathcal{E}(\lambda)$ is an element in $\mathcal{S}^{p' \times q'}(\mathbb{D})$, where $q' = \dim\big(\mathbb{C}^{n+q} \ominus \mathcal{D}_V\big)$, $p' = \dim\big(\mathbb{C}^{n+p} \ominus \mathcal{R}_V\big)$ and $\Sigma(\lambda)$ is the characteristic function of a very special unitary colligation from $\mathbb{C}^n \oplus \mathbb{C}^{p'} \oplus \mathbb{C}^q$ onto $\mathbb{C}^n \oplus \mathbb{C}^p \oplus \mathbb{C}^{q'}$ (which is not of the type referred to in (1) because the input and output spaces have been enlarged to $\mathbb{C}^{p'} \oplus \mathbb{C}^q$ and $\mathbb{C}^p \oplus \mathbb{C}^{q'}$, respectively).

The description of the characteristic functions of unitary colligations that extend a given isometric colligation originates in the work of Arov and Grossman [1983; 1992]. See also [Kheifets 1988a; 1988b; 1990] for applications to the Abstract Interpolation Problem of [Katsnelson et al. 1987], and [Dym and Freydin 1997a; 1997b] for an application of these methods to the BIP problem in the setting of upper triangular operators. A useful discussion of applications of unitary colligations and of the Arov–Grossman formula may be found in [Arocena 1994].

Analogues of the problem of Katsnelson, Kheifets, and Yuditskii in the setting of Section 11. Here we exhibit a connection between the formulas that emerge from the analysis in Section 11 and the formulas that emerge by adapting the strategy of [Katsnelson et al. 1987], as outlined in the preceding subsection, to the setting of Section 11. A full analysis of these calculations will appear elsewhere.

Let P be a nonnegative solution of the GLS equation (11.1). Then Lemma 11.4 guarantees that

$$\Delta_\omega(\omega) = G(\omega)^* P G(\omega) + \rho_\omega(\omega) C_2^* C_2 = H(\omega) P H(\omega)^* + \rho_\omega(\omega) C_1^* C_1. \quad (12.6)$$

Clearly $\Delta_\omega(\omega)$ is positive semidefinite for every point $\omega \in \Omega_+$. Fix such a point and let

$$W_1 = \begin{bmatrix} P^{1/2} G(\omega) \\ \rho_\omega(\omega)^{1/2} C_2 \end{bmatrix} \quad \text{and} \quad W_2 = \begin{bmatrix} -P^{1/2} H(\omega)^* \\ \rho_\omega(\omega)^{1/2} C_1 \end{bmatrix}.$$

Then, in view of (12.6), $V : W_1x \to W_2x$ is an isometry from

$$\mathcal{D}_V = \{W_1x : x \in \mathbb{C}^n\} \subset \mathbb{C}^n \oplus \mathbb{C}^q$$

onto

$$\mathcal{R}_V = \{W_2x : x \in \mathbb{C}^n\} \subset \mathbb{C}^n \oplus \mathbb{C}^p.$$

Let $k = \dim \mathcal{D}_V = \dim \mathcal{R}_V$. Then $q' = \dim \mathcal{D}_V^\perp = n + q - k$, $p' = \dim \mathcal{R}_V^\perp = n+p-k$, and $k \le n$, with equality if and only if $\Delta_\omega(\omega) > 0$. For the time being we shall assume only that $k \ge 1$ and shall let

$$W_1^\perp \in \mathbb{C}^{(n+q)\times q'} \quad \text{and} \quad W_2^\perp \in \mathbb{C}^{(n+p)\times p'}$$

be isometric matrices whose columns span $\mathcal{D}_V^\perp$ and $\mathcal{R}_V^\perp$, respectively.

The next step is to define the matrix

$$P = \begin{array}{c} \begin{bmatrix} U_{11} & U_{12} & U_{13} \\ U_{21} & U_{22} & U_{23} \\ U_{31} & U_{32} & U_{33} \end{bmatrix} \begin{array}{l} \}n \\ \}p \\ \}q' \end{array} \\ \begin{array}{ccc} \underbrace{\phantom{U_{31}}}_{n} & \underbrace{\phantom{U_{32}}}_{q} & \underbrace{\phantom{U_{33}}}_{p'} \end{array} \end{array} .$$

by the formulas

$$\begin{bmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{bmatrix} = W_2(W_1^*W_1)^{-1}W_1^*, \qquad [U_{31}\ \ U_{32}] = (W_1^\perp)^*, \qquad \begin{bmatrix} U_{13} \\ U_{23} \end{bmatrix} = W_2^\perp,$$

and $U_{33} = 0$.

