EMS Surv. Math. Sci. 3 (2016), 1–105
DOI 10.4171/EMSS/14

EMS Surveys in
Mathematical Sciences

CMV matrices, a matrix version of Baxter's theorem, scattering and de Branges spaces

Harry Dym and David P. Kimsey

Abstract. In this survey we establish bijective correspondences between the following classes of objects: (1) β_{-1} and $\{\beta_n\}_{n=0}^{\infty}$, with $\beta_n \in \mathbb{C}^{p\times p}$ for $n = -1, 0, \ldots$, β_{-1} unitary, $\|\beta_j\| < 1$ for $j \geq 0$ and $\sum_{j=0}^{\infty} \|\beta_j\| < \infty$; (2) A unitary matrix $\beta_{-1} \in \mathbb{C}^{p\times p}$ and a spectral density Δ belonging to the Wiener algebra $\mathcal{W}^{p\times p}$ with $\Delta(\zeta) \succ 0$ for all ζ on the unit circle $\mathbb{T}$; (3) CMV matrices based on a unitary matrix $\beta_{-1} \in \mathbb{C}^{p\times p}$ and a spectral density Δ that meets the constraints in (2); (4) scattering matrices that belong to the Wiener algebra $\mathcal{W}^{p\times p}$; (5) a class of solutions of an associated matricial Nehari problem.

The bijective correspondence between summable sequences of contractions and positive spectral densities in the Wiener algebra $\mathcal{W}^{p\times p}$ (i.e., between class (1) and class (2)) is known as Baxter's theorem and was established by Baxter when $p = 1$ and Geronimo when $p \geq 1$. The connections between CMV matrices, the solutions of a related Nehari problem and an inverse scattering problem seem to be new when $p > 1$. There is partial overlap of the connection between the considered Nehari problem and a discrete analogue of an inverse scattering problem considered by Krein and Melik-Adamjan. de Branges spaces of vector-valued polynomials are used to ease a number of computations.

Mathematics Subject Classification (*2010*). 42C05; 47B36, 47A57, 47B35, 46E22, 47B32.

Keywords. Matrix orthogonal polynomials, Schur parameters CMV matrices, Nehari problem, scattering matrices, reproducing kernel Hilbert spaces, de Branges spaces, Baxter's theorem.

H. Dym, Department of Mathematics, The Weizmann Institute of Science, 234 Herzl Street, Rehovot 7610001, Israel

E-mail: harry.dym@weizmann.ac.il

D. P. Kimsey, Department of Mathematics, Ben Gurion University of the Negev, PO Box 653, Beer Sheva 8410501, Israel

E-mail: dpkimsey@gmail.com

Contents

1. Introduction

The main objective of this paper is to establish a bijective correspondence between:

(1) A class of unitary matrices $\beta_{-1} \in \mathbb{C}^{p\times p}$ and $p \times p$ mvf's (matrix-valued functions) $\Delta(\zeta) = \sum_{n=-\infty}^{\infty} \zeta^n \Delta_n$ on the unit circle $\mathbb{T}$ that are subject to the constraints:

$$\sum_{n=-\infty}^{\infty} \|\Delta_n\| < \infty, \tag{1.1}$$

$$\Delta(\zeta) \succ 0 \quad \text{for} \quad \zeta \in \mathbb{T} \tag{1.2}$$

and

$$\frac{1}{2\pi}\int_0^{2\pi} \Delta(e^{i\theta})d\theta = I_p. \tag{1.3}$$

(2) A class of infinite sequences $\{\beta_n\}_{n=0}^{\infty}$ of $p \times p$ strict contractions (which are usually called Schur parameters) that are subject to the constraint

$$\sum_{n=0}^{\infty} \|\beta_n\| < \infty \tag{1.4}$$

and a unitary matrix $\beta_{-1} \in \mathbb{C}^{p \times p}$.

(3) A class of matrices (commonly called CMV matrices) that play a key role in the matrix representation of the operator of multiplication by ζ in $L_2^{p\times p}(\mathbb{T}, \Delta)$.

(4) A class of scattering matrices (see Definition 10.3) that belong to the Wiener algebra $\mathcal{W}^{p\times p}$.

(5) A set of solutions of an associated matricial Nehari problem.

The constraint (1.1) means that Δ belongs to the Wiener algebra $\mathcal{W}^{p\times p}$. For ease of future reference, we shall say that Δ meets the constraint

(D1) if (1.1), (1.2) and (1.3) are in force.

Remark 1.1. It follows from a theorem of Gohberg and Krein (see, e.g., Corollary 10.4 in Chapter XXX of Gohberg, Goldberg and Kaashoek [23], applied to $\Delta(\zeta)$ and $\Delta(\bar{\zeta})$), that the first two conditions in (D1) hold if and only if

$$\Delta(\zeta) = Q(\zeta)^* Q(\zeta) = R(\zeta) R(\zeta)^* \quad \text{for } \zeta \in \mathbb{T} \tag{1.5}$$

where $Q^{\pm 1}$ and $R^{\pm 1}$ belong to the algebra $\mathcal{W}_+^{p\times p}$ of mvf's $F(\zeta) = \sum_{n=-\infty}^{\infty} \zeta^n F_n$ belonging to $\mathcal{W}^{p\times p}$ with $F_n = 0_{p\times p}$ for $n < 0$.

The supplementary constraint

(D2) $Q(0) \succ 0$ and $R(0) \succ 0$

on the factors in (1.5) insures uniqueness.

The bijection between the classes described in (1) and (2) when $p = 1$ was established by Baxter [5] and is usually referred to as Baxter's theorem (see, e.g., Simon [31] and Bingham [6, 7]). Simon [31] also refers to Stahl [33] and Nuttall and Singh [29] for additional treatments of the "hard direction" of Baxter's theorem. The extension of Baxter's theorem to the case $p > 1$ was first shown by Geronimo [21].

CMV matrices were introduced by Cantero, Moral and Velázquez in [8] (see Simon [32] for a good survey) when $p = 1$ and $\beta_{-1} = 1$. In this case, the CMV matrix based on a probability measure σ on $\mathbb{T}$ is the unitary operator $\mathfrak{A} : \ell_2 \to \ell_2$ given by

$$\mathfrak{A} = V^* M_\zeta V,$$

where $V : \ell_2 \to L_2(\mathbb{T}, \sigma)$ given by $V\mathrm{e}_n = \chi_n$ and M_ζ denotes the operator of multiplication by ζ in $L_2(\mathbb{T}, \sigma)$, $\{\mathrm{e}_n\}_{n=0}^{\infty}$ denotes the canonical orthonormal basis of ℓ_2 and $\{\chi_n\}_{n=0}^{\infty}$ is the CMV basis which will defined in Section 7. Connections between the classes given in (3), (4) and (5) when $p = 1$ are discussed by Golinskii, Kheifets, Peherstorfer and Yuditskii [25]. They focus on the case of square summable Schur parameters.

The bijection between scalar probability measures σ on the unit circle $\mathbb{T}$ and square summable sequences $\{\beta_n\}_{n=0}^{\infty}$, where $|\beta_n| < 1$ for $n = 0, 1, \ldots$, is classical and goes back to Schur [30], Verblunsky [36] and Szegő [34]. Szegő showed that if the Lebesgue decomposition of a scalar probability measure σ is given by

$$d\sigma(\theta) = w(e^{i\theta})\frac{d\theta}{2\pi} + d\sigma_s(\theta)$$

and the density w with respect to the normalized Lebesgue measure satisfies the Szegő condition

$$\frac{1}{2\pi}\int_0^{2\pi} \ln w(e^{i\theta})d\theta > -\infty, \tag{1.6}$$

then the Schur parameters of σ are square summable:

$$\sum_{n=0}^{\infty} |\beta_n|^2 < \infty. \tag{1.7}$$

Conversely, if $\{\beta_n\}_{n=0}^{\infty}$, where $|\beta_n| < 1$ for $n = 0, 1, \ldots$, is square summable, then there exists exactly one probability measure σ on $\mathbb{T}$ so that (1.6) holds and the Schur parameters of σ are given by $\{\beta_n\}_{n=0}^{\infty}$.

There is also a correspondence between $p \times p$ probability measures σ which satisfy a natural analog of (1.6) and sequences of $p \times p$ strict contractions which satisfy $\sum_{n=0}^{\infty} \|\beta_n\|^2 < \infty$ (see, e.g., Damanik, Pushnitski and Simon [9]). Orthogonal polynomials and CMV matrices based on $p \times p$ probability measures are also studied in [9] (see also Simon [32]).

At first glance, the focus on densities Δ in the Wiener algebra $\mathcal{W}^{p\times p}$ may seem overly restrictive. The choice was made initially in order to minimize technical details. But an even stronger case for this restriction is that it fits naturally with the setting of summable Schur parameters, as confirmed by the equivalences between the classes noted earlier.

There is a vast literature on matrix and scalar orthogonal polynomials on the unit circle (see, e.g., Simon [31], Damanik, Pushnitski and Simon [9], Geronimo [21, 22] and Delsarte, Genin and Kamp [10–13]). But going from one source to another is often difficult because of widely different notation and normalizations on the orthogonal polynomials. To minimize this difficulty, a serious attempt has been made to make this presentation self-contained and easily accessible. To this end, three appendices are included with expository material on special properties of scalar orthogonal polynomials, a proof of Baxter's inequality in the matrix case adapted from Findley [20] and a related Nehari problem in the Wiener setting. We have tried to make the proofs as transparent as possible by exploiting the theory of J-inner mvf's and RKHS's (reproducing kernel Hilbert spaces) whenever possible.

Outline of the paper. Section 2 is devoted to matrix orthogonal polynomials; Section 3 to reverse matrix polynomials; Section 4 to the Schur algorithm; Section 5 to an auxiliary pair of orthogonal matrix polynomials; Section 6 to RKHS's; Section 7 to CMV matrices; Section 8 to convergence results; Section 13 to recalling results on a related Nehari problem; Section 14 to explicit formulas for a rational case. The remaining sections are devoted to establishing the equivalences between the classes (1)–(5) discussed earlier; the table below exhibits the locations of the key results.

For the injective map from	See
class (1) to class (2)	Theorem 9.6
class (2) to class (1)	Theorem 11.1
class (2) to class (3)	Theorem 12.1
class (3) to class (4)	Definition 10.3
class (4) to class (5)	Theorem 16.1
class (5) to class (1)	Theorem 15.1
class (5) to class (3)	Corollary 15.3

Table 1.

Remark 1.2. The direct passage from class (2) to class (1) is not needed to establish the equivalence between the classes (1)–(5). It is included because it depends upon a direct construction and not upon a weak compactness of matrix-valued probability measures argument which appears in the construction of a spectral measure when $\sum_{n=0}^{\infty} \|\beta_n\|^2 < \infty$ (see, e.g., Damanik, Pushnitski, and Simon [9]).

Notation.

$\mathbb{D} = \{\lambda \in \mathbb{C} : |\lambda| < 1\}, \mathbb{T} = \{\lambda \in \mathbb{C} : |\lambda| = 1\}$ and $\overline{\mathbb{D}} = \{\lambda \in \mathbb{C} : |\lambda| \leq 1\}$.

$\mathbb{C}^{p\times q}$ = matrices of size $p \times q$ with complex-valued entries.

A^* denotes the Hermitian transpose of $A \in \mathbb{C}^{p\times q}$ and $A^{-*} = (A^{-1})^* = (A^*)^{-1}$ when appropriate.

$A \succeq B$ and $A \succ B$ if $A - B$ is positive semidefinite and positive definite, respectively, for matrices $A, B \in \mathbb{C}^{p\times p}$.

$$j_p = \begin{bmatrix} I_p & 0 \\ 0 & -I_p \end{bmatrix} \in \mathbb{C}^{2p\times 2p}.$$

$\|\cdot\|$ denotes the operator norm.

$F_j = \dfrac{1}{2\pi}\displaystyle\int_0^{2\pi} e^{-ij\theta} F(e^{i\theta})d\theta$, $j = 0, \pm 1, \ldots$, denote the Fourier coefficients of the mvf F.

$\displaystyle\sum_{n=-\infty}^{\infty} e^{in\theta} F_n$ denotes the Fourier series of F.

$F^{\#}(\lambda) = F(1/\overline{\lambda})^*$.

$L_2^{p\times q}(\mathbb{T}, \Delta) = \{$measurable $\mathbb{C}^{p\times q}$ mvf's $F : \|F\|_\Delta^2 < \infty\}$, where

$$\|F\|_\Delta^2 = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{trace}\{F(e^{i\theta})^*\Delta(e^{i\theta})F(e^{i\theta})\}d\theta.$$

$$\langle F, \widetilde{F}\rangle_\Delta = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{trace}\{\widetilde{F}(e^{i\theta})^*\Delta(e^{i\theta})F(e^{i\theta})\}d\theta \quad \text{for } F, \widetilde{F} \in L_2^{p\times p}(\mathbb{T}, \Delta).$$

$$\langle F, \widetilde{F}\rangle_{\mathrm{st}} = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{trace}\{\widetilde{F}(e^{i\theta})^* F(e^{i\theta})\}d\theta \quad \text{for } F, \widetilde{F} \in L_2^{p\times p}(\mathbb{T}, I_p).$$

$$[F, \widetilde{F}]_\Delta = \frac{1}{2\pi}\int_0^{2\pi} \widetilde{F}(e^{i\theta})^*\Delta(e^{i\theta})F(e^{i\theta})d\theta \in \mathbb{C}^{p\times p} \quad \text{for } F, \widetilde{F} \in L_2^{p\times p}(\mathbb{T}, \Delta).$$

$$[F, \widetilde{F}]_{\mathrm{st}} = \frac{1}{2\pi}\int_0^{2\pi} \widetilde{F}(e^{i\theta})^* F(e^{i\theta})d\theta \in \mathbb{C}^{p\times p} \quad \text{for } F, \widetilde{F} \in L_2^{p\times p}(\mathbb{T}, I_p).$$

$H_2^{p\times q} = \{$holomorphic $p\times q$ mvf's F on $\mathbb{D} : \sum_{n=0}^{\infty} \|F_n\|^2 < \infty\}$.

$(H_2^{p\times q})^\perp = \{$holomorphic $p\times q$ mvf's F on $\mathbb{C} \setminus \overline{\mathbb{D}} : \sum_{n=-\infty}^{-1} \|F_n\|^2 < \infty\}$.

$\mathcal{R}^{p\times p} = \{p\times p$ rational mvf's$\}$.

$\mathcal{W}^{p\times p} = \{p\times p$ mvf's F on $\mathbb{T} : \|F\|_{\mathcal{W}}^2 = \sum_{n=-\infty}^{\infty} \|F_n\| < \infty\}$.

$\mathcal{W}_+^{p\times p} = \{F \in \mathcal{W}^{p\times p} : F_n = 0$ for $n < 0\}$.

$\mathcal{W}_-^{p\times p} = \{F \in \mathcal{W}^{p\times p} : F_n = 0$ for $n > 0\}$.

$L_2^p = L_2^{p\times 1}$, $H_2^p = H_2^{p\times 1}$ and $(H_2^p)^\perp = (H_2^{p\times 1})^\perp$.

$\mathfrak{p}$ denotes the orthogonal projection of $L_2^p(\mathbb{T}, I_p)$ onto H_2^p and $\mathfrak{q} = I - \mathfrak{p}$.

$\mathcal{S}^{p\times p}$ denotes the Schur class of $p\times p$ mvf's, which are holomorphic on $\mathbb{D}$ and satisfy $\|S(\lambda)\| \le 1$ for all $\lambda \in \mathbb{D}$.

$\mathcal{S}_{\mathrm{in}}^{p\times p}$ denotes the set of inner mvf's $S \in \mathcal{S}^{p\times p}$ for which $\|S(\zeta)\| = 1$ for all a.e. $\zeta \in \mathbb{T}$.

$\mathcal{C}^{p\times p}$ denotes the Carathéodory class of $p\times p$ mvf's C which are holomorphic on $\mathbb{D}$ and satisfy $C(\lambda) + C(\lambda)^* \succeq 0$.

2. Orthogonal matrix polynomials

If the density Δ satisfies (D1), then the block Toeplitz matrices

$$T_n[\Delta] \stackrel{\text{def}}{=} \begin{bmatrix} \Delta_0 & \cdots & \Delta_{-n} \\ \vdots & \ddots & \vdots \\ \Delta_n & \cdots & \Delta_0 \end{bmatrix} \quad \text{for } n = 0, 1, \dots,$$

based on the Fourier coefficients $\Delta_j = \frac{1}{2\pi}\int_0^{2\pi} e^{-ij\theta}\Delta(e^{i\theta})d\theta$ of Δ, are positive definite for $n = 0, 1, \dots$. Therefore, they are invertible,

$$\Gamma_n = (T_n[\Delta])^{-1} = \begin{bmatrix} \gamma_{00}^{(n)} & \cdots & \gamma_{0n}^{(n)} \\ \vdots & \ddots & \vdots \\ \gamma_{n0}^{(n)} & \cdots & \gamma_{nn}^{(n)} \end{bmatrix} \succ 0 \quad \text{for } n = 0, 1, \dots$$

and

$$\{\gamma_{jk}^{(n)}\}^* = \gamma_{kj}^{(n)} \quad \text{for } 0 \le j, k \le n.$$

Let

$$E_n^+(\lambda) = \sum_{j=0}^{n} \lambda^j \gamma_{j0}^{(n)} \{\gamma_{00}^{(n)}\}^{-1/2} \tag{2.1}$$

and

$$E_n^-(\lambda) = \sum_{j=0}^{n} \lambda^j \gamma_{jn}^{(n)} \{\gamma_{nn}^{(n)}\}^{-1/2}. \tag{2.2}$$

Theorem 2.1. *If Δ satisfies* (D1) *and the matrix polynomials $\{E_n^\pm\}_{n=0}^\infty$ are defined by* (2.1) *and* (2.2)*, then*:

$$\frac{1}{2\pi}\int_0^{2\pi} e^{im\theta} E_m^+(e^{i\theta})^* \Delta(e^{i\theta}) e^{-in\theta} E_n^+(e^{i\theta}) d\theta = \begin{cases} I_p & \text{if } m = n \\ 0_{p\times p} & \text{if } m \ne n. \end{cases} \tag{2.3}$$

$$\frac{1}{2\pi}\int_0^{2\pi} E_m^-(e^{i\theta})^* \Delta(e^{i\theta}) E_n^-(e^{i\theta}) d\theta = \begin{cases} I_p & \text{if } m = n \\ 0_{p\times p} & \text{if } m \ne n. \end{cases} \tag{2.4}$$

$$[E_n^-, E_m^+]_\Delta = \begin{cases} \{\gamma_{00}^{(m)}\}^{-\frac{1}{2}} \gamma_{0m}^{(m)} \{\gamma_{mm}^{(m)}\}^{-\frac{1}{2}} & \text{if } m = n \\ 0_{p\times p} & \text{if } m < n. \end{cases} \tag{2.5}$$

Proof. If $V_n(\lambda) = \sum_{j=0}^n \lambda^j \gamma_{0,n-j}^{(n)}$, then

$$\begin{aligned} \frac{1}{2\pi}\int_0^{2\pi} e^{ik\theta}\Delta(e^{i\theta})V_n(e^{i\theta})^* d\theta &= \sum_{j=0}^n \left\{ \frac{1}{2\pi}\int_0^{2\pi} e^{i(k-j)\theta}\Delta(e^{i\theta})d\theta \right\} \gamma_{n-j,0}^{(n)} \\ &= \sum_{j=0}^n \Delta_{j-k}\gamma_{n-j,0}^{(n)} \end{aligned}$$

$$
\begin{aligned}
&= \sum_{j=0}^{n} \Delta_{n-k-j}\gamma_{j0}^{(n)} \\
&= \begin{bmatrix} \Delta_{n-k} & \cdots & \Delta_{-k} \end{bmatrix} \begin{bmatrix} \gamma_{00}^{(n)} \\ \vdots \\ \gamma_{n0}^{(n)} \end{bmatrix} \\
&= \begin{cases} I_p & \text{if } k = n \\ 0_{p\times p} & \text{if } k = 0,\dots,n-1 \end{cases}
\end{aligned} \tag{2.6}
$$

and hence

$$
\frac{1}{2\pi}\int_0^{2\pi} V_m(e^{i\theta})\Delta(e^{i\theta})V_n(e^{i\theta})^* d\theta = \begin{cases} \gamma_{00}^{(n)} & \text{if } m = n \\ 0_{p\times p} & \text{if } m \neq n. \end{cases} \tag{2.7}
$$

Similarly, if $W_n(\lambda) = \sum_{j=0}^{n} \lambda^j \gamma_{jn}^{(n)}$, then

$$
\begin{aligned}
\frac{1}{2\pi}\int_0^{2\pi} e^{-ik\theta}\Delta(e^{i\theta})W_n(e^{i\theta})d\theta &= \sum_{j=0}^{n} \left\{ \frac{1}{2\pi}\int_0^{2\pi} e^{-i(k-j)\theta}\Delta(e^{i\theta})d\theta \right\} \gamma_{jn}^{(n)} \\
&= \sum_{j=0}^{n} \Delta_{k-j}\gamma_{jn}^{(n)} \\
&= \begin{bmatrix} \Delta_{k} & \cdots & \Delta_{k-n} \end{bmatrix} \begin{bmatrix} \gamma_{0n}^{(n)} \\ \vdots \\ \gamma_{nn}^{(n)} \end{bmatrix} \\
&= \begin{cases} I_p & \text{if } k = n \\ 0_{p\times p} & \text{if } k = 0,\dots,n-1 \end{cases}
\end{aligned} \tag{2.8}
$$

and hence

$$
\frac{1}{2\pi}\int_0^{2\pi} W_m(e^{i\theta})^*\Delta(e^{i\theta})W_n(e^{i\theta}) d\theta = \begin{cases} \gamma_{nn}^{(n)} & \text{if } m = n \\ 0_{p\times p} & \text{if } m \neq n. \end{cases} \tag{2.9}
$$

Formulas (2.3) and (2.4) follow from (2.7) and (2.9), respectively, and the identifications $\lambda^n E_n^+(1/\bar{\lambda})^* = (\gamma_{00}^{(n)})^{-1/2}V_n(\lambda)$ and $E_n^-(\lambda) = W_n(\lambda)(\gamma_{nn}^{(n)})^{-1/2}$ for $n = 0, 1, \dots$. Statement (iii) is an easy consequence of (2.8). □

The "orthonormality" exhibited in Theorem 2.1 leads easily to the following recursion (see, e.g., formulas (13.12) and (13.13) in [15] for help if need be):

$$
\begin{bmatrix} E_{n+1}^-(\lambda) & E_{n+1}^+(\lambda) \end{bmatrix} = \begin{bmatrix} \lambda E_n^-(\lambda) & E_n^+(\lambda) \end{bmatrix} \Lambda_{n+1}, \tag{2.10}
$$

where

$$\Lambda_{n+1} = \begin{bmatrix} \{\gamma_{nn}^{(n)}\}^{-1/2} & 0 \\ 0 & \{\gamma_{00}^{(n)}\}^{-1/2} \end{bmatrix} \begin{bmatrix} I_p & \gamma_{n+1,0}^{(n+1)}\{\gamma_{00}^{(n+1)}\}^{-1} \\ \gamma_{0,n+1}^{(n+1)}\{\gamma_{n+1,n+1}^{(n+1)}\}^{-1} & I_p \end{bmatrix}$$
$$\times \begin{bmatrix} \{\gamma_{n+1,n+1}^{(n+1)}\}^{1/2} & 0 \\ 0 & \{\gamma_{00}^{(n+1)}\}^{1/2} \end{bmatrix}$$
$$= \begin{bmatrix} \{\gamma_{nn}^{(n)}\}^{-1/2}\{\gamma_{n+1,n+1}^{(n+1)}\}^{1/2} & \{\gamma_{nn}^{(n)}\}^{-1/2}\gamma_{n+1,0}^{(n+1)}\{\gamma_{00}^{(n+1)}\}^{-1/2} \\ \{\gamma_{00}^{(n)}\}^{-1/2}\gamma_{0,n+1}^{(n+1)}\{\gamma_{n+1,n+1}^{(n+1)}\}^{-1/2} & \{\gamma_{00}^{(n)}\}^{-1/2}\{\gamma_{00}^{(n+1)}\}^{1/2} \end{bmatrix}. \tag{2.11}$$

Remark 2.2. The diagonal entries in the recursion (4.9) in [19] are incorrect and should be replaced by (2.10).

3. Reverse matrix polynomials

A number of useful formulas are obtained almost for free from the fact that Δ meets the conditions (D1) and (D2) if and only if the mvf

$$\widetilde{\Delta}(\zeta) = \Delta(\zeta^{-1})$$

meets the conditions (D1) and (D2) and the observation that

$$T_n[\widetilde{\Delta}] = \begin{bmatrix} \Delta_0 & \cdots & \Delta_n \\ \vdots & \ddots & \vdots \\ \Delta_{-n} & \cdots & \Delta_0 \end{bmatrix} = Z_n T_n[\Delta] Z_n \succ 0,$$

where

$$Z_n = \begin{bmatrix} 0 & & I_p \\ & \iddots & \\ I_p & & 0 \end{bmatrix} \quad \text{is of size } (n+1)p \times (n+1)p.$$

Consequently,

$$\widetilde{\Gamma}_n = (T_n[\widetilde{\Delta}])^{-1} = Z_n \Gamma_n Z_n = \begin{bmatrix} \gamma_{nn}^{(n)} & \cdots & \gamma_{n0}^{(n)} \\ \vdots & \ddots & \vdots \\ \gamma_{0n}^{(n)} & \cdots & \gamma_{00}^{(n)} \end{bmatrix},$$

i.e.,

$$\widetilde{\gamma}_{jk}^{(n)} = \gamma_{n-j,n-k}^{(n)} \quad \text{for } 0 \le j,k \le n. \tag{3.1}$$

Correspondingly,

$$\begin{aligned}\widetilde{E}_n^+(\lambda) &= \sum_{k=0}^{n} \lambda^k \widetilde{\gamma}_{k0}^{(n)} \{\widetilde{\gamma}_{00}^{(n)}\}^{-1/2} \\ &= \sum_{k=0}^{n} \lambda^k \gamma_{n-k,n}^{(n)} \{\gamma_{nn}^{(n)}\}^{-1/2} \\ &= \sum_{k=0}^{n} \lambda^{n-k} \gamma_{kn}^{(n)} \{\gamma_{nn}^{(n)}\}^{-1/2} = \lambda^n E_n^-(1/\lambda) \end{aligned} \tag{3.2}$$

and

$$\begin{aligned}\widetilde{E}_n^-(\lambda) &= \sum_{k=0}^{n} \lambda^k \widetilde{\gamma}_{kn}^{(n)} \{\widetilde{\gamma}_{nn}^{(n)}\}^{-1/2} \\ &= \sum_{k=0}^{n} \lambda^k \gamma_{n-k,0}^{(n)} (\gamma_{00}^{(n)})^{-1/2} \\ &= \sum_{k=0}^{n} \lambda^{n-k} \gamma_{k0}^{(n)} (\gamma_{00}^{(n)})^{-1/2} = \lambda^n E_n^+(1/\lambda). \end{aligned} \tag{3.3}$$

Moreover, if

$$\widetilde{\Delta}(\zeta) = \widetilde{Q}(\zeta)^* \widetilde{Q}(\zeta) = \widetilde{R}(\zeta)\widetilde{R}(\zeta)^*$$

with

$$\widetilde{Q}^{\pm 1} \in \mathcal{W}_+^{p\times p}, \quad \widetilde{R}^{\pm 1} \in \mathcal{W}_+^{p\times p}, \quad \widetilde{Q}(0) \succ 0, \quad \text{and} \quad \widetilde{R}(0) \succ 0,$$

then

$$\Delta(\zeta) = \widetilde{Q}(\zeta^{-1})^* \widetilde{Q}(\zeta^{-1}) = \widetilde{R}(\zeta^{-1})\widetilde{R}(\zeta^{-1})^*.$$

Thus, by the uniqueness of factorizations with factors subject to the stated conditions, it follows that

$$\widetilde{Q}(\zeta^{-1})^* = R(\zeta) \quad \text{and} \quad \widetilde{R}(\zeta^{-1}) = Q(\zeta)^* \quad \text{for } \zeta \in \mathbb{T}. \tag{3.4}$$

4. The Schur algorithm

In Theorem 4.2 below we shall present an algorithm for generating a sequence of strict contractions $\beta_0, \beta_1, \ldots$ in $\mathbb{C}^{p\times p}$ from a density Δ that meets the constraint (D1). This treatment is partially adapted from [14]. We begin, however, with some notation and a preliminary lemma.

Let

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

and, for $\beta \in \mathbb{C}^{p\times p}$ with $\|\beta\| < 1$, let

$$H(\beta) = \begin{bmatrix} I_p & \beta \\ \beta^* & I_p \end{bmatrix} \begin{bmatrix} (I_p - \beta\beta^*)^{-1/2} & 0 \\ 0 & (I_p - \beta^*\beta)^{-1/2} \end{bmatrix}. \tag{4.1}$$

It is readily checked that

$$H(\beta)^* j_p H(\beta) = H(\beta) j_p H(\beta)^* = j_p, \tag{4.2}$$

$$H(\alpha) = H(\beta) \Longleftrightarrow \alpha = \beta \quad \text{for } \|\alpha\|, \|\beta\| < 1, \tag{4.3}$$

$$H(\beta)^{-1} = H(-\beta), \tag{4.4}$$

and

$$\det H(\beta) = 1, \tag{4.5}$$

since

$$\det \begin{bmatrix} I_p & \beta \\ \beta^* & I_p \end{bmatrix} = \det(I_p - \beta\beta^*) = \det(I_p - \beta^*\beta).$$

Lemma 4.1. *If F_n and G_n belong to $\mathcal{W}_+^{p\times p}$ and*

$$\begin{bmatrix} F_n(\lambda) & G_n(\lambda) \end{bmatrix} j_p \begin{bmatrix} F_n(\lambda)^* \\ G_n(\lambda)* \end{bmatrix} \succ 0 \quad \textit{for } \lambda \in \overline{\mathbb{D}}, \tag{4.6}$$

then:

(i) *$F_n(\lambda)$ is invertible for every point $\lambda \in \overline{\mathbb{D}}$.*

(ii) *$F_n^{-1} \in \mathcal{W}_+^{p\times p}$.*

(iii) *The mvf*

$$S_n(\lambda) = -F_n(\lambda)^{-1} G_n(\lambda)$$

belongs to $\mathcal{S}^{p\times p} \cap \mathcal{W}_+^{p\times p}$.

(iv) *The matrix $\beta_n = S_n(0)$ is a strict contraction, i.e., $\|\beta_n\| < 1$.*

(v) *The mvf's F_{n+1} and G_{n+1} that are defined by the formula*

$$\begin{bmatrix} F_{n+1}(\lambda) & G_{n+1}(\lambda) \end{bmatrix} = \frac{1}{\lambda} \begin{bmatrix} F_n(\lambda) & G_n(\lambda) \end{bmatrix} H(\beta_n) \begin{bmatrix} \lambda I_p & 0 \\ 0 & I_p \end{bmatrix} \tag{4.7}$$

both belong to $\mathcal{W}_+^{p\times p}$, and

$$\begin{bmatrix} F_{n+1}(\lambda) & G_{n+1}(\lambda) \end{bmatrix} j_p \begin{bmatrix} F_{n+1}(\lambda)^* \\ G_{n+1}(\lambda)^* \end{bmatrix} \succ 0 \quad \textit{for } \lambda \in \overline{\mathbb{D}}. \tag{4.8}$$

Proof. Assertions (i)–(iv) are easy consequences of the inequality

$$F_n(\lambda)F_n(\lambda)^* \succ G_n(\lambda)G_n(\lambda)^* \succeq 0,$$

which follows from (4.6).

In view of (4.7),

$$\begin{aligned} F_{n+1}(\lambda) &= (F_n(\lambda) + G_n(\lambda)\beta_n^*)(I_p - \beta_n\beta_n^*)^{-1/2} \\ &= F_n(\lambda)(I_p + F_n(\lambda)^{-1}G_n(\lambda)\beta_n^*)(I_p - \beta_n\beta_n^*)^{-1/2} \end{aligned} \tag{4.9}$$

which is clearly invertible in $\overline{\mathbb{D}}$ and belongs to $\mathcal{W}_+^{p\times p}$, whereas

$$G_{n+1}(\lambda) = \left\{\frac{F_n(\lambda)\beta_n + G_n(\lambda)}{\lambda}\right\}(I_p - \beta_n^*\beta_n)^{-1/2} \tag{4.10}$$

is holomorphic in $\mathbb{D}$ since $F_n(0)\beta_n + G_n(0) = 0_{p\times p}$ and hence belongs to $\mathcal{W}_+^{p\times p}$. Moreover,

$$\begin{aligned} \begin{bmatrix} F_{n+1}(\zeta) & G_{n+1}(\zeta)\end{bmatrix} j_p \begin{bmatrix} F_{n+1}(\zeta)^* \\ G_{n+1}(\zeta)^*\end{bmatrix} &= \begin{bmatrix} F_n(\zeta) & G_n(\zeta)\end{bmatrix} j_p \begin{bmatrix} F_n(\zeta)^* \\ G_n(\zeta)^*\end{bmatrix} \\ &\succ 0 \quad \text{for } \zeta \in \mathbb{T}. \end{aligned}$$

Therefore, the Poisson formula

$$F_{n+1}(\lambda)^{-1}G_{n+1}(\lambda) = \frac{1}{2\pi}\int_0^{2\pi} \left(\frac{1-|\lambda|^2}{|e^{i\theta}-\lambda|^2}\right) F_{n+1}(e^{i\theta})^{-1}G_{n+1}(e^{i\theta})d\theta$$

for $\lambda \in \mathbb{D}$ is applicable and yields the bound

$$\begin{aligned} \|F_{n+1}(\lambda)^{-1}G_{n+1}(\lambda)\| &\leq \frac{1}{2\pi}\int_0^{2\pi} \left(\frac{1-|\lambda|^2}{|e^{i\theta}-\lambda|^2}\right) \|F_{n+1}(e^{i\theta})^{-1}G_{n+1}(e^{i\theta})\|d\theta. \\ &\leq \frac{1}{2\pi}\int_0^{2\pi} \left(\frac{1-|\lambda|^2}{|e^{i\theta}-\lambda|^2}\right) d\theta \\ &= 1. \end{aligned}$$

Thus, (4.8) holds. □

Theorem 4.2. *If a density Δ satisfies (D1) and $C(\lambda) = I_p + 2\sum_{n=1}^{\infty}\lambda^n\Delta_n$, then:*

(i) *The mvf's*

$$F_0(\lambda) = C(\lambda) + I_p \quad \text{and} \quad G_0(\lambda) = C(\lambda) - I_p$$

both belong to $\mathcal{W}_+^{p\times p}$ and

$$F_0(\lambda)F_0(\lambda)^* - G_0(\lambda)G_0(\lambda)^* = 2\{C(\lambda) + C(\lambda)^*\} \succ 0, \quad \text{for } \lambda \in \overline{\mathbb{D}}.$$

(ii) *There exists a sequence of strict contractions* $\{\beta_n\}_{n=0}^{\infty}$ *given by*

$$\beta_n = -F_n(0)^{-1}G_n(0) \quad \textit{for} \;\; n = 0, 1, \ldots, \tag{4.11}$$

where

$$\begin{bmatrix} F_{n+1}(\lambda) & G_{n+1}(\lambda) \end{bmatrix} = \frac{1}{\lambda}\begin{bmatrix} F_n(\lambda) & G_n(\lambda) \end{bmatrix} H(\beta_n) \begin{bmatrix} \lambda I_p & 0 \\ 0 & I_p \end{bmatrix} \quad \textit{for} \;\; n \geq 0.$$

Proof. The first assertion in Statement (i) is self-evident. The second assertion in Statement (i) follows by noting that

$$\frac{C(\lambda) + C(\lambda)^*}{2} = \begin{cases} \dfrac{1}{2\pi} \displaystyle\int_0^{2\pi} \dfrac{1 - |\lambda|^2}{|e^{i\theta} - \lambda|^2} \Delta(e^{i\theta}) d\theta & \text{if } |\lambda| < 1 \\ \Delta(\lambda) & \text{if } |\lambda| = 1. \end{cases}$$

Statement (ii) follows from Lemma 4.1 and Statement (i). □

Definition 4.3. The sequence of strict contractions $\{\beta_n\}_{n=0}^{\infty}$ in Theorem 4.2 will be called the *Schur parameters* corresponding to the density Δ.

Remark 4.4. In the setting of Theorem 4.2, $\beta_0 = 0_{p\times p}$ since $G_0(0) = 0_{p\times p}$.

Corollary 4.5. *If* $\{\beta_n\}_{n=0}^{\infty}$ *are the Schur parameters corresponding to a density* Δ *which satisfies* (D1) *and*

$$S_n(\lambda) = -F_n(\lambda)^{-1}G_n(\lambda) \quad \textit{for} \;\; n = 0, 1, \ldots, \tag{4.12}$$

then

$$S_{n+1}(\lambda) = (I_p - \beta_n\beta_n^*)^{1/2}(I_p - S_n(\lambda)\beta_n^*)^{-1}\left\{\frac{S_n(\lambda) - \beta_n}{\lambda}\right\}(I_p - \beta_n^*\beta_n)^{-1/2} \tag{4.13}$$

and

$$\beta_{n+1} = (I_p - \beta_n\beta_n^*)^{-1/2} \lim_{\lambda \downarrow 0} \left\{\frac{S_n(\lambda) - \beta_n}{\lambda}\right\}(I_p - \beta_n^*\beta_n)^{-1/2} \tag{4.14}$$

for $n = 0, 1, \ldots$.

Proof. This is immediate from (4.9), (4.10) and (4.12). □

5. Orthogonal matrix polynomials generated by a sequence of strict contractions

Inversion in Wiener algebras. It is well known that:

(1) If $f \in \mathcal{W}^{1\times 1}$, then $f^{-1} \in \mathcal{W}^{1\times 1}$ if and only if $f(\zeta) \neq 0$ for $\zeta \in \mathbb{T}$.

(2) If $f \in \mathcal{W}_+^{1\times 1}$, then $f^{-1} \in \mathcal{W}_+^{1\times 1}$ if and only if $f(\lambda) \neq 0$ for $\lambda \in \overline{\mathbb{D}}$.

(3) If $f \in \mathcal{W}_-^{1\times 1}$, then $f^{-1} \in \mathcal{W}_-^{1\times 1}$ if and only if $f(\lambda) \neq 0$ for $\lambda \in \mathbb{C} \setminus \mathbb{D}$, and
$$\lim_{\lambda\to\infty} f(\lambda) \neq 0.$$

The first assertion is a theorem of Wiener; for a proof based on Gelfand theory, see, e.g., Theorem 1.10.6 in Arveson [4]. Item (2) is given as a exercise on p. 30 of [4] (see, also, Theorem W_+ on p. 176 of Krein [27] for a continuous analog). Item (3) is an easy consequence of (2), since
$$f \in \mathcal{W}_+^{1\times 1} \iff f^{\#} \in \mathcal{W}_-^{1\times 1}$$
and
$$\lambda \in \overline{\mathbb{D}} \iff 1/\overline{\lambda} \in (\mathbb{C} \setminus \mathbb{D}) \cup \{\infty\}.$$

Items (1)–(3) carry over easily to the matrix case, since a mvf F is invertible at λ if and only if $\det F(\lambda) \neq 0$, and in that case
$$F(\lambda)^{-1} = \frac{G(\lambda)}{\det F(\lambda)},$$
where $G = (g_{jk})_{j,k=1}^p$, and
$$g_{jk}(\lambda) = (-1)^{j+k} \times kj \text{ minor of } F(\lambda).$$

Theorem 5.1. *The following statements hold*:

(1) *If* $F \in \mathcal{W}^{p\times p}$, $F^{-1} \in \mathcal{W}^{p\times p}$ *if and only if* $\det F(\zeta) \neq 0$ *for* $\zeta \in \mathbb{T}$.

(2) *If* $F \in \mathcal{W}_+^{p\times p}$, *then* $F^{-1} \in \mathcal{W}_+^{p\times p}$ *if and only if* $\det F(\lambda) \neq 0$ *for* $\lambda \in \overline{\mathbb{D}}$.

(3) *If* $F \in \mathcal{W}_-^{p\times p}$, *then* $F^{-1} \in \mathcal{W}_-^{p\times p}$ *if and only if* $\det F(\lambda) \neq 0$ *for* $\lambda \in \mathbb{C} \setminus \mathbb{D}$, *and*
$$\lim_{\lambda\to\infty} \det F(\lambda) \neq 0.$$

Given $\beta_0, \ldots, \beta_n \in \mathbb{C}^{p\times p}$ with $\|\beta_k\| < 1$ for $k = 0, \ldots, n$, let
$$\vartheta_k(\lambda) = H(\beta_k) \begin{bmatrix} \lambda I_p & 0 \\ 0 & I_p \end{bmatrix} \quad \text{for } k = 0, \ldots, n, \tag{5.1}$$
$$\Theta_k(\lambda) = \vartheta_0(\lambda) \cdots \vartheta_k(\lambda) \quad \text{for } k = 0, \ldots, n \tag{5.2}$$
and
$$\begin{bmatrix} \lambda F_k^-(\lambda) & F_k^+(\lambda) \end{bmatrix} = \begin{bmatrix} I_p & I_p \end{bmatrix} \Theta_k(\lambda) \quad \text{for } k = 0, \ldots, n. \tag{5.3}$$

A recursion relation for the sequences of mvf's $\{F_k^+\}_{k=0}^n$ and $\{F_k^-\}_{k=0}^n$ follows readily from (5.2) and is given by
$$\begin{bmatrix} F_k^-(\lambda) & F_k^+(\lambda) \end{bmatrix} = \begin{bmatrix} \lambda F_{k-1}^-(\lambda) & F_{k-1}^+(\lambda) \end{bmatrix} H(\beta_k) \tag{5.4}$$
for $k = 1, \ldots, n$.

It will be convenient to write

$$\Theta_k(\lambda) = \begin{bmatrix} \theta_{11}^{(k)}(\lambda) & \theta_{12}^{(k)}(\lambda) \\ \theta_{21}^{(k)}(\lambda) & \theta_{22}^{(k)}(\lambda) \end{bmatrix} \quad \text{for } k = 0, \dots, n \tag{5.5}$$

and let

$$X_k = (I_p - \beta_0\beta_0^*)^{-1/2} \cdots (I_p - \beta_k\beta_k^*)^{-1/2} \quad \text{for } k = 0, \dots, n \tag{5.6}$$

and

$$Y_k = (I_p - \beta_0^*\beta_0)^{-1/2} \cdots (I_p - \beta_k^*\beta_k)^{-1/2} \quad \text{for } k = 0, \dots, n. \tag{5.7}$$

Theorem 5.2. *If* $\beta_0, \dots, \beta_n \in \mathbb{C}^{p\times p}$ *with* $\beta_0 = 0_{p\times p}$ *and* $\|\beta_k\| < 1$ *for* $k = 0, \dots, n$, *then*:

(i) $\Theta_k(\lambda)$ *is a matrix polynomial of degree* $k+1$.

(ii) $\Theta_k(0) = \left[\begin{smallmatrix} 0 & 0 \\ 0 & Y_k \end{smallmatrix}\right]$.

(iii) $\lambda^{-k-1}\Theta_k(\lambda) \to \left[\begin{smallmatrix} X_k & 0 \\ 0 & 0 \end{smallmatrix}\right]$ *as* $\lambda \to \infty$.

(iv) $\Theta_k(\lambda)^* j_p \Theta_k(\lambda) \preceq j_p$ *if* $0 \le |\lambda| \le 1$ *with equality when* $|\lambda| = 1$.

(v) $\Theta_k(\lambda) j_p \Theta_k(\lambda)^* \preceq j_p$ *if* $0 \le |\lambda| \le 1$ *with equality when* $|\lambda| = 1$.

(vi) $\Theta_k(\lambda)^* j_p \Theta_k(\lambda) \succeq j_p$ *if* $1 \le |\lambda| < \infty$ *with equality when* $|\lambda| = 1$.

(vii) $\Theta_k(\lambda) j_p \Theta_k(\lambda)^* \succeq j_p$ *if* $1 \le |\lambda| < \infty$ *with equality when* $|\lambda| = 1$.

(viii) $\det \Theta_k(\lambda) = \lambda^{(k+1)p}$.

(ix) $(\theta_{22}^{(k)})^{\pm 1} \in \mathcal{W}_+^{p\times p}$.

(x) $(\zeta^{-k-1}\theta_{11}^{(k)})^{\pm 1} \in \mathcal{W}_-^{p\times p}$.

Proof. Statements (i)–(iii) are clear from (5.2). Statement (iv) and (v) are verified by using (4.2) to obtain

$$\begin{aligned} \vartheta_k(\lambda)^* j_p \vartheta_k(\lambda) &= \begin{bmatrix} \lambda I_p & 0 \\ 0 & I_p \end{bmatrix}^* H(\beta_k)^* j_p H(\beta_k) \begin{bmatrix} \lambda I_p & 0 \\ 0 & I_p \end{bmatrix} \\ &= \begin{bmatrix} \lambda I_p & 0 \\ 0 & I_p \end{bmatrix}^* j_p \begin{bmatrix} \lambda I_p & 0 \\ 0 & I_p \end{bmatrix} \\ &= \begin{bmatrix} |\lambda|^2 I_p & 0 \\ 0 & -I_p \end{bmatrix} \\ &\preceq j_p \quad \text{if } \lambda \in \overline{\mathbb{D}}. \end{aligned} \tag{5.8}$$

and

$$\vartheta_k(\lambda) j_p \vartheta_k(\lambda)^* \preceq j_p \quad \text{for } \lambda \in \overline{\mathbb{D}},$$

respectively.

The inequality

$$\vartheta_k(\lambda)^* j_p \vartheta_k(\lambda) \succeq j_p \quad \text{for } 1 \le |\lambda| < \infty \tag{5.9}$$

can be verified in much the same way as (5.8). Statements (vi) and (vii) follow directly from (5.9) and

$$\vartheta_k(\lambda) j_p \vartheta_k(\lambda)^* \succeq j_p \quad \text{for } 1 \le |\lambda| < \infty,$$

respectively.

In view of (4.5),

$$\det \vartheta_k(\lambda) = \det H(\beta_k)\lambda^p = \lambda^p$$

and (viii) follows easily from (5.2).

To verify Statement (ix), note that the 22 block of the inequality in Statement (iv) implies that

$$\theta_{22}^{(k)}(\lambda)^*\theta_{22}^{(k)}(\lambda) \succeq I_p + \theta_{12}^{(k)}(\lambda)^*\theta_{12}^{(k)}(\lambda) \quad \text{for } \lambda \in \overline{\mathbb{D}} \tag{5.10}$$

and hence $\theta_{22}^{(k)}$ is invertible for $\lambda \in \overline{\mathbb{D}}$. Therefore, in view of item (2) of Theorem 5.1, $(\theta_{22}^{(k)})^{-1} \in \mathcal{W}_+^{p\times p}$. The conclusion $(\theta_{22}^{(k)}) \in \mathcal{W}_+^{p\times p}$ is immediate from (i).

To verify (x), note that the 11 block of the inequality in Statement (vi) is

$$\theta_{11}^{(k)}(\lambda)^*\theta_{11}^{(k)}(\lambda) - \theta_{21}^{(k)}(\lambda)^*\theta_{21}^{(k)}(\lambda) \succeq I_p \quad \text{for } |\lambda| \ge 1$$

and hence $\theta_{11}^{(k)}(\lambda)$ is invertible for every point $|\lambda| \ge 1$. In view of item (iii),

$$\lim_{\lambda\to\infty} \det \lambda^{-k-1}\theta_{11}^{(k)}(\lambda) = \det X_k \ne 0.$$

Thus, in view of item (3) of Theorem 5.1, $(\zeta^{-k-1}\theta_{11}^{(k)})^{-1} \in \mathcal{W}_-^{p\times p}$. The conclusion $(\zeta^{-k-1}\theta_{11}^{(k)}) \in \mathcal{W}_-^{p\times p}$ is immediate from (i). □

In view of item (iv) of Theorem 5.2, the mvf Θ_k generated by $\beta_0, \dots, \beta_n$ is a j_p-inner mvf for $k = 0, \dots, n$. The equality on $\mathbb{T}$ extends to

$$\Theta_k^\#(\lambda) j_p \Theta_k(\lambda) = j_p = \Theta_k(\lambda) j_p \Theta_k^\#(\lambda) \quad \text{for } \lambda \in \mathbb{C} \tag{5.11}$$

and $k = 0, \dots, n$.

Linear fractional transformations.

Theorem 5.3. *If $\mathcal{E} \in \mathcal{S}^{p\times p}$ and $\beta_0, \dots, \beta_k \in \mathbb{C}^{p\times p}$ with $\|\beta_k\| < 1$ for $k = 0, \dots, n$, then $\theta_{21}^{(k)}\mathcal{E} + \theta_{22}^{(k)}$ is invertible on $\overline{\mathbb{D}}$ and*

$$T_{\Theta_k}[\mathcal{E}] \overset{\text{def}}{=} (\theta_{11}^{(k)}\mathcal{E} + \theta_{12}^{(k)})(\theta_{21}^{(k)}\mathcal{E} + \theta_{22}^{(k)})^{-1}$$

maps $\mathcal{S}^{p\times p}$ into $\mathcal{S}^{p\times p}$ for $k = 0, \dots, n$.

Proof. We will first show that $\theta_{21}^{(k)}\mathcal{E} + \theta_{22}^{(k)}$ is invertible. The 22 block of the inequality in item (v) of Theorem 5.2 implies that

$$\theta_{21}^{(k)}(\lambda)\theta_{21}^{(k)}(\lambda)^* - \theta_{22}^{(k)}(\lambda)\theta_{22}^{(k)}(\lambda)^* \preceq -I_p \quad \text{for } \lambda \in \overline{\mathbb{D}}.$$

Thus, $\theta_{22}^{(k)}(\lambda)^{-1}\theta_{21}^{(k)}(\lambda)$ is a strict contraction for every $\lambda \in \overline{\mathbb{D}}$ and hence

$$\|\theta_{22}^{(k)}(\lambda)^{-1}\theta_{21}^{(k)}(\lambda)\mathcal{E}(\lambda)\| < 1 \quad \text{for } \lambda \in \overline{\mathbb{D}}.$$

Consequently, $\theta_{21}^{(k)}\mathcal{E} + \theta_{22}^{(k)}$ is invertible for every $\lambda \in \overline{\mathbb{D}}$.

Moreover, since $\theta_{ij}^{(k)}$ are polynomials for $i, j = 1, 2$, we have that

$$(\theta_{21}^{(k)}\mathcal{E} + \theta_{22}^{(k)})^{-1} \quad \text{and} \quad (\theta_{11}^{(k)}\mathcal{E} + \theta_{12}^{(k)})$$

are holomorphic on $\mathbb{D}$, and thus $T_{\Theta_k}[\mathcal{E}]$ is holomorphic on $\mathbb{D}$.

It remains to check that $I_p - T_{\Theta_k}[\mathcal{E}]^* T_{\Theta_k}[\mathcal{E}] \succeq 0$. But this follows from item (iv) of Theorem 5.2:

$$\begin{aligned} I_p - T_{\Theta_k}[\mathcal{E}]^* T_{\Theta_k}[\mathcal{E}] &= (\theta_{21}^{(k)}\mathcal{E} + \theta_{22}^{(k)})^{-*} \begin{bmatrix} \mathcal{E}^* & I_p \end{bmatrix} \Theta_k^*(-j_p)\Theta_k \begin{bmatrix} \mathcal{E} \\ I_p \end{bmatrix} \\ &\qquad\qquad\qquad\qquad \times (\theta_{21}^{(k)}\mathcal{E} + \theta_{22}^{(k)})^{-1} \\ &\succeq (\theta_{21}^{(k)}\mathcal{E} + \theta_{22}^{(k)})^{-*} \begin{bmatrix} \mathcal{E}^* & I_p \end{bmatrix} (-j_p) \begin{bmatrix} \mathcal{E} \\ I_p \end{bmatrix} (\theta_{21}^{(k)}\mathcal{E} + \theta_{22}^{(k)})^{-1}, \end{aligned}$$

since

$$\begin{bmatrix} \mathcal{E}^* & I_p \end{bmatrix} (-j_p) \begin{bmatrix} \mathcal{E} \\ I_p \end{bmatrix} = I_p - \mathcal{E}^*\mathcal{E} \succeq 0 \quad \text{on } \mathbb{D}.$$ □

Parametrization of $\boldsymbol{\Theta_k}$. The mvf's $\Theta_k(\lambda)$, $k = 0, \dots, n$, defined by (5.1) and (5.3) are completely determined by the given sequence $\beta_0, \dots, \beta_n \in \mathbb{C}^{p\times p}$ with $\|\beta_k\| < 1$ for $k = 0, \dots, n$, the inequality (5.10) implies that

$$\sigma_k \overset{\text{def}}{=} T_{\Theta_k}[0_{p\times p}] = (\theta_{12}^{(k)})(\theta_{22}^{(k)})^{-1} \quad \text{for } k = 0, \dots, n$$

is strictly contractive on $\overline{\mathbb{D}}$ and hence that the mvf

$$C_k = (I_p - \sigma_k)(I_p + \sigma_k)^{-1} \quad \text{for } k = 0, \dots, n \tag{5.12}$$

belongs to the Carathéodory class $\mathcal{C}^{p\times p}$ and

$$(I_p + C_k)^{\pm 1} \in \mathcal{W}_+^{p\times p} \quad \text{for } k = 0, \dots, n.$$

Moreover, in view of Theorem 5.2, $\theta_{11}^{(k)}(0) = 0_{p\times p}$. Therefore, $\lambda^{-1}\theta_{11}^{(k)}(\lambda)$ is a matrix polynomial of degree at most k.

Lemma 5.4. *If $\{F_k^{\pm}\}_{k=0}^n$ are the matrix polynomials defined by* (5.3) *in terms of the strict contractions $\beta_0, \dots, \beta_n \in \mathbb{C}^{p\times p}$, then*

$$F_k^+(\lambda) = 2\{I_p + C_k(\lambda)\}^{-1}\theta_{22}^{(k)}(\lambda) \quad \text{for } \lambda \in \overline{\mathbb{D}}, \tag{5.13}$$

$$F_k^-(\lambda) = 2\{I_p + C_k^{\#}(\lambda)\}^{-1}\lambda^{-1}\theta_{11}^{(k)}(\lambda) \quad \text{for } \lambda \in \mathbb{C}\setminus\mathbb{D}, \tag{5.14}$$

$$\theta_{12}^{(k)}(\lambda) = \sigma_k(\lambda)\theta_{22}^{(k)}(\lambda) = \left\{\frac{I_p - C_k(\lambda)}{2}\right\} F_k^+(\lambda) \quad \text{for } \lambda \in \overline{\mathbb{D}}, \tag{5.15}$$

$$\begin{aligned}(\theta_{21}^{(k)})^{\#}(\lambda) &= (\theta_{11}^{(k)})^{\#}(\lambda)\sigma_k(\lambda)\\ &= (\lambda F_k^-)^{\#}(\lambda)\left\{\frac{I_p - C_k(\lambda)}{2}\right\} \quad \text{for } \lambda \in \mathbb{C}\setminus\mathbb{D}\end{aligned} \tag{5.16}$$

and

$$\Theta_k(\zeta) = \frac{1}{2}\begin{bmatrix}\{I_p + C_k^{\#}(\zeta)\}\zeta F_k^-(\zeta) & \{I_p - C_k(\zeta)\}F_k^+(\zeta)\\ \{I_p - C_k^{\#}(\zeta)\}\zeta F_k^-(\zeta) & \{I_p + C_k(\zeta)\}F_k^+(\zeta)\end{bmatrix} \tag{5.17}$$

for $\zeta \in \mathbb{T}$.

Proof. The proof is broken into steps.

1. Verification of (5.13) *and* (5.15). In view of (5.3),

$$\begin{aligned}F_k^+(\lambda) &= \theta_{12}^{(k)}(\lambda) + \theta_{22}^{(k)}(\lambda)\\ &= \{\sigma_k(\lambda) + I_p\}\theta_{22}^{(k)}(\lambda)\\ &= 2\{I_p + C_k(\lambda)\}^{-1}\theta_{22}^{(k)}(\lambda) \quad \text{for } \lambda \in \overline{\mathbb{D}}.\end{aligned}$$

Thus (5.13) holds. To verify (5.15), use (5.13) to write

$$\begin{aligned}\theta_{12}^{(k)}(\lambda) &= \theta_{12}^{(k)}(\lambda)\theta_{22}^{(k)}(\lambda)^{-1}\theta_{22}^{(k)}(\lambda)\\ &= \sigma_k(\lambda)\theta_{22}^{(k)}(\lambda)\\ &= \sigma_k(\lambda)\left\{\frac{I_p + C_k(\lambda)}{2}\right\} F_k^+(\lambda) \quad \text{for } \lambda \in \overline{\mathbb{D}}.\end{aligned}$$

2. Verification of (5.14) *and* (5.16). In view of (5.11),

$$(\theta_{22}^{(k)})^{\#}(\lambda)^{-1}(\theta_{12}^{(k)})^{\#}(\lambda) = \theta_{21}^{(k)}(\lambda)\theta_{11}^{(k)}(\lambda)^{-1} \quad \text{for } \lambda \in \mathbb{C}\setminus\mathbb{D}.$$

Therefore, in view of (5.3),

$$\begin{aligned}\lambda F_k^-(\lambda) &= \theta_{11}^{(k)}(\lambda) + \theta_{21}^{(k)}(\lambda)\\ &= \{I_p + \theta_{21}^{(k)}(\lambda)\theta_{11}^{(k)}(\lambda)^{-1}\}\theta_{11}^{(k)}(\lambda)\\ &= \{I_p + (\theta_{22}^{(k)})^{\#}(\lambda)^{-1}(\theta_{12}^{(k)})^{\#}(\lambda)\}\theta_{11}^{(k)}(\lambda)\\ &= 2\{I_p + C_k^{\#}(\lambda)\}^{-1}\theta_{11}^{(k)}(\lambda) \quad \text{for } \lambda \in \mathbb{C}\setminus\mathbb{D}.\end{aligned}$$

Thus, (5.14) holds. To verify (5.16), use (5.11) to obtain

$$\theta_{11}^{\#}(\lambda)\theta_{12}(\lambda) = \theta_{21}^{\#}(\lambda)\theta_{22}(\lambda)$$

and hence

$$\begin{aligned}(\theta_{21}^{(k)})^{\#}(\lambda) &= (\theta_{21}^{(k)})^{\#}(\lambda)\theta_{22}^{(k)}(\lambda)\theta_{22}^{(k)}(\lambda)^{-1} \\ &= \theta_{11}^{\#}(\lambda)\theta_{12}(\lambda)\theta_{22}^{(k)}(\lambda)^{-1} \\ &= \theta_{11}^{\#}(\lambda)\sigma_k(\lambda) \\ &= (\lambda F_k^{-})^{\#}(\lambda)\left\{\frac{I_p - C_k(\lambda)}{2}\right\} \quad \text{for } \lambda \in \mathbb{C} \setminus \mathbb{D}.\end{aligned}$$

3. Verification of (5.17). In view of (5.5), Assertion (5.17) follows directly from (5.13)–(5.16). □

Theorem 5.5. *If $\beta_0, \dots, \beta_n \in \mathbb{C}^{p\times p}$ with $\beta_0 = 0_{p\times p}$ and $\|\beta_k\| < 1$ for $k = 1, \dots, n$, then the $p \times p$ mvf's C_k, F_k^- and F_k^+, for $k = 0, \dots, n$, enjoy the following properties:*

(1) *F_k^+ is a matrix polynomial of degree at most k and $F_k^+(0) = Y_k$ is invertible.*

(2) *F_k^- is a matrix polynomial of degree k and*

$$\lim_{\lambda\to\infty} \lambda^{-k} F_k^-(\lambda) = X_k.$$

(3) $(F_k^+)^{\pm 1} \in \mathcal{W}_+^{p\times p}$.

(4) $(\zeta^{-k} F_k^-)^{\pm 1} \in \mathcal{W}_-^{p\times p}$.

(5) $F_k^-(\zeta)^* \left\{\frac{C_k(\zeta)+C_k(\zeta)^*}{2}\right\} F_k^-(\zeta) = I_p$ *for* $\zeta \in \mathbb{T}$.

(6) $F_k^+(\zeta)^* \left\{\frac{C_k(\zeta)+C_k(\zeta)^*}{2}\right\} F_k^+(\zeta) = I_p$ *for* $\zeta \in \mathbb{T}$.

Proof. The proof is divided into steps.

1. Verification of (1) *and* (2). Since

$$\begin{bmatrix}\lambda F_k^-(\lambda) & F_k^+(\lambda)\end{bmatrix} = \begin{bmatrix}I_p & I_p\end{bmatrix}\Theta_k(\lambda) \tag{5.18}$$

$$= \begin{bmatrix}I_p & I_p\end{bmatrix}\Theta_{k-1}(\lambda)H(\beta_k)\begin{bmatrix}\lambda I_p & 0 \\ 0 & I_p\end{bmatrix} \tag{5.19}$$

and $\Theta_{k-1}(\lambda)$ is a matrix polynomial of degree at most k, it is clear that $F_k^-(\lambda)$ and $F_k^+(\lambda)$ are matrix polynomials of degree at most k and $F_k^+(0)$ is invertible. The assertions $F_k^+(0) = Y_k$ and $\lambda^{-k}F_k^-(\lambda) \to X_k$ as $\lambda \to \infty$ follow from items (ii) and (iii) in Theorem 5.2, respectively.

2. Verification of (3) *and* (4). The assertions $F_k^+ \in \mathcal{W}_+^{p\times p}$ and $\zeta^{-k}F_k^- \in \mathcal{W}_-^{p\times p}$ are automatic since F_k^+ is a matrix polynomial and F_k^- is a matrix polynomial of degree k. In view of (5.13) and item (ix) in Theorem 5.2, $F_k^+(\lambda)$ is invertible for $\lambda \in \mathbb{D}$. Thus, it follows from item (2) of Theorem 5.1 that $(F_k^+)^{-1} \in \mathcal{W}_+^{p\times p}$. Similarly, in view of (5.14) and item (x) of Theorem 5.2, $\zeta^{-k}F_k^-$ is invertible for all $\mathbb{C}\setminus\mathbb{D}$. Thus, it follows from item (3) of Theorem 5.1 that $(\zeta^{-k}F_k^-)^{-1} \in \mathcal{W}_-^{p\times p}$.

3. Verification of (5) *and* (6). Both formulas are straightforward computations based on the formula given in item (iv) in Theorem 5.2 when $|\lambda| = 1$. Thus, for example, the 22 block yields the identity

$$\theta_{12}^{(k)}(\zeta)^*\theta_{12}^{(k)}(\zeta) - \theta_{22}^{(k)}(\zeta)^*\theta_{22}^{(k)}(\zeta) = -I_p \quad \text{for } \zeta \in \mathbb{T},$$

which implies that

$$F_k^+(\zeta)^*\left(\frac{\{I_p - C_k(\zeta)^*\}\{I_p - C_k(\zeta)\} - \{I_p + C_k(\zeta)^*\}\{I_p + C_k(\zeta)\}}{4}\right)F_k^+(\zeta) = -I_p \tag{5.20}$$

for $\zeta \in \mathbb{T}$. Thus, (6) follows directly since

$$\{I_p - C_k(\zeta)^*\}\{I_p - C_k(\zeta)\} - \{I_p + C_k(\zeta)^*\}\{I_p + C_k(\zeta)\} = -2\{C_k(\zeta) + C_k(\zeta)^*\}$$

for $\zeta \in \mathbb{T}$. The verification of (5) is similar, but is based on the formula

$$\Theta_k(\zeta)^* j_p \Theta_k(\zeta) = j_p \quad \text{for } \zeta \in \mathbb{T}.$$

□

Theorem 5.6. *If $\{\beta_n\}_{n=0}^\infty$ are the Schur parameters based on a density Δ which satisfies* (D1), *then:*

(1) $\frac{1}{\lambda^{n+1}}\{C(\lambda) - C_n(\lambda)\}$ is a holomorphic mvf on $\mathbb{D}$.

(2) $\frac{1}{2\pi}\int_0^{2\pi} e^{ik\theta}\Delta(e^{i\theta})d\theta = \frac{1}{2\pi}\int_0^{2\pi} e^{ik\theta}\left\{\frac{C_n(e^{i\theta})+C_n(e^{i\theta})^*}{2}\right\}d\theta$ if $|k| \le n$.

(3) $\frac{1}{2\pi}\int_0^{2\pi} F_k^-(e^{i\theta})^*\Delta(e^{i\theta})F_n^-(e^{i\theta})d\theta = \begin{cases} 0_{p\times p} & \text{if } k \ne n \\ I_p & \text{if } k = n. \end{cases}$

(4) $\frac{1}{2\pi}\int_0^{2\pi} \{e^{-ik\theta}F_k^+(e^{i\theta})\}^*\Delta(e^{i\theta})\{e^{-in\theta}F_n^+(e^{i\theta})\}d\theta = \begin{cases} 0_{p\times p} & \text{if } k \ne n \\ I_p & \text{if } k = n. \end{cases}$

Proof. The proof is divided into steps.

1. Verification of (1). By definition,

$$\begin{bmatrix} C(\lambda) + I_p & C(\lambda) - I_p \end{bmatrix}\Theta_n(\lambda) = \lambda^{n+1}\begin{bmatrix} F_{n+1}(\lambda) & G_{n+1}(\lambda) \end{bmatrix}$$

and by direct computation,

$$\begin{bmatrix} C_n(\lambda) + I_p & C_n(\lambda) - I_p \end{bmatrix}\Theta_n(\lambda) = \begin{bmatrix} \{C_n(\lambda) + C_n^\#(\lambda)\}\lambda F_n^-(\lambda) & 0 \end{bmatrix}.$$

Thus, if we subtract the second formula from the first we get

$$\begin{bmatrix} C(\lambda) - C_n(\lambda) & C(\lambda) - C_n(\lambda) \end{bmatrix} \Theta_n(\lambda) = \begin{bmatrix} * & \lambda^{n+1} G_{n+1}(\lambda) \end{bmatrix}.$$

Therefore,

$$\begin{aligned} \{C(\lambda) - C_n(\lambda)\} \begin{bmatrix} \lambda F_n^-(\lambda) & F_n^+(\lambda) \end{bmatrix} &= \{C(\lambda) - C_n(\lambda)\} \begin{bmatrix} I_p & I_p \end{bmatrix} \Theta_n(\lambda) \\ &= \begin{bmatrix} * & \lambda^{n+1} G_{n+1}(\lambda) \end{bmatrix}. \end{aligned}$$

Thus,

$$\{C(\lambda) - C_n(\lambda)\} F_n^+(\lambda) = \lambda^{n+1} G_{n+1}(\lambda)$$

and, upon calculating $*$,

$$\{C(\lambda) - C_n(\lambda)\} \lambda F_n^-(\lambda) = \lambda^{n+1} F_{n+1}^-(\lambda) - \{C_n(\lambda) + C_n^\#(\lambda)\} \lambda F_n^-(\lambda).$$

Since $(\theta_{22}^{(n)})^{\pm 1} \in \mathcal{W}_+^{p\times p}$, $(I_p + C_n)^{\pm 1} \in \mathcal{W}_+^{p\times p}$ and $F_n^+ = 2(I_p + C_n)^{-1}\theta_{22}^{(n)}$, it is clear that $(F_n^+)^{\pm 1} \in \mathcal{W}_+^{p\times p}$ and hence

$$C - C_n \in \zeta^{n+1} \mathcal{W}_+^{p\times p}.$$

Thus, (1) holds.

2. Verification of (2). In view of (1),

$$\frac{1}{2\pi} \int_0^{2\pi} e^{ik\theta} \{C(e^{i\theta}) - C_n(e^{i\theta})\} d\theta = 0_{p\times p} \quad \text{if} \quad k \geq -n$$

and

$$\frac{1}{2\pi} \int_0^{2\pi} e^{ik\theta} \{C(e^{i\theta})^* - C_n(e^{i\theta})^*\} d\theta = 0_{p\times p} \quad \text{if} \quad k \leq n.$$

Therefore, both formulas are in force if $|k| \leq n$ and hence

$$\begin{aligned} \frac{1}{2\pi} \int_0^{2\pi} e^{ik\theta} \Delta(e^{i\theta}) d\theta &= \frac{1}{2\pi} \int_0^{2\pi} e^{ik\theta} \left\{ \frac{C(e^{i\theta}) + C(e^{i\theta})^*}{2} \right\} d\theta \\ &= \frac{1}{2\pi} \int_0^{2\pi} e^{ik\theta} \left\{ \frac{C_n(e^{i\theta}) + C_n(e^{i\theta})}{2} \right\} d\theta \quad \text{if} \quad |k| \leq n. \end{aligned}$$

3. Verification of (3). In view of items (4) and (5) of Theorem 5.5,

$$\begin{aligned}
\frac{1}{2\pi}\int_0^{2\pi} & F_k^-(e^{i\theta})^*\Delta(e^{i\theta})F_n^-(e^{i\theta})d\theta \\
&= \frac{1}{2\pi}\int_0^{2\pi} F_k^-(e^{i\theta})^*\left\{\frac{C_n(e^{i\theta})+C_n(e^{i\theta})^*}{2}\right\}F_n^-(e^{i\theta})d\theta \qquad (5.21)\\
&= \frac{1}{2\pi}\int_0^{2\pi} F_k^-(e^{i\theta})^*F_n^-(e^{i\theta})^{-*}d\theta \\
&= \frac{1}{2\pi}\int_0^{2\pi} e^{i(n-k)\theta}\{e^{-ik\theta}F_k^-(e^{i\theta})\}^*\{e^{-in\theta}F_n^-(e^{i\theta})\}^{-*}d\theta \\
&= \begin{cases} 0_{p\times p} & \text{if } k=0,\dots,n-1 \\ I_p & \text{if } k=n. \end{cases}
\end{aligned}$$

This proves (3) for $k \le n$. If $k > n$, then (3) follows from

$$[F_n^-, F_k^-]_\Delta = \{[F_k^-, F_n^-]_\Delta\}^* = 0_{p\times p}.$$

4. Verification of (4). In view of items (3) and (6) of Theorem 5.5,

$$\begin{aligned}
\frac{1}{2\pi}\int_0^{2\pi} & \{e^{-ik\theta}F_k^+(e^{i\theta})\}^*\Delta(e^{i\theta})\{e^{-in\theta}F_n^+(e^{i\theta})\}d\theta \\
&= \frac{1}{2\pi}\int_0^{2\pi} \{e^{-ik\theta}F_k^+(e^{i\theta})\}^*\left\{\frac{C_n(e^{i\theta})+C_n(e^{i\theta})^*}{2}\right\}e^{-in\theta}F_n^+(e^{i\theta})d\theta \\
&= \frac{1}{2\pi}\int_0^{2\pi} e^{i(k-n)\theta}F_k^+(e^{i\theta})^*F_n^+(e^{i\theta})^{-*}d\theta \\
&= \begin{cases} 0_{p\times p} & \text{if } k=0,\dots,n-1 \\ I_p & \text{if } k=n. \end{cases}
\end{aligned}$$

This proves (4) for $0 \le k \le n$. The proof of (4) for $k > n$ follows from

$$[\zeta^{-n}F_n^+, \zeta^{-k}F_k^+]_\Delta = \{[\zeta^{-k}F_k^+, \zeta^{-n}F_n^+]_\Delta\}^* = 0_{p\times p}. \qquad \square$$

Lemma 5.7. *If* $\{\beta_n\}_{n=0}^\infty$ *are the Schur parameters based on a density* Δ *which satisfies* (D1)*, then*

$$F_n^+(\lambda)F_n^+(\lambda)^* - \{\lambda F_n^-(\lambda)\}\{\lambda F_n^-(\lambda)\}^* \succeq 0 \tag{5.22}$$

for all $\lambda \in \overline{\mathbb{D}}$ *and* $n = 0, 1, \dots$ *with equality when* $\lambda \in \mathbb{T}$.