It is readily checked that the formulas $W_1(W_1^*W_1)^{-1}W_1^*$, $W_2(W_2^*W_2)^{-1}W_2^*$ and $W_2(W_1^*W_1)^{-1}W_1^*$ are meaningful single-valued mappings when the inverses are interpreted as inverse images even if $W_1^*W_1 = W_2^*W_2$ is not invertible and furthermore, since

$$I_{n+q} - W_1(W_1^*W_1)^{-1}W_1^* = W_1^\perp(W_1^\perp)^*$$

and

$$I_{n+p} - W_2(W_2^*W_2)^{-1}W_2^* = W_2^\perp(W_2^\perp)^*,$$

that U is unitary. This matrix U corresponds to the special unitary colligation singled out in the last subsection.

The next step is to define the mvf

$$\Sigma(\lambda) = \begin{array}{c} \begin{bmatrix} \Sigma_{11}(\lambda) & \Sigma_{12}(\lambda) \\ \Sigma_{21}(\lambda) & \Sigma_{22}(\lambda) \end{bmatrix} \begin{array}{l} \}p \\ \}q' \end{array} \\ \begin{array}{cc} \underbrace{\phantom{\Sigma_{21}(\lambda)}}_{p'} & \underbrace{\phantom{\Sigma_{22}(\lambda)}}_{q} \end{array} \end{array}$$

$$= \begin{bmatrix} U_{23} & U_{22} \\ 0 & U_{32} \end{bmatrix} + \psi_\omega(\lambda) \begin{bmatrix} U_{21} \\ U_{31} \end{bmatrix} \left(I - \psi_\omega(\lambda)U_{11}\right)^{-1}[U_{13}\ \ U_{12}], \tag{12.7}$$

where

$$\psi_\omega(\lambda) = \frac{\delta_\omega(\lambda)}{\rho_\omega(\lambda)}.$$

This mvf is the natural analogue of the characteristic function of U in this setting. It is well defined for $\lambda \in \Omega_+$ since

$$\sigma(\lambda) = \frac{b(\lambda)}{a(\lambda)} \quad \text{and} \quad \psi_\omega(\lambda) = \frac{a(\omega)}{a(\omega)^*}\left(\frac{\sigma(\lambda)-\sigma(\omega)}{1-\sigma(\lambda)\sigma(\omega)^*}\right)$$

are contractive for $\lambda \in \Omega_+$.

From now on we shall assume that $\Delta_\omega(\omega) > 0$, even though many of the formulas are meaningful without this restriction. Then, with the help of the identity

$$\rho_\omega(\lambda)\Delta_\omega(\omega) + \delta_\omega(\lambda)G(\omega)^* P H(\omega)^* = \rho_\omega(\omega)\Delta_\omega(\lambda), \tag{12.8}$$

it is readily checked that

$$\big(I_n - \psi_\omega(\lambda)U_{11}\big)^{-1}U_{12} = -\rho_\omega(\omega)^{-1/2}\rho_\omega(\lambda)P^{1/2}H(\omega)^*\Delta_\omega(\lambda)^{-1}C_2^*, \tag{12.9}$$

$$U_{21}\big(I_n - \psi_\omega(\lambda)U_{11}\big)^{-1} = \rho_\omega(\omega)^{-1/2}\rho_\omega(\lambda)C_1\Delta_\omega(\lambda)^{-1}G(\omega)^*P^{1/2}, \tag{12.10}$$

$$\big(I_n - \psi_\omega(\lambda)U_{11}\big)^{-1} = I_n - \rho_\omega(\omega)^{-1}\delta_\omega(\lambda)P^{1/2}H(\omega)^*\Delta_\omega(\lambda)^{-1}G(\omega)^*P^{1/2}. \tag{12.11}$$

Therefore, by direct substitution of these last three formulas into the entries of formula (12.7), we obtain

$$\Sigma_{12}(\lambda) = U_{22} + \psi_\omega(\lambda)U_{21}\big(I - \psi_\omega(\lambda)U_{11}\big)^{-1}U_{12} = \rho_\omega(\lambda)C_1\Delta_\omega(\lambda)^{-1}C_2^*,$$

$$\begin{aligned}
\Sigma_{11}(\lambda) &= U_{23} + \psi_\omega(\lambda)U_{21}\big(I_n - \psi_\omega(\lambda)U_{11}\big)^{-1}U_{13} \\
&= [\psi_\omega(\lambda)U_{21}\big(I_n - \psi_\omega(\lambda)U_{11}\big)^{-1} \vdots I_p]\begin{bmatrix} U_{13} \\ U_{23}\end{bmatrix} \\
&= [\rho_\omega(\omega)^{-1/2}\delta_\omega(\lambda)C_1\Delta_\omega(\lambda)^{-1}G(\omega)^*P^{1/2} \vdots I_p]W_2^{\perp},
\end{aligned}$$

$$\begin{aligned}
\Sigma_{22}(\lambda) &= U_{32} + \psi_\omega(\lambda)U_{31}\big(I_n - \psi_\omega(\lambda)U_{11}\big)^{-1}U_{12} \\
&= [U_{31}\ \ U_{32}]\begin{bmatrix} \psi_\omega(\lambda)\big(I_n - \psi_\omega(\lambda)U_{11}\big)^{-1}U_{12} \\ I_{q'}\end{bmatrix} \\
&= (W_1^{\perp})^*\begin{bmatrix} -\rho_\omega(\omega)^{-1/2}\delta_\omega(\lambda)P^{1/2}H(\omega)^*\Delta_\omega(\lambda)^{-1}C_2^* \\ I_q\end{bmatrix},
\end{aligned}$$

$$\begin{aligned}
\Sigma_{21}(\lambda) &= \psi_\omega(\lambda)U_{31}\big(I_n - \psi_\omega(\lambda)U_{11}\big)^{-1}U_{13} \\
&= \psi_\omega(\lambda)(W_1^{\perp})^* \\
&\quad \times \begin{bmatrix} I_n \\ 0\end{bmatrix}\big(I_n - \rho_\omega(\omega)^{-1}\delta_\omega(\lambda)P^{1/2}H(\omega)^*\Delta_\omega(\lambda)^{-1}G(\omega)^*P^{1/2}\big)[I_n\ \ 0]W_2^{\perp}.
\end{aligned}$$

If $\operatorname{rank} P = r$ and $r < n$, we can write the isometric matrices $W_1^\perp$ and $W_2^\perp$ as follows (I wish to thank Vladimir Bolotnikov for suggesting this decomposition, which led to some improved formulas in this subsection):

$$W_1^\perp = \begin{array}{c} \begin{bmatrix} X & Y_1 \\ 0 & Z_1 \end{bmatrix} \\ \underbrace{\qquad}_{n-r}\ \underbrace{\qquad}_{q+r-k} \end{array} \begin{array}{l} \}n \\ \}q \\ \\ \end{array} \quad \text{and} \quad W_2^\perp = \begin{array}{c} \begin{bmatrix} X & Y_2 \\ 0 & Z_2 \end{bmatrix} \\ \underbrace{\qquad}_{n-r}\ \underbrace{\qquad}_{p+r-k} \end{array} \begin{array}{l} \}n \\ \}p \\ \\ \end{array}, \tag{12.12}$$

where the columns of X are an orthonormal basis for $\ker P$ and hence are independent of ω. We shall keep this notation for $r = n$ also.

THEOREM 12.1. *If $\omega \in \Omega_+$ and $\operatorname{rank} P = r$, the block entries in the "characteristic function" $\Sigma(\lambda)$ defined by the formulas* (12.7) *can be expressed by*

$$\Sigma_{11}(\lambda) = [\underbrace{\ 0\ }_{n-r}\ \ \underbrace{\widetilde{\Sigma}_{11}(\lambda)}_{p+r-k}\,]\}p\ , \qquad \Sigma_{12}(\lambda) = \rho_\omega(\lambda) C_1 \Delta_\omega(\lambda)^{-1} C_2^*,$$

$$\Sigma_{21}(\lambda) = \begin{bmatrix} \psi_\omega(\lambda) I_{n-r} & 0 \\ 0 & \underbrace{\widetilde{\Sigma}_{21}(\lambda)}_{p+r-k} \end{bmatrix} \}q+r-k\ , \quad \Sigma_{22}(\lambda) = \begin{bmatrix} 0 \\ \underbrace{\widetilde{\Sigma}_{22}(\lambda)}_{q} \end{bmatrix} \begin{array}{l} \}n-r \\ \}q+r-k \end{array},$$