Proof. Assertion (5.22) follows from item (iv) in Theorem 5.2, since

$$\begin{bmatrix}\lambda F_n^-(\lambda) & F_n^+(\lambda)\end{bmatrix} = \begin{bmatrix} I_p & I_p\end{bmatrix}\Theta_n(\lambda). \qquad \square$$

In the following theorem we will make use of the notation

$$\delta_{mn} = \begin{cases} 1 & \text{if } m = n \\ 0 & \text{if } m \neq n. \end{cases}$$

Theorem 5.8. *Suppose Δ is a density which satisfies (D1) and $\{P_n^{\pm}\}_{n=0}^{\infty}$ are sequences of $p \times p$ matrix polynomials which satisfy the following conditions*:

(i) P_n^- *is of degree n for $n = 0, 1, \ldots$.*

(ii) $P_n^+(0)$ *is invertible for $n = 0, 1, \ldots$.*

(iii) $\frac{1}{2\pi}\int_0^{2\pi} P_m^-(e^{i\theta})^* \Delta(e^{i\theta}) P_n^-(e^{i\theta}) d\theta = \delta_{mn} I_p$.

(iv) $\frac{1}{2\pi}\int_0^{2\pi} e^{im\theta} P_m^+(e^{i\theta})^* \Delta(e^{i\theta}) e^{-in\theta} P_n^+(e^{i\theta}) d\theta = \delta_{mn} I_p$.

Then for $n = 0, 1, \ldots$ there exist $p \times p$ unitary matrices M_n and N_n so that

$$P_n^+(\lambda) = E_n^+(\lambda) M_n \tag{5.23}$$

and

$$P_n^-(\lambda) = E_n^-(\lambda) N_n. \tag{5.24}$$

Proof. Since P_n^- is a matrix polynomial of degree n and the matrix coefficient of λ^n is invertible and E_j^- is a matrix polynomial of degree j and the matrix coefficient of λ^j is invertible, it is readily checked that

$$P_n^-(\lambda) = \sum_{j=0}^{n} E_j^-(\lambda) W_{jn},$$

where

$$\begin{aligned} W_{jn} &= \frac{1}{2\pi}\int_0^{2\pi} E_j^-(e^{i\theta})^* \Delta(e^{i\theta}) P_n^-(e^{i\theta}) d\theta \\ &= 0_{p\times p} \end{aligned}$$

for $j = 0, \ldots, n-1$. Thus,

$$P_n^-(\lambda) = E_n^-(\lambda) W_{nn}.$$

Moreover, W_{nn} is unitary since

$$\begin{aligned} I_p &= \frac{1}{2\pi}\int_0^{2\pi} P_n^-(e^{i\theta})^* \Delta(e^{i\theta}) P_n^-(e^{i\theta}) d\theta \\ &= W_{nn}^* \left\{ \int_0^{2\pi} E_n^-(e^{i\theta})^* \Delta(e^{i\theta}) E_n^-(e^{i\theta}) d\theta \right\} W_{nn} \\ &= W_{nn}^* W_{nn}. \end{aligned}$$

Thus, (5.26) holds with $M_n = W_{nn}$ for $n = 0, 1, \ldots$. The formula (5.25) is established in much the same way from the formula

$$\lambda^n (P_n^+)^{\#}(\lambda) = \sum_{j=0}^{n} \lambda^j (E_j^+)^{\#}(\lambda) Z_{jn}. \qquad \square$$

Corollary 5.9. *If $\{E_n^{\pm}\}_{n=0}^{\infty}$ are the matrix polynomials that are defined by (2.1) and (2.2) in terms of the Fourier coefficients of a density Δ that meets the constraint (D1) and $\{F_n^{\pm}\}_{n=0}^{\infty}$ are the matrix polynomials defined by (5.3) in terms of the Schur parameters of Δ, then there exist two sequences of unitary matrices $\{U_n\}_{n=0}^{\infty}$ and $\{V_n\}_{n=0}^{\infty}$ in $\mathbb{C}^{p\times p}$ such that*

$$F_n^+(\lambda) = E_n^+(\lambda)U_n \quad \text{for } n = 0, 1, \dots \tag{5.25}$$

and

$$F_n^-(\lambda) = E_n^-(\lambda)V_n \quad \text{for } n = 0, 1, \dots . \tag{5.26}$$

Moreover,

$$(I_p - \beta_0^*\beta_0)^{-1/2} \cdots (I_p - \beta_n^*\beta_n)^{-1/2} = \{\gamma_{00}^{(n)}\}^{1/2}U_n \tag{5.27}$$

and

$$(I_p - \beta_0\beta_0^*)^{-1/2} \cdots (I_p - \beta_n\beta_n^*)^{-1/2} = \{\gamma_{nn}^{(n)}\}^{1/2}V_n. \tag{5.28}$$

Proof. In view of items (1) and (2) in Theorem 5.5 and items (3) and (4) in Theorem 5.6, the matrix polynomials $\{F_n^{\pm}\}_{n=0}^{\infty}$ satisfy the hypotheses of Theorem 5.8. Thus, assertions (5.25) and (5.26) hold.

Item (ii) of Theorem 4.2 implies that

$$F_n^+(0) = \theta_{22}^{(n)}(0) = (I_p - \beta_0^*\beta_0)^{-1/2} \cdots (I_p - \beta_n^*\beta_n)^{-1/2}.$$

Formula (5.25) implies that $F_n^+(0) = E_n^+(0)U_n = \{\gamma_{00}^{(n)}\}^{1/2}U_n$, whence (5.27) holds. Assertion (5.28) is proved in much the same way. □

Theorem 5.10. *Let Δ and $\widetilde{\Delta}$ be densities which satisfy (D1). If $\{\beta_n\}_{n=0}^{\infty}$ and $\{\tilde{\beta}_n\}_{n=0}^{\infty}$ are the Schur parameters of Δ and $\widetilde{\Delta}$, respectively, and $\beta_n = \tilde{\beta}_n$ for $n = 0, 1, \dots$, then*

$$\Delta(\zeta) = \widetilde{\Delta}(\zeta) \quad \textit{for } \zeta \in \mathbb{T}.$$

Proof. Let $\{F_n^{\pm}\}_{n=0}^{\infty}$ and $\{\widetilde{F}_n^{\pm}\}_{n=0}^{\infty}$ denote the sequences of matrix polynomials given by (5.2) corresponding to the Schur parameters of Δ and $\widetilde{\Delta}$, respectively. If $\beta_n = \tilde{\beta}_n$ for $n = 0, 1, \dots$ and Θ_n and $\widetilde{\Theta}_n$ are defined by (5.2) and correspond to $\{\beta_n\}_{n=0}^{\infty}$ and $\{\tilde{\beta}_n\}_{n=0}^{\infty}$, respectively, then

$$\Theta_n = \widetilde{\Theta}_n \quad \text{for } n = 0, 1, \dots .$$

In view of the recursion (5.4),

$$F_n^+ = \widetilde{F}_n^+ \quad \text{and} \quad F_n^- = \widetilde{F}_n^- \quad \text{for } n = 0, 1, \dots .$$

Consequently,

$$C_n = \widetilde{C}_n \quad \text{for } n = 0, 1, \dots ,$$

where C_n and $\widetilde{C}_n$ are the mvf's defined by (5.12) which correspond to $\{\beta_n\}_{n=0}^{\infty}$ and $\{\tilde{\beta}_n\}_{n=0}^{\infty}$, respectively. In view of item (2) in Theorem 5.6, we have that

$$\Delta_n = \widetilde{\Delta}_n \quad \text{for } n = 0, 1, \dots$$

and hence

$$\Delta(\zeta) = \widetilde{\Delta}(\zeta) \quad \text{for } \zeta \in \mathbb{T}. \qquad \square$$

Theorem 5.11. *If U_n and V_n are as in (5.25) and (5.26), respectively, then*

$$\begin{aligned}\beta_{n+1} &= V_n^*\{\gamma_{nn}^{(n)}\}^{-1/2}\gamma_{n+1,0}^{(n+1)}\{\gamma_{00}^{(n+1)}\}^{-1}\{\gamma_{00}^{(n)}\}^{1/2}U_n\\ &= V_n^*\{\gamma_{nn}^{(n)}\}^{1/2}\{\gamma_{n+1,n+1}^{(n+1)}\}^{-1}\gamma_{n+1,0}^{(n+1)}\{\gamma_{00}^{(n)}\}^{-1/2}U_n, \qquad (5.29)\\ (I_p-\beta_{n+1}^*\beta_{n+1})^{1/2} &= U_{n+1}^*\{\gamma_{00}^{(n+1)}\}^{-1/2}\{\gamma_{00}^{(n)}\}^{1/2}U_n\\ &= U_n^*\{\gamma_{00}^{(n)}\}^{1/2}\{\gamma_{00}^{(n+1)}\}^{-1/2}U_{n+1}, \qquad (5.30)\\ \textit{and}\quad (I_p-\beta_{n+1}\beta_{n+1}^*)^{1/2} &= V_{n+1}^*\{\gamma_{n+1,n+1}^{(n+1)}\}^{-1/2}\{\gamma_{nn}^{(n)}\}^{1/2}V_n\\ &= V_n^*\{\gamma_{nn}^{(n)}\}^{1/2}\{\gamma_{n+1,n+1}^{(n+1)}\}^{-1/2}V_{n+1}. \qquad (5.31)\end{aligned}$$

Proof. In view of formulas (5.25) and (5.26), the recursion

$$\begin{bmatrix}F_{n+1}^-(\lambda) & F_{n+1}^+(\lambda)\end{bmatrix} = \begin{bmatrix}\lambda F_n^-(\lambda) & F_n^+(\lambda)\end{bmatrix}H(\beta_{n+1})$$

can be rewritten as

$$\begin{bmatrix}E_{n+1}^-(\lambda) & E_{n+1}^+(\lambda)\end{bmatrix} = \begin{bmatrix}\lambda E_n^-(\lambda) & E_n^+(\lambda)\end{bmatrix}\begin{bmatrix}V_n & 0\\ 0 & U_n\end{bmatrix}H(\beta_{n+1})\begin{bmatrix}V_{n+1}^* & 0\\ 0 & U_{n+1}^*\end{bmatrix}.$$

Thus, in view of (2.10),

$$\begin{aligned}H(\beta_{n+1}) = &\begin{bmatrix}V_n^* & 0\\ 0 & U_n^*\end{bmatrix}\begin{bmatrix}\gamma_{nn}^{(n)} & 0\\ 0 & \gamma_{00}^{(n)}\end{bmatrix}^{-1/2}\\ &\times\begin{bmatrix}I_p & \gamma_{n+1,0}^{(n+1)}\{\gamma_{00}^{(n+1)}\}^{-1}\\ \gamma_{0,n+1}^{(n+1)}\{\gamma_{n+1,n+1}^{(n+1)}\}^{-1} & I_p\end{bmatrix}\\ &\times\begin{bmatrix}\gamma_{n+1,n+1}^{(n+1)} & 0\\ 0 & \gamma_{00}^{(n+1)}\end{bmatrix}^{1/2}\begin{bmatrix}V_{n+1} & 0\\ 0 & U_{n+1}\end{bmatrix}. \qquad (5.32)\end{aligned}$$

Consequently, both formulas in (5.30) and (5.31) drop out easily from the 11 and 22 blocks of (5.32). Both formulas in (5.29) can be obtained from the 12 and 21 blocks of (5.32) with the help of (5.30) and (5.31). □

6. The reproducing kernel Hilbert space $\mathcal{B}(\mathfrak{F}_n)$

Let $\{F_k^\pm\}_{k=0}^n$ be the matrix polynomials defined by (5.3) in terms of the strict contractions $\beta_0,\ldots,\beta_n\in\mathbb{C}^{p\times p}$. In view of Lemma 5.7, item (3) of Theorem 5.5 and (4.6) the $p\times p$ blocks $F_n^-(\lambda)$ and $F_n^+(\lambda)$ of the $p\times 2p$ mvf

$$\mathfrak{F}_n(\lambda) = \begin{bmatrix}\lambda F_n^-(\lambda) & F_n^+(\lambda)\end{bmatrix}$$

meet the conditions

$$\det F_n^+(\lambda)\neq 0 \quad \text{for } \lambda\in\overline{\mathbb{D}} \qquad (6.1)$$

and

$$(F_n^+)^{-1} F_n^- \in \mathcal{S}_{\rm in}^{p\times p}. \tag{6.2}$$

Thus, in the terminology of [3] adapted to $\mathbb{D}$, $\mathfrak{F}_n(\lambda)$ is a *de Branges matrix* and the space

$$\mathcal{B}(\mathfrak{F}_n) = \{p \times 1 \text{ vvf's } f\colon\, (F_n^+)^{-1} f \in H_2^p \text{ and } (\zeta F_n^-)^{-1} f \in (H_2^p)^\perp\} \tag{6.3}$$

endowed with the inner product

$$\langle f, g\rangle_{\mathcal{B}(\mathfrak{F}_n)} \overset{\text{def}}{=} \frac{1}{2\pi}\int_0^{2\pi} g(e^{i\theta})^*\{F_n^+(e^{i\theta})F_n^+(e^{i\theta})^*\}^{-1} f(e^{i\theta})d\theta \tag{6.4}$$

is a *de Branges space*.

Lemma 6.1. *If* $\{F_k^\pm\}_{k=0}^n$ *are the matrix polynomials defined by* (5.3) *in terms of the strict contractions* $\beta_0, \dots, \beta_n \in \mathbb{C}^{p\times p}$*, then*

$$f \in \mathcal{B}(\mathfrak{F}_n) \iff f \in H_2^p \ominus \zeta^{n+1} H_2^p. \tag{6.5}$$

Proof. By Theorem 5.5, $(F_n^\pm)^{\pm 1} \in \mathcal{W}_+^{p\times p}$ and $(\zeta^{-n} F_n^-)^{\pm 1} \in \mathcal{W}_-^{p\times p}$. Therefore,

$$(F_n^+)^{-1} f \in H_2^p \text{ if and only if } f \in H_2^p$$

and

$$(\zeta F_n^-)^{-1} \in (H_2^p)^\perp \text{ if and only if } \zeta^{-n-1} f \in (H_2^p)^\perp.$$

Thus, (6.5) holds. □

It will be convenient to let $\rho_\omega(\lambda) = 1 - \lambda\overline{\omega}$. This function plays an important role because

$$\frac{I_p}{\rho_\omega(\lambda)} \text{ is a RK (reproducing kernel) for } H_2^p \text{ if } |\omega| < 1.$$

This statement means that

$$\frac{I_p}{\rho_\omega} u = \frac{u}{\rho_\omega} \in H_2^p$$

and

$$\left\langle f, \frac{u}{\rho_\omega}\right\rangle_{\rm st} = u^* f(\omega)$$

for every choice of $u \in \mathbb{C}^p$, $\omega \in \mathbb{D}$ and $f \in H_2^p$. This can be shown by Cauchy's formula. Analogously,

$$-\frac{I_p}{\rho_\omega(\lambda)} \text{ is a RK for } (H_2^p)^\perp \text{ if } |\omega| > 1.$$

Theorem 6.2. *If $\{F_k^{\pm}\}_{k=0}^n$ are the matrix polynomials defined by (5.3) in terms of the strict contractions $\beta_0, \dots, \beta_n \in \mathbb{C}^{p\times p}$, then $\mathcal{B}(\mathfrak{F}_n)$ is a RKHS with RK*

$$K_\omega^n(\lambda) = -\frac{\mathfrak{F}_n(\lambda) j_p \mathfrak{F}_n(\omega)^*}{\rho_\omega(\lambda)} = \begin{cases} \dfrac{F_n^+(\lambda)F_n^+(\omega)^* - \lambda\overline{\omega}F_n^-(\lambda)F_n^-(\omega)^*}{\rho_\omega(\lambda)} & \text{if } \lambda\overline{\omega} \neq 1 \\ \zeta\{(F_n^-)'(\zeta)F_n^-(\zeta)^* - (F_n^+)'(\zeta)F_n^+(\zeta)^* + \bar{\zeta}F_n^-(\zeta)F_n^-(\zeta)^*\} & \text{if } \zeta \in \mathbb{T}, \end{cases} \tag{6.6}$$

where $\zeta = \omega = \lambda$.

Proof. There are two facts to verify for every choice of $\omega \in \mathbb{C}$, $u \in \mathbb{C}^p$ and $f \in \mathcal{B}(\mathfrak{F}_n)$:

$$K_\omega^n u \in \mathcal{B}(\mathfrak{F}_n) \tag{6.7}$$

and

$$\langle f, K_\omega^n u\rangle_{\mathcal{B}(\mathfrak{F}_n)} = u^* f(\omega). \tag{6.8}$$

The justification is broken into steps:

1. Verification of (6.7). Since $K_0^n(\lambda) = F_n^+(\lambda)F_n^+(0)^*$, the assertion is clear if $\omega = 0$. The proof for $\omega \neq 0$ rests on the identity

$$F_n^+(\zeta)F_n^+(\zeta)^* = F_n^-(\zeta)F_n^-(\zeta)^* \quad \text{for } \zeta \in \mathbb{T}, \tag{6.9}$$

which extends to

$$F_n^+(\lambda)(F_n^+)^\#(\lambda) = F_n^-(\lambda)(F_n^-)^\#(\lambda) \quad \text{for } \lambda \neq 0, \tag{6.10}$$

or, equivalently to

$$F_n^+(1/\overline{\omega})F_n^+(\omega)^* = F_n^-(1/\overline{\omega})F_n^-(\omega)^* \quad \text{for } \omega \neq 0. \tag{6.11}$$

Thus,

$$K_\omega^n(\lambda) = \frac{\{F_n^+(\lambda) - F_n^+(1/\overline{\omega})\}F_n^+(\omega)^* + \{F_n^-(1/\overline{\omega}) - \lambda\overline{\omega}F_n^-(\lambda)\}F_n^-(\omega)^*}{\rho_\omega(\lambda)}$$

is a matrix polynomial of degree at most n if $\omega \neq 0$.

2. Verification of (6.8) *when* $|\omega| < 1$. In view of (6.4),

$$\begin{aligned}\langle f, K_\omega^n u\rangle_{\mathcal{B}(\mathfrak{F}_n)} &= \langle (F_n^+)^{-1}f, (F_n^+)^{-1}K_\omega^n u\rangle_{\mathrm{st}} \\ &= \left\langle (F_n^+)^{-1}f, \frac{F_n^+(\omega)^*}{\rho_\omega}u - \frac{(F_n^+)^{-1}(\zeta\overline{\omega})F_n^-F_n^-(\omega)^*}{\rho_\omega}u\right\rangle_{\mathrm{st}} \\ &= \left\langle (F_n^+)^{-1}f, \frac{F_n^+(\omega)^*}{\rho_\omega}u\right\rangle_{\mathrm{st}} - \left\langle (\zeta F_n^-)^{-1}f, \frac{\overline{\omega}}{\rho_\omega}F_n^-(\omega)^*u\right\rangle_{\mathrm{st}}.\end{aligned}$$

The second inner product is equal to zero since $u/\rho_\omega \in H_2^p$ if $\omega \in \mathbb{D}$ and $(\zeta F_n^-)^{-1}f \in (H_2^p)^\perp$.

This completes the proof, since

$$\begin{aligned}\left\langle (F_n^+)^{-1} f, \frac{F_n^+(\omega)^*}{\rho_\omega} u \right\rangle_{\mathrm{st}} &= u^* F_n^+(\omega) F_n^+(\omega)^{-1} f(\omega)\\ &= u^* f(\omega),\end{aligned}$$

because I_p/ρ_ω is a RK for H_2^p if $\omega \in \mathbb{D}$.

3. Verification of (6.8) *when* $|\omega| > 1$. If $|\omega| > 1$ and $u \in \mathbb{C}^p$, then

$$\left\langle (F_n^+)^{-1} f, \frac{F_n^+(\omega)^*}{\rho_\omega} u \right\rangle_{\mathrm{st}} = 0.$$

Thus,

$$\begin{aligned}\langle f, K_\omega^n u\rangle_{\mathcal{B}(\mathfrak{F}_n)} &= -\left\langle (\zeta F_n^-)^{-1} f, \frac{\overline{\omega}}{\rho_\omega} F_n^-(\omega)^* u \right\rangle_{\mathrm{st}}\\ &= u^* f(\omega),\end{aligned}$$

since $-I_p/\rho_\omega$ is a RK for $(H_2^p)^\perp$ if $|\omega| > 1$.

4. Verification of (6.8) *when* $|\omega| = 1$. Given $\omega \in \mathbb{T}$, we can construct a sequence $\{\omega_k\}_{k=0}^\infty$, with $|\omega_k| > 1$ for $k = 0, 1, \ldots$ and

$$\lim_{k\uparrow\infty} \omega_k = \omega.$$

If $u \in \mathbb{C}^p$, then using Step 3 we have

$$u^* f(\omega_k) = \langle f, K_{\omega_k}^n u\rangle_{\mathcal{B}(\mathfrak{F}_n)} \quad \text{for } k = 0, 1, \ldots.$$

Thus, as f is a vector polynomial,

$$u^* f(\omega) = \lim_{k\uparrow\infty} u^* f(\omega_k) = \lim_{k\uparrow\infty} \langle f, K_{\omega_k}^n\rangle_{\mathcal{B}(\mathfrak{F}_n)} = \langle f, K_\omega^n\rangle_{\mathcal{B}(\mathfrak{F}_n)}. \qquad \square$$

Theorem 6.3. *If $\{F_n^\pm\}_{n=0}^\infty$ and $\{E_n^\pm\}_{n=0}^\infty$ are the matrix polynomials defined by* (5.3) *in terms of the Schur parameters $\{\beta_n\}_{n=0}^\infty$ and the Fourier coefficients, respectively, of a density Δ which satisfies* (D1) *and* (D2)*, then*:

(1) *The sequence of spaces $\{\mathcal{B}(\mathfrak{F}_n)\}_{n=0}^\infty$ is ordered by inclusion, i.e.,*

$$\mathcal{B}(\mathfrak{F}_n) \subseteq \mathcal{B}(\mathfrak{F}_{n+1}) \subseteq L_2^p(\mathbb{T}, \Delta) \quad \textit{for } n = 0, 1, \ldots \tag{6.12}$$

and the inclusions are isometries.

(2) *The orthogonal projection $P_{\mathcal{B}(\mathfrak{F}_n)}$ of $\rho_\omega^{-1} Q^{-1} Q(\omega)^* u$ onto $\mathcal{B}(\mathfrak{F}_n)$ is*

$$P_{\mathcal{B}(\mathfrak{F}_n)} \frac{Q^{-1} Q(\omega)^{-*}}{\rho_\omega} u = K_\omega^n u$$

for $n = 0, 1, \ldots$ and $\omega \in \mathbb{D}$.

(3) *For every choice of* $\omega \in \mathbb{D}$ *and* $u \in \mathbb{C}^p$

$$\kappa_n \overset{\text{def}}{=} \left\| \frac{Q^{-1}Q(\omega)^{-*}}{\rho_\omega} u - K_\omega^n u \right\|_\Delta^2 = u^* \frac{Q(\omega)^{-1}Q(\omega)^{-*}}{\rho_\omega(\omega)} u - u^* K_\omega^n(\omega) u. \quad (6.13)$$

(4) $\kappa_n \to 0$ *as* $n \uparrow \infty$.

Proof. The proof is broken into steps.

1. Verification of (1). Assertion (1) is clear from item (2) of Theorem 5.6 and the characterization of $\mathcal{B}(\mathfrak{F}_n)$ as a $p \times 1$ vector polynomials of degree at most n with inner product

$$\langle f, g \rangle_{\mathcal{B}(\mathfrak{F}_n)} = \frac{1}{2\pi} \int_0^{2\pi} g(e^{i\theta})^* \Delta(e^{i\theta}) f(e^{i\theta}) d\theta.$$

2. Verification of (2). If $\alpha, \omega \in \mathbb{D}$ and $u, v \in \mathbb{C}^p$, then

$$\begin{aligned} v^* \left(P_{\mathcal{B}(\mathfrak{F}_n)} \frac{Q^{-1}Q(\omega)^*}{\rho_\omega} u \right)(\alpha) &= \left\langle \frac{Q^{-1}Q(\omega)^{-*}}{\rho_\omega} u, K_\alpha^n v \right\rangle_\Delta \\ &= \left\langle \frac{Q(\omega)^{-*}}{\rho_\omega} u, Q K_\alpha^n v \right\rangle_{\text{st}} \\ &= \overline{\left\langle Q K_\alpha^n v, \frac{Q(\omega)^{-*}}{\rho_\omega} u \right\rangle_{\text{st}}} \\ &= \left\{ u^* Q(\omega)^{-1} Q(\omega) K_\alpha^n(\omega) v \right\}^* = v^* K_\omega^n(\alpha) u. \end{aligned}$$

Since both sides are polynomials the equality is valid for every point $\alpha \in \mathbb{C}$.

3. Verification of (3). If $u \in \mathbb{C}^p$, $\omega \in \mathbb{D}$ $f = \rho_\omega^{-1} Q^{-1} Q(\omega)^{-*} u$ and $P \overset{\text{def}}{=} P_{\mathcal{B}(\mathfrak{F}_n)}$, then, since P is an orthogonal projection,

$$\begin{aligned} \left\| \frac{Q^{-1}Q(\omega)^{-*}}{\rho_\omega} u - K_\omega^n u \right\|_\Delta^2 &= \|(I-P)f\|_\Delta^2 = \|f\|_\Delta^2 - \|Pf\|_\Delta^2 \\ &= \frac{u^* Q(\omega)^{-1} Q(\omega)^{-*} u}{1 - \rho_\omega(\omega)} - u^* K_\omega(\omega) u. \end{aligned}$$

4. Verification of (4). Let

$$\mathcal{Q}_\omega(\lambda) = \frac{Q(\lambda)^{-1} Q(\omega)^{-*}}{\rho_\omega(\lambda)}.$$

If $\omega \in \mathbb{D}$ and $u \in \mathbb{C}^p$, then

$$\mathcal{Q}_\omega(\zeta) u = \sum_{j=0}^{\infty} \zeta^j \xi_j \quad \text{where} \quad \sum_{j=0}^{\infty} \|\xi_j\| < \infty.$$

Let $f_n(\zeta) = \sum_{j=0}^{j} \zeta^j \xi_j$. In view of (2),

$$\|\mathcal{Q}_\omega u - K_\omega^n u\|_\Delta^2 = \min_{f\in\mathcal{B}(\mathfrak{F}_n)} \|\mathcal{Q}_\omega u - f\|_\Delta^2.$$

Therefore,

$$\begin{aligned}\|\mathcal{Q}_\omega u - K_\omega^n u\|_\Delta^2 &\le \|\mathcal{Q}_\omega u - f_n\|_\Delta^2 \le \kappa\|\mathcal{Q}_\omega u - f_n\|_{st}^2 \\ &= \kappa \sum_{j=n+1}^{\infty} \|\xi_j\|^2 \to 0 \quad \text{as } n\uparrow\infty. \qquad \square\end{aligned}$$

Let

$$D_n(\lambda) = \frac{C_n(\lambda) + C_n(\lambda)^*}{2}$$

and

$$\langle f, g\rangle_{D_n} = \frac{1}{2\pi}\int_0^{2\pi} g(e^{i\theta})^* D_n(e^{i\theta}) f(e^{i\theta}) d\theta$$

for $f, g \in L_2^p(\mathbb{T}, \Delta)$.

Theorem 6.4. *If $\{F_n^\pm\}_{n=0}^\infty$ and $\{E_n^\pm\}_{n=0}^\infty$ are matrix polynomials defined in terms of the $\{\beta_n\}_{n=0}^\infty$ the Schur parameters and the Fourier coefficients, respectively, of a density Δ that satisfies* (D1) *and* (D2)*, then*:

(1) *The $p\times 2p$ mvf*

$$\mathfrak{E}_n(\lambda) = \begin{bmatrix}\lambda E_n^-(\lambda) & E_n^+(\lambda)\end{bmatrix},$$

is a de Branges matrix,

$$\mathcal{B}(\mathfrak{E}_n) = \mathcal{B}(\mathfrak{F}_n) \quad \textit{for } n = 0, 1, \ldots$$

and

$$\begin{aligned}K_\omega^n(\lambda) &= -\frac{\mathfrak{E}_n(\lambda) j_p \mathfrak{E}_n(\omega)^*}{\rho_\omega(\lambda)} \\ &= \begin{cases}\dfrac{E_n^+(\lambda)E_n^+(\omega)^* - \lambda\overline{\omega}E_n^-(\lambda)E_n^-(\omega)^*}{\rho_\omega(\lambda)} & \textit{if } \lambda\overline{\omega}\neq 1 \\ \zeta\{(E_n^-)'(\zeta)E_n^-(\zeta)^* - (E_n^+)'(\zeta)E_n^+(\zeta)^* + \bar{\zeta}E_n^-(\zeta)E_n^-(\zeta)^*\}, \end{cases}\end{aligned} \tag{6.14}$$

where $\zeta = \lambda = \omega \in \mathbb{T}$.

(2) $\langle f, f\rangle_{\mathcal{B}(\mathfrak{E}_n)} = \langle f, f\rangle_{\mathcal{B}(\mathfrak{F}_n)} = \langle f, f\rangle_{D_n} = \langle f, f\rangle_\Delta$ *for* $n = 0, 1, \ldots$.

(3) *The RK $K_\omega^n(\lambda)$ can also be expressed as*

$$K_\omega^n(\lambda) = \sum_{j,k=0}^{n} \lambda^j \gamma_{jk}^{(n)} \overline{\omega}^k \tag{6.15}$$

and

$$K_\omega^n(\lambda) = \sum_{j=0}^{n} E_j^-(\lambda) E_j^-(\omega)^*. \tag{6.16}$$

Proof. The proof of (1) and (2) is immediate from (6.6) using the identities (5.25) and (5.26).

In order to show (6.15), it suffices to check that $Z^n_\omega(\lambda) = \lambda^j \gamma^{(n)}_{jk} \overline{\omega}^k$ is a RK for $\mathcal{B}(\mathfrak{E}_n) = \mathcal{B}(\mathfrak{F}_n)$. In view of (6.5), Z^n_ω clearly belongs to $\mathcal{B}(\mathfrak{E}_n)$. If $u \in \mathbb{C}^p$ and $f(\lambda) = \lambda^m f_m$ for $0 \le m \le n$, then, in view of item (6) of Theorem 5.5 and item (2) of Theorem 5.6,

$$
\begin{aligned}
\langle f, Z^n_\omega u\rangle_{\mathcal{B}(\mathfrak{E}_n)} &= \frac{1}{2\pi}\int_0^{2\pi} u^* Z^n_\omega(e^{i\theta})^* E^+_n(e^{i\theta})^{-*} E^+_n(e^{i\theta})^{-1}\{e^{im\theta} f_n\} d\theta \\
&= \frac{1}{2\pi}\int_0^{2\pi} u^* Z^n_\omega(e^{i\theta})^* \left\{\frac{C_n(e^{i\theta}) + C_n(e^{i\theta})^*}{2}\right\}\{e^{im\theta} f_n\} d\theta \\
&= \frac{1}{2\pi}\int_0^{2\pi} u^* Z^n_\omega(e^{i\theta})^* \Delta(e^{i\theta})\{e^{im\theta} f_n\} d\theta \\
&= \sum_{j,k=0}^{n} \omega^k u^* \gamma^{(n)}_{kj}\left\{\frac{1}{2\pi}\int_0^{2\pi} e^{i(m-j)\theta}\Delta(e^{i\theta}) d\theta\right\} f_m \\
&= u^*\left(\sum_{k=0}^{n}\begin{bmatrix}\gamma^{(n)}_{k0} & \cdots & \gamma^{(n)}_{kn}\end{bmatrix}\begin{bmatrix}\Delta_{-m}\\ \vdots \\ \Delta_{n-m}\end{bmatrix}\right) f_m \\
&= u^*\omega^m f_m \\
&= u^* f(\omega).
\end{aligned}
$$

Since $\langle\cdot,\cdot\rangle$ is linear in the first argument,

$$
u^* f(\omega) = \langle f, Z^n_\omega u\rangle_{\mathcal{B}(\mathfrak{E}_n)} \quad \text{for} \quad f \in \mathcal{B}(\mathfrak{E}_n).
$$

Thus, $Z^n_\omega(\lambda)$ is a RK for $\mathcal{B}(\mathfrak{E}_n)$. Since there is only one RK for a RKHS, (6.15) holds.

We will now show formula (6.16). Since E^-_j is a matrix polynomial of degree j with invertible top coefficient, there exist matrices $A_1, \dots, A_n$ belonging to $\mathbb{C}^{p\times p}$ such that

$$
K^n_\omega(\lambda) = \sum_{j=0}^{n} E^-_j(\lambda) A_j.
$$

Thus, as

$$
u^* E^-_k(\omega) = \langle E^-_k, K^n_\omega u\rangle_{\mathcal{B}(\mathfrak{E}_n)} \quad \text{for } u \in \mathbb{C}^p,
$$

(2.4) can be used to check that $A_j = E^-_j(\omega)^*$ for $j = 0, \dots, n$. □

Corollary 6.5. *If Δ is a density that satisfies* (D1) *and* (D2)*, then*

$$\gamma_{jk}^{(n)} = \gamma_{jn}^{(n)}\{\gamma_{nn}^{(n)}\}^{-1}\gamma_{nk}^{(n)} \quad \textit{for } j,k = 0,\dots,n, \tag{6.17}$$

$$0 \prec \gamma_{00}^{(n-1)} = \gamma_{00}^{(n)} - \gamma_{0n}^{(n)}\{\gamma_{nn}^{(n)}\}^{-1}\gamma_{n0}^{(n)} \prec \gamma_{00}^{(n)} \prec Q(0)^{-1}Q(0)^{-*} \tag{6.18}$$

and

$$0 \prec \gamma_{n-1,n-1}^{(n-1)} = \gamma_{nn}^{(n)} - \gamma_{n0}^{(n)}\{\gamma_{00}^{(n)}\}^{-1}\gamma_{0n}^{(n)} \prec \gamma_{nn}^{(n)} \prec R(0)^{-*}R(0)^{-1} \tag{6.19}$$

for $n = 1,2,\dots$.

Proof. The identity (6.17) follows readily from comparing the expressions for $K_\omega^n(\lambda)$ given in (6.15) and (6.16) since $E_j^-(\lambda) = \sum_{m=0}^{j}\lambda^m\gamma_{mj}^{(j)}\{\gamma_{jj}\}^{-1/2}$.

In view of formulas (6.13), (6.14) and (6.16),

$$\begin{aligned} 0 \prec E_n^-(0)E_n^-(0)^* &= K_0^n(0) - K_0^{n-1}(0) \\ &= E_n^+(0)E_n^+(0)^* - E_{n-1}^+(0)E_{n-1}^+(0)^* \prec Q(0)^{-1}Q(0)^{-*} \end{aligned}$$

for $n = 1,2,\dots$. The statements in (6.18) follow easily, since $E_n^+(0) = \{\gamma_{00}^{(n)}\}^{1/2}$ and $E_n^-(0) = \gamma_{0n}^{(n)}\{\gamma_{nn}^{(n)}\}^{-1/2}$.