where

$$\tilde{\Sigma}_{11}(\lambda) = Z_2 + \rho_\omega(\omega)^{-1/2} \delta_\omega(\lambda) C_1 \Delta_\omega(\lambda)^{-1} G(\omega)^* P^{1/2} Y_2,$$

$$\tilde{\Sigma}_{21}(\lambda) = \psi_\omega(\lambda) Y_1^* \big(I_n - \rho_\omega(\omega)^{-1} \delta_\omega(\lambda) P^{1/2} H(\omega)^* \Delta_\omega(\lambda)^{-1} G(\omega)^* P^{1/2} \big) Y_2,$$

$$\tilde{\Sigma}_{22}(\lambda) = Z_1^* - \rho_\omega(\omega)^{-1/2} \delta_\omega(\lambda) Y_1^* P^{1/2} H(\omega)^* \Delta_\omega(\lambda)^{-1} C_2^*,$$

and the entries Y_1, Z_1, Y_2, Z_2 from the second block columns of $W_1^\perp$ and $W_2^\perp$ depend upon ω.

Thus, upon writing $\varepsilon(\lambda) \in \mathcal{S}^{p' \times q'}$ in the block form

$$\varepsilon(\lambda) = \begin{bmatrix} \varepsilon_{11}(\lambda) & \varepsilon_{12}(\lambda) \\ \underbrace{\varepsilon_{21}(\lambda)}_{n-r} & \underbrace{\varepsilon_{22}(\lambda)}_{q+r-k} \end{bmatrix} \begin{array}{l} \}n-r \\ \}p+r-k \end{array},$$

it is readily checked that

$$\begin{aligned} &\Sigma_{21}(\lambda) + \Sigma_{11}(\lambda) \varepsilon(\lambda) \big(I_{q'} - \Sigma_{21}(\lambda) \varepsilon(\lambda) \big)^{-1} \Sigma_{22}(\lambda) \\ &\qquad = \Sigma_{21}(\lambda) + \tilde{\Sigma}_{11}(\lambda) \tilde{\varepsilon}(\lambda) \big(I_{q'-(n-r)} - \tilde{\Sigma}_{21}(\lambda) \tilde{\varepsilon}(\lambda) \big)^{-1} \tilde{\Sigma}_{22}(\lambda), \end{aligned}$$

where

$$\tilde{\varepsilon}(\lambda) = \varepsilon_{22}(\lambda) + \varepsilon_{21}(\lambda) \psi_\omega(\lambda) \big(I_{n-r} - \psi_\omega(\lambda) \varepsilon_{11}(\lambda) \big)^{-1} \varepsilon_{12}(\lambda).$$

This leads easily to the following result:

THEOREM 12.2. *If* $\operatorname{rank} P = r$, *then*

$$\left\{\Sigma_{12} + \Sigma_{11}\varepsilon(I_{q'} - \Sigma_{21}\varepsilon)^{-1}\Sigma_{22} : \varepsilon \in \mathcal{S}^{p'\times q'}\right\} = \left\{\Sigma_{12} + \tilde{\Sigma}_{11}\tilde{\varepsilon}(I_{q''} - \tilde{\Sigma}_{21}\tilde{\varepsilon})^{-1}\tilde{\Sigma}_{22} : \tilde{\varepsilon} \in \mathcal{S}^{p''\times q''}\right\},$$

where $p'' = p' - n + r$ *and* $q'' = q' - n + r$.

The formulas

$$\begin{bmatrix} -H(\omega)P^{1/2} & \rho_\omega(\omega)^{1/2}C_1^* \end{bmatrix} W_2^\perp = 0 \quad \text{and} \quad (W_1^\perp)^* \begin{bmatrix} P^{1/2}G(\omega) \\ \rho_\omega(\omega)^{1/2}C_2 \end{bmatrix} = 0$$

permit additional simplifications.