The statements in (6.19) follow by applying (6.18) to the reverse polynomials $\widetilde{E}_n^+$, $\widetilde{E}_n^-$ and the identity

$$\widetilde{\gamma}_{jk}^{(n)} = \gamma_{n-j,n-k}^{(n)} \quad \text{for } 0 \le j,k \le n. \qquad \square$$

Theorem 6.6. *The Schur parameters $\{\beta_n\}_{n=0}^\infty$ corresponding to a density Δ that meets the constraints* (D1) *and* (D2) *are subject to the bounds*

$$\|\beta_n\| \le \|\gamma_{0n}^{(n)}\{\gamma_{nn}^{(n)}\}^{-1/2}\| \quad \textit{for } n = 1,2,\dots \tag{6.20}$$

and

$$\sum_{j=n}^{\infty}\|\beta_j\|^2 \le \operatorname{trace}\{Q(0)^{-1}Q(0)^{-*} - \gamma_{00}^{(n-1)}\} \quad \textit{for } n = 1,2,\dots. \tag{6.21}$$

Proof. Since U_n is unitary, formulas (5.7) and (5.27) imply that

$$\beta_n^*\beta_n = Y_n^{-1}\left\{\gamma_{00}^{(n)} - \gamma_{00}^{(n-1)}\right\}Y_n^{-*}$$

and hence, with the help of (6.18), that

$$\begin{aligned} \|\beta_n\|^2 = \|\beta_n^*\beta_n\| &\le \|Y_n^{-1}\|\,\|\gamma_{00}^{(n)} - \gamma_{00}^{(n-1)}\|\,\|Y_n^{-*}\| \le \|\gamma_{00}^{(n)} - \gamma_{00}^{(n-1)}\|. \\ &= \|\gamma_{0n}^{(n)}\{\gamma_{nn}^{(n)}\}^{-1}\gamma_{n0}^{(n)}\| = \|\gamma_{0n}^{(n)}\{\gamma_{nn}^{(n)}\}^{-1/2}\|^2. \end{aligned} \tag{6.22}$$

The inequality (6.21) is obtained from the preceding sequence of inequalities by noting that

$$\|\gamma_{00}^{(n)} - \gamma_{00}^{(n-1)}\| \le \operatorname{trace}\{\gamma_{00}^{(n)} - \gamma_{00}^{(n-1)}\},$$

since $\gamma_{00}^{(n)} - \gamma_{00}^{(n-1)} \succ 0$, and hence that

$$\sum_{j=n}^{n+k}\|\beta_j\|^2 \le \operatorname{trace}\{\gamma_{00}^{(n+k)} - \gamma_{00}^{(n-1)}\} \le \operatorname{trace}\{Q(0)^{-1}Q(0)^{-*} - \gamma_{00}^{(n-1)}\}. \qquad \square$$

7. CMV matrices

In this section, we will show how to generate a unitary operator $\mathfrak{A}$ on ℓ_2^p that has a factorization in terms of a unitary matrix $\beta_{-1} \in \mathbb{C}^{p\times p}$ and the Schur parameters $\{\beta_n\}_{n=0}^{\infty}$ of a density Δ that satisfies (D1). If $\beta_{-1} = I_p$, then $\mathfrak{A}$ is the matrix representation of the operator of multiplication by ζ in $L_2^{p\times p}(\mathbb{T}, \Delta)$ with respect to an orthonormal basis that will be constructed in terms of the matrix polynomials $\{F_n^{\pm}\}_{n=0}^{\infty}$ that are defined in terms of the Schur parameters $\{\beta_n\}_{n=0}^{\infty}$ of Δ. This matrix is completely specified by the Schur parameters of Δ. If $\beta_{-1} = I_p$ and $p = 1$, then this construction is due to Cantero, Moral and Velásquez [8]. The case when $\beta_{-1} = I_p$ and $p > 1$ was considered first in Simon [32] (see also Simon, Damanik and Pushnitski [9]).

Let $\{\Psi_n\}_{n=0}^{\infty}$ be a sequence of mvf's belonging to $L_2^{p\times p}(\mathbb{T}, \Delta)$ and $\{A_n\}_{n=0}^{\infty}$ be a sequence of matrices belonging to $\mathbb{C}^{p\times p}$. We will write

$$F(\zeta) = \sum_{n=0}^{\infty} \Psi_n(\zeta) A_n \quad \text{for } F \in L_2^{p\times p}(\mathbb{T}, \Delta)$$

if

$$\lim_{n\uparrow\infty} \left[F - \sum_{j=0}^{n} \Psi_j A_j, F - \sum_{j=0}^{n} \Psi_j A_j \right]_{\Delta} = 0_{p\times p}, \tag{7.1}$$

i.e.,

$$\lim_{n\uparrow\infty} \int_0^{2\pi} \left\{ F(e^{i\theta}) - \sum_{j=0}^{n} \Psi_j(e^{i\theta}) A_j \right\}^* \Delta(e^{i\theta}) \{F(e^{i\theta}) - \sum_{j=0}^{n} \Psi_j(e^{i\theta}) A_j\} d\theta = 0_{p\times p}.$$

Definition 7.1. A sequence of $p \times p$ mvf's $\{\Psi_n\}_{n=0}^{\infty}$ in $L_2^{p\times p}(\mathbb{T}, \Delta)$ will be called an "orthonormal basis" for $L_2^{p\times p}(\mathbb{T}, \Delta)$ if:

(i) $[\Psi_m, \Psi_n]_{\Delta} = \begin{cases} I_p & \text{if } m = n \\ 0_{p\times p} & \text{if } m \neq n. \end{cases}$

(ii) There exists a sequence $\{A_n\}_{n=0}^{\infty}$ of $p \times p$ matrices such that (7.1) holds for each $F \in L_2^{p\times p}(\mathbb{T}, \Delta)$.

Let $\{F_n^{\pm}\}_{n=0}^{\infty}$ denote the matrix polynomials given by (5.3) that are defined in terms of the Schur parameters $\{\beta_n\}_{n=0}^{\infty}$ of a density Δ which meets the constraint (D1) and set

$$\begin{bmatrix} \chi_{2k}(\zeta) \\ \chi_{2k+1}(\zeta) \end{bmatrix} = \zeta^{-k} \begin{bmatrix} F_{2k}^{+}(\zeta) \\ F_{2k+1}^{-}(\zeta) \end{bmatrix} \quad \text{for } k = 0, 1, \ldots \tag{7.2}$$

and

$$\begin{bmatrix} y_{2k}(\zeta) \\ y_{2k+1}(\zeta) \end{bmatrix} = \begin{bmatrix} \zeta^{-k} F_{2k}^{-}(\zeta) \\ \zeta^{-k-1} F_{2k+1}^{+}(\zeta) \end{bmatrix} \quad \text{for } k = 0, 1, \ldots. \tag{7.3}$$

Definition 7.2. Let Δ be a density which meets the constraint (D1) and $\beta_{-1} \in \mathbb{C}^{p\times p}$ be unitary. The *CMV matrix based on* Δ *and* β_{-1} is the operator $\mathfrak{A} : \ell_2^p \to \ell_2^p$ given by

$$\mathrm{e}_m^* \mathfrak{A}\, \mathrm{e}_n = \begin{cases} [\zeta\chi_n, \chi_m]_{\Delta}\beta_{-1} & \text{if } m = 0, 1, \ldots \text{ and } n = 0 \\ [\zeta\chi_n, \chi_m]_{\Delta} & \text{if } m = 0, 1, \ldots \text{ and } n = 1, 2, \ldots. \end{cases} \tag{7.4}$$

If $\beta_{-1} = I_p$ in Definition 7.2, then

$$\mathfrak{A} = V^{-1} M_\zeta V, \tag{7.5}$$

where $V : \ell_2^p \to L_2^{p\times p}(\mathbb{T}, \Delta)$ is given by $V\,\mathrm{e}_n = \chi_n$ and M_ζ denotes the operator of multiplication by ζ in $L_2^{p\times p}(\mathbb{T}, \Delta)$.

Theorem 7.3. *If $\{F_n^\pm\}_{n=0}^\infty$ are the matrix polynomials generated by (5.3) in terms of the Schur parameters of a density Δ that satisfies (D1) and $\{\chi_n\}_{n=0}^\infty$ and $\{y_n\}_{n=0}^\infty$ are given by (7.2) and (7.3), respectively, then*:

(i) $\{\chi_n\}_{n=0}^\infty$ *is an "orthonormal basis" of $L_2^{p\times p}(\mathbb{T}, \Delta)$.*

(ii) $\{y_n\}_{n=0}^\infty$ *is an "orthonormal basis" of $L_2^{p\times p}(\mathbb{T}, \Delta)$.*

Proof. The proof of (i) is broken into steps. The proof of (ii) is similar.

1. Verification of the orthonormality of $\{\chi_n\}_{n=0}^\infty$. If $m = n$, then, it follows readily from items (3) and (4) in Theorem 5.6,

$$\begin{aligned}[\chi_m, \chi_m]_\Delta &= \begin{cases}[F_m^+, F_m^+]_\Delta & \text{if } m \text{ is even}\\ [F_m^-, F_m^-]_\Delta & \text{if } m \text{ is odd}\end{cases}\\ &= I_p.\end{aligned}$$

If $m \neq n$, then we may assume, without loss of generality, that $m > n$. It follows from (2.8) and (5.26) that

$$[F_m^-, \zeta^k I_p]_\Delta = 0_{p\times p} \quad \text{for } k = 0, \dots, m-1.$$

If $m = 2j + 1$, then

$$[\chi_{2j+1}, \zeta^{k-j} I_p]_\Delta = [\zeta^{-j} F_{2j+1}^-, \zeta^{k-j} I_p]_\Delta = 0_{p\times p} \quad \text{for } k = 0, \dots, 2j,$$

i.e.,

$$[\chi_{2j+1}, \zeta^i I_p]_\Delta = 0_{p\times p} \quad \text{for } -j \le i \le j.$$

Therefore,

$$[\chi_{2j+1}, \chi_i]_\Delta = 0_{p\times p} \quad \text{for } i = 0, \dots, 2j,$$

i.e.,

$$[\chi_m, \chi_n]_\Delta = 0_{p\times p} \quad \text{for } m > n \text{ when } m \text{ is odd.}$$

It follows from (2.6) and (5.25) that

$$[F_m^+, \zeta^{m-k} I_p]_\Delta = 0_{p\times p} \quad \text{for } k = 0, \dots, m-1.$$

Thus, if $m = 2j$ and $j > 0$, then

$$[\chi_{2j}, \zeta^{j-k} I_p]_\Delta = [\zeta^{-j} F_{2j}^+, \zeta^{j-k} I_p]_\Delta = 0_{p\times p} \quad \text{for } k = 0, \dots, 2j-1.$$

Therefore,

$$[\chi_{2j}, \chi_i]_\Delta = 0_{p\times p} \quad \text{for } i = 0, \dots, 2j-1,$$

i.e.,

$$[\chi_m, \chi_n]_\Delta = 0_{p\times p} \quad \text{for } m > n \text{ when } m \text{ is even.}$$

2. *$\{\chi_n\}_{n=0}^{\infty}$ is a basis for $L_2^{p\times p}(\mathbb{T},\Delta)$.* It follows from (5.25) and (2.1) that

$$F_k^+(\zeta) = A_k + \text{matrix linear combination}\{\zeta,\dots,\zeta^k\}, \tag{7.6}$$

where $A_k \in \mathbb{C}^{p\times p}$ is invertible. Similarly, by using (5.26) and (2.2),

$$F_k^-(\zeta) = \zeta^k B_k + \text{matrix linear combination}\{1,\dots,\zeta^{k-1}\}. \tag{7.7}$$

The proof that item (ii) in Definition 7.1 holds follows easily from (7.6), (7.7) and the fact that mvf's of the form $\sum_{k=-n}^{n}\zeta^k A_k$ are dense in $L_2^{p\times p}(\mathbb{T},\Delta)$. □

Corollary 7.4. *If Δ is a density that satisfies (D1) and $F \in L_2^{p\times p}(\mathbb{T},\Delta)$, then*

$$F(\zeta) = \sum_{n=0}^{\infty}\chi_n(\zeta)[F,\chi_n]_\Delta = \sum_{n=0}^{\infty} y_n(\zeta)[F,y_n]_\Delta. \tag{7.8}$$

Proof. Both formulas in (7.8) follow immediately from Theorem 7.3. □

Lemma 7.5. *If Δ is a density that satisfies (D1), then*

$$\zeta\chi_n(\zeta) = \sum_{m=0}^{\infty}\chi_m(\zeta)\mathrm{e}_m^*\mathfrak{A}\mathfrak{B}^*\mathrm{e}_n \quad \textit{for } \zeta\in\mathbb{T}, \tag{7.9}$$

where

$$\mathfrak{B} = \begin{bmatrix} \beta_{-1} & 0 & 0 & \cdots \\ 0 & I_p & 0 & \cdots \\ 0 & 0 & I_p & \\ \vdots & \vdots & & \ddots \end{bmatrix}. \tag{7.10}$$

Proof. By (7.8), the coefficient in the expansion

$$\zeta\chi_n(\zeta) = \sum_{m=0}^{\infty}\chi_m(\zeta)[\zeta\chi_n,\chi_m]_\Delta$$

can be evaluated as

$$[\zeta\chi_n,\chi_m]_\Delta = e_m^*\mathfrak{A}\mathfrak{B}^*e_n.$$ □

It will be convenient to introduce the unitary matrices

$$A_n = \begin{bmatrix} -\beta_n^* & (I_p-\beta_n^*\beta_n)^{1/2} \\ (I_p-\beta_n\beta_n^*)^{1/2} & \beta_n \end{bmatrix} \quad \text{for } n = 0,1,\dots. \tag{7.11}$$

Theorem 7.6. *The CMV matrix $\mathfrak{A}$ based on a density Δ that satisfies* (D1) *and a unitary matrix $\beta_{-1} \in \mathbb{C}^{p\times p}$ admits the factorization*:

$$\mathfrak{A} = \mathfrak{A}_{\mathrm{odd}}\mathfrak{A}_{\mathrm{even}}\mathfrak{B}, \tag{7.12}$$

where

$$\mathfrak{A}_{\mathrm{odd}} = \begin{bmatrix} A_1 & 0 & 0 & \cdots \\ 0 & A_3 & 0 & \cdots \\ 0 & 0 & A_5 & \\ \vdots & \vdots & & \ddots \end{bmatrix}, \quad \mathfrak{A}_{\mathrm{even}} = \begin{bmatrix} I_p & 0 & 0 & \cdots \\ 0 & A_2 & 0 & \cdots \\ 0 & 0 & A_4 & \\ \vdots & \vdots & & \ddots \end{bmatrix}, \tag{7.13}$$

$\mathfrak{B}$ is given by (7.10) *and the blocks $\{A_n\}_{n=0}^{\infty}$ are defined in terms of the Schur parameters of Δ by* (7.11).

Proof. It follows from (7.4) that

$$\mathrm{e}_m^* \,\mathfrak{A}\mathfrak{B}^* \,\mathrm{e}_n = [\zeta\chi_n, \chi_m]_\Delta.$$

In view of Corollary 7.4, the mvf χ_n can be expressed as

$$\chi_n(\zeta) = \sum_{k=0}^{\infty} y_k(\zeta)P_k,$$

where

$$P_k = [\chi_n, y_k]_\Delta \quad \text{for } k = 0, 1, \dots,$$

and hence

$$\begin{aligned} [\zeta\chi_n, \chi_m]_\Delta &= \sum_{k=0}^{\infty} [\zeta y_k, \chi_m]_\Delta P_k \\ &= \sum_{k=0}^{\infty} [\zeta y_k, \chi_m]_\Delta [\chi_n, y_k]_\Delta. \end{aligned} \tag{7.14}$$

The rest of the proof is broken into steps and is devoted to showing

$$\begin{aligned} \mathrm{e}_m^* \,\mathfrak{A}_{\mathrm{odd}} \,\mathrm{e}_k &\overset{\mathrm{def}}{=} [\zeta y_k, \chi_m]_\Delta \\ &= \begin{cases} -\beta_{2n+1}^* & \text{if } k = m = 2n \\ \beta_{2n+1} & \text{if } k = m = 2n+1 \\ (I_p - \beta_{2n+1}\beta_{2n+1}^*)^{1/2} & \text{if } m = 2n+1 \text{ and } k = 2n \\ (I_p - \beta_{2n+1}^*\beta_{2n+1})^{1/2} & \text{if } m = 2n \text{ and } k = 2n+1 \\ 0_{p\times p} & \text{otherwise} \end{cases} \end{aligned} \tag{7.15}$$

and

$$
\mathrm{e}_k^* \mathfrak{A}_{\mathrm{even}}\, \mathrm{e}_n \stackrel{\mathrm{def}}{=} [\chi_n, y_k]_\Delta
= \begin{cases} I_p & \text{if } k = n = 0 \\ -\beta_{2m+2}^* & \text{if } k = n = 2m+1 \\ \beta_{2m+2} & \text{if } k = n = 2m+2 \\ (I_p - \beta_{2m+2}\beta_{2m+2}^*)^{1/2} & \text{if } k = 2m+2 \text{ and } n = 2m+1 \\ (I_p - \beta_{2m+2}^*\beta_{2m+2})^{1/2} & \text{if } k = 2m+1 \text{ and } n = 2m+2 \\ 0_{p\times p} & \text{otherwise} \end{cases} \tag{7.16}
$$

1. Verification of (7.15). The recursion (5.4) can be rewritten as

$$
\zeta F_k^-(\zeta) = F_{k+1}^-(\zeta)(I_p - \beta_{k+1}\beta_{k+1}^*)^{1/2} - F_k^+(\zeta)\beta_{k+1}^* \tag{7.17}
$$

and

$$
F_k^+(\zeta) = F_{k+1}^+(\zeta)(I_p - \beta_{k+1}^*\beta_{k+1})^{1/2} - \zeta F_k^-(\zeta)\beta_{k+1}. \tag{7.18}
$$

Since $H(\beta_{k+1})^{-1} = H(-\beta_{k+1})$, the recursion (5.4) can also be written as

$$
F_{k+1}^-(\zeta) = \zeta F_k^-(\zeta)(I_p - \beta_{k+1}\beta_{k+1}^*)^{1/2} + F_{k+1}^+(\zeta)\beta_{k+1}^* \tag{7.19}
$$

and

$$
F_{k+1}^+(\zeta) = F_k^+(\zeta)(I_p - \beta_{k+1}^*\beta_{k+1})^{1/2} + F_{k+1}^-(\zeta)\beta_{k+1}. \tag{7.20}
$$

If $k = 2n$, then (7.17) can be reexpressed as

$$
\zeta y_{2n}(\zeta) = \chi_{2n+1}(\zeta)(I_p - \beta_{2n+1}\beta_{2n+1}^*)^{1/2} - \chi_{2n}(\zeta)\beta_{2n+1}^* \tag{7.21}
$$

and hence, with the help of Theorem 7.3, it is easily checked that

$$
[\zeta y_{2n}, \chi_{2n}]_\Delta = -\beta_{2n+1}^*, \tag{7.22}
$$

$$
[\zeta y_{2n}, \chi_{2n+1}]_\Delta = (I_p - \beta_{2n+1}\beta_{2n+1}^*)^{1/2}, \tag{7.23}
$$

and

$$
[\zeta y_{2n}, \chi_m]_\Delta = 0_{p\times p} \quad \text{for } m = 0, \dots, 2n-1, 2n+2, \dots. \tag{7.24}
$$

If $k = 2n$, then (7.20) can be reexpressed as

$$
\zeta y_{2n+1}(\zeta) = \chi_{2n}(\zeta)(I_p - \beta_{2n+1}^*\beta_{2n+1})^{1/2} + \chi_{2n+1}(\zeta)\beta_{2n+1} \tag{7.25}
$$

and hence by another application of Theorem 7.3, it is easily checked that

$$
[\zeta y_{2n+1}, \chi_{2n+1}]_\Delta = \beta_{2n+1}, \tag{7.26}
$$

$$
[\zeta y_{2n+1}, \chi_{2n}]_\Delta = (I_p - \beta_{2n+1}^*\beta_{2n+1})^{1/2} \tag{7.27}
$$

and

$$
[\zeta y_{2n+1}, \chi_m]_\Delta = 0_{p\times p} \quad \text{for } m = 0, \dots, 2n-1, 2n+2, \dots. \tag{7.28}
$$

Formulas (7.22)–(7.24) and (7.26)–(7.28) serve to justify (7.15).

2. Verification of (7.16). Since $\chi_0(\zeta) = y_0(\zeta) = I_p$ it follows easily from Theorem 7.3 that

$$\begin{aligned}[\chi_0, y_m]_\Delta &= [y_0, y_m]_\Delta \\ &= \begin{cases} I_p & \text{if } m = 0 \\ 0_{p\times p} & \text{if } m > 0 \end{cases}\end{aligned} \tag{7.29}$$

and

$$\begin{aligned}[\chi_m, y_0]_\Delta &= [\chi_m, \chi_0]_\Delta \\ &= \begin{cases} I_p & \text{if } m = 0 \\ 0_{p\times p} & \text{if } m > 0. \end{cases}\end{aligned} \tag{7.30}$$

If $k = 2n - 1$, then (7.19) can be reexpressed as

$$y_{2n}(\zeta) = \chi_{2n-1}(\zeta)(I - \beta_{2n}\beta_{2n}^*)^{1/2} + \chi_{2n}(\zeta)\beta_{2n}^* \quad \text{for } n = 1, 2, \ldots. \tag{7.31}$$

Using (7.31) and Theorem 7.3, it is easily checked that

$$[\chi_{2n}, y_{2n}]_\Delta = \beta_{2n} \quad \text{for } n = 1, 2, \ldots, \tag{7.32}$$

$$[\chi_{2n-1}, y_{2n}]_\Delta = (I_p - \beta_{2n}\beta_{2n}^*)^{1/2}, \tag{7.33}$$

and

$$[\chi_m, y_{2n}]_\Delta = 0_{p\times p} \quad \text{for } m = 0, \ldots, 2n-2, 2n+1, \ldots. \tag{7.34}$$

If $k = 2n - 1$ in (7.18), then multiplying both sides by ζ^{-n} we obtain

$$y_{2n-1}(\zeta) = \chi_{2n}(\zeta)(I - \beta_{2n}^*\beta_{2n})^{1/2} - \chi_{2n-1}(\zeta)\beta_{2n} \quad \text{for } n = 1, 2, \ldots. \tag{7.35}$$

Using (7.35) and Theorem 7.3, it is easily checked that

$$[\chi_{2n-1}, y_{2n-1}]_\Delta = -\beta_{2n}^* \quad \text{for } n = 1, 2, \ldots, \tag{7.36}$$

$$[\chi_{2n}, y_{2n-1}]_\Delta = (I_p - \beta_{2n}^*\beta_{2n})^{1/2}, \tag{7.37}$$

and

$$[\chi_m, y_{2n-1}]_\Delta = 0_{p\times p} \quad \text{for } m = 0, \ldots, 2n-2, 2n+1, \ldots. \tag{7.38}$$

Formulas (7.29), (7.30), (7.32)–(7.34) and (7.36)–(7.38) serve to justify (7.16). □

Let $c_n = (I_p - \beta_n^*\beta_n)^{1/2}$ and $d_n = (I_p - \beta_n\beta_n^*)^{1/2}$ for $n = 1, 2, \ldots$. The next formula is presented to convey some idea of the structure of the CMV matrix $\mathfrak{A}$:

$$\mathfrak{A} = \begin{bmatrix} -\beta_1^*\beta_{-1} & -c_1\beta_2^* & c_1c_2 & 0 & 0 & \cdots \\ d_1\beta_{-1} & -\beta_1\beta_2^* & \beta_1c_2 & 0 & 0 & \cdots \\ 0 & -\beta_3^*d_2 & -\beta_3^*\beta_2 & -c_3\beta_4^* & c_3c_4 & \cdots \\ 0 & d_3d_2 & d_3\beta_2 & -\beta_3\beta_4^* & \beta_3c_4 & \cdots \\ 0 & 0 & 0 & -\beta_5^*d_4 & -\beta_5^*\beta_4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \tag{7.39}$$

We will now introduce an alternative CMV matrix $\mathfrak{C}: \ell_2^p \to \ell_2^p$ based on a density Δ that satisfies (D1) and a unitary matrix $\beta_{-1} \in \mathbb{C}^{p\times p}$ by interchanging the role of the basis $\{\chi_j\}_{j=0}^{\infty}$ and the basis $\{y_j\}_{j=0}^{\infty}$. Let

$$e_m^* \,\mathfrak{C}\, e_n = \begin{cases} \beta_{-1}[\zeta y_n, y_m]_\Delta & \text{if } m=0 \text{ and } n=0,1,\dots \\ [\zeta y_n, y_m]_\Delta & \text{if } m=1,2,\dots \text{ and } n=0,1,\dots. \end{cases} \tag{7.40}$$

If $\beta_{-1} = I_p$, then

$$\mathfrak{C} = \widetilde{V}^{-1} M_\zeta \widetilde{V}, \tag{7.41}$$

where $\widetilde{V} e_n = y_n$ for $n = 0, 1, \dots$, or, equivalently,

$$e_m^* \,\mathfrak{C}\, e_n = [\zeta y_n, y_m]_\Delta. \tag{7.42}$$

Theorem 7.7. *The alternative CMV matrix $\mathfrak{C}$ based on a density Δ that satisfies (D1) and a unitary matrix $\beta_{-1} \in \mathbb{C}^{p\times p}$ has the factorization:*

$$\mathfrak{C} = \mathfrak{B}\mathfrak{A}_{\text{even}}\mathfrak{A}_{\text{odd}}, \tag{7.43}$$

where $\mathfrak{B}$, $\mathfrak{A}_{\text{even}}$ and $\mathfrak{A}_{\text{odd}}$ are as in Theorem 7.6.

Proof. It follows from (7.40) that

$$e_m^* \,\mathfrak{B}^* \mathfrak{C}\, e_n = [\zeta y_n, y_m]_\Delta.$$

In view of Corollary 7.4, the mvf y_n can be expressed as

$$y_n(\zeta) = \sum_{k=0}^{\infty} \chi_k(\zeta) P_k,$$

where

$$P_k = [y_n, \chi_k]_\Delta \quad \text{for } k = 0, 1, \dots,$$

and hence

$$\begin{aligned} [\zeta y_n, y_m]_\Delta &= \sum_{k=0}^{\infty} [\zeta y_n, \chi_k P_k]_\Delta \\ &= \sum_{k=0}^{\infty} P_k^* [\zeta y_n, \chi_k]_\Delta \\ &= \sum_{k=0}^{\infty} [\chi_k, y_m]_\Delta [\zeta y_n, \chi_k]_\Delta \\ &= \sum_{k=0}^{\infty} [\chi_k, y_m]_\Delta [\zeta y_n, \chi_k]_\Delta. \end{aligned} \tag{7.44}$$

Assertion (7.43) follows directly from (7.44) using the identifications made in (7.15) and (7.16). $\square$

The next formula is presented to convey some idea of the structure of the alternative CMV matrix $\mathfrak{C}$:

$$\mathfrak{C} = \begin{bmatrix} -\beta_{-1}\beta_1^* & \beta_{-1}c_1 & 0 & 0 & 0 & \cdots \\ -\beta_2^* d_1 & -\beta_2^*\beta_1 & -c_2\beta_3^* & c_2c_3 & 0 & \cdots \\ d_2d_1 & d_2\beta_1 & -\beta_2\beta_3^* & \beta_2c_3 & 0 & \cdots \\ 0 & 0 & -\beta_4^* d_3 & -\beta_4^*\beta_3 & -c_4\beta_5^* & \cdots \\ 0 & 0 & d_4d_3 & \beta_4\beta_3 & -\beta_4\beta_5^* & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \tag{7.45}$$

Theorem 7.8. *Suppose Δ and $\widetilde{\Delta}$ are densities which both meet the constraint (D1) and β_{-1} and $\tilde{\beta}_{-1}$ are unitary matrices belonging to $\mathbb{C}^{p\times p}$. If the CMV matrix $\mathfrak{A}$ based on Δ and β_{-1} and the CMV matrix $\widetilde{\mathfrak{A}}$ based on $\widetilde{\Delta}$ and $\tilde{\beta}_{-1}$ coincide, i.e.,*

$$\mathfrak{A} = \widetilde{\mathfrak{A}}, \tag{7.46}$$

then $\beta_{-1} = \tilde{\beta}_{-1}$ and $\Delta = \widetilde{\Delta}$.

Proof. Let $\{\beta_n\}_{n=0}^{\infty}$ and $\{\tilde{\beta}_n\}_{n=0}^{\infty}$ be the Schur parameters corresponding to Δ and $\widetilde{\Delta}$, respectively. Let $c_n = (I_p - \beta_n^*\beta_n)^{1/2}$, $d_n = (I_p - \beta_n\beta_n^*)^{1/2}$, $\tilde{c}_n = (I_p - \tilde{\beta}_n^*\tilde{\beta}_n)^{1/2}$ and $\tilde{d}_n = (I_p - \tilde{\beta}_n\tilde{\beta}_n^*)^{1/2}$ for $n = 0, 1, \dots$. Then, since

$$d_1\beta_{-1} = \mathrm{e}_1^*\,\mathfrak{A}\,\mathrm{e}_0 = \mathrm{e}_1^*\,\widetilde{\mathfrak{A}}\,\mathrm{e}_0 = \tilde{d}_1\tilde{\beta}_{-1},$$

and d_1 and $\tilde{d}_1$ are both positive definite, whereas β_{-1} and $\tilde{\beta}_{-1}$ are both unitary, the uniqueness of the polar decomposition implies that

$$d_1 = \tilde{d}_1 \quad \text{and} \quad \beta_{-1} = \tilde{\beta}_{-1}.$$

We will now show $\beta_n = \tilde{\beta}_n$ for $n = 0, 1, \dots$ by induction. First note that we have $\beta_0 = \tilde{\beta}_0 = 0_{p\times p}$ by construction (see Remark 4.4). In view of $\beta_{-1} = \tilde{\beta}_{-1}$ and

$$-\beta_1^*\beta_{-1} = \mathrm{e}_0^*\,\mathfrak{A}\,\mathrm{e}_0 = \mathrm{e}_0^*\,\widetilde{\mathfrak{A}}\,\mathrm{e}_0 = -\tilde{\beta}_1^*\tilde{\beta}_{-1},$$
$$\beta_1 = \tilde{\beta}_1.$$

If $\beta_m = \tilde{\beta}_m$ for $m = 2n$, then the formulas

$$\beta_{2n+1}^* d_{2n} = \mathrm{e}_{2n}^*\,\mathfrak{A}\,\mathrm{e}_{2n-1} = \mathrm{e}_{2n}^*\,\widetilde{\mathfrak{A}}\,\mathrm{e}_{2n-1} = \tilde{\beta}_{2n+1}^*\tilde{d}_{2n},$$

clearly imply that $\beta_{2n+1} = \beta_{2n+1}$. If $\beta_m = \tilde{\beta}_m$ for $m = 2n+1$, then the formulas

$$c_{2n+1}\beta_{2n+2}^* = \mathrm{e}_{2n}^*\,\mathfrak{A}\,\mathrm{e}_{2n+1} = \mathrm{e}_{2n}^*\,\mathfrak{A}\,\mathrm{e}_{2n+1} = \tilde{c}_{2n+1}\tilde{\beta}_{2n+2}^*$$

clearly imply that $\beta_{2n+2} = \beta_{2n+2}$.

Finally, as $\beta_n = \tilde{\beta}_n$ for $n = 0, 1, \dots$, the proof that $\Delta = \widetilde{\Delta}$ on $\mathbb{T}$ is completed by invoking Theorem 5.10. □

Theorem 7.9. *Suppose Δ and $\widetilde{\Delta}$ are densities which both meet the constraint (D1). If the alternative CMV matrix $\mathfrak{C}$ based on Δ and β_{-1} and the alternative CMV matrix $\widetilde{\mathfrak{C}}$ based on $\widetilde{\Delta}$ and $\tilde{\beta}_{-1}$ coincide, i.e.,*

$$\mathfrak{C} = \widetilde{\mathfrak{C}}, \tag{7.47}$$

then $\beta_{-1} = \tilde{\beta}_{-1}$ and $\Delta(\zeta) = \widetilde{\Delta}(\zeta)$ for $\zeta \in \mathbb{T}$.

Proof. The proof is completed in much the same way as Theorem 7.8. □

8. Convergence results

We begin with four lemmas.

Lemma 8.1. *If $\{A_n\}_{n=0}^{\infty}$ is a sequence of $p \times p$ positive definite matrices and A is a $p \times p$ positive definite matrix, then*

$$\lim_{n\uparrow\infty} \|A_n - A\| = 0 \Longleftrightarrow \lim_{n\uparrow\infty} \|A_n^{1/2} - A^{1/2}\| = 0. \tag{8.1}$$

Proof. In view of the well known formula, see, e.g., (17.39) in [17],

$$A^{1/2} - \{A_n\}^{1/2} = \frac{\sin(\pi/2)}{2\pi} \int_0^\infty x^{1/2}(xI_p + A_n)^{-1}(A_n - A)(xI_p + A)^{-1} dx.$$

Thus,

$$\begin{aligned} \|A^{1/2} - \{A_n\}^{1/2}\| &\le \left(\frac{1}{2\pi} \int_0^\infty x^{1/2} \|(xI_p + A)^{-1}\| \|(xI_p + A_n)^{-1}\| dx \right) \\ &\qquad\qquad\qquad\qquad\qquad \times \|A_n - A\| \\ &\le \kappa \|A_n - A\|, \end{aligned}$$

for some constant $\kappa > 0$, which justifies the implication $\Longrightarrow$ in (8.1). The converse implication follows from the fact that

$$\begin{aligned} \|A_n - A\| &= \|A_n^{1/2}(A_n^{1/2} - A^{1/2}) + (A_n^{1/2} - A^{1/2})A^{1/2}\| \\ &\le (\|A_n^{1/2}\| + \|A^{1/2}\|)\|A_n^{1/2} - A^{1/2}\| \\ &\le (\|A_n^{1/2} - A^{1/2}\| + 2\|A^{1/2}\|)\|A_n^{1/2} - A^{1/2}\|. \end{aligned}$$

□

Lemmas 8.2 and 8.3 are well known results, see, e.g., Delsarte, Genin and Kamp [13].

Lemma 8.2. *If $P_1, \dots, P_n$ are $p \times p$ Hermitian matrices and $P_j \succeq I_p$ for $j = 1, \dots, n$, then*

$$\|P_1 \cdots P_n - I_p\| \le \|P_1\| \cdots \|P_n\| - 1. \tag{8.2}$$

Proof. If $A_1 = U^* D U$, where U is unitary and $D = \mathrm{diag}(\mu_1, \dots, \mu_p)$ with

$$\mu_1 \ge \cdots \ge \mu_p \ge 1,$$

then

$$\begin{aligned}\|P_1 - I_p\| = \|U^*DU - I_p\| &= \|U^*(D - I_p)U\| \\ &= \|D - I_p\| = \mu_1 - 1 \\ &= \|P_1\| - 1.\end{aligned}$$

Thus, (8.2) holds for $n = 1$. If $n > 1$ and the inequality is valid for $n - 1$, then

$$\begin{aligned}\|P_1P_2\cdots P_n - I_p\| &\leq \|P_1(P_2\cdots P_n - I_p)\| + \|P_1 - I_p\| \\ &= \|P_1\|(\|P_2\|\cdots\|P_n\| - 1) + \|P_1\| - 1,\end{aligned}$$

which coincides with (8.2). □

Lemma 8.3. *If $\{A_n\}_{n=0}^{\infty}$ is a sequence of $p \times p$ positive semidefinite matrices and $B_n = (I_p + A_0)\cdots(I_p + A_n)$, then*

$$\lim_{n\uparrow\infty} B_n = B \text{ and } B \text{ is invertible} \tag{8.3}$$

if and only if

$$\sum_{n=0}^{\infty} \|A_n\| < \infty. \tag{8.4}$$

Proof. If $m > n$, then

$$\begin{aligned}\|B_m - B_n\| &\leq \|B_n\|\|(I_p + A_{n+1})\cdots(I_p + A_m) - I_p\| \\ &\leq \|B_n\|\{\|I_p + A_{n+1}\|\cdots\|I_p + A_m\| - 1\}.\end{aligned}$$

Therefore, since

$$\|I_p + A_j\| = 1 + \|A_j\| \leq \exp\|A_j\|,$$

it is readily checked that

$$\|B_m - B_n\| \leq \exp\left(\sum_{j=1}^{n}\|A_j\|\right)\left\{\exp\left(\sum_{j=n+1}^{m}\|A_j\|\right) - 1\right\}.$$

Thus, if (8.4) is in force, $\{B_n\}_{n=0}^{\infty}$ tends to a limit $B \in \mathbb{C}^{p\times p}$ by the Cauchy convergence criterion. Moreover, as

$$1 \leq \det B_n \leq \det B_{n+1} \leq \det B,$$

B is invertible.

Conversely, if (8.3) is in force, then (8.4) holds, since

$$\begin{aligned}\det B \geq \det B_n &= \det(I_p + A_0)\cdots\det(I_p + A_n)\\ &\geq \prod_{j=0}^{n}(1+\|A_j\|)\\ &\geq 1+\sum_{j=0}^{n}\|A_j\|.\end{aligned}$$

□

Lemma 8.4. *If $\beta_n \in \mathbb{C}^{p\times p}$ and $\|\beta_j\| \leq \rho < 1$ for $j = 1, 2, \ldots$, then*

$$1 \leq \prod_{j=1}^{n}(1+\|\beta_j\|) \leq \exp\left\{\sum_{j=1}^{n}\|\beta_j\|\right\} \tag{8.5}$$

and

$$\begin{aligned}1 &\leq \prod_{j=1}^{n}(1-\|\beta_j\|)^{-1}\\ &\leq \prod_{j=1}^{n}\frac{1+\|\beta_j\|}{1-\|\beta_j\|} \leq \exp\left\{\frac{2}{1-\rho}\sum_{j=1}^{n}\|\beta_j\|\right\}.\end{aligned} \tag{8.6}$$

If $\sum_{j=0}^{\infty}\|\beta_n\| < \infty$, then

$$\prod_{j=1}^{n}(1+\|\beta_j\|) \quad \textit{and} \quad \prod_{j=1}^{n}(1+\|\beta_j\|)$$

converge to finite positive limits as $n \uparrow \infty$.

Proof. The bounds in (8.5) and the lower bound in (8.6) are self-evident. The upper bound in (8.6) follows from the observation that

$$\begin{aligned}\frac{1}{1-\|\beta_j\|} \leq \frac{1+\|\beta_j\|}{1-\|\beta_j\|} &= 1+\frac{2\|\beta_j\|}{1-\|\beta_j\|}\\ &\leq 1+\frac{2}{1-\rho}\|\beta_j\| \leq \exp\left\{\frac{2}{1-\rho}\|\beta_j\|\right\}.\end{aligned}$$

Finally, the asserted existence of the finite positive limits follows from the monotonicity of the two sequences and the bounds in (8.5) and (8.6). □

Lemma 8.5. *If* $F \in \mathcal{W}^{p\times p}$ *and* $\|I_p - F\|_{\mathcal{W}} \le \varepsilon < 1$, *then*

(1) F *is invertible in* $\mathcal{W}^{p\times p}$.

(2) $1-\varepsilon \le \|F\|_{\mathcal{W}} \le 1+\varepsilon$.

(3) $\frac{1}{1+\varepsilon} \le \|F^{-1}\|_{\mathcal{W}} \le \frac{1}{1-\varepsilon}$.

Proof. The identity

$$F(\zeta) = I_p + F(\zeta) - I_p$$

implies that

$$\|I_p\|_{\mathcal{W}} - \|I_p - F\|_{\mathcal{W}} \le \|F\|_{\mathcal{W}} \le \|I_p\|_{\mathcal{W}} + \|F - I_p\|_{\mathcal{W}},$$

which is equivalent to (2).

Next, if $\zeta \in \mathbb{T}$ and $u \in \mathbb{C}^p$, then

$$\begin{aligned}\|F(\zeta)u\| &= \|u - (I_p - F(\zeta)u\| \\ &\ge \|u\| - \|I_p - F(\zeta)\|_{\mathcal{W}}\|u\| \\ &= (1-\varepsilon)\|u\|\end{aligned}$$

Therefore, $F(\zeta)$ is invertible for $\zeta \in \mathbb{T}$ and hence, by item (1) of Theorem 5.1, (1) holds.

Finally, the lower bound in (3) follows from the inequalities

$$\begin{aligned}1 = \|F^{-1}F\|_{\mathcal{W}} &\le \|F^{-1}\|_{\mathcal{W}}\|I_p + (F - I_p)\|_{\mathcal{W}} \\ &\le \|F^{-1}\|_{\mathcal{W}}(1+\varepsilon),\end{aligned}$$

whereas the upper bound follows from the inequalities

$$\begin{aligned}\|F^{-1}\|_{\mathcal{W}} &= \|I_p + F^{-1} - I_p\|_{\mathcal{W}} \le \|I_p\|_{\mathcal{W}} + \|F^{-1}(I_p - F)\|_{\mathcal{W}} \\ &\le 1 + \varepsilon\|F^{-1}\|_{\mathcal{W}}.\end{aligned}$$

□

Corollary 8.6. *If* $\{G_n\}_{n=0}^{\infty}$ *is a sequence in* $\mathcal{W}^{p\times p}$ *such that* $G_n^{-1} \in \mathcal{W}^{p\times p}$ *and*

$$\lim_{n\uparrow\infty} \|G_n - G\|_{\mathcal{W}} = 0 \quad \text{and} \quad G^{-1} \in \mathcal{W}^{p\times p}, \tag{8.7}$$

then

$$\lim_{n\uparrow\infty} \|G_n^{-1} - G^{-1}\|_{\mathcal{W}} = 0. \tag{8.8}$$

Proof. In view of (8.7), for any $\varepsilon < 1$, there exists a positive integer n_ε such that

$$\|G^{-1}G_n - I_p\|_{\mathcal{W}} \le \varepsilon < 1 \quad \text{for } n \ge n_\varepsilon.$$

Thus, it follows from item (3) of Lemma 8.5,

$$\|G_n^{-1}G\|_{\mathcal{W}} \le \frac{1}{1-\varepsilon} \quad \text{for } n \ge n_\varepsilon.$$

Consequently,

$$\begin{aligned}\|G_n^{-1} - G^{-1}\|_{\mathcal{W}} &= \|G_n^{-1}(G - G_n)G^{-1}\|_{\mathcal{W}} \\ &= \|G_n^{-1}GG^{-1}(G - G_n)G^{-1}\|_{\mathcal{W}} \\ &\le \|G_n^{-1}G\|_{\mathcal{W}}\|G^{-1}\|_{\mathcal{W}}^2\|G - G_n\|_{\mathcal{W}} \\ &\to 0 \ \text{ as } n \uparrow \infty. \end{aligned}$$

□

Lemma 8.7. *If the matrix polynomials $\{F_n^{\pm}\}_{n=0}^{\infty}$ are defined by (5.3) in terms of a given sequence of strict contractions $\{\beta_n\}_{n=0}^{\infty}$ with $\beta_0 = 0_{p\times p}$ and $\sum_{n=0}^{\infty}\|\beta_n\| < \infty$ and X_n and Y_n are given by (5.6) and (5.7), respectively, then:*

(i) *$X_n \to X_\infty$ as $n \uparrow \infty$, where $X_\infty \in \mathbb{C}^{p\times p}$ is nonsingular.*

(ii) *$Y_n \to Y_\infty$ as $n \uparrow \infty$, where $Y_\infty \in \mathbb{C}^{p\times p}$ is nonsingular.*

(iii) *$X_n^{-1} \to X_\infty^{-1}$ as $n \uparrow \infty$.*

(iv) *$Y_n^{-1} \to Y_\infty^{-1}$ as $n \uparrow \infty$.*

Proof. The proof is broken into steps.

1. Verification of (i) *and* (ii). If

$$A_j = (I_p - \beta_n\beta_n^*)^{-1/2} - I_p \quad \text{for } \ j = 0, 1, \ldots,$$

then

$$X_n = (I_p + A_0)\cdots(I_p + A_n).$$

Since $A_n \succ 0$ and

$$\begin{aligned}\|A_j\| &= \frac{1}{\{1 - \|\beta_j\|^2\}^{1/2}} - 1 = \frac{1}{\{(1 - \|\beta_j\|)(1 + \|\beta_j\|)\}^{1/2}} - 1 \\ &\le \frac{1}{1 - \|\beta_j\|} - 1 \\ &= \frac{\|\beta_j\|}{1 - \|\beta_j\|},\end{aligned}$$

it is readily seen that $\sum_{n=0}^{\infty}\|A_n\| < \infty$. Therefore, (i) follows from Lemma 8.3. The proof of (ii) is similar.

2. Verification of (iii) *and* (iv). Assertion (iii) follows readily from Corollary 8.6 applied to the sequence $\{X_n\}_{n=0}^{\infty}$. The verification of (iv) is similar. □

Lemma 8.8. *If the matrix polynomials $\{E_n^{\pm}\}_{n=0}^{\infty}$ are defined by (2.1) and (2.2) in terms of the Fourier coefficients of a density Δ that satisfies (D1) and (D2), then*

$$\lim_{n\uparrow\infty}\{\gamma_{00}^{(n)}\}^{1/2} = Q(0)^{-1} \succ 0 \tag{8.9}$$

and

$$\lim_{n\uparrow\infty}\{\gamma_{nn}^{(n)}\}^{1/2} = R(0)^{-1} \succ 0. \tag{8.10}$$

Proof. If $u \in \mathbb{C}^p$ and $\omega = 0$ in (6.13), then, in view of Theorem 6.3,

$$\lim_{n\uparrow\infty} u^*(Q(0)^{-2} - F_n^+(0)F_n^+(0)^*)u = 0.$$

Therefore, since

$$\gamma_{00}^{(n)} = F_n^+(0)F_n^+(0)^* = E_n^+(0)E_n^+(0)^* \quad \text{for } n = 0, 1, \ldots,$$

$$\lim_{n\uparrow\infty} u^*(Q(0)^{-2} - \gamma_{00}^{(n)})u = 0.$$

In view of (8.1),

$$\lim_{n\uparrow\infty} u^*(Q(0)^{-1} - \{\gamma_{00}^{(n)}\}^{1/2})u,$$

i.e., (8.9) holds.

We will now prove (8.10). It follows from (8.9) that

$$\lim_{n\uparrow\infty} \{\gamma_{nn}^{(n)}\}^{1/2} = \widetilde{Q}(0)^{-1} \succ 0.$$

Taking advantage of the identification (3.4), (8.10) is readily obtained. □

Corollary 8.9. *If the matrix polynomials $\{E_n^\pm\}_{n=0}^\infty$ are defined by (2.1) and (2.2) in terms of the Fourier coefficients of a density Δ that satisfies (D1) and (D2), then*

$$\det Q(0) = \det R(0). \tag{8.11}$$

Proof. This follows easily from (8.9) and (8.10), since $\det \gamma_{00}^{(n)} = \det \gamma_{nn}^{(n)}$. □

Lemma 8.10. *If the matrix polynomials $\{E_n^\pm\}_{n=0}^\infty$ are defined by (2.1) and (2.2) in terms of the Fourier coefficients of a density Δ that satisfies (D1) and (D2) and $u \in \mathbb{C}^p$, then*

$$\lim_{n\uparrow\infty} \|(Q^{-1} - E_n^+)u\|_\Delta = 0, \tag{8.12}$$

$$\lim_{n\uparrow\infty} \|(Q^{-1}Q(0)^{-1} - F_n^+F_n^+(0)^*)u\|_\Delta = 0, \tag{8.13}$$

$$\lim_{n\uparrow\infty} \|(R^{-*} - \zeta^{-n}E_n^-)u\|_\Delta = 0 \tag{8.14}$$

and

$$\lim_{n\uparrow\infty} \|(R^{-*}R(0)^{-1} - \zeta^{-n}F_n^-X_n^*)u\|_\Delta = 0. \tag{8.15}$$

Proof. If $u \in \mathbb{C}^p$, then (8.12) follows from item (4) in Theorem 6.3 with $\omega = 0$ and (8.9). Indeed,

$$\|(Q^{-1} - E_n^+)u\|_\Delta = \|(Q^{-1}E_n^+(0)^* - E_n^+E_n^+(0)^*)E_n^+(0)^{-*}u\|_\Delta.$$

Assertion (8.13) follows directly from item (4) in Theorem 6.3 with $\omega = 0$.

If $\widetilde{\Delta}(\zeta) = \Delta(\bar{\zeta}) = \widetilde{Q}(\zeta)\widetilde{Q}(\zeta)$, where $\widetilde{Q}^\pm \in \mathcal{W}_+^{p\times p}$ then, it follows from (8.12) that

$$\lim_{n\uparrow\infty} \|(\widetilde{Q}^{-1} - \widetilde{E}_n^+)u\|_{\widetilde{\Delta}} = 0, \tag{8.16}$$

where $\widetilde{E}_n^+$ is given by (3.2). Assertion (8.14) drops out easily from (8.16) in view of the identifications (3.4) and (3.2).

Assertion (8.15) can be obtained from (8.13) in a similar fashion. □

9. The Schur parameters of a positive spectral density in the Wiener algebra are summable (the harder half of Baxter's theorem)

In this section the convergence results established in Section 8 will be strengthened with the help of a matrix version of Baxter's inequality. The main conclusion in this section, Theorem 9.6, is verified in complete detail. This result was obtained earlier by Geronimo [21] via a matrix generalization of Baxter's inequality due to Hirschman [26].

Corollary 9.1. *If $A \in \mathbb{C}^{p\times p}$ and $\|I_p - A\| \le \varepsilon < 1$, then*:

(1) *A is invertible.*

(2) $1-\varepsilon \le \|A\| \le 1+\varepsilon$.

(3) $\frac{1}{1+\varepsilon} \le \|A^{-1}\| \le \frac{1}{1-\varepsilon}$.

The following theorem depends heavily on the matrix extension of Baxter's inequality, which is presented in Appendix C.

Theorem 9.2. *If $\{E_n^{\pm}\}_{n=0}^{\infty}$ are matrix polynomials based on the Fourier coefficients of a density Δ that satisfies* (D1) *and* (D2)*, then*

$$\lim_{n\uparrow\infty} \|Q^{-1}Q_0^{-1} - E_n^{+}\{\gamma_{00}^{(n)}\}^{1/2}\|_{\mathcal{W}} = 0 \tag{9.1}$$

and

$$\lim_{n\uparrow\infty} \|R^{-*}R_0^{-1} - \zeta^{-n}E_n^{-}\{\gamma_{nn}^{(n)}\}^{1/2}\|_{\mathcal{W}} = 0. \tag{9.2}$$

Moreover,

(1) $\|Q^{-1} - E_n^{+}\|_{\mathcal{W}} \to 0$ *as* $n \uparrow \infty$.

(2) $\|R^{-*} - \zeta^{-n}E_n^{-}\|_{\mathcal{W}} \to 0$ *as* $n \uparrow \infty$.

(3) $\|Q - (E_n^{+})^{-1}\|_{\mathcal{W}} \to 0$ *as* $n \uparrow \infty$.

(4) $\|R^{*} - (\zeta^{-n}E_n^{-})^{-1}\|_{\mathcal{W}} \to 0$ *as* $n \uparrow \infty$.

(5) $\|C_n - C\|_{\mathcal{W}} \to 0$ *as* $n \uparrow \infty$.

Proof. The proof of (9.1) and (9.2) drops out from Theorem C.1. The rest of the proof is broken into steps.

1. Verification of (1) *and* (2). It follows from Theorem C.1 that there exists a constant κ so that

$$\|E_n^{+}\{\gamma_{00}^{(n)}\}^{1/2}\|_{\mathcal{W}} \le \kappa$$

for $n \ge 0$. Thus, as (1) holds if and only if $\|Q^{-1}Q_0^{-1} - E_n^{+}Q_0^{-1}\|_{\mathcal{W}} \to 0$ as $n \uparrow \infty$ and

$$\|Q^{-1}Q_0^{-1} - E_n^{+}Q_0^{-1}\|_{\mathcal{W}} \le \kappa_n + \ell_n$$

where

$$\kappa_n = \|Q^{-1}Q_0^{-1} - E_n^{+}\{\gamma_{00}^{(n)}\}^{1/2}\|_{\mathcal{W}}$$

and

$$\begin{aligned}\ell_n &= \|E_n^+(\{\gamma_{00}^{(n)}\}^{1/2} - Q_0^{-1})\|_{\mathcal{W}} \\ &\le \|E_n^+\|_{\mathcal{W}}\|\{\gamma_{00}^{(n)}\}^{1/2} - Q_0^{-1}\| \\ &\le \kappa\|\{\gamma_{00}^{(n)}\}^{-1/2}\|\|\{\gamma_{00}^{(n)}\}^{1/2} - Q_0^{-1}\| \\ &\le \kappa\|Q_0\|\|\{\gamma_{00}^{(n)}\}^{1/2} - Q_0^{-1}\|\end{aligned}$$

we obtain (1) by using (8.9), (9.1) and the monotonicity in (6.18).

The proof of (2) is similar.

2. Verification of (3) *and* (4). Assertion (3) follows from item (3) of Lemma 8.5 and (1). Assertion (4) is proved similarly.

3. Verification of (5). In view of item (6) of Theorem 5.5 and (5.22),

$$\left\{\frac{C_n(\zeta) + C_n(\zeta)^*}{2}\right\} = F_n^+(\zeta)^{-*}F_n^+(\zeta)^{-1} = E_n^+(\zeta)^{-*}E_n^+(\zeta)^{-1}.$$

Thus, using (1) we have

$$\lim_{n\uparrow\infty}\left\|\left\{\frac{C_n + C_n^*}{2}\right\} - Q^*Q\right\|_{\mathcal{W}} = 0. \tag{9.3}$$

But as $Q(\zeta)^*Q(\zeta) = \Delta(\zeta) = \{C(\zeta) + C(\zeta)^*\}/2$ and C_n and C belong to $\mathcal{W}_+^{p\times p}$, (5) can be obtained from (9.3). □

In view of formulas (5.17), (5.25) and (5.26), it is readily checked that the mvf

$$\begin{aligned}\Xi_n(\zeta) &\overset{\text{def}}{=} \Theta_n(\zeta)\begin{bmatrix}V_n^* & 0\\ 0 & U_n^*\end{bmatrix}\begin{bmatrix}\zeta^{-n-1}I_p & 0\\ 0 & I_p\end{bmatrix}, \\ &= \frac{1}{2}\begin{bmatrix}\{I_p + C_n(\zeta)^*\}\zeta^{-n}E_n^-(\zeta) & \{I_p - C_n(\zeta)\}E_n^+(\zeta)\\ \{I_p - C_n(\zeta)^*\}\zeta^{-n}E_n^-(\zeta) & \{I_p + C_n(\zeta)\}E_n^+(\zeta)\end{bmatrix}.\end{aligned}$$

Corollary 9.3. *If Δ is a density that satisfies* (D1) *and* (D2)*, then*

$$\lim_{n\uparrow\infty}\|\Xi_n - \Xi_\infty\|_{\mathcal{W}} = 0, \tag{9.4}$$

where

$$\Xi_\infty(\zeta) = \frac{1}{2}\begin{bmatrix}\{I_p + C(\zeta)^*\}R(\zeta)^{-*} & \{I_p - C(\zeta)\}Q(\zeta)^{-1}\\ \{I_p - C(\zeta)^*\}R(\zeta)^{-*} & \{I_p + C(\zeta)\}Q(\zeta)^{-1}\end{bmatrix} \tag{9.5}$$

for $\zeta \in \mathbb{T}$.

Proof. Assertion (9.4) follows easily from items (3)–(5) in Theorem 9.2. □

Lemma 9.4. *If Δ is a density which satisfies* (D1) *and* (D2) *and if the mvf Ξ_∞ defined in* (9.5) *is written in block form as*

$$\Xi_\infty = \begin{bmatrix} \Xi_\infty^{(11)} & \Xi_\infty^{(12)} \\ \Xi_\infty^{(21)} & \Xi_\infty^{(22)} \end{bmatrix}$$

and

$$\Omega(\zeta) \stackrel{\text{def}}{=} \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix} = Z_2 \Xi_\infty(\zeta)^* Z_2, \quad \textit{with } Z_2 = \begin{bmatrix} 0 & I_p \\ I_p & 0 \end{bmatrix}, \tag{9.6}$$

then:

(i) *Ω is j_p-unitary on $\mathbb{T}$.*

(ii) *$\mathcal{E}\Omega_{12} + \Omega_{22}$ is invertible on $\mathbb{T}$ for every contractive matrix $\mathcal{E} \in \mathbb{C}^{p\times p}$.*

(iii) *The identity*

$$(\mathcal{E}\Omega_{12} + \Omega_{22})^{-1}(\mathcal{E}\Omega_{11} + \Omega_{21}) = (\Xi_\infty^{(11)}\mathcal{E} + \Xi_\infty^{(12)})(\Xi_\infty^{(21)}\mathcal{E} + \Xi_\infty^{(22)})^{-1}$$

holds on $\mathbb{T}$ for every contractive matrix $\mathcal{E} \in \mathbb{C}^{p\times p}$.

Proof. The verification of (i) is easy, since $Z^* = Z$ and $Z^* j_p Z = -j_p$. The verification of (ii) follows easily from the identity

$$\Omega_{22}(\zeta)^*\Omega_{22}(\zeta) = \Omega_{12}(\zeta)^*\Omega_{12}(\zeta) + I_p,$$

which is the 22 block of $\Omega(\zeta)^* j_p \Omega(\zeta) = j_p$. Finally, (iii) holds if and only if

$$(\mathcal{E}\Omega_{11} + \Omega_{21})(\Xi_\infty^{(21)}\mathcal{E} + \Xi_\infty^{(22)}) - (\mathcal{E}\Omega_{12} + \Omega_{22})(\Xi_\infty^{(11)}\mathcal{E} + \Xi_\infty^{(12)}) = 0_{p\times p} \tag{9.7}$$

on $\mathbb{T}$. But the left-hand side of (9.7) can be rewritten as

$$\begin{aligned}
&\begin{bmatrix} \mathcal{E} & I_p \end{bmatrix} \begin{bmatrix} \Omega_{11} \\ \Omega_{21} \end{bmatrix} \begin{bmatrix} \Xi_\infty^{(21)} & \Xi_\infty^{(22)} \end{bmatrix} \begin{bmatrix} \mathcal{E} \\ I_p \end{bmatrix} - \begin{bmatrix} \mathcal{E} & I_p \end{bmatrix} \begin{bmatrix} \Omega_{12} \\ \Omega_{22} \end{bmatrix} \begin{bmatrix} \Xi_\infty^{(11)} & \Xi_\infty^{(12)} \end{bmatrix} \begin{bmatrix} \mathcal{E} \\ I_p \end{bmatrix} \\
&\qquad = \begin{bmatrix} \mathcal{E} & I_p \end{bmatrix} \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix} j_p \begin{bmatrix} \Xi_\infty^{(21)} & \Xi_\infty^{(22)} \\ \Xi_\infty^{(11)} & \Xi_\infty^{(12)} \end{bmatrix} \begin{bmatrix} \mathcal{E} \\ I_p \end{bmatrix} \\
&\qquad = \begin{bmatrix} \mathcal{E} & I_p \end{bmatrix} Z \Xi_\infty^* Z j_p Z \Xi_\infty \begin{bmatrix} \mathcal{E} \\ I_p \end{bmatrix} \\
&\qquad = \begin{bmatrix} I_p & \mathcal{E} \end{bmatrix} \Xi_\infty^* (-j_p) \Xi_\infty \begin{bmatrix} \mathcal{E} \\ I_p \end{bmatrix} \\
&\qquad = -\begin{bmatrix} I_p & \mathcal{E} \end{bmatrix} j_p \begin{bmatrix} \mathcal{E} \\ I_p \end{bmatrix} = 0_{p\times p}
\end{aligned}$$

on $\mathbb{T}$. □

Remark 9.5. If $Q(\zeta)^{-1} = \sum_{j=0}^{\infty} \zeta^j L_j$ and $R(\zeta)^{-1} = \sum_{j=0}^{\infty} \zeta^j M_j$, then

$$Q_0 L_0 = R_0 M_0 = I_p$$

and the convergence in $\mathcal{W}^{p\times p}$ indicated in (9.1) and (9.2) is equivalent to

$$\lim_{n\uparrow\infty} \left\{ \sum_{j=0}^{n} \left\| L_j L_0 - \sum_{j=0}^{n} \gamma_{j0}^{(n)} \right\| + \sum_{j=n+1}^{\infty} \|L_j L_0\| \right\} = 0 \tag{9.8}$$

and

$$\lim_{n\uparrow\infty} \left\{ \sum_{j=0}^{n} \left\| M_j^* M_0 - \sum_{j=0}^{n} \gamma_{n-j,n}^{(n)} \right\| + \sum_{j=n+1}^{\infty} \|M_j^* M_0\| \right\} = 0, \tag{9.9}$$

respectively.

The proof of the next theorem is modeled on the proof of (x) implies (xii) in Theorem 5.2.2 of Simon [31], which treats the scalar case $p = 1$. Simon credits his proof to a "clever argument of Baxter [5]".

Theorem 9.6. *If $\{\beta_n\}_{n=0}^{\infty}$ are the Schur parameters based on a density Δ which satisfies* (D1)*, then*

$$\sum_{n=0}^{\infty} \|\beta_n\| < \infty. \tag{9.10}$$

Proof. The proof is broken into steps.

1. The matrix polynomials $\{E_n^{\pm}\}_{n=0}^{\infty}$ defined by (2.1) *and* (2.2) *obey the recursion*:

$$E_n^+(\lambda)E_n^+(0)^{-1} - E_{n-1}^+(\lambda)E_{n-1}^+(0)^{-1} = \lambda E_{n-1}^-(\lambda)\beta_n' \quad \textit{for } n = 1, 2, \dots \tag{9.11}$$

with

$$\beta_n' = V_n \beta_n U_{n-1}^* E_{n-1}^+(0)^{-1} \quad \textit{for } n = 1, 2, \dots. \tag{9.12}$$

The recursion (5.18) implies that

$$F_n^+(\lambda)(I_p - \beta_n^*\beta_n)^{1/2} = \lambda F_{n-1}^-(\lambda)\beta_n + F_{n-1}^+(\lambda),$$

which, upon invoking the formulas

$$F_n^+(\lambda) = E_n^+(\lambda)U_n, \quad F_n^-(\lambda) = E_n^-(\lambda)V_n$$

and

$$\begin{aligned} U_n(I_p - \beta_n^*\beta_n)^{1/2} &= \{\gamma_{00}^{(n)}\}^{-1/2}\{\gamma_{00}^{(n-1)}\}^{1/2}U_{n-1} \\ &= E_n^+(0)^{-1}E_{n-1}^+(0)U_{n-1}, \end{aligned}$$

can be rewritten as (9.11).

2. The inequality

$$2\|R_0\{\gamma_{n-1,n-1}^{(n-1)}\}^{1/2}\beta_n'\| - \|R^* E_{n-1}^- \beta_n'\|_{\mathcal{W}} \\ \le \|R^* E_n^+ E_n^+(0)^{-1}\|_{\mathcal{W}} - \|R^* E_{n-1}^+ E_{n-1}^+(0)^{-1}\|_{\mathcal{W}} \quad (9.13)$$

holds for $n = 1, 2, \ldots$.

The inequality

$$R(\zeta)^* E_{n-1}^+(\zeta) E_{n-1}^+(0)^{-1} = R(\zeta)^* E_n^+(\zeta) E_n^+(0)^{-1} - R(\zeta)^* \zeta E_{n-1}^-(\zeta)\beta_n' \quad (9.14)$$

is easily obtained from (9.11), or, equivalently, in terms of the notation

$$A(\zeta) = \sum_{j=-\infty}^{n} \zeta^j A_j = R(\zeta)^* E_n^+(\zeta) E_n^+(0)^{-1}$$

and

$$B(\zeta) = \sum_{j=-\infty}^{n} \zeta^j B_j = R(\zeta)^* \zeta E_{n-1}^-(\zeta)\beta_n',$$

can be restated as

$$R(\zeta)^* E_{n-1}^+(\zeta) E_{n-1}^+(0)^{-1} = \sum_{j=-\infty}^{n} \zeta^j A_j - \sum_{j=-\infty}^{n} \zeta^j B_j.$$

Thus, as E_{n-1}^+ is a polynomial of degree at most $n-1$,

$$A_n = B_n = R_0^*\{\gamma_{n-1,n-1}^{(n-1)}\}^{1/2}\beta_n' \quad \text{for } n = 1, 2, \ldots$$

and hence that

$$\|R^* E_{n-1}^+ E_{n-1}^+(0)^{-1}\|_{\mathcal{W}} \le \sum_{j=-\infty}^{n} \|A_j\| + \sum_{j=-\infty}^{n} \|B_j\| - 2\|B_n\| \\ = \|R^* E_n^+ E_n^+(0)^{-1}\|_{\mathcal{W}} + \|R^* E_{n-1}^- \beta_n'\|_{\mathcal{W}} - 2\|B_n\|.$$

Consequently, (9.13) holds.

3. If $0 < \varepsilon < 1$, $\rho_n = R_0\{\gamma_{n-1,n-1}^{(n-1)}\}^{1/2}$ *and* $\Psi_n(\zeta) = R(\zeta)^* \zeta^{-n} E_n^-(\zeta)$, *then there exists a positive integer* n_ε *such that*

$$\|\rho_n - I_p\| \le \varepsilon \quad \textit{and} \quad \|\Psi_{n-1} - I_p\|_{\mathcal{W}} \le \varepsilon \quad \textit{if } n \ge n_\varepsilon. \quad (9.15)$$

The existence of n_ε follows from items (2) and (4) in Theorem 9.2.

4. $\|\rho_n \beta_n'\| \ge (1-\varepsilon)\|\beta_n'\|$ *if* $n \ge n_\varepsilon$. In view of the bound (9.15) and Lemma 8.5, ρ_n is invertible and $(\|\rho_n^{-1}\|)^{-1} \ge 1 - \varepsilon$. Therefore,

$$\|\beta_n'\| = \|\rho_n^{-1}\rho_n\beta_n'\| \le \|\rho_n^{-1}\|\|\rho_n\beta_n'\|,$$

i.e.,

$$\|\rho_n\beta_n'\| \ge \|\rho_n^{-1}\|^{-1}\|\beta_n'\| \ge (1-\varepsilon)\|\beta_n'\| \quad \text{if } n \ge n_\varepsilon.$$

5. $\|R^*E_{n-1}^-\beta_n'\|_{\mathcal{W}} \le (1+\varepsilon)\|\beta_n'\|$ *if* $n \ge n_\varepsilon$. Clearly,

$$\begin{aligned}\|R^*E_{n-1}^-\beta_n'\|_{\mathcal{W}} &\le \|R^*E_{n-1}^-\|_{\mathcal{W}}\|\beta_n'\|\\ &= \|\Psi_{n-1}\|_{\mathcal{W}}\|\beta_n'\|\\ &= \|\Psi_{n-1} - I_p + I_p\|_{\mathcal{W}}\|\beta_n'\|\\ &\le \{\|\Psi_{n-1} - I_p\|_{\mathcal{W}} + 1\}\|\beta_n'\|,\end{aligned}$$

which is equivalent to the asserted inequality when $n \ge n_\varepsilon$ by the second bound in (9.15).

6. If $n \ge n_\varepsilon$, *then*

$$2\|D_n\beta_n'\| - \|R^*E_{n-1}^-\beta_n'\|_{\mathcal{W}} \ge (1-3\varepsilon)\|Q_0^{-1}\|^{-1}\|\beta_n\|$$

for $n = 0, 1, \ldots$.

By Steps 4 and 5, if $n \ge n_\varepsilon$, then

$$2\|\rho_n\beta_n'\| - \|R^*E_{n-1}^-\beta_n'\|_{\mathcal{W}} \ge (1-3\varepsilon)\|\beta_n'\|$$

Moreover, since V_n and U_{n-1} are unitary matrices,

$$\begin{aligned}\|\beta_n\| &= \|V_n^*\beta_n'E_{n-1}^+(0)U_{n-1}\|\\ &\le \|V_n^*\beta_n'\|\|E_n^+(0)U_{n-1}\|\\ &= \|\beta_n'\|\|E_n^+(0)\|\\ &\le \|\beta_n'\|\|Q_0^{-1}\|,\end{aligned}$$

since $E_n^+(0) = \{\gamma_{00}^{(n)}\}^{1/2} \preceq Q_0^{-1}$. The asserted conclusion drops out easily by combining the two inequalities.

7. Verifying (9.10). If we let $\kappa_\varepsilon = \frac{1}{1-3\varepsilon}\|Q_0^{-1}\|$, then it follows from Steps 2 and 6 that

$$\begin{aligned}\sum_{j=n}^{n+k}\|\beta_j\| &\le \kappa_\varepsilon\{\|R^*E_{n+k}^+E_{n+k}^+(0)^{-1}\|_{\mathcal{W}} - \|R^*E_{n-1}^+E_{n-1}^+(0)^{-1}\|_{\mathcal{W}}\}\\ &\le \kappa_\varepsilon\|R^*\|_{\mathcal{W}}\|E_{n+k}^+E_{n+k}^+(0)^{-1}\|_{\mathcal{W}}\\ &\le \kappa_\varepsilon\|R^*\|_{\mathcal{W}}\left\{\|E_{n+k}^+E_{n+k}^+(0)^{-1} - Q^{-1}Q_0\|_{\mathcal{W}} + \|Q^{-1}Q_0\|_{\mathcal{W}}\right\}.\end{aligned}$$

Thus,

$$\sum_{j=n}^{\infty}\|\beta_j\| \le \kappa_\varepsilon\|R^*\|_{\mathcal{W}}\|Q^{-1}Q_0\|_{\mathcal{W}} \quad \text{if } n \ge n_\varepsilon.$$

and hence (9.10) holds. □

10. Asymptotics for CMV matrices

Throughout this section we will assume that Δ is a density which satisfies (D1) and (D2). In view of Theorem 9.6, the Schur parameters $\{\beta_n\}_{n=0}^{\infty}$ corresponding to Δ satisfy

$$\sum_{n=0}^{\infty} \|\beta_n\| < \infty.$$

Theorem 10.1. *If $\{\beta_n\}_{n=0}^{\infty}$ are the Schur parameters of a density Δ which satisfies (D1) and (D2) and $u \in \mathbb{C}^p$, then*

$$\lim_{n\uparrow\infty} \|(Q^{-1}Q(0)^{-1} - F_n^+ Y_\infty)u\|_\Delta = 0 \tag{10.1}$$

and

$$\lim_{n\uparrow\infty} \|(R^{-*}R(0)^{-1} - \zeta^{-n} F_n^- X_\infty^*)u\|_\Delta = 0, \tag{10.2}$$

where X_∞ and Y_∞ are nonsingular matrices given in Lemma 8.7.

Proof. Assertion (10.1) follows directly from (8.13), $Y_n \to Y_\infty$ as $n \uparrow \infty$, the identification given in item (1) of Theorem 5.5 and

$$\begin{aligned}\|(Q^{-1}Q(0)^{-*} - F_n^+ Y_\infty)u\|_\Delta &\le \|(Q^{-1}Q(0)^{-*} - F_n^+ Y_n)u\|_\Delta + \|F_n^+(Y_n - Y_\infty)u\|_\Delta \\ &= \|(Q^{-1}Q(0)^{-*} - F_n^+ Y_n)u\|_\Delta + \|(Y_n - Y_\infty)u\|.\end{aligned}$$

Assertion (10.2) is shown in much the same way using (8.15) and the identification given in item (2) of Theorem 5.5. □

Theorem 10.2. *If Δ is a density that satisfies (D1) and (D2) and $\{U_n\}_{n=0}^{\infty}$ and $\{V_n\}_{n=0}^{\infty}$ are sequences of unitary matrices given in (5.25) and (5.26), respectively, then there exist unitary matrices U_∞ and V_∞ such that*

$$\lim_{n\uparrow\infty} U_n = U_\infty \tag{10.3}$$

and

$$\lim_{n\uparrow\infty} V_n = V_\infty, \tag{10.4}$$

respectively.