Upon making use of the identities

$$\delta_\omega(\lambda)G(\omega)^* + \rho_\omega(\lambda)H(\omega) = \rho_\omega(\omega)H(\lambda)$$

and

$$\delta_\omega(\lambda)H(\omega)^* + \rho_\omega(\lambda)G(\omega) = \rho_\omega(\omega)G(\lambda),$$

and then letting ω tend to Ω_0, we obtain:

$$\begin{aligned}
\tilde{\Sigma}_{11}(\lambda) &= \left(I_p - \rho_\omega(\lambda)C_1\Delta_\omega(\lambda)^{-1}C_1^*\right)Z_2, \\
\tilde{\Sigma}_{21}(\lambda) &= \psi_\omega(\lambda)Y_1^*Y_2 + \rho_\omega(\lambda)Z_1^*C_2\Delta_\omega(\lambda)^{-1}C_1^*Z_2, \\
\tilde{\Sigma}_{22}(\lambda) &= Z_1^*\left(I_q - \rho_\omega(\lambda)C_2\Delta_\omega(\lambda)^{-1}C_2^*\right).
\end{aligned}$$

If $\Delta_\omega(\omega) > 0$ for $\omega \in \Omega_0$, then $n = r$ and $Y_1 = Y_2 = 0$, and we may choose $Z_1 = I_q$ and $Z_2 = I_p$. Then the formulas in Theorem 12.1 agree with the formula for $\Sigma(\lambda)$ in Theorem 11.7, which was derived by an entirely different strategy.

Other methods, other problems. The analysis presented here is based largely on [Dym 1989a; 1989b; 1994b] and subsequent extensions. There are many other approaches. The books [Ball et al. 1990; Foiaş and Frazho 1990; Helton et al. 1987] reflect three other schools of thought, and each contains an extensive bibliography and notes to the literature. Additional sources may be found in [Dubovoj et al. 1992; Dym 1994a]; see also [Arov 1993; Ivanchenko and Sakhnovich 1994].

References

[Alpay and Dym 1984] D. Alpay and H. Dym, "Hilbert spaces of analytic functions, inverse scattering and operator models, I", *Integral Equations Operator Theory* **7**:5 (1984), 589–641.

[Alpay and Dym 1992] D. Alpay and H. Dym, "On reproducing kernel spaces, the Schur algorithm, and interpolation in a general class of domains", pp. 30–77 in *Operator theory and complex analysis* (Sapporo, 1991), edited by T. Ando and I. Gohberg, Oper. Theory Adv. Appl. **59**, Birkhäuser, Basel, 1992.

[Alpay and Dym 1993a] D. Alpay and H. Dym, “On a new class of reproducing kernel spaces and a new generalization of the Iohvidov laws”, *Linear Algebra Appl.* **178** (1993), 109–183.

[Alpay and Dym 1993b] D. Alpay and H. Dym, “On a new class of structured reproducing kernel spaces”, *J. Funct. Anal.* **111**:1 (1993), 1–28.

[Alpay and Dym 1996] D. Alpay and H. Dym, “On a new class of realization formulas and their application”, *Linear Algebra Appl.* **241/243** (1996), 3–84.

[Arocena 1994] R. Arocena, “Unitary colligations and parametrization formulas”, *Ukraïn. Mat. Zh.* **46**:3 (1994), 147–154.

[Aronszajn 1950] N. Aronszajn, “Theory of reproducing kernels”, *Trans. Amer. Math. Soc.* **68** (1950), 337–404.

[Arov 1973] D. Z. Arov, “Realization of matrix-valued functions according to Darlington”, *Izv. Akad. Nauk SSSR Ser. Mat.* **37** (1973), 1299–1331. In Russian.

[Arov 1993] D. Z. Arov, “The generalized bitangent Carathéodory–Nevanlinna–Pick problem and (j, J_0)-inner matrix functions”, *Izv. Ross. Akad. Nauk Ser. Mat.* **57**:1 (1993), 3–32. In Russian; translation in *Russian Acad. Sci. Izv. Math.* **42**:1 (1994), 1–26.

[Arov and Grossman 1983] D. Z. Arov and L. Z. Grossman, “Scattering matrices in the theory of extensions of isometric operators”, *Dokl. Akad. Nauk SSSR* **270**:1 (1983), 17–20. In Russian; translation in *Soviet Math. Dokl.* **27**:3 (1983), 518–521.

[Arov and Grossman 1992] D. Z. Arov and L. Z. Grossman, “Scattering matrices in the theory of unitary extension of isometric operators”, *Math. Nachr.* **157** (1992), 105–123.

[Ball 1975] J. A. Ball, “Models for noncontractions”, *J. Math. Anal. Appl.* **52**:2 (1975), 235–254.

[Ball and Helton 1986] J. A. Ball and J. W. Helton, “Interpolation problems of Pick–Nevanlinna and Loewner types for meromorphic matrix functions: parametrization of the set of all solutions”, *Integral Equations Operator Theory* **9**:2 (1986), 155–203.

[Ball et al. 1990] J. A. Ball, I. Gohberg, and L. Rodman, *Interpolation of rational matrix functions*, Operator Theory: Advances and Applications **45**, Birkhäuser Verlag, Basel, 1990.

[Bruinsma 1991] P. Bruinsma, “Degenerate interpolation problems for Nevanlinna pairs”, *Indag. Math.* (*N.S.*) **2**:2 (1991), 179–200.

[de Branges 1963] L. de Branges, “Some Hilbert spaces of analytic functions, I”, *Trans. Amer. Math. Soc.* **106** (1963), 445–468.

[de Branges and Rovnyak 1966] L. de Branges and J. Rovnyak, “Canonical models in quantum scattering theory”, pp. 295–392 in *Perturbation theory and its application in quantum mechanics* (Madison, 1965), edited by C. H. Wilcox, Wiley, New York, 1966.

[Dubovoj 1984] V. K. Dubovoĭ, “Indefinite metric in Schur’s interpolation problem for analytic functions, IV”, *Teor. Funktsiĭ Funktsional. Anal. i Prilozhen.* **42** (1984), 46–57. In Russian.

[Dubovoj et al. 1992] V. K. Dubovoj, B. Fritzsche, and B. Kirstein, *Matricial version of the classical Schur problem*, Teubner-Texte zur Mathematik **129**, B. G. Teubner, Stuttgart, 1992.

[Dym 1989a] H. Dym, *J contractive matrix functions, reproducing kernel Hilbert spaces and interpolation*, CBMS Regional Conference Series in Mathematics **71**, Amer. Math. Soc., Providence, 1989.

[Dym 1989b] H. Dym, "On reproducing kernel spaces, J unitary matrix functions, interpolation and displacement rank", pp. 173–239 in *The Gohberg anniversary collection*, vol. 2, edited by H. Dym, S. Goldberg, M. A. Kaashoek, P. Lancaster, Oper. Theory Adv. Appl. **41**, Birkhäuser, Basel, 1989.

[Dym 1994a] H. Dym, "Review of 'The commutant lifting approach to interpolation problems' by C. Foias and A. E. Frazho", *Bull. Amer. Math. Soc.* (*N.S.*) **31** (1994), 125–140.

[Dym 1994b] H. Dym, "Shifts, realizations and interpolation, redux", pp. 182–243 in *Nonselfadjoint operators and related topics* (Beer Sheva, 1992), edited by A. Feintuch and I. Gohberg, Oper. Theory Adv. Appl. **73**, Birkhäuser, Basel, 1994.

[Dym 1996] H. Dym, "More on maximum entropy interpolants and maximum determinant completions of associated Pick matrices", *Integral Equations Operator Theory* **24**:2 (1996), 188–229.

[Dym and Freydin 1997a] H. Dym and B. Freydin, "Bitangential interpolation for upper triangular operators", pp. 143–164 in *Topics in interpolation theory*, edited by H. Dym, B. Fritzsche, V. Katsnelson, and B. Kirstein, Oper. Theory Adv. Appl. **95**, Birkhäuser, Basel, 1997.

[Dym and Freydin 1997b] H. Dym and B. Freydin, "Bitangential interpolation for triangular operators when the Pick operator is strictly positive", pp. 104–142 in *Topics in interpolation theory*, edited by H. Dym, B. Fritzsche, V. Katsnelson, and B. Kirstein, Oper. Theory Adv. Appl. **95**, Birkhäuser, Basel, 1997.

[Foiaş and Frazho 1990] C. Foiaş and A. E. Frazho, *The commutant lifting approach to interpolation problems*, Oper. Theory Adv. Appl. **44**, Birkhäuser, Basel, 1990.

[Helton et al. 1987] J. W. Helton, J. A. Ball, C. R. Johnson, and J. N. Palmer, *Operator theory, analytic functions, matrices, and electrical engineering*, CBMS Regional Conference Series in Mathematics, Amer. Math. Soc., 1987.