Proof. In view of (5.27),

$$U_n = \{\gamma_{00}^{(n)}\}^{-1/2} Y_n,$$

whence (10.3) follows easily from (8.9) and item (ii) of Lemma 8.7, which is applicable due to Theorem 9.6. The fact that U_∞ is unitary is self-evident.

The verification of (10.4) is similar. □

Definition 10.3. Given the CMV matrix $\mathfrak{A}$ based on the unitary matrix $\beta_{-1} \in \mathbb{C}^{p\times p}$ and a density Δ which satisfies (D1) and (D2), the scattering matrix Φ is given by

$$\Phi(\zeta) = \beta_{-1} R(\zeta)^* Q(\zeta)^{-1} \quad \text{for } \zeta \in \mathbb{T}. \tag{10.5}$$

Remark 10.4. The definition of Φ in formula (10.5) is motivated by asymptotics which appear in Theorem 10.5.

In Theorem 10.5 we shall let

$$\begin{bmatrix} \Psi_{2k}(\zeta) \\ \Psi_{2k+1}(\zeta) \end{bmatrix} = \begin{bmatrix} \beta_{-1} R(\zeta)^* \chi_{2k}(\zeta) \\ R(\zeta)^* \chi_{2k+1}(\zeta) \end{bmatrix} \quad \text{for } \zeta \in \mathbb{T} \tag{10.6}$$

and

$$\begin{bmatrix} \widetilde{\Psi}_{2k}(\zeta) \\ \widetilde{\Psi}_{2k+1}(\zeta) \end{bmatrix} = \begin{bmatrix} Q(\zeta) y_{2k}(\zeta) \\ Q(\zeta) y_{2k+1}(\zeta) \end{bmatrix} \quad \text{for } \zeta \in \mathbb{T}, \tag{10.7}$$

where $\{\chi_n\}_{n=0}^{\infty}$ and $\{y_n\}_{n=0}^{\infty}$ are given by (7.2) and (7.3), respectively.

Theorem 10.5. *If $\mathfrak{A}$ is the CMV matrix and $\mathfrak{C}$ is the alternative CMV matrix based on a density Δ which satisfies* (D1) *and* (D2) *and the unitary matrix $\beta_{-1} \in \mathbb{C}^{p\times p}$, $\Phi = \beta_{-1} R^* Q^{-1}$ and $u \in \mathbb{C}^p$, then the following asymptotics hold in $L_2^p(\mathbb{T}, I_p)$ norm:*

$$\Psi_{2n}(\zeta) = \zeta^{-n}\Phi(\zeta) Q(0)^{-1} Y_\infty^{-1} + o(1), \tag{10.8}$$

$$\Psi_{2n+1}(\zeta) = \zeta^{n+1} R(0)^{-1} X_\infty^{-*} + o(1), \tag{10.9}$$

$$\widetilde{\Psi}_{2n}(\zeta) = \zeta^{n}\Phi(\zeta)^* \beta_{-1} R(0)^{-*} X_\infty^{-*} u + o(1) \tag{10.10}$$

and

$$\widetilde{\Psi}_{2n+1}(\zeta) = \zeta^{-n-1} Q(0)^{-1} Y_\infty^{-1} + o(1), \tag{10.11}$$

where X_∞ and Y_∞ are as in Lemma 8.7 and $\{\Psi_n\}_{n=0}^{\infty}$ and $\{\widetilde{\Psi}_n\}_{n=0}^{\infty}$ are given by (10.6) *and* (10.7)*, respectively. Moreover, if $\beta_{-1} = I_p$, then*

$$\begin{bmatrix} \Psi_0(\zeta) & \Psi_1(\zeta) & \cdots \end{bmatrix} \mathfrak{A} = \zeta \begin{bmatrix} \Psi_0(\zeta) & \Psi_1(\zeta) & \cdots \end{bmatrix} \tag{10.12}$$

and

$$\begin{bmatrix} \widetilde{\Psi}_0(\zeta) & \widetilde{\Psi}_1(\zeta) & \cdots \end{bmatrix} \mathfrak{C} = \zeta \begin{bmatrix} \widetilde{\Psi}_0(\zeta) & \widetilde{\Psi}_1(\zeta) & \cdots \end{bmatrix}. \tag{10.13}$$

Proof. Let $\{\beta_n\}_{n=0}^{\infty}$ denote the Schur parameters of Δ. In view of Theorem 9.6, $\sum_{n=0}^{\infty} \|\beta_n\| < \infty$. Thus, X_∞ and Y_∞ exist due to Lemma 8.7.

The proof is broken into steps.

1. Verification of (10.8) *and* (10.9). It follows from (10.1) that

$$\lim_{n\uparrow\infty} \|(R^* Q^{-1} Q(0)^{-1} - R^* F_{2n}^{+} Y_\infty) u\|_{\text{st}} = 0.$$

Thus, in view of (7.2) and the unitarity of β_{-1},

$$\lim_{n\uparrow\infty} \|(\zeta^{-n}\beta_{-1} R^* Q^{-1} Q(0)^{-1} - \beta_{-1} R^* \chi_{2n} Y_\infty) u\|_{\text{st}} = 0.$$

Therefore, since $\Phi(\zeta) = \beta_{-1} R(\zeta)^* Q(\zeta)^{-1}$, (10.8) holds.

We will now verify (10.9). It follows from (10.2) that

$$\lim_{n\uparrow\infty} \|(R(0)^{-1} X_\infty^{-*} \zeta^{-(2n+1)} R^* F_{2n+1}^{-}) u\|_{\text{st}} = 0.$$

Thus, in view of (7.2) and the unitarity of β_{-1},

$$\lim_{n\uparrow\infty} \|(R(0)^{-1} X_\infty^{-*} - \zeta^{-n-1} R^* \chi_{2n+1}) u\|_{\text{st}} = 0.$$

2. *Verification of* (10.10) *and* (10.11). Assertions (10.10) and (10.11) are proved in a similar manner to (10.8) and (10.9), respectively, using (7.3).

3. *Verification of* (10.12) *and* (10.13). If $\beta_{-1} = I_p$, then, in view of (7.8) with $F(\zeta) = \Psi_n(\zeta)$,

$$\begin{aligned}\begin{bmatrix}\Psi_0(\zeta) & \Psi_1(\zeta) & \cdots\end{bmatrix}\mathfrak{A}\,\mathrm{e}_n &= \sum_{m=0}^{\infty}\Psi_m(\zeta)[\zeta\chi_n,\chi_m]_\Delta \\ &= R^*\sum_{m=0}^{\infty}\chi_m(\zeta)[\zeta\chi_n,\chi_m]_\Delta \\ &= R^*\zeta\chi_n(\zeta) \\ &= \zeta\Psi_n(\zeta).\end{aligned}$$

The proof of (10.13) is similar. □

11. Generating a positive spectral density from a summable sequence of strict contractions (the easier half of Baxter's theorem)

In this section we shall show that each sequence $\beta_0, \beta_1, \dots$ of $p\times p$ matrices with

$$\beta_0 = 0_{p\times p}, \quad \|\beta_n\| < 1 \quad \text{and} \quad \sum_{n=0}^{\infty}\|\beta_n\| < \infty$$

can be identified as the Schur parameters of exactly one density Δ which meets the constraints in (D1). This result is known, see, e.g., [12] and [21], and will provide a converse to Theorem 9.6. Given $\{\beta_n\}_{n=0}^{\infty}$, with $\beta_0 = 0_{p\times p}$ and $\|\beta_n\| < 1$ for $n = 0, 1, \dots$, define Θ_n, as in (5.2), for $n = 0, 1, \dots$ and the matrix polynomials $\{F_n^{\pm}\}_{n=0}^{\infty}$ via (5.3). Let

$$D_n(\zeta) = F_n^+(\zeta)^{-*}F_n^+(\zeta)^{-1} = F_n^-(\zeta)^{-*}F_n^-(\zeta)^{-1} \quad \text{for } \zeta \in \mathbb{T}. \tag{11.1}$$

In view of Theorem 5.5, it is easily seen that the mvf D_n satisfies (1.1) and (1.2) for $n = 0, 1, \dots$.

Theorem 11.1. *Let $\{\beta_n\}_{n=1}^{\infty}$ be a given sequence of $p\times p$ strict contractions, i.e., $\|\beta_n\| < 1$ for $n = 1, 2, \dots$. If $\sum_{n=1}^{\infty}\|\beta_n\| < \infty$, then there exists exactly one density Δ for which* (D1) *is in force with Schur parameters equal to the given sequence $\{\beta_n\}_{n=0}^{\infty}$, where $\beta_0 = 0_{p\times p}$. Moreover,*

$$\lim_{n\uparrow\infty}\|\Delta - (F_n^+)^{-*}(F_n^+)^{-1}\|_{\mathcal{W}} = 0 \tag{11.2}$$

and

$$\lim_{n\uparrow\infty}\|\Delta - (F_n^-)^{-*}(F_n^-)^{-1}\|_{\mathcal{W}} = 0. \tag{11.3}$$

Before proving Theorem 11.1, we first need some preliminary results.

Lemma 11.2. *If the matrix polynomials $\{F_n^\pm\}_{n=0}^\infty$ are defined by (5.3) in terms of a given sequence of strict contractions $\{\beta_n\}_{n=0}^\infty$ with $\beta_0 = 0_{p\times p}$, then:*

(i) *For all $0 \le j, k \le n$,*

$$[F_j^-, F_k^-]_{D_n} = \begin{cases} I_p & \text{if } j = k \\ 0_{p\times p} & \text{if } j \ne k. \end{cases}$$

(ii) *For all $0 \le j, k \le n$,*

$$[\zeta^{-j} F_j^+, \zeta^{-k} F_k^+]_{D_n} = \begin{cases} I_p & \text{if } j = k \\ 0_{p\times p} & \text{if } j \ne k. \end{cases}$$

Proof. In view of the identification given in item (5) of Theorem 5.5,

$$D_n(\zeta) = \frac{C_n(\zeta) + C_n(\zeta)^*}{2} \quad \text{for } \zeta \in \mathbb{T}, \tag{11.4}$$

where C_n is the mvf given by (5.12). In view of (11.4) and the chain of equalities, beginning at (5.21), carried out in Step 2 of the proof of Theorem 5.6, (i) holds. The proof of (ii) is carried out in much the same way using the chain of equalities in Step 4 of the proof of Theorem 5.6. □

Lemma 11.3. *The matrix polynomials $\{F_n^\pm\}_{n=0}^\infty$ defined by (5.3) in terms of a given sequence of strict contractions $\{\beta_n\}_{n=0}^\infty$ with $\beta_0 = 0_{p\times p}$ and*

$$\|\beta_n\| \le \rho < 1 \quad \text{for } n = 0, 1, \ldots$$

are subject to the bounds

$$\exp\left\{-\left(\frac{2}{1-\rho}\right)\sum_{j=1}^n \|\beta_j\|\right\} \le \|F_n^\pm(\zeta)\| \le \exp\left\{\frac{2}{1-\rho}\sum_{j=1}^n \|\beta_j\|\right\} \tag{11.5}$$

for $\zeta \in \mathbb{T}$ and $n = 0, 1, \ldots$.

Proof. The proof is broken into steps.

1. Verification of the bounds

$$\left\{\frac{1-\|\beta_{n+1}\|}{1+\|\beta_{n+1}\|}\right\}\|F_n^-(\zeta)\| \le \|F_{n+1}^-(\zeta)\| \le \|F_n^-(\zeta)\|\left\{\frac{1+\|\beta_{n+1}\|}{1-\|\beta_{n+1}\|}\right\} \tag{11.6}$$

for $\zeta \in \mathbb{T}$ and $n = 0, 1, \ldots$.

The recursion (5.4) implies

$$\begin{aligned} F_{n+1}^-(\zeta) &= (\zeta F_n^-(\zeta) + F_n^+(\zeta)\beta_{n+1}^*)(I_p - \beta_{n+1}\beta_{n+1}^*)^{-1/2} \\ &= F_n^-(\zeta)\{\zeta I_p + F_n^-(\zeta)^{-1}F_n^+(\zeta)\beta_{n+1}^*\}(I_p - \beta_{n+1}\beta_{n+1}^*)^{-1/2} \end{aligned} \tag{11.7}$$

and hence, as $F_n^-(\zeta)^{-1}F_n^+(\zeta)$ is unitary for $\zeta\in\mathbb{T}$,

$$\|F_{n+1}^-(\zeta)\|\le\|F_n^-(\zeta)\|(1+\|\beta_{n+1}\|)\|(I_p-\beta_{n+1}\beta_{n+1}^*)^{-1/2}\|.$$

Let $\beta_{n+1}=USV^*$ be the singular value decomposition for β_{n+1}, where U and V are unitary, $S=\mathrm{diag}(s_1,\dots,s_p)$ and $s_1\ge\cdots\ge s_p\ge 0$. Then,

$$(I_p-\beta_{n+1}\beta_{n+1}^*)^{-1/2}=U(I_p-S^2)^{-1/2}U^*$$

and hence

$$\|(I_p-\beta_{n+1}\beta_{n+1}^*)^{-1/2}\|=\frac{1}{\sqrt{1-s_1^2}}=\frac{1}{\sqrt{1-\|\beta_{n+1}\|^2}}.$$

Thus,

$$\begin{aligned}\|F_{n+1}^-(\zeta)\|&\le\|F_n^-(\zeta)\|\left\{\frac{1+\|\beta_{n+1}\|}{\sqrt{1-\|\beta_{n+1}\|^2}}\right\}\\&=\|F_n^-(\zeta)\|\left\{\frac{(1+\|\beta_n\|)^2}{1-\|\beta_{n+1}\|^2}\right\}^{1/2}\\&\le\|F_n^-(\zeta)\|\left\{\frac{1+\|\beta_{n+1}\|}{1-\|\beta_{n+1}\|}\right\},\end{aligned}$$

which justifies the upper bound in (11.6).

The equality (11.7) implies that

$$F_n^-(\zeta)=F_{n+1}^-(\zeta)(I_p-\beta_{n+1}\beta_{n+1}^*)^{1/2}(\zeta I_p-F_n^-(\zeta)^{-1}F_n^+(\zeta)\beta_{n+1}^*)^{-1}\tag{11.8}$$

for $\zeta\in\mathbb{T}$. Since

$$\begin{gathered}(I_p-\beta_{n+1}\beta_{n+1}^*)^{1/2}=U(I-S^2)^{1/2}U^*,\\\|(I_p-\beta_{n+1}\beta_{n+1}^*)^{1/2}\|<1\le 1+\|\beta_{n+1}\|.\end{gathered}$$

Thus, upon invoking the inequality

$$\|(I_p-A)^{-1}\|\le\frac{1}{1-\|A\|}$$

for $A\in\mathbb{C}^{p\times p}$ with $\|A\|<1$ and the unitarity of $F_n^-(\zeta)^{-1}F_n^+(\zeta)$,

$$\|F_n^-(\zeta)\|\le\|F_{n+1}^-(\zeta)\|\left\{\frac{1+\|\beta_{n+1}\|}{1-\|\beta_{n+1}\|}\right\},$$

which justifies the lower bound in (11.6).

2. Verification of (11.5). The bounds

$$\prod_{j=0}^{n}\left\{\frac{1-\|\beta_{j+1}\|}{1+\|\beta_{j+1}\|}\right\}\|F_0^-(\zeta)\| \le \|F_{n+1}^-(\zeta)\| \le \|F_0^-(\zeta)\|\prod_{j=0}^{n}\left\{\frac{1+\|\beta_{j+1}\|}{1-\|\beta_{j+1}\|}\right\} \tag{11.9}$$

follow from (11.6). In view of $F_0^-(\zeta) = I_p$ and (8.6), the bounds for $\|F_{n+1}^-(\zeta)\|$ advertised in (11.5) are readily obtained from (11.9). The bounds for $\|F_{n+1}^+(\zeta)\|$ in (11.5) follow from (11.1). □

Theorem 11.4. *If the matrix polynomials $\{F_n^\pm\}_{n=0}^\infty$ are defined by* (5.3) *in terms of a given sequence of strict contractions $\{\beta_n\}_{n=0}^\infty$ with $\beta_0 = 0_{p\times p}$ and $\sum_{n=0}^\infty \|\beta_n\| < \infty$, then:*

(i) *There exists a mvf $A \in \mathcal{W}_+^{p\times p}$ such that $A^{-1} \in \mathcal{W}_+^{p\times p}$,*

$$\|F_n^+ - A\|_{\mathcal{W}} \to 0 \quad \textit{as } n \uparrow \infty \tag{11.10}$$

and

$$\|(F_n^+)^{-1} - A^{-1}\|_{\mathcal{W}} \to 0 \quad \textit{as } n \uparrow \infty. \tag{11.11}$$

(ii) *There exists a mvf $B \in \mathcal{W}_+^{p\times p}$ such that $B^{-1} \in \mathcal{W}_+^{p\times p}$,*

$$\|\zeta^n (F_n^-)^* - B\|_{\mathcal{W}} \to 0 \quad \textit{as } n \uparrow \infty \tag{11.12}$$

and

$$\|\zeta^{-n}(F_n^-)^{-*} - B^{-1}\|_{\mathcal{W}} \to 0 \quad \textit{as } n \uparrow \infty. \tag{11.13}$$

Proof. The proof is broken into steps. Let $\{X_n\}_{n=0}^\infty$ and $\{Y_n\}_{n=0}^\infty$ be given by (5.6) and (5.7), respectively.

1. If $k_n = \|F_n^+ Y_n^{-1}\|_{\mathcal{W}}$ and $\ell_n = \|F_n^- X_n^{-1}\|_{\mathcal{W}}$, then $\{k_n\}_{n=0}^\infty$ and $\{\ell_n\}_{n=0}^\infty$ are bounded. Using the recursion (5.4), it is readily checked that

$$F_{n+1}^-(\zeta)X_{n+1}^{-1} = \{\zeta F_n^-(\zeta) + F_n^+(\zeta)\beta_{n+1}^*\}X_n^{-1} \tag{11.14}$$

and

$$F_{n+1}^+(\zeta)Y_{n+1}^{-1} = \{\zeta F_n^-(\zeta)\beta_{n+1} + F_n^+(\zeta)\}Y_n^{-1}. \tag{11.15}$$

Consequently,

$$\|F_{n+1}^- X_{n+1}^{-1}\|_{\mathcal{W}} \le \|F_n^- X_n^{-1}\|_{\mathcal{W}} + \|F_n^+ \beta_{n+1}^* X_n^{-1}\|_{\mathcal{W}}$$

and

$$\|F_{n+1}^+ Y_{n+1}^{-1}\|_{\mathcal{W}} \le \|F_n^- \beta_{n+1} Y_n^{-1}\|_{\mathcal{W}} + \|F_n^+ Y_n^{-1}\|_{\mathcal{W}}.$$

Thus, if we let

$$a_n = \|F_n^- X_n^{-1}\|_{\mathcal{W}} + \|F_n^+ Y_n^{-1}\|_{\mathcal{W}} \quad \text{for } n = 0, 1, \dots,$$

then

$$\begin{aligned} a_{n+1} &\le a_n + \|F_n^+\beta_{n+1}^* X_n^{-1}\|_{\mathcal{W}} + \|F_n^-\beta_{n+1}Y_n^{-1}\|_{\mathcal{W}} \\ &\le a_n + \|F_n^+Y_n^{-1}\|_{\mathcal{W}}\|Y_n\|\|\beta_{n+1}\|\|X_n^{-1}\| + \|F_n^-X_n^{-1}\|_{\mathcal{W}}\|X_n\|\|\beta_{n+1}\|\|Y_n^{-1}\| \\ &\le a_n(1+\kappa\|\beta_{n+1}\|), \end{aligned} \tag{11.16}$$

where, in view of Lemma 8.7,

$$\kappa = \sup_{n=0,1,\dots}\{\|Y_n\|\|X_n^{-1}\|, \|Y_n^{-1}\|\|X_n\|\} < \infty.$$

Iterating the bound (11.16), since $a_0 = 1$, it is easily seen that,

$$\begin{aligned} k_{n+1} + \ell_{n+1} = a_{n+1} &\le a_0 \prod_{j=0}^{n}(1+\kappa\|\beta_{j+1}\|) \\ &\le \exp\left\{\kappa\sum_{j=0}^{n}\|\beta_{j+1}\|\right\}, \end{aligned}$$

which serves to justify the boundedness of the sequences $\{k_n\}_{n=0}^{\infty}$ and $\{\ell_n\}_{n=0}^{\infty}$, since $\sum_{j=0}^{\infty}\|\beta_j\| < \infty$.

2. $\{F_n^+Y_n^{-1}\}_{n=0}^{\infty}$ is a Cauchy sequence in $\mathcal{W}_+^{p\times p}$. If $m > n \ge 0$, then, in view of (11.15),

$$\begin{aligned} \|F_m^+Y_m^{-1} - F_n^+Y_n^{-1}\|_{\mathcal{W}} &\le \sum_{j=n}^{m-1}\|F_j^-\beta_{j+1}Y_j^{-1}\|_{\mathcal{W}} \\ &= \sum_{j=n}^{m-1}\|F_j^-X_j^{-1}X_j\beta_{j+1}Y_j^{-1}\|_{\mathcal{W}} \\ &\le \sum_{j=n}^{m-1}\|F_j^-X_j^{-1}\|_{\mathcal{W}}\|X_j\|\|\beta_{j+1}\|\|Y_j^{-1}\| \\ &\le \tilde{\kappa}\sum_{j=n}^{m-1}\|\beta_{j+1}\|, \end{aligned}$$

where, in view of Step 1 and Lemma 8.7,

$$\tilde{\kappa} = \sup_{j=0,1,\dots}\{a_j\|X_j\|\|Y_j^{-1}\|\} < \infty,$$

Thus, as

$$\sum_{j=n}^{m-1}\|\beta_{j+1}\| \to 0 \quad \text{as } m,n\uparrow\infty,$$

Step 2 holds.

3. $\{F_n^+\}_{n=0}^{\infty}$ *is a Cauchy sequence in* $\mathcal{W}_+^{p\times p}$. This follows from the sequence of inequalities

$$\begin{aligned}\|F_m^+ - F_n^+\|_{\mathcal{W}} &= \|F_m^+ Y_m^{-1} Y_m - F_n^+ Y_n^{-1} Y_n\|_{\mathcal{W}} \\ &= (F_m^+ Y_m^{-1} - F_n^+ Y_n^{-1}) Y_m + F_n^+ Y_n^{-1}(Y_m - Y_n)\|_{\mathcal{W}} \\ &\le \|F_m^+ Y_m^{-1} - F_n^+ Y_n^{-1}\|_{\mathcal{W}} \|Y_m\| + \|F_n^+ Y_n^{-1}\|_{\mathcal{W}} \|Y_m - Y_n\|\end{aligned}$$

and Step 2. Therefore, there exists $A \in \mathcal{W}_+^{p\times p}$ such that (11.10) holds.

4. Verification of $A^{-1} \in \mathcal{W}_+^{p\times p}$. In view of Lemma 11.3,

$$\exp\left\{-\left(\frac{2}{1-\rho}\right)\sum_{j=1}^{\infty}\|\beta_j\|\right\} \le \|F_n^-(\zeta)\| \le \exp\left\{\frac{2}{1-\rho}\sum_{j=1}^{\infty}\|\beta_j\|\right\},$$

where ρ as in the statement of Lemma 11.3. Consequently, there exists a subsequence $\{n_k\}_{k=0}^{\infty}$ of $\{0, 1, \ldots\}$ and a mvf P on $\overline{\mathbb{D}}$ such that

$$\lim_{k\uparrow\infty} \|F_{n_k}^+(\lambda)^{-1} - P(\lambda)\| = 0 \quad \text{at each point } \lambda \in \overline{\mathbb{D}}. \tag{11.17}$$

We claim that $P(\lambda) = A(\lambda)^{-1}$. In view of (11.17),

$$\begin{aligned}\|I_p - A(\lambda)P(\lambda)\| &= \|F_{n_k}^+(\lambda)F_{n_k}^+(\lambda)^{-1} - A(\lambda)P(\lambda)\| \\ &\le \|F_{n_k}^+(\lambda)\|\|F_{n_k}^+(\lambda)^{-1} - P(\lambda)\| + \|F_{n_k}^+(\lambda) - A(\lambda)\|\|P(\lambda)\| \\ &\to 0 \quad \text{as } k \uparrow \infty.\end{aligned}$$

Therefore, as $A(\lambda)$ is invertible for all $\lambda \in \overline{\mathbb{D}}$, it follows from item (2) of Theorem 5.1 that $A^{-1} \in \mathcal{W}_+^{p\times p}$.

5. Verification of (11.11). Since

$$\begin{aligned}\|(F_n^+)^{-1} - A^{-1}\|_{\mathcal{W}} &= \|(F_n^+)^{-1}(A - F_n^+)A^{-1}\|_{\mathcal{W}} \\ &\le \|(F_n^+)^{-1}\|_{\mathcal{W}}\|A - F_n^+\|_{\mathcal{W}}\|A^{-1}\|_{\mathcal{W}},\end{aligned}$$

it suffices to show that $\|(F_n^+)^{-1}\|_{\mathcal{W}}$ is bounded. But, in view of (11.10),

$$\lim_{n\uparrow\infty} \|F_n^+ A^{-1} - I_p\|_{\mathcal{W}} = 0. \tag{11.18}$$

Thus, using Lemma 8.5 we have any $0 < \varepsilon < 1$,

$$\|A(F_n^+)^{-1}\|_{\mathcal{W}} < \frac{1}{1-\varepsilon} \quad \text{for } n \text{ sufficiently large} \tag{11.19}$$

and hence

$$\begin{aligned}\|(F_n^+)^{-1}\|_{\mathcal{W}} &= \|A^{-1}A(F_n^+)^{-1}\|_{\mathcal{W}} \\ &\le \|A^{-1}\|_{\mathcal{W}}\|A(F_n^+)^{-1}\|_{\mathcal{W}} \\ &\le \left(\frac{1}{1-\varepsilon}\right)\|A^{-1}\|_{\mathcal{W}} \quad \text{for } n \text{ sufficiently large.}\end{aligned}$$

6. Verification of (ii). The verification of (ii) is carried out in much the same way as the verification of (i) in Steps 1–5. We will outline the major steps. In view of (11.14) and (11.15),

$$\|\zeta^{-(n+1)}F_{n+1}^{-}X_{n+1}^{-1}\|_{\mathcal{W}} \le \|\zeta^{-n}F_n^{-}X_n^{-1}\|_{\mathcal{W}} + \|F_n^{+}\beta_{n+1}^{*}X_n^{-1}\|_{\mathcal{W}}$$

and
$$\|F_{n+1}^{+}Y_{n+1}^{-1}\|_{\mathcal{W}} \le \|\zeta^{-n}F_n^{-}\beta_{n+1}Y_n^{-1}\|_{\mathcal{W}} + \|F_n^{+}Y_n^{-1}\|_{\mathcal{W}},$$

which can be used to show that $\{\zeta^{-n}F_n^{-}X_n^{-1}\}_{n=0}^{\infty}$ is a Cauchy sequence in $\mathcal{W}_{-}^{p\times p}$. Consequently, one can show that $\{\zeta^{-n}F_n^{-}\}_{n=0}^{\infty}$ is also a Cauchy sequence in $\mathcal{W}_{-}^{p\times p}$. Thus, there exists a mvf $B\in\mathcal{W}_{+}^{p\times p}$ such that (11.12) holds. The verification of $B^{-1}\in\mathcal{W}_{+}^{p\times p}$ and (11.13) is completed in a way similar to Step 3 and 4, respectively. □

We are now ready to prove Theorem 11.1.

Proof of Theorem 11.1. The proof is broken into steps.

1. There exists a density Δ *which satisfies* (1.1), (1.2), (11.2) *and* (11.3). In view of items (i) and (ii) in Theorem 11.4, there exist mvf's A and B with $A^{\pm 1}$ and $B^{\pm 1}$ both belonging to $\mathcal{W}_{+}^{p\times p}$ such that (11.10)–(11.13) hold. If

$$\Delta(\zeta) = A(\zeta)^{-*}A(\zeta)^{-1} \quad \text{for } \zeta\in\mathbb{T}, \tag{11.20}$$

then (11.2) follows from (11.11) and the bound

$$\begin{aligned}&\|A^{-*}A^{-1} - (F_n^{+})^{-*}(F_n^{+})^{-1}\|_{\mathcal{W}}\\ &\qquad \le \|A^{-*}\|_{\mathcal{W}}\|A^{-1}-(F_n^{+})^{-1}\|_{\mathcal{W}} + \|A^{-*}-(F_n^{+})^{-*}\|_{\mathcal{W}}\|(F_n^{+})^{-1}\|_{\mathcal{W}}.\end{aligned}$$

The limit (11.3) follows from (11.2) and (11.1).

By an argument based on (11.13) that is similar to verification of (11.20), it follows that

$$\lim_{n\uparrow\infty}\|B^{-1}B^{-*} - (\zeta^{-n}F_n^{-})^{-*}(\zeta^{-n}F_n^{-})^{-1}\|_{\mathcal{W}} = 0$$

and hence that, in addition to (11.20), Δ admits the second factorization

$$\Delta(\zeta) = B(\zeta)^{-1}B(\zeta)^{-*} \quad \text{for } \zeta\in\mathbb{T}. \tag{11.21}$$

2. The matrix polynomials $\{F_n^{\pm}\}_{n=0}^{\infty}$ *satisfy*

$$[F_j^{-}, F_k^{-}]_{\Delta} = \begin{cases} I_p & \text{if } j = k\\ 0_{p\times p} & \text{if } j\neq k,\end{cases} \tag{11.22}$$

$$[\zeta^{-j}F_j^{+}, \zeta^{-k}F_k^{+}]_{\Delta} = \begin{cases} I_p & \text{if } j = k\\ 0_{p\times p} & \text{if } j\neq k\end{cases} \tag{11.23}$$

and (1.3).

If $0 \le j, k \le n$, then item (i) in Lemma 11.2 guarantees that

$$\frac{1}{2\pi}\int_0^{2\pi} F_k^-(e^{i\theta})^* D_n(e^{i\theta}) F_j^-(e^{i\theta})d\theta = [F_j^-, F_k^-]_{D_n} \tag{11.24}$$
$$= \begin{cases} I_p & \text{if } j = k \\ 0_{p\times p} & \text{if } j \neq k. \end{cases}$$

Thus, as

$$\|(F_k^-)^*(D_n - \Delta)F_j^-\|_{\mathcal{W}} \le \|F_k^-\|_{\mathcal{W}}\|D_n - \Delta\|_{\mathcal{W}}\|F_j^-\|_{\mathcal{W}}$$

tends to 0 as $n \uparrow \infty$, the limit as $n \uparrow \infty$ can be brought inside the integral to obtain

$$\frac{1}{2\pi}\int_0^{2\pi} F_k^-(e^{i\theta})^* \Delta(e^{i\theta}) F_j^-(e^{i\theta})d\theta = \begin{cases} I_p & \text{if } j = k \\ 0_{p\times p} & \text{if } j \neq k. \end{cases} \tag{11.25}$$

Thus, (11.22) holds.

The proof of (11.23) is completed in much the same way from item (ii) in Lemma 11.2. In view of $F_0^-(\zeta) = I_p$, (1.3) follows from (11.25).

3. The Schur parameters of Δ are equal to $\{\beta_n\}_{n=0}^{\infty}$. Let $\{\alpha_n\}_{n=0}^{\infty}$ denote the Schur parameters of Δ. In view of (11.22) and (11.23), we may use Theorem 5.8 to obtain sequences of unitary matrices $\{M_n\}_{n=0}^{\infty}$ and $\{N_n\}_{n=0}^{\infty}$ so that

$$F_n^+(\lambda) = E_n^+(\lambda)M_n \tag{11.26}$$

and

$$F_n^-(\lambda) = E_n^-(\lambda)N_n. \tag{11.27}$$

Consequently, the matrix polynomials $\{F_n^{\pm}\}_{n=0}^{\infty}$ generated by $\{\beta_n\}_{n=0}^{\infty}$,

$$\begin{bmatrix} F_{n+1}^-(\lambda) & F_{n+1}^+(\lambda) \end{bmatrix} = \begin{bmatrix} \lambda F_n^-(\lambda) & F_n^+(\lambda) \end{bmatrix} H(\beta_{n+1})$$

can be written as

$$\begin{bmatrix} E_{n+1}^-(\lambda) & E_{n+1}^+(\lambda) \end{bmatrix} = \begin{bmatrix} \lambda E_n^-(\lambda) & E_n^+(\lambda) \end{bmatrix} \begin{bmatrix} N_n & 0 \\ 0 & M_n \end{bmatrix} H(\beta_{n+1}) \begin{bmatrix} N_{n+1}^* & 0 \\ 0 & M_{n+1}^* \end{bmatrix}.$$

Thus, as the last recursion can also be written in terms of the Schur parameters $\{\alpha_n\}_{n=0}^{\infty}$ of Δ as

$$\begin{bmatrix} E_{n+1}^-(\lambda) & E_{n+1}^+(\lambda) \end{bmatrix} = \begin{bmatrix} \lambda E_n^-(\lambda) & E_n^+(\lambda) \end{bmatrix} \begin{bmatrix} V_n & 0 \\ 0 & U_n \end{bmatrix} H(\alpha_{n+1}) \begin{bmatrix} V_{n+1}^* & 0 \\ 0 & U_{n+1}^* \end{bmatrix},$$

$$\begin{bmatrix} N_n & 0 \\ 0 & M_n \end{bmatrix} H(\beta_{n+1}) \begin{bmatrix} N_{n+1}^* & 0 \\ 0 & M_{n+1}^* \end{bmatrix} = \begin{bmatrix} V_n & 0 \\ 0 & U_n \end{bmatrix} H(\alpha_{n+1}) \begin{bmatrix} V_{n+1}^* & 0 \\ 0 & U_{n+1}^* \end{bmatrix}$$

must hold for $n = 0, 1, \dots$. Therefore, by the uniqueness of the polar decomposition, $Z_1 = V_1$, and, continuing by induction,

$$Z_n = V_n \quad \text{for } n = 0, 1, \dots.$$

In much the same way, using (5.30), one can show that

$$M_n = U_n \quad \text{for } n = 0, 1, \ldots .$$

Therefore,

$$H(\alpha_{n+1}) = H(\beta_{n+1}) \quad \text{for } n = 0, 1, \ldots$$

and hence

$$\alpha_{n+1} = \beta_{n+1} \quad \text{for } n = 0, 1, \ldots .$$

4. Δ is unique. The asserted uniqueness has been established in Theorem 5.10. □

12. Generating a positive spectral density from a unitary operator of the form (7.12)

Let β_{-1} be a $p \times p$ unitary matrix and $\{\beta_n\}_{n=1}^{\infty}$ be a sequence of $p \times p$ strict contractions which satisfy

$$\sum_{n=1}^{\infty} \|\beta_n\| < \infty.$$

Let $\{\mathfrak{u}_n\}_{n=1}^{\infty}$ be a sequence of $2p \times 2p$ unitary matrices given by

$$\mathfrak{u}_n = \begin{bmatrix} -\beta_n^* & (I_p - \beta_n^*\beta_n)^{1/2} \\ (I_p - \beta_n\beta_n^*)^{1/2} & \beta_n \end{bmatrix} \quad \text{for } n = 1, 2, \ldots .$$

and

$$\mathfrak{U} = \mathfrak{U}_{\mathrm{odd}}\mathfrak{U}_{\mathrm{even}}\mathfrak{V}, \tag{12.1}$$

where

$$\mathfrak{U}_{\mathrm{odd}} = \begin{bmatrix} u_1 & 0 & 0 & \cdots \\ 0 & u_3 & 0 & \cdots \\ 0 & 0 & u_5 & \\ \vdots & \vdots & & \ddots \end{bmatrix}, \quad \mathfrak{U}_{\mathrm{even}} = \begin{bmatrix} I_p & 0 & 0 & \cdots \\ 0 & u_2 & 0 & \cdots \\ 0 & 0 & u_4 & \\ \vdots & \vdots & & \ddots \end{bmatrix}$$

and

$$\mathfrak{V} = \begin{bmatrix} \beta_{-1} & 0 & 0 & \cdots \\ 0 & I_p & 0 & \cdots \\ 0 & 0 & I_p & \\ \vdots & \vdots & & \ddots \end{bmatrix}.$$

Since $\mathfrak{u}_n$ is unitary for $n = 1, 2, \ldots$, it is readily checked that $\mathfrak{U} : \ell_2^p \to \ell_2^p$ is unitary.