[Ivanchenko and Sakhnovich 1994] T. S. Ivanchenko and L. A. Sakhnovich, "An operator approach to the Potapov scheme for the solution of interpolation problems", pp. 48–86 in *Matrix and operator valued functions*, edited by I. Gohberg and L. Sakhnovich, Oper. Theory Adv. Appl. **72**, Birkhäuser, Basel, 1994.

[Kailath 1980] T. Kailath, *Linear systems*, Prentice-Hall Information and System Sciences Series, Prentice-Hall, Englewood Cliffs, NJ, 1980.

[Katsnelson and Kirstein 1997] V. E. Katsnelson and B. Kirstein, "On the theory of matrix-valued functions belonging to the Smirnov class", pp. 283–350 in *Topics in interpolation theory*, edited by H. Dym, B. Fritzsche, V. Katsnelson, and B. Kirstein, Oper. Theory Adv. Appl. **95**, Birkhäuser, Basel, 1997.

[Katsnelson et al. 1987] V. È. Katsnel'son, A. Y. Kheĭfets, and P. M. Yuditskiĭ, "An abstract interpolation problem and the theory of extensions of isometric operators", pp. 83–96, 146 in *Operators in function spaces and problems in function theory*, edited by V. A. Marchenko, Naukova Dumka, Kiev, 1987. In Russian; translation in *Topics in interpolation theory*, edited by Harry Dym et al., Oper. Theory Adv. Appl. **95**, Birkhäuser, Basel, 1997, pp. 283–298.

[Kheifets 1988a] A. Y. Kheĭfets, "The Parseval equality in an abstract problem of interpolation, and the coupling of open systems, I", *Teor. Funktsiĭ Funktsional. Anal. i Prilozhen.* **49** (1988), 112–120, 126. In Russian; translation in *J. Sov. Math.* **49**:4 (1990) 1114–1120.

[Kheifets 1988b] A. Y. Kheĭfets, "The Parseval equality in an abstract problem of interpolation, and the coupling of open systems, II", *Teor. Funktsiĭ Funktsional. Anal. i Prilozhen.* **50** (1988), 98–103, iv. In Russian; translation in *J. Sov. Math.* **49**:6 (1990) 1307–1310.

[Kheifets 1990] A. Y. Kheĭfets, "A generalized bitangential Schur–Nevanlinna–Pick problem and its related Parseval equality", *Teor. Funktsiĭ Funktsional. Anal. i Prilozhen.* **54** (1990), 89–96. In Russian; translation in *J. Sov. Math.* **58**:4 (1992), 358–364.

[Kheifets 1998] A. Y. Kheifets, "Abstract interpolation problem and applications", pp. 351–379 in *Holomorphic Spaces*, edited by S. Axler et al., Math. Sci. Res. Inst. Publications **33**, Cambridge Univ. Press, New York, 1998.

[Kheifets and Yuditskii 1994] A. Y. Kheĭfets and P. M. Yuditskiĭ, "An analysis and extension of V. P. Potapov's approach to interpolation problems with applications to the generalized bi-tangential Schur–Nevanlinna–Pick problem and J-inner-outer factorization", pp. 133–161 in *Matrix and operator valued functions*, edited by I. Gohberg and L. A. Sakhnovich, Oper. Theory Adv. Appl. **72**, Birkhäuser, Basel, 1994.

[Lev-Ari and Kailath 1986] H. Lev-Ari and T. Kailath, "Triangular factorization of structured Hermitian matrices", pp. 301–324 in *I. Schur methods in operator theory and signal processing*, edited by I. Gohberg, Oper. Theory Adv. Appl. **18**, Birkhäuser, Basel, 1986.

[Nudelman 1993] A. A. Nudelman, "Some generalizations of classical interpolation problems", pp. 171–188 in *Operator extensions, interpolation of functions and related topics* (Timişoara, 1992), edited by D. T. A. Gheondea and F.-H. Vascilescu, Oper. Theory Adv. Appl. **61**, Birkhäuser, Basel, 1993.

[Redheffer 1960] R. M. Redheffer, "On a certain linear fractional transformation", *J. Math. and Phys.* **39** (1960), 269–286.

[Rovnyak 1968] J. Rovnyak, "Characterization of spaces $\mathcal{H}(M)$", 1968. Available at http://wsrv.clas.virginia.edu/~jlr5m/papers/papers.html.

Harry Dym
Department of Theoretical Mathematics
The Weizmann Institute of Science
Rehovot 76100
Israel