Theorem 12.1. *Let $\{\beta_n\}_{n=1}^{\infty}$ be a sequence of $p \times p$ strict contractions with*

$$\sum_{n=1}^{\infty} \|\beta_n\| < \infty.$$

If $\mathfrak{U}$ is the unitary operator on ℓ_2^p given by (12.1), then there is exactly one density Δ meeting the constraint (D1) and one unitary matrix $\beta_{-1} \in \mathbb{C}^{p\times p}$ with the property that $\mathfrak{U}$ is the CMV matrix based on the mvf Δ and β_{-1} (see Definition 7.2).

Proof. Let $\beta_0 = 0_{p\times p}$. If $\{\beta_n\}_{n=0}^{\infty}$ satisfies $\sum_{n=0}^{\infty} \|\beta_n\| < \infty$, then, using Theorem 11.1, there exists exactly one density Δ which satisfies (D1) and the Schur parameters of Δ are given by $\{\beta_n\}_{n=0}^{\infty}$. Let $\mathfrak{A}$ denote the CMV based matrix based on Δ and β_{-1}. It follows from Theorem 7.6 that $\mathfrak{A} = \mathfrak{U}$. The fact that there is only exactly one density Δ which meets (D1) and exactly one unitary matrix β_{-1} so that $\mathfrak{U}$ is the CMV matrix based on Δ and β_{-1} follows from Theorem 7.8. $\square$

13. A Nehari problem

In this section a number of important connections with a Nehari problem in $\mathcal{W}^{p\times p}$ are summarized. Most of the facts follow from the fundamental study of the Nehari problem in a general setting by Adamjan, Arov and Krein [2]. For the convenience of the reader proofs that are adapted mostly from [2] to the present simpler setting are presented in Appendix C. However, Theorem 13.5 and Corollary 13.6 are based on the work of Treil and Volberg [35].

Let $\mathfrak{W}^{p\times p}$ denote the set of $\Phi \in \mathcal{W}^{p\times p}$ on $\mathbb{T}$ for which the Hankel operator

$$\widehat{\Gamma}_{\Phi} = \mathfrak{q} M_{\Phi}|_{H_2^p}$$

is strictly contractive, i.e., $\|\widehat{\Gamma}_{\Phi}\| < 1$, and let $\mathcal{N}(\Phi)$ denote the set of all mvf's $\Psi \in \mathcal{W}^{p\times p}$ with $\|\Psi(\zeta)\| \le 1$ for every point $\zeta \in \mathbb{T}$ for which

$$\Phi - \Psi \in \mathcal{W}_+^{p\times p}. \tag{13.1}$$

It is readily checked that

$$\Psi \in \mathcal{N}(\Phi) \Longleftrightarrow \mathfrak{q}\Psi f = \mathfrak{q}\Phi f \quad \text{for every } f \in H_2^p.$$

In order to simplify the typography we shall abuse notation a little and shall allow operators that act on $p \times 1$ vvf's to act on $p \times k$ mvf's with the understanding that they act column by column. Thus, for example, if

$$F = \begin{bmatrix} f_1 & \cdots & f_k \end{bmatrix} \in H_2^{p\times k},$$

then

$$\widehat{\Gamma}_{\Phi} F \quad \text{is interpreted as} \quad \begin{bmatrix} \widehat{\Gamma}_{\Phi} f_1 & \cdots & \widehat{\Gamma}_{\Phi} f_k \end{bmatrix}. \tag{13.2}$$

The main result will be to parameterize $\mathcal{N}(\Phi)$ in terms of a linear fractional transformation

$$T_\Theta[\mathcal{E}] = (\theta_{11}\mathcal{E} + \theta_{12})(\theta_{21}\mathcal{E} + \theta_{22})^{-1}$$

based on a $2p \times 2p$ mvf

$$\Theta = \begin{bmatrix} \theta_{11} & \theta_{12} \\ \theta_{21} & \theta_{22} \end{bmatrix} \tag{13.3}$$

with blocks

$$\theta_{11} \in \mathcal{W}_-^{p\times p},\ \zeta\theta_{12} \in \mathcal{W}_-^{p\times p},\ \zeta^{-1}\theta_{21} \in \mathcal{W}_+^{p\times p} \quad \text{and} \quad \theta_{22} \in \mathcal{W}_+^{p\times p} \tag{13.4}$$

that will be defined in terms of the $p \times p$ positive definite matrices

$$M = [(I - \widehat{\Gamma}_\Phi^* \widehat{\Gamma}_\Phi)^{-1} I_p, I_p]_{\text{st}} \quad \text{and} \quad N = [(I - \widehat{\Gamma}_\Phi \widehat{\Gamma}_\Phi^*)^{-1} \zeta^{-1} I_p, \zeta^{-1} I_p]_{\text{st}} \tag{13.5}$$

when $\Phi \in \mathfrak{W}^{p\times p}$.

Theorem 13.1. *If $\Phi \in \mathfrak{W}^{p\times p}$, then there exists exactly one mvf Θ of the form indicated in (13.3) such that*

$$\begin{bmatrix} I & -\widehat{\Gamma}_\Phi \\ -\widehat{\Gamma}_\Phi^* & I \end{bmatrix} \begin{bmatrix} \zeta^{-1}\theta_{11} & \theta_{12} \\ \zeta^{-1}\theta_{21} & \theta_{22} \end{bmatrix} = \begin{bmatrix} \zeta^{-1} N^{-1/2} & 0 \\ 0 & M^{-1/2} \end{bmatrix}. \tag{13.6}$$

Moreover,

(1) $\theta_{11}^\#(0) = N^{1/2}$, $\theta_{12}^\#(0) = 0_{p\times p}$, $\theta_{21}(0) = 0_{p\times p}$ *and* $\theta_{22}(0) = M^{1/2}$.

(2) Θ *is* j_p *unitary on* $\mathbb{T}$, *i.e.,*

$$\Theta(\zeta)^* j_p \Theta(\zeta) = \Theta(\zeta) j_p \Theta(\zeta)^* = j_p \quad \text{for } \zeta \in \mathbb{T}. \tag{13.7}$$

(3) $T_\Theta[\tau]$ *is unitary on* $\mathbb{T}$ *for every mvf* $\tau \in \mathcal{W}^{p\times p}$ *such that* $\tau(\zeta)\tau(\zeta)^* = I_p$ *for* $\zeta \in \mathbb{T}$.

(4) $\|\theta_{11}(\zeta)^{-1}\theta_{12}(\zeta)\| \le \delta < 1$ *and* $\|\theta_{22}(\zeta)^{-1}\theta_{21}(\zeta)\| \le \varepsilon < 1$ *for* $\zeta \in \mathbb{T}$.

(5) $(\theta_{21}\mathcal{E} + \theta_{22})^{\pm 1} \in \mathcal{W}_+^{p\times p}$ *if* $\mathcal{E} \in \mathcal{S}^{p\times p} \cap \mathcal{W}_+^{p\times p}$.

(6) $(\theta_{11} + \theta_{12}\mathcal{E}^*)^{\pm 1} \in \mathcal{W}_-^{p\times p}$ *if* $\mathcal{E} \in \mathcal{S}^{p\times p} \cap \mathcal{W}_+^{p\times p}$.

Proof. The proofs of items (1)–(4) are presented in Subsection D.2; (5) and (6) are verified in Subsection D.5. □

Let $\Phi(\zeta) = \sum_{n=-\infty}^{\infty} \zeta^n \Phi_n$ belong to $\mathfrak{W}^{p\times p}$ with Φ_0 in the matrix ball

$$\{N^{-1/2}\beta M^{-1/2} + C_0 : \beta \in \mathbb{C}^{p\times p} \quad \text{and} \quad \beta^*\beta \preceq I_p\}, \tag{13.8}$$

with center

$$C_0 = -(\widehat{\Gamma}_\Phi \mathfrak{p} \zeta^{-1} (I - \widehat{\Gamma}_\Phi^* \widehat{\Gamma}_\Phi)^{-1} M^{-1})_0. \tag{13.9}$$

We now introduce a second Hankel operator

$$\widehat{G}_\Phi^{(\beta)} f = \widetilde{\mathfrak{q}} \Phi f = \zeta \mathfrak{q} \zeta^{-1} \Phi f \quad \text{for} \quad f \in H_2^p, \tag{13.10}$$

in which $\widetilde{\mathfrak{q}}$ denotes the orthogonal projection of L_2^p onto $\zeta(H_2^p)^\perp$ and $\beta \in \mathbb{C}^{p\times p}$ is a contraction.

The operators

$$(G_\Phi^{(\beta)}\xi)_j = \sum_{k=0}^{\infty} \gamma_{j+k}\xi_k \quad \text{for } j = 0, 1, \dots \tag{13.11}$$

and

$$(\Gamma_\Phi\xi)_j = \sum_{k=0}^{\infty} \gamma_{j+k+1}\xi_k \quad \text{for } j = 0, 1, \dots, \tag{13.12}$$

with

$$\gamma_j = \Phi_{-j} \quad \text{for } j = 0, 1, \dots, \tag{13.13}$$

are the counterparts in ℓ_2^p of the operators $\widehat{G}_\Phi^{(\beta)}$ and $\widehat{\Gamma}_\Phi$, respectively.

It is readily checked that

$$\|\widehat{G}_\Phi^{(\beta)}\| = \|G_\Phi^{(\beta)}\| \quad \text{and} \quad \|\widehat{\Gamma}_\Phi\| = \|\Gamma_\Phi\|.$$

Moreover,

$$G_\Phi^{(\beta)}\xi = Y\gamma_0 Y^*\xi + T\Gamma_\Phi + YY^*\Gamma_\Phi T^*\xi \quad \text{for } \xi \in \ell_2^p \tag{13.14}$$

and

$$\gamma_0 = -Y^*\Gamma_\Phi T^*(I - \Gamma_\Phi^*\Gamma_\Phi)^{-1}YM^{-1} + N^{-1/2}\beta M^{-1/2}, \tag{13.15}$$

where

$$\begin{aligned} Y\xi &= \mathrm{col}(\xi, 0, \dots, 0) \in \ell_2^p \quad \text{for } \xi \in \mathbb{C}^p, \\ Y^*\xi &= \xi_0 \in \mathbb{C}^p \quad \text{for } \xi = \mathrm{col}(\xi_0, \xi_1, \dots) \in \ell_2^p, \\ T\xi &= \mathrm{col}(0, \xi_0, \xi_1, \dots) \in \ell_2^p \quad \text{for } \xi = \mathrm{col}(\xi_0, \xi_1, \dots) \in \ell_2^p, \end{aligned}$$

and

$$T^*\xi = \mathrm{col}(\xi_1, \xi_2, \dots) \in \ell_2^p \quad \text{for } \xi = \mathrm{col}(\xi_0, \xi_1, \dots) \in \ell_2^p.$$

Theorem 13.2. *If $\Phi \in \mathfrak{W}^{p\times p}$, Θ is specified by (13.6) and $\beta \in \mathbb{C}^{p\times p}$ with $\beta^*\beta \preceq I_p$, then*

$$(\theta_{21}\beta + \theta_{22})^{\pm 1} \in \mathcal{W}_+^{p\times p}, \tag{13.16}$$

$$(\theta_{11}\beta + \theta_{12})^{\pm 1} \in \mathcal{W}_-^{p\times p} \tag{13.17}$$

and

$$\dim\ker\{I - (\widehat{G}_\Phi^{(\beta)})^*\widehat{G}_\Phi^{(\beta)}\} = \dim\ker\{I_p - \beta^*\beta\}. \tag{13.18}$$

If $\beta^*\beta = I_p$, then

$$\widehat{G}_\Phi^{(\beta)}(\theta_{21}\beta + \theta_{22}) = (\theta_{11}\beta + \theta_{12}), \tag{13.19}$$

$$(\widehat{G}_\Phi^{(\beta)})^*(\theta_{11}\beta + \theta_{12}) = (\theta_{21}\beta + \theta_{22}), \tag{13.20}$$

$$\|\widehat{G}_\Phi^{(\beta)}\| = 1 \tag{13.21}$$

and

$$\{(\theta_{21}\beta + \theta_{22})u : u \in \mathbb{C}^p\} = \ker\{I - (\widehat{G}_\Phi^{(\beta)})^*\widehat{G}_\Phi^{(\beta)}\} \tag{13.22}$$

is a p dimensional subspace of H_2^p.

Proof. See Subsection D.4. □

Let $\mathcal{X}^{p\times p}$ denote the class of $p \times p$ mvf's $X \in \mathcal{W}^{p\times p}$ which are unitary on $\mathbb{T}$ and admit a factorization of the form

$$X(\zeta) = X_-(\zeta)X_+(\zeta)^{-1} \quad \text{for } \zeta \in \mathbb{T}, \tag{13.23}$$

with $(X_-)^{\pm 1} \in \mathcal{W}_-^{p\times p}$ and $(X_+)^{\pm 1} \in \mathcal{W}_+^{p\times p}$.

Theorem 13.3. *If $\Phi \in \mathfrak{W}^{p\times p}$ and Θ is specified by (13.6), then*

$$\mathcal{N}(\Phi) = \{T_\Theta[\mathcal{E}] : \mathcal{E} \in \mathcal{S}^{p\times p} \cap \mathcal{W}_+^{p\times p}\} \tag{13.24}$$

and

$$\mathcal{N}(\Phi) \cap \mathcal{X}^{p\times p} = \{T_\Theta[\mathcal{E}] : \mathcal{E} \in \mathbb{C}^{p\times p} \textit{ and is unitary}\}. \tag{13.25}$$

Proof. See Subsection D.5 for the proof of the inclusion

$$\{T_\Theta[\mathcal{E}] : \mathcal{E} \in \mathcal{W}_+^{p\times p} \cap \mathcal{S}^{p\times p}\} \subseteq \mathcal{N}(\Phi)$$

in (13.24), Subsection D.7 for the proof of the inclusion

$$\mathcal{N}(\Phi) \subseteq \{T_\Theta[\mathcal{E}] : \mathcal{E} \in \mathcal{W}_+^{p\times p} \cap \mathcal{S}^{p\times p}\}$$

in (13.24) and the verification of (13.25). □

Corollary 13.4. *If $\Phi \in \mathfrak{W}^{p\times p}$, Θ is specified by (13.6) and*

$$\Psi = T_\Theta[\beta] \quad \textit{for some unitary matrix } \beta \in \mathbb{C}^{p\times p},$$

then (the Fourier coefficients of Ψ)

$$\Psi_{-k} = \begin{cases} N^{-1/2}\beta M^{-1/2} - (\Gamma_\Phi T^*\mathbf{d})_0 M^{-1/2} & \textit{if } k = 0 \\ \Phi_{-k} & \textit{if } k = 1, 2, \ldots, \end{cases} \tag{13.26}$$

where

$$\mathbf{d} = \begin{bmatrix} (\theta_{22})_0 \\ (\theta_{22})_1 \\ \vdots \end{bmatrix}.$$

Moreover, if

$$\widehat{G}_\Psi^{(\beta)} = \widetilde{\mathfrak{q}}\Psi|_{H_2^p},$$

then

$$(\widehat{G}_\Psi^{(\beta)} f)(\zeta) = (\Psi_0 + \sum_{j=1}^{\infty} \zeta^{-j}\gamma_j) f_0 + \zeta(\widehat{\Gamma}_\Phi \mathfrak{p}\zeta^{-1} f)(\zeta) \tag{13.27}$$

for $f = \sum_{j=0}^{\infty} \zeta^j f_j$ belonging to H_2^p.

Proof. See Subsection D.8. □

Theorem 13.5. *The set $\mathfrak{X}^{p\times p}$ is a subset of the set $\mathfrak{W}^{p\times p}$, i.e., if $X\in\mathfrak{X}^{p\times p}$, then the Hankel operator $\widehat{\Gamma}_X=\mathfrak{q}X|_{H_2^p}$ is strictly contractive.*

Proof. Let

$$\Delta_X(\zeta)=X_-(\zeta)^*X_-(\zeta)\quad\text{for }\ \zeta\in\mathbb{T}.$$

Then, in view of the presumed unitarity of X on $\mathbb{T}$,

$$\Delta_X(\zeta)=X_+(\zeta)^*X_+(\zeta)\quad\text{for }\ \zeta\in\mathbb{T}.$$

Moreover, there exist a pair of positive constants $0<a<b$ such that

$$aI_p\preceq\Delta_X(\zeta)\preceq bI_p\quad\text{for }\ \zeta\in\mathbb{T}.$$

Therefore, the averages

$$A_I(\Delta_X^{\pm1})=\frac{1}{|I|}\int_I\Delta_X(e^{i\theta})^{\pm1}d\theta$$

over any subinterval I of $[0,2\pi]$ of length $|I|$ are subject to the bounds

$$A_I(\Delta_X)\preceq bI_p\quad\text{and}\quad A_I(\Delta_X^{-1})\preceq a^{-1}I_p.$$

Consequently,

$$\begin{aligned}\|A_I(\Delta_X)^{1/2}A_I(\Delta_X^{-1})^{1/2}\|&\le\|A_I(\Delta_X)^{1/2}\|\|A_I(\Delta_X^{-1})^{1/2}\|\\&\le(b/a)^{1/2},\end{aligned}$$

i.e., Δ_X meets the Treil–Volberg matrix Muckenhoupt condition (A2) in [35], and hence, by the main result of [35], the angle between the "past"

$$\mathfrak{z}_-(\Delta_X)=\text{cls}\{\zeta^j\xi:\ j\le-1\ \text{ and }\ \xi\in\mathbb{C}^p\}$$

in $L_2^p(\mathbb{T},\Delta_X)$ and the "future"

$$\mathfrak{z}_+(\Delta_X)=\text{cls}\{\zeta^j\xi:\ j\ge 0\ \text{ and }\ \xi\in\mathbb{C}^p\}$$

in $L_2^p(\mathbb{T},\Delta_X)$ is strictly positive:

$$\sup\{|\langle f_-,f_+\rangle_{\Delta_X}|:\ f_-\in\mathfrak{z}_-(\Delta_X),\ f_+\in\mathfrak{z}_+(\Delta_X)\ \text{ and }\ \|f_-\|_{\Delta_X}=\|f_+\|_{\Delta_X}=1\}<1. \tag{13.28}$$

But

$$\begin{aligned}\langle f_-,f_+\rangle_{\Delta_X}&=\langle X_-f_-,X_-f_+\rangle_{\text{st}}\\&=\langle X_-f_-,XX_+f_+\rangle_{\text{st}}.\end{aligned}$$

Therefore, since $g_-=X_-f_-$ belongs to $(H_2^p)^\perp$, $g_+=X_+f_+$ belongs to H_2^p and $\|f_\pm\|_{\Delta_X}=\|g_+\|_{\text{st}}$,

$$\langle f_-,f_+\rangle_{\Delta_X}=\langle g_-,Xg_+\rangle_{\text{st}}=\langle g_-,\widehat{\Gamma}_Xg_+\rangle_{\text{st}}. \tag{13.29}$$

Furthermore, since

$$X_+ \text{ maps } \mathfrak{z}_+(\Delta_X) \text{ bijectively onto } H_2^p$$

and

$$X_- \text{ maps } \mathfrak{z}_-(\Delta_X) \text{ bijectively onto } (H_2^p)^\perp$$

it follows readily from (13.28) and (13.29) that

$$\|\widehat{\Gamma}_X\| \le \rho < 1. \qquad \square$$

Corollary 13.6. *If $\Delta(\zeta) = Q(\zeta)^* Q(\zeta) = R(\zeta) R(\zeta)^*$ for $\zeta \in \mathbb{T}$ with $Q^{\pm 1} \in \mathcal{W}_+^{p\times p}$ and $R^{\pm 1} \in \mathcal{W}_+^{p\times p}$, then the Hankel operator $\widehat{\Gamma}_F = \mathfrak{q} F|_{H_2^p}$ with symbol*

$$F(\zeta) = R(\zeta)^* Q(\zeta)^{-1} \quad \textit{for } \zeta \in \mathbb{T}$$

is strictly contractive.

Proof. It suffices to show that $F \in \mathcal{X}^{p\times p}$, which is self-evident. $\square$

14. Explicit formulas for the rational case

This section is adapted from [16], where all rational solutions of a matricial Nehari problem based on a Hankel operator $\Gamma : H_2^p(\Pi_+) \to H_2^p(\Pi_-)$, where Π_+ and Π_- denote the right half plane and left half plane, respectively.

Let $\mathcal{R}^{p\times p}$ denote the set of $p \times p$ rational mvf's. Let $\Phi(\zeta) = \sum_{n=-\infty}^{\infty} \zeta^n \Phi_n$ belong to $\mathcal{W}^{p\times p}$ and suppose $\Phi_-(\zeta) = \sum_{n=-\infty}^{-1} \zeta^n \Phi_n$ belongs to $\mathcal{R}^{p\times p}$ and admits a minimal realization of the form

$$\Phi_-(\lambda) = C(\lambda I_n - A)^{-1} B, \tag{14.1}$$

where $A \in \mathbb{C}^{p\times p}$ with $\sigma(A) \subset \mathbb{D}$, $B \in \mathbb{C}^{n\times p}$ and $C \in \mathbb{C}^{p\times n}$. Let

$$F_o(\lambda) = C(\lambda I_n - A)^{-1}, \quad F_c(\lambda) = B^*(I_n - \lambda A^*)^{-1}, \tag{14.2}$$

$$P_o = \frac{1}{2\pi}\int_0^{2\pi} F_o(e^{i\theta})^* F_o(e^{i\theta}) d\theta \quad \text{and} \quad P_c = \frac{1}{2\pi}\int_0^{2\pi} F_c(e^{i\theta})^* F_c(e^{i\theta}) d\theta. \tag{14.3}$$

If $\|\widehat{\Gamma}_\Phi\| < 1$, then $I_n - P_o P_c$ and $I_n - P_c P_o$ are invertible (see Lemma 14.10). In addition, let

$$\mathcal{M}_o = \{F_o(\lambda) u : u \in \mathbb{C}^p\} \quad \text{and} \quad \mathcal{M}_c = \{F_c(\lambda) u : u \in \mathbb{C}^p\} \tag{14.4}$$

be endowed with the inner product

$$\langle F_o u, F_o v\rangle_{\mathcal{M}_o} = v^* P_o u \quad \text{and} \quad \langle F_c u, F_c v\rangle_{\mathcal{M}_c} = v^* P_c u,$$

respectively.

The assumption that the realization given in (14.1) is minimal means that the pair (C, A) is observable and the pair (A, B) is controllable, i.e.,

$$\bigcap_{j=0}^{n-1} \ker CA^j = \{0\} \quad \text{and} \quad \bigcap_{j=0}^{n-1} \ker B^*(A^*)^j = \{0\}.$$

Thus, if $F_o(\lambda)u = 0$ for every $\lambda \in \mathbb{C} \setminus \sigma(A)$, then $u = 0$. Similarly, if $F_c(\lambda)u = 0$ for every $\lambda \in \mathbb{C} \setminus \sigma^\#(A)$, where

$$\sigma^\#(A) = \{1/\bar{\lambda} : \lambda \in \sigma(A) \setminus \{0\}\},$$

then $u = 0$. Consequently, the n columns of $F_o(\lambda)$ and $F_c(\lambda)$ form a basis for $\mathcal{M}_o$ and $\mathcal{M}_c$, respectively. Moreover, P_o and P_c are both positive definite matrices.

Let $\mathcal{N}_{\mathcal{R}}(\Phi)$ denote the set of all mvf's $\Psi \in \mathcal{W}^{p\times p} \cap \mathcal{R}^{p\times p}$ with

$$\Psi_-(\zeta) = \sum_{n=-\infty}^{0} \zeta^n \Phi_n \in \mathcal{R}^{p\times p}$$

and $\|\Psi(\zeta)\| \leq 1$ for every point $\zeta \in \mathbb{T}$. The main result of this section is devoted to obtaining explicit formulas for the blocks θ_{jk}, $j, k = 1, 2$, in the linear fractional transformation $T_\Theta[\mathcal{E}]$ in Theorem 13.3.

Theorem 14.1. *If $\Phi \in \mathfrak{W}^{p\times p}$ and Φ_- has a minimal realization given by (14.1), then*

$$\mathcal{N}_{\mathcal{R}}(\Phi) = \{(\theta_{11}\mathcal{E} + \theta_{12})(\theta_{21}\mathcal{E} + \theta_{22})^{-1} : \mathcal{E} \in \mathcal{S}^{p\times p} \cap \mathcal{R}^{p\times p}\}, \tag{14.5}$$

where

$$\theta_{11}(\lambda) = \lambda F_o(\lambda)(I_n - P_cP_o)^{-1}P_cC^*\{C(I_n - P_cP_o)^{-1}P_cC^*\}^{-1/2}, \tag{14.6}$$

$$\theta_{12}(\lambda) = F_o(\lambda)P_c(I_n - P_oP_c)^{-1}P_oB\{B^*(I_n - P_oP_c)^{-1}P_oB\}^{-1/2}, \tag{14.7}$$

$$\theta_{21}(\lambda) = \lambda F_c(\lambda)P_o(I_n - P_cP_o)^{-1}P_cC^*\{C(I_n - P_cP_o)^{-1}P_cC^*\}^{-1/2}, \tag{14.8}$$

and
$$\theta_{22}(\lambda) = F_c(\lambda)(I_n - P_oP_c)^{-1}P_oB\{B^*(I_n - P_oP_c)^{-1}P_oB\}^{-1/2}. \tag{14.9}$$

Moreover,

$$\begin{aligned}\Theta(\lambda)\begin{bmatrix}\lambda^{-1}I_p & 0\\ 0 & I_p\end{bmatrix} &= \begin{bmatrix}\lambda^{-1}\theta_{11}(\lambda) & \theta_{12}(\lambda)\\ \lambda^{-1}\theta_{21}(\lambda) & \theta_{22}(\lambda)\end{bmatrix}\\ &= \begin{bmatrix}C & 0\\ 0 & B^*\end{bmatrix}\begin{bmatrix}\lambda I_n - A & 0\\ 0 & I_n - \lambda A^*\end{bmatrix}^{-1}\begin{bmatrix}P_c^{-1} & -I_n\\ -I_n & P_o^{-1}\end{bmatrix}^{-1}\\ &\quad\times\begin{bmatrix}C^*N^{-1/2} & 0\\ 0 & BM^{-1/2}\end{bmatrix},\end{aligned} \tag{14.10}$$

where

$$N = C(I_n - P_cP_o)^{-1}P_cC^* \succ 0$$

and
$$M = B^*(I_n - P_oP_c)^{-1}P_oB \succ 0.$$

The proof of Theorem 14.1 will be deferred until the end of the section.

Lemma 14.2. *If $\Phi \in \mathfrak{W}^{p\times p}$ and Φ_- has a minimal realization given by* (14.1)*, then the spaces $\mathcal{M}_o$ and $\mathcal{M}_c$ are both n-dimensional RKHS's with RK's*

$$K_\omega^o(\lambda) = F_o(\lambda)P_o^{-1}F_o(\omega)^* \quad \text{for } \lambda,\omega \in \mathbb{C} \setminus \sigma(A) \tag{14.11}$$

and

$$K_\omega^c(\lambda) = F_c(\lambda)P_c^{-1}F_c(\omega)^* \quad \text{for } \lambda,\omega \in \mathbb{C} \setminus \sigma^\#(A). \tag{14.12}$$

Proof. Since the realizations in (14.2) are minimal, $\mathcal{M}_o$ and $\mathcal{M}_c$ are both n dimensional spaces, to prove the assertion for $\mathcal{M}_o$, it suffices to show that:

(1) $K_\omega^o u \in \mathcal{M}_o$ for every $u \in \mathbb{C}^p$ and $\omega \in \mathbb{C} \setminus \sigma(A)$.

(2) $\langle f, K_\omega^o u\rangle_{\mathcal{M}_o} = u^* f(\omega)$ for every $u \in \mathbb{C}^p$ and $f \in \mathcal{M}_o$.

If $u \in \mathbb{C}^p$ and $\omega \in \mathbb{C} \setminus \sigma(A)$, then $K_\omega^o(\lambda)u = F_o(\lambda)v$ for $v = P_o^{-1}F_o(\omega)^*u$. Thus, (1) holds. Next if $f = F_o v$, where $v \in \mathbb{C}^p$, then

$$\begin{aligned}
\langle f, K_\omega^o u\rangle_{\mathcal{M}_o} &= \langle F_o v, K_\omega^o u\rangle_{\mathcal{M}_o} \\
&= \langle F_o v, F_o P_o^{-1} F_o(\omega)^* u\rangle_{\mathcal{M}_o} \\
&= u^* F_o(\omega) P_o^{-1} P_o v \\
&= u^* F_o(\omega) v \\
&= u^* f(\omega).
\end{aligned}$$

Thus, (2) holds.

The verification for $\mathcal{M}_c$ is completed in much the same way. □

Lemma 14.3. *If $\Phi \in \mathfrak{W}^{p\times p}$ and Φ_- has a minimal realization given by* (14.1)*, then*:

(i) *P_0 is the only solution of the Stein equation*

$$C^*C = P_o - A^* P_o A. \tag{14.13}$$

(ii) *P_c is the only solution of the Stein equation*

$$BB^* = P_c - A P_c A^*. \tag{14.14}$$

Proof. Since $\sigma(A) \subset \mathbb{D}$,

$$\begin{aligned}
P_o &= \frac{1}{2\pi}\int_0^{2\pi} F_o(e^{i\theta})^* F_o(e^{i\theta}) d\theta \\
&= \frac{1}{2\pi}\int_0^{2\pi} (I_n - e^{i\theta}A^*)^{-1} C^*C (I_n - e^{-i\theta}A)^{-1} d\theta \\
&= \frac{1}{2\pi}\int_0^{2\pi} \left\{\sum_{j=0}^\infty (e^{i\theta}A^*)^j\right\} C^*C \left\{\sum_{k=0}^\infty (e^{-i\theta}A)^k\right\} d\theta \\
&= C^*C + A^*\left\{\frac{1}{2\pi}\int_0^{2\pi} \left(\sum_{j=1}^\infty e^{i(j-1)\theta}(A^*)^{j-1}\right) C^*C \left(\sum_{k=1}^\infty e^{-i(k-1)\theta}A^{k-1}\right) d\theta\right\} A \\
&= C^*C + A^* P_o A.
\end{aligned}$$

Thus, P_o is a solution of (14.13). P_0 is the only solution of (14.13) because $\sigma(A) \subset \mathbb{D}$ and hence

$$\sigma(A) \cap \sigma^{\#}(A) = \emptyset,$$

(see, e.g., Theorem 18.2 in [17]).

The verification of (ii) is similar. □

Remark 14.4. Since $\sigma(A) \subset \mathbb{D}$,

$$P_o = \sum_{j=0}^{\infty} (A^*)^j C^* C A^j \tag{14.15}$$

and

$$P_c = \sum_{j=0}^{\infty} (A)^j B B^* (A^*)^j. \tag{14.16}$$

The recipe for P_o and P_c given in (14.15) and (14.16) follows easily from (14.13) and (14.14), respectively.

Lemma 14.5. *If $\Phi \in \mathfrak{W}^{p\times p}$ and Φ_- has a minimal realization given by* (14.1)*, then*:

(i) *The mvf*

$$\theta_o(\lambda) = I_p - (\lambda - 1) F_o(\lambda) P_o^{-1} (I_n - A^*)^{-1} C^*$$

is inner with respect to $\mathbb{C} \setminus \mathbb{D}$.

(ii) *The mvf*

$$\theta_c(\lambda) = I_p - (1 - \lambda) F_c(\lambda) P_c^{-1} (I_n - A)^{-1} B$$

is inner with respect to $\mathbb{D}$.

Proof. With the help of (14.13), it is readily checked

$$-I_p + \theta_o(\lambda)\theta_o(\omega)^* = (1 - \lambda\overline{\omega}) F_o(\lambda) P_o^{-1} F_o(\omega)^* \tag{14.17}$$

and hence that θ_o is inner with respect to $\mathbb{C} \setminus \overline{\mathbb{D}}$.

Similarly, with the help of (14.14), it is readily checked that

$$I_p - \theta_c(\lambda)\theta_c(\omega)^* = (1 - \lambda\overline{\omega}) F_c(\lambda) P_c^{-1} F_c(\omega)^* \tag{14.18}$$

and hence θ_c is inner with respect to $\mathbb{D}$. □

Lemma 14.6. *If $\Phi \in \mathfrak{W}^{p\times p}$ and Φ_- has a minimal realization given by* (14.1) *and $f \in H_2^p$, then*

$$(\widehat{\Gamma}_\Phi f)(\lambda) = F_o(\lambda) \left\{ \frac{1}{2\pi} \int_0^{2\pi} F_c(e^{i\theta})^* f(e^{i\theta}) d\theta \right\} \tag{14.19}$$

for every point $\lambda \in \mathbb{C} \setminus \sigma(A)$.

Proof. If $A = \mathrm{diag}(\omega_1, \dots, \omega_n)$ and $\mathfrak{e}_j$, $j = 1, \dots, n$, denotes the standard basis of $\mathbb{C}^n$, then

$$(\lambda I_n - A)^{-1} = \sum_{j=1}^{n} \frac{\mathfrak{e}_j \mathfrak{e}_j^*}{\lambda - \omega_j}.$$

It is readily seen that

$$\begin{aligned}
(\widehat{\Gamma}_\Phi f)(\lambda) &= (\mathfrak{q}\Phi f)(\lambda) \\
&= \sum_{j=1}^{n} \mathfrak{q} C \, \frac{\mathfrak{e}_j \mathfrak{e}_j^*}{\lambda - \omega_j} B f \\
&= \mathfrak{q} \left\{ \sum_{j=1}^{n} C \mathfrak{e}_j \mathfrak{e}_j^* B \left(\frac{f - f(\omega_j)}{\lambda - \omega_j} \right) + \sum_{j=1}^{n} C \, \mathfrak{e}_j \mathfrak{e}_j^* B \left(\frac{f(\omega_j)}{\lambda - \omega_j} \right) \right\} \\
&= \sum_{j=1}^{n} C \left(\frac{\mathfrak{e}_j \mathfrak{e}_j^*}{\lambda - \omega_j} \right) B f(\omega_j),
\end{aligned}$$

since

$$\frac{f(\lambda) - f(\omega_j)}{\lambda - \omega_j} \in H_2^p \quad \text{and} \quad \frac{f(\omega_j)}{\lambda - \omega_j} \in (H_2^p)^\perp$$

when $\omega_j \in \mathbb{D}$. Therefore,

$$(\widehat{\Gamma}_\Phi f)(\lambda) = C(\lambda I_n - A)^{-1} \sum_{k=1}^{n} \mathfrak{e}_k \mathfrak{e}_k^* B f(\omega_k).$$

Thus, upon invoking Cauchy's formula for H_2^p,

$$f(\omega_k) = \frac{1}{2\pi} \int_0^{2\pi} \frac{f(e^{i\theta})}{1 - \omega_k e^{-i\theta}} d\theta,$$

it follows that

$$\begin{aligned}
(\widehat{\Gamma}_\Phi f)(\lambda) &= C(\lambda I_n - A)^{-1} \left\{ \frac{1}{2\pi} \int_0^{2\pi} (I_n - A e^{-i\theta})^{-1} B f(e^{i\theta}) d\theta \right\}. \\
&= C(\lambda I_n - A)^{-1} \left\{ \frac{1}{2\pi} \int_0^{2\pi} (B^*(I_n - e^{-i\theta} A^*)^{-1})^* f(e^{i\theta}) d\theta \right\} \\
&= F_o(\lambda) \left\{ \frac{1}{2\pi} \int_0^{2\pi} F_c(\lambda)^* f(e^{i\theta}) d\theta \right\}.
\end{aligned}$$

This completes the proof of (14.19) when A is diagonal. The preceding argument can easily be adjusted to obtain the same conclusion when A is diagonalizable. Therefore, since the $n \times n$ diagonalizable matrices are dense in $\mathbb{C}^{n \times n}$, (14.19) holds for any $A \in \mathbb{C}^{n \times n}$. $\square$

Lemma 14.7. *If $\Phi \in \mathfrak{W}^{p\times p}$ and Φ_- has a minimal realization given by* (14.1) *and $h(\zeta) = \sum_{n=0}^{\infty} \zeta^{-n} h_n$ belongs to $\zeta(H_2^p)^\perp$, then*

$$\left(\mathfrak{p}\, \frac{h}{1-\zeta\omega}\right)(\lambda) = \begin{cases} \dfrac{h(1/\omega)}{1-\lambda\omega} & \text{if } \omega \in \mathbb{D} \setminus \{0\} \\ h_0 & \text{if } \omega = 0. \end{cases} \tag{14.20}$$

Proof. If $h \in \zeta(H_2^p)^\perp$, $u \in \mathbb{C}^p$ and $\lambda \in \mathbb{D}$, then

$$\begin{aligned} u^* \left(\mathfrak{p}\, \frac{f}{\rho_\omega}\right)(\lambda) &= \left\langle \frac{h}{\rho_\omega}, \frac{u}{\rho_\lambda} \right\rangle \\ &= \overline{\left\langle \frac{h^\# u}{\rho_\lambda}, \frac{1}{\rho_\omega} \right\rangle} \\ &= \overline{\left(\frac{h^\# u}{\rho_\lambda}\right)(\omega)} \\ &= \frac{u^* h(1/\overline{\omega})}{\rho_\omega(\lambda)} \\ &= \frac{u^* h(1/\overline{\omega})}{1-\lambda\overline{\omega}} \quad \text{if } \omega \in \mathbb{D} \setminus \{0\}. \end{aligned}$$

If $\omega = 0$, then (14.20) follows from the evaluation $(\mathfrak{p}h)(\zeta) = h_0$. $\square$

Lemma 14.8. *If $\Phi \in \mathfrak{W}^{p\times p}$ and Φ_- has a minimal realization given by* (14.1) *and $g \in (H_2^p)^\perp$, then*

$$(\widehat{\Gamma}_\Phi^* g)(\lambda) = F_c(\lambda) \left\{ \frac{1}{2\pi} \int_0^{2\pi} F_o(e^{i\theta})^* g(e^{i\theta}) d\theta \right\} \tag{14.21}$$

for $\lambda \in \mathbb{C} \setminus \sigma^\#(A)$.

Proof. If $g \in (H_2^p)^\perp$ and $A = \mathrm{diag}(\omega_1, \ldots, \omega_n)$, then

$$\begin{aligned} (\widehat{\Gamma}_\Phi^*\, g)(\lambda) &= (\mathfrak{p} B^* (\overline{\zeta} I_n - A^*)^{-1} C^* g)(\lambda) \\ &= (\mathfrak{p} B^* \sum_{j=1}^n \frac{\mathfrak{e}_j \mathfrak{e}_j^*}{1-\zeta\overline{\omega}_j} C^* \zeta g)(\lambda) \end{aligned}$$

Let $h(\zeta) = \zeta g(\zeta)$ and write $h(\zeta) = \sum_{n=0}^{\infty} \zeta^{-n} h_n$. Using Lemma 14.7 we get

$$\begin{aligned} (\widehat{\Gamma}_\Phi^*\, g)(\lambda) &= \sum_{j=1}^n B^* \left(\frac{\mathfrak{e}_j \mathfrak{e}_j^*}{1-\lambda\overline{\omega}_j} \right) \sum_{k=1}^n \mathfrak{e}_k \mathfrak{e}_k^* C^* h(\overline{\omega}_k) \\ &= B^* (I_n - \lambda A^*)^{-1} \sum_{k=1}^n \mathfrak{e}_k \mathfrak{e}_k^* C^* h(1/\overline{\omega}_k) \end{aligned}$$

$$= F_c(\lambda)\sum_{k=1}^{n}\mathfrak{e}_k\mathfrak{e}_k^* C^* h(\overline{\omega}_k). \tag{14.22}$$

Formula (14.21) is obtained from (14.22) using Cauchy's formula. □

Lemma 14.9. *If $\Phi \in \mathfrak{W}^{p\times p}$ and Φ_- has a minimal realization given by* (14.1)*, then*:

(i) *The formulas $\widehat{\Gamma}_\Phi F_c u = F_o P_c u$ and $\widehat{\Gamma}_\Phi^* F_o v = F_c P_o v$ hold for every choice of $u, v \in \mathbb{C}^n$.*

(ii) *If $g \in (H_2^p)^\perp$ and $\langle g, F_o u\rangle_{\rm st} = 0$ for every $u \in \mathbb{C}^p$, then $\widehat{\Gamma}_\Phi^* g = 0$.*

(iii) *If $f \in H_2^p$ and $\langle f, F_c u\rangle_{\rm st} = 0$ for every $u \in \mathbb{C}^p$, then $\widehat{\Gamma}_\Phi f = 0$.*

Proof. Assertions (i)–(iii) follow easily from (14.19) and (14.21), respectively. □

Lemma 14.10. *If $\Phi \in \mathfrak{W}^{p\times p}$ and Φ_- has a minimal realization given by* (14.1)*, $P_c^{1/2}P_oP_c^{1/2} = UDU^*$, where $U \in \mathbb{C}^{n\times n}$ is a unitary matrix with columns $u_1, \dots, u_n$ and $D = \mathrm{diag}(s_1^2, \dots, s_n^2)$ and $s_1 \ge \cdots \ge s_n > 0$,*

$$f_j = F_cP_c^{-1/2}u_j \quad \textit{and} \quad g_j = \left(\frac{1}{s_j}\right)F_oP_c^{1/2}u_j \quad \textit{for}\ \ j = 1, \dots, n,$$

then:

(1) $\widehat{\Gamma}_\Phi f_j = s_j g_j$ *and* $\widehat{\Gamma}_\Phi^* g_j = s_j f_j$ *for* $j = 1, \dots, n$.

(2) $\langle f_j, f_k\rangle_{\rm st} = \langle g_j, g_k\rangle_{\rm st} = \begin{cases} 0 & \textit{if}\ \ j \ne k \\ 1 & \textit{if}\ \ j = k. \end{cases}$

(3) $\langle f_j, g_k\rangle_{\rm st} = 0$ *for* $j, k = 1, \dots, n$.

(4) $I_n - P_cP_o$ *and* $I_n - P_oP_c$ *are invertible matrices.*

Proof. Assertion (1)–(4) are an easy consequence of the formulas advertised in item (i) of Lemma 14.9. We will now check that $I_n - P_cP_o$ is invertible. In view of

$$I_n - P_cP_o = P_c^{1/2}(I_n - P_c^{1/2}P_oP_c^{1/2})P_c^{-1/2},$$

$I_n - P_cP_o$ is invertible if and only if $I_n - P_c^{1/2}P_oP_c^{1/2}$ is invertible. In view of (1),

$$\begin{aligned}(I - \widehat{\Gamma}_\Phi^*\widehat{\Gamma}_\Phi)f_j &= f_j - \widehat{\Gamma}_\Phi^* s_j g_j \\ &= (1 - s_j^2)f_j \quad \text{for}\ \ j = 1, \dots, n.\end{aligned}$$

Thus, as $I - \widehat{\Gamma}_\Phi^*\widehat{\Gamma}_\Phi$ is a positive operator and f_j is an eigenvector corresponding to the eigenvalue $1 - s_j^2$, $1 - s_j > 0$ for $j = 1, \dots, n$. Thus, $I_n - P_cP_o$ is invertible, and consequently, $I_n - P_oP_c$ is also invertible. □

Lemma 14.11. *If $\Phi \in \mathfrak{W}^{p\times p}$ and Φ_- has a minimal realization given by (14.1), then*

$$\{(I - \widehat{\Gamma}_\Phi^* \widehat{\Gamma}_\Phi)^{-1} I_p\}(\lambda) = F_c(\lambda)(I_n - P_o P_c)^{-1} P_o B \tag{14.23}$$

for $\lambda \in \mathbb{C} \setminus \sigma(A)$, and

$$\{(I - \widehat{\Gamma}_\Phi \widehat{\Gamma}_\Phi^*)^{-1} \zeta^{-1} I_p\}(\lambda) = F_o(\lambda)(I_n - P_c P_o)^{-1} P_c C^* \tag{14.24}$$

for $\mathbb{C} \setminus \sigma^\#(A)$. Moreover, the positive definite matrices M and N defined in (13.5) can be written as

$$M = B^*(I_n - P_o P_c)^{-1} P_o B \tag{14.25}$$

and

$$N = C(I_n - P_c P_o)^{-1} P_c C^*. \tag{14.26}$$

Proof. The proof is broken into steps.

1. Verification of (14.23) *and* (14.24). It is readily seen that (14.23) is equivalent to

$$\sum_{j=0}^{\infty} \{(\widehat{\Gamma}_\Phi^* \widehat{\Gamma}_\Phi)^j I_p\}(\lambda) = \sum_{j=0}^{\infty} F_c(\lambda)(P_o P_c)^j P_o B. \tag{14.27}$$

But

$$\sum_{j=0}^{k} \{(\widehat{\Gamma}_\Phi^* \widehat{\Gamma}_\Phi)^j I_p\}(\lambda) = \sum_{j=0}^{k} F_c(\lambda)(P_o P_c)^j P_o B \tag{14.28}$$

can be checked by induction on k using formulas (14.19) and (14.21). Thus, (14.27) holds and so must (14.23).

The verification of (14.24) is similar.

2. Verification of (14.25) *and* (14.26). In view of (14.23),

$$\begin{aligned}
M &\overset{\text{def}}{=} [(I - \widehat{\Gamma}_\Phi^* \widehat{\Gamma}_\Phi)^{-1} I_p, I_p]_{\text{st}} \\
&= [F_c(I_n - P_o P_c)^{-1} P_o B, I_p]_{\text{st}} \\
&= [F_c, I_p]_{\text{st}} (I_n - P_o P_c)^{-1} P_o B \\
&= B^* \left\{ \frac{1}{2\pi} \int_0^{2\pi} B^*(I_n - e^{i\theta} A^*)^{-1} d\theta \right\} (I_n - P_o P_c)^{-1} P_o B \\
&= B^* \sum_{k=0}^{\infty} \left\{ \frac{1}{2\pi} \int_0^{2\pi} e^{ik\theta} \{A^*\}^k d\theta \right\} (I_n - P_o P_c)^{-1} P_o B \\
&= B^*(I_n - P_o P_c)^{-1} P_o B.
\end{aligned}$$

The verification of (14.26) is similar using (14.24). □

Theorem 14.12. *If $\Phi \in \mathfrak{W}^{p\times p}$ and Φ_- has a minimal realization given by* (14.1)*, then*:

(i) *The n-dimensional RKHS $\mathcal{M}_o$ can be identified as*

$$\mathcal{M}_o = (H_2^p)^\perp \ominus \theta_o(H_2^p)^\perp$$

with RK

$$K_\omega^o(\lambda) = -\frac{I_p - \theta_o(\lambda)\theta_o(\omega)^*}{\rho_\omega(\lambda)} \quad \text{for } \lambda, \omega \in \mathbb{C} \setminus \sigma^\#(A).$$

(ii) *The n-dimensional RKHS $\mathcal{M}_c$ can be identified as*

$$\mathcal{M}_c = (H_2^p) \ominus \theta_c(H_2^p)$$

with RK

$$K_\omega^c(\lambda) = \frac{I_p - \theta_c(\lambda)\theta_c(\omega)^*}{\rho_\omega(\lambda)} \quad \text{for } \lambda, \omega \in \mathbb{C} \setminus \sigma(A).$$

Proof. The formula for $K_\omega^o(\lambda)$ and $K_\omega^c(\lambda)$ follow from (14.17) and (14.18), respectively. The identifications for $\mathcal{M}_o$ and $\mathcal{M}_c$ follow from Lemma 14.9 and Theorem 14.13. □

Theorem 14.13. *If $\Phi \in \mathfrak{W}^{p\times p}$ and Φ_- has a minimal realization given by* (14.1)*, then*:

(i) *The Hankel operator $\widehat{\Gamma}_\Phi$ maps $\mathcal{M}_c$ injectively onto $\mathcal{M}_o$ and*

$$\ker \widehat{\Gamma}_\Phi = H_2^p \ominus \mathcal{M}_c. \tag{14.29}$$

(ii) *The Hankel operator $\widehat{\Gamma}_\Phi^*$ maps $\mathcal{M}_o$ injectively onto $\mathcal{M}_c$ and*

$$\ker \widehat{\Gamma}_\Phi^* = (H_2^p)^\perp \ominus \mathcal{M}_o. \tag{14.30}$$

Proof. Assertion (i) follows easily from the definition of $\mathcal{M}_c$ given in (14.4) and the formula (14.19), since (C, A) is an observable pair: if $\widehat{\Gamma}_\Phi f = 0$ for $f \in \mathcal{M}_c$, then $f = F_c u$ for some $u \in \mathbb{C}^p$ and

$$\widehat{\Gamma}_\Phi f = \widehat{\Gamma}_\Phi u = F_o P_c u = 0.$$

Thus, as (C, A) is an observable pair and P_c is invertible, $u = 0$. Therefore, $\widehat{\Gamma}_\Phi$ maps the n-dimensional space $\mathcal{M}_c$ injectively onto the n-dimensional space $\mathcal{M}_o$. Finally, (14.29) follows from (14.19).

Assertion (ii) is proved in much the same way since

$$(\widehat{\Gamma}_\Phi^* F_o u)(\lambda) = F_c(\lambda) P_o u$$

and, as (A, B) is a controllable pair, (B^*, A^*) is an observable pair. □

Proof of Theorem 14.1. In view of (13.24), it suffices to justify the formulas advertised in (14.6)–(14.9). In view of (14.23) and (14.25),

$$\begin{aligned}\theta_{22}(\lambda) &= F_c(\lambda)(I_n - P_oP_c)^{-1}P_oBM^{-1/2}\\ &= F_c(\lambda)(I_n - P_oP_c)^{-1}P_oB\{B^*(I_n - P_oP_c)^{-1}P_oB\}^{-1/2}.\end{aligned} \tag{14.31}$$

The formula for θ_{12} can be obtained from (D.15) using (14.31) and item (i) of Lemma 14.9. The verification of the formulas for θ_{11} and θ_{21} are similar.

It is readily checked that

$$\begin{aligned}\Theta(\lambda)\begin{bmatrix}\lambda^{-1}I_p & 0\\ 0 & I_p\end{bmatrix} &= \begin{bmatrix}C & 0\\ 0 & B^*\end{bmatrix}\begin{bmatrix}\lambda I_n - A & 0\\ 0 & I_n - \lambda A^*\end{bmatrix}^{-1}\\ &\quad\times\begin{bmatrix}(I_n - P_cP_o)^{-1}P_c & P_c(I_n - P_oP_c)^{-1}P_o\\ P_o(I_n - P_cP_o)^{-1}P_c & (I_n - P_oP_c)^{-1}P_o\end{bmatrix}\\ &\quad\times\begin{bmatrix}C^*N^{-1/2} & 0\\ 0 & BM^{-1/2}\end{bmatrix}.\end{aligned} \tag{14.32}$$

As

$$\begin{bmatrix}P_c^{-1} & -I_n\\ -I_n & P_o^{-1}\end{bmatrix}^{-1} = \begin{bmatrix}(I_n - P_cP_o)^{-1}P_c & P_c(I_n - P_oP_c)^{-1}P_o\\ P_o(I_n - P_cP_o)^{-1}P_c & (I_n - P_oP_c)^{-1}P_o\end{bmatrix},$$

(14.10) follows directly from (14.32). □

Remark 14.14. The matrix

$$\widetilde{P} = \begin{bmatrix}P_o^{-1} & -I_n\\ -I_n & P_o^{-1}\end{bmatrix}$$

appearing in (14.10) is positive definite. By a Schur complement argument,

$$\begin{aligned}\widetilde{P} &= \begin{bmatrix}I_n & -P_o\\ 0 & I_n\end{bmatrix}\begin{bmatrix}P_c^{-1} - P_o & 0\\ 0 & P_o^{-1}\end{bmatrix}\begin{bmatrix}I_n & 0\\ -P_o & I_n\end{bmatrix}\\ &= \begin{bmatrix}I_n & -P_o\\ 0 & I_n\end{bmatrix}\begin{bmatrix}P_c^{-1/2}\{I_n - P_c^{1/2}P_oP_c^{1/2}\}P_c^{-1/2} & 0\\ 0 & P_o^{-1}\end{bmatrix}\begin{bmatrix}I_n & 0\\ -P_o & I_n\end{bmatrix}.\end{aligned}$$

We have already observed in the proof of Lemma 14.10 that the eigenvalues of the positive definite matrix $P_c^{1/2}P_oP_c^{1/2}$ lie in $(0, 1)$. Thus,

$$I_n - P_c^{1/2}P_oP_c^{1/2} \succ 0$$

and consequently $\widetilde{P} \succ 0$.

15. An inverse scattering problem

In this section there is partial overlap of the connection between the considered Nehari problem and a discrete analogue of an inverse scattering problem considered by Krein and Melik-Adamjan [28].

Theorem 15.1. *If $\Phi \in \mathfrak{W}^{p\times p}$, Θ is defined by (13.6) and $\Phi = T_\Theta[\mathcal{E}]$ for a unitary matrix $\mathcal{E} \in \mathbb{C}^{p\times p}$, then there exists exactly one factorization*

$$\Phi(\zeta) = UR(\zeta)^* Q(\zeta)^{-1} \quad \textit{for all points } \zeta \in \mathbb{T} \tag{15.1}$$

with the following properties:

(1) $Q^{\pm 1} \in \mathcal{W}_+^{p\times p}$ *and* $R^{\pm 1} \in \mathcal{W}_+^{p\times p}$.

(2) $Q(0) \succ 0$ *and* $R(0) \succ 0$.

(3) $Q(\zeta)^* Q(\zeta) = R(\zeta)R(\zeta)^*$ *for all points* $\zeta \in \mathbb{T}$.

(4) *The integral*
$$\frac{1}{2\pi}\int_0^{2\pi} Q(e^{i\theta})^* Q(e^{i\theta})d\theta = I_p.$$

(5) $U \in \mathbb{C}^{p\times p}$ *is a unitary matrix.*

Proof. The proof is broken into steps.

1. Φ admits at least one factorization of the form (15.1). Let

$$Q(\lambda) = \{\theta_{21}(\lambda)\mathcal{E} + \theta_{22}(\lambda)\}ZV \quad \text{for } \lambda \in \overline{\mathbb{D}}, \tag{15.2}$$

where

$$Z = \left\{\frac{1}{2\pi}\int_0^{2\pi} [\theta_{21}(e^{i\theta})\mathcal{E} + \theta_{22}(e^{i\theta})]^*[\theta_{21}(e^{i\theta})\mathcal{E} + \theta_{22}(e^{i\theta})]d\theta\right\}^{-1/2}$$

and $V \in \mathbb{C}^{p\times p}$ is a unitary matrix such that

$$Q(0) = \theta_{22}(0)ZV \succ 0,$$

and

$$R(\lambda) = V^* Z\{\mathcal{E}^*\theta_{11}^\#(\lambda) + \theta_{12}^\#(\lambda)\}U \quad \text{for } \lambda \in \overline{\mathbb{D}}, \tag{15.3}$$

where $U \in \mathbb{C}^{p\times p}$ is a unitary matrix such that

$$R(0) = V^* Z\mathcal{E}^*\theta_{11}^\#(0)U \succ 0.$$

By Theorem 13.1, $(\theta_{11}\mathcal{E}+\theta_{12})^{\pm 1} \in \mathcal{W}_-^{p\times p}$ and $(\theta_{21}\mathcal{E}+\theta_{22})^{\pm 1} \in \mathcal{W}_+^{p\times p}$. Consequently, in view of (15.2) and (15.3), $Q^{\pm 1} \in \mathcal{W}_+^{p\times p}$ and $R^{\pm 1} \in \mathcal{W}_+^{p\times p}$. As

$$R(\zeta)^* Q(\zeta)^{-1} = U^* T_\Theta[\mathcal{E}] \quad \text{for } \zeta \in \mathbb{T},$$
$$\Phi(\zeta) = UR(\zeta)^* Q(\zeta)^{-1} \quad \text{for } \zeta \in \mathbb{T}$$

and the factorization above satisfies properties (1)–(5).

2. Φ *admits exactly one factorization of the form* (15.1) *that meets constraints* (1)–(5). If

$$\Phi = UR^*Q^{-1} = \widetilde{U}\widetilde{R}^*\widetilde{Q}^{-1} \quad \text{on } \mathbb{T},$$

are two factorizations such that (1)–(5) hold, then

$$\widetilde{R}^{-*}\widetilde{U}^*UR^* = \widetilde{Q}^{-1}Q \quad \text{on } \mathbb{T}. \tag{15.4}$$

Therefore, since $\widetilde{Q}^{-1}Q \in \mathcal{W}_+^{p\times p}$ and $\widetilde{R}^{-*}\widetilde{U}^*UR^* \in \mathcal{W}_-^{p\times p}$,

$$\widetilde{Q}(\zeta)^{-1}Q(\zeta) = K \quad \text{for } K \in \mathbb{C}^{p\times p}.$$

However, since

$$I_p = \frac{1}{2\pi}\int_0^{2\pi} Q(e^{i\theta})^*Q(e^{i\theta})d\theta = \frac{1}{2\pi}\int_0^{2\pi} \widetilde{Q}(e^{i\theta})^*\widetilde{Q}(e^{i\theta})d\theta,$$

K must be unitary, and, as $Q(0) = \widetilde{Q}(0)K$, the uniqueness of the polar decomposition for the positive definite matrices $Q(0)$ and $\widetilde{Q}(0)$ forces

$$K = I_p$$

and, consequently,

$$Q(\zeta) = \widetilde{Q}(\zeta) \quad \text{for } \zeta \in \mathbb{T}.$$

Therefore,

$$R(0)U^*\widetilde{U} = \widetilde{R}(0),$$

and as $R(0) \succ 0$ and $\widetilde{R}(0) \succ 0$, another application of the uniqueness of the polar decomposition leads to the conclusion

$$U = \widetilde{U}.$$

Thus,

$$R(\zeta) = \widetilde{R}(\zeta) \quad \text{for } \zeta \in \mathbb{T}$$

and the proof of Step 2 is complete. □

Definition 15.2. If $\mathfrak{A}$ is a CMV matrix based on a density Δ that satisfies (D1) and (D2) and a unitary matrix $\beta_{-1} \in \mathbb{C}^{p\times p}$, then we will write $\mathfrak{A} \in \mathbf{W}$.

Corollary 15.3. *If $\Phi \in \mathfrak{W}^{p\times p}$, Θ is defined by* (13.6) *and $\Phi = T_\Theta[\mathcal{E}]$ for a unitary matrix $\mathcal{E} \in \mathbb{C}^{p\times p}$, then there exists exactly one CMV matrix $\mathfrak{A} \in \mathbf{W}$ whose scattering matrix is Φ.*

Proof. In view of Theorem 15.1, there exists exactly one factorization

$$\Phi(\zeta) = UR(\zeta)^*Q(\zeta)^{-1} \quad \text{for } \zeta \in \mathbb{T}, \tag{15.5}$$

where Q, R and U satisfy properties (1)–(5) in Theorem 15.1. In view of Theorem 9.6, the CMV matrix $\mathfrak{A}$ based on $\Delta(\zeta) = Q(\zeta)^*Q(\zeta)$ and $\beta_{-1} = U$ belongs to the class $\mathbf{W}$. Moreover, in view of Definition 10.3, Φ is the scattering matrix of $\mathfrak{A}$. The asserted uniqueness of $\mathfrak{A}$ follows from the uniqueness of the factorization (15.5). □

Remark 15.4. If, in the proof of Theorem 15.1, $Z = I_p$, then in formulas (15.2) and (15.3),

$$V = I_p \quad \text{and} \quad U = \mathcal{E}.$$

16. A mvf of the form (10.5) is a solution of a related Nehari problem

In the following theorem, we shall let $T_\Theta[\mathcal{E}]$ be as in (13.3).

Theorem 16.1. *Suppose $\Phi \in \mathcal{W}^{p\times p}$ and $\beta_{-1} \in \mathbb{C}^{p\times p}$ is a unitary matrix such that*:

(i) $\Phi = \beta_{-1} R^* Q^{-1}$ *is unitary on* $\mathbb{T}$.

(ii) $Q^{\pm 1}, R^{\pm 1} \in \mathcal{W}_+^{p\times p}$, $Q(0) \succ 0$ *and* $R(0) \succ 0$.

(iii) $\frac{1}{2\pi}\int_0^{2\pi} Q(e^{i\theta})^* Q(e^{i\theta}) d\theta = I_p$.

Then

(1) $\|\widehat{\Gamma}_\Phi\| < 1$.

(2) *There exists exactly one mvf* $\Theta \in \mathcal{W}^{2p\times 2p}$ *which satisfies* (13.6).

(3) $\Phi = T_\Theta[\mathcal{E}]$ *for some unitary matrix* $\mathcal{E} \in \mathbb{C}^{p\times p}$.

Proof. In view of the hypotheses (i)–(iii), Φ and R^*Q^{-1} both belong the class $\mathcal{X}^{p\times p}$. It follows from Corollary 13.6 that $\|\widehat{\Gamma}_{R^*Q^{-1}}\| < 1$. But, since $\|\widehat{\Gamma}_\Phi\| = \|\widehat{\Gamma}_{R^*Q^{-1}}\|$, (2) follows immediately. Assertion (2) has already been observed in Theorem 13.1. Assertion (3) follows directly from (13.25). □

A. Scalar results

In this short appendix a number of formulas that have been established earlier for $p \geq 1$ are reviewed in the special case that $p = 1$ in terms of the notation introduced in Section 3. This leads to simplifications and helps to ease comparisons with the extensive literature that is available for classical scalar orthogonal polynomials on $\mathbb{T}$.

Theorem A.1. *If $\{E_n^\pm\}_{n=0}^\infty$ are the polynomials defined by (2.1) and (2.2), respectively, in terms of the Fourier coefficients of a density $\Delta \in \mathcal{W}^{1\times 1}$ which satisfies (D1) and $\{F_n^\pm\}_{n=0}^\infty$ are the polynomials defined by (5.3) in terms of the Schur parameters $\{\beta_n\}_{n=0}^\infty$, then:*

(1) $\gamma_{jk}^{(n)} = \overline{\gamma_{kj}^{(n)}} = \gamma_{n-k,n-j}^{(n)}$ *for* $j, k = 0, \dots, n$.

(2) $E_n^-(\lambda) = F_n^-(\lambda)$ *and* $E_n^+(\lambda) = F_n^+(\lambda)$.

(3) $\lambda^n (E_n^+)^\#(\lambda) = E_n^-(\lambda)$ *and* $\lambda^n (E_n^-)^\#(\lambda) = E_n^+(\lambda)$.

(4) $\beta_n = \gamma_{n0}^{(n)}\{\gamma_{00}^{(n)}\}^{-1} = \gamma_{n0}^{(n)}\{\gamma_{nn}^{(n)}\}^{-1}$.

Proof. If $p = 1$, then, $\Delta(\zeta) = \overline{\Delta(\zeta)}$ for $\zeta \in \mathbb{T}$, and consequently

$$\Delta_j = \overline{\Delta_{-j}} \quad \text{for} \quad j = 0, \pm 1, \dots.$$

Therefore, the Toeplitz matrices $T_n[\Delta]$ and $T_n[\widetilde{\Delta}]$ satisfy

$$T_n[\widetilde{\Delta}] = \overline{T}_n[\Delta] \quad \text{and} \quad \widetilde{\Gamma}_n = \overline{\Gamma}_n = \Gamma_n^T,$$

i.e.,

$$\gamma_{jk}^{(n)} = \overline{\gamma_{kj}^{(n)}} = \gamma_{n-k,n-j}^{(n)}.$$

Thus, (1) holds.

Next, formulas (5.28) and (5.27) imply that the unitary 1×1 matrices U_n and V_n are both positive. Therefore $U_n = V_n = 1$ and hence (2) holds, thanks to the two formulas in (5.25).

Finally, (3) is a straightforward computation and (4) follows from (1) and formulas (5.29) and (5.30), since $U_n = V_n = 1$ and the terms in the indicated formulas commute. □

Remark A.2. Let $\{\phi_n\}_{n=0}^{\infty}$ be the sequence of polynomials constructed by the Gram-Schmidt procedure in (1.1.1) of Simon [31] with respect to a density $\Delta \in \mathcal{W}^{1\times 1}$ that satisfies (D1). In Theorem 1.5.2 of [31], Simon constructed a sequence $\{\alpha_n\}_{n=0}^{\infty}$ such that

$$|\alpha_n| < 1 \quad \text{for } n = 0, 1, \ldots$$

and

$$\begin{bmatrix}\phi_{n+1}(\lambda) & \widehat{\phi}_{n+1}(\lambda)\end{bmatrix} = \begin{bmatrix}\lambda\phi_n(\lambda) & \widehat{\phi}_n(\lambda)\end{bmatrix} H(-\alpha_n) \quad \text{for } n = 0, 1, \ldots, \tag{A.1}$$

where $H(\alpha_n)$ is given by (4.6) and $\widehat{\phi}_n(\lambda) = \lambda^n \phi_n^{\#}(\lambda)$.

Let $\{F_n^-\}_{n=0}^{\infty}$ denote the polynomials defined by (5.3) in terms of the Schur parameters $\{\beta_n\}_{n=0}^{\infty}$ of $\Delta \in \mathcal{W}^{1\times 1}$. Since $\{F_n^-(\lambda)\}_{n=0}^{\infty}$ and $\{\phi_n\}_{n=0}^{\infty}$ are orthonormal sequences of polynomials with respect to Δ with positive leading coefficients, it follows from Theorem 5.8,

$$\phi_n(\lambda) = F_n^-(\lambda) \quad \text{for } n = 0, 1, \ldots. \tag{A.2}$$

Comparing (A.1) and (5.3), it follows easily that

$$\alpha_n = -\beta_{n+1} \quad \text{for } n = 0, 1, \ldots. \tag{A.3}$$

B. Dictionary for matrix polynomials and Schur parameters

To ease the comparison between results formulated in this paper and those that are presented in the basic references [9], [10] and [21], a dictionary of notation is presented below.

Let Δ be a $p \times p$ density which satisfies (D1). A sequence of $p \times p$ matrix polynomials $\{A_n(\lambda)\}_{n=0}^{\infty}$ will be called *LMOP* (left matrix orthogonal polynomials) with respect to Δ if the leading matrix coefficient of $A_n(\lambda)$ is an invertible matrix and

$$\frac{1}{2\pi}\int_0^{2\pi} A_m(e^{i\theta})\Delta(e^{i\theta})A_n(e^{i\theta})^* = \delta_{mn} I_p \quad \text{for } m, n = 0, 1, \ldots.$$

Similarly, a sequence of $p \times p$ matrix polynomials $\{B_n(\lambda)\}_{n=0}^{\infty}$ will be called *RMOP* (right matrix orthogonal polynomials) with respect to Δ if the leading matrix coefficient of $B_n(\lambda)$ is an invertible matrix and

$$\frac{1}{2\pi}\int_0^{2\pi} B_m(e^{i\theta})^*\Delta(e^{i\theta})B_n(e^{i\theta}) = \delta_{mn}I_p \quad \text{for } m,n = 0,1,\dots.$$

The acronyms LMOP$_n$ (resp. RMOP$_n$) stand for the left (resp. right) orthogonal matrix polynomials of degree n with respect to Δ; LC$_n$ denotes the leading coefficient; SP denotes the Schur parameters with respect to Δ.

Object	This paper		[9]		[10]		[21]
LMOP$_n$	$\lambda^n F_n^+\left(\frac{1}{\bar\lambda}\right)^*$	$=$	$\varphi_n^L(\lambda)$	$=$	$P_n(\lambda)$	$=$	$\phi^L(\lambda,n)$
LC$_n$	Y_n		κ_n^L		M_n		$J(n,n)$
RMOP$_n$	$F_n^-(\lambda)$	$=$	$\varphi_n^R(\lambda)$	$=$	$Q_n(\lambda)$	$=$	$\phi^R(\lambda,n)$
LC$_n$	X_n		κ_n^R		N_n		$K(n,n)$
SP$_n$	$-\beta_{n+1}$	$=$	α_n	$=$	ω_n	$=$	τ_n

Table 2.

Warning. In [10] and [21] Schur parameters are denoted by $\{E_n\}_{n=0}^{\infty}$. We have chosen new notation to avoid confusion with the matrix orthogonal polynomials E_n^+ and E_n^- that used in this paper.

The asserted equalities in rows 1 and 3 follow by repeated applications of the following elementary fact:

Lemma B.1. *If $T,S,Y \in \mathbb{C}^{p\times p}$, Y is invertible, $TY^{-1} \succ 0$, $SY^{-1} \succ 0$ and $T = US$ with U unitary, then $T = S$.*

Proof. Under the given assumptions

$$TY^{-1} \succ 0 \quad \text{and} \quad TY^{-1}U^* \succ 0.$$

Therefore, by the uniqueness of the polar decomposition of a matrix, $U = I_p$. □

The verification of the equalities in the above table is divided into steps.

1. Verification of the equalities for the LMOP and RMOP. By definition,

$$F_0^+(\lambda) = \varphi_0^L(\lambda) = P_0(\lambda) = \phi^L(\lambda,0) = I_p \tag{B.1}$$

and

$$F_0^-(\lambda) = \varphi_0^R(\lambda) = Q_0(\lambda) = \phi^R(\lambda,0) = I_p. \tag{B.2}$$

In [9],

$$\kappa_{n+1}^L\{\kappa_n^L\}^{-1} \succ 0 \quad \text{and} \quad \{\kappa_n^R\}^{-1}\kappa_{n+1}^R \succ 0 \quad \text{for } n = 0,1,\dots. \tag{B.3}$$

Similarly, in [10],

$$M_{n+1}\{M_n\}^{-1} \succ 0 \quad \text{and} \quad \{N_n\}^{-1}N_{n+1} \succ 0 \quad \text{for } n = 0,1,\dots, \tag{B.4}$$

whereas, in [21],

$$J(n+1,n+1)^* J(n,n)^{-*} \succ 0 \quad \text{for } n = 0,1,\ldots \tag{B.5}$$

and

$$K(n,n)^{-*} K(n+1,n+1)^* \succ 0 \quad \text{for } n = 0,1,\ldots. \tag{B.6}$$

We will now verify the equalities in row 1 of the table. The verification of the equalities in row 3 is carried out in a similar way using the fact that $F_n^+(0)$ is invertible and hence the leading coefficient of $\lambda^n F_n^+(1/\overline{\lambda})^*$ is invertible. In view of Theorem 5.8, there exist a sequence of $p \times p$ unitary matrices $\{H_n\}_{n=0}^\infty$ such that

$$\varphi_n^R(\lambda) H_n^* = E_n^-(\lambda) \quad \text{for } n = 0,1,\ldots.$$

Consequently,

$$\kappa_n^R H_n^* = \{\gamma_{nn}^{(n)}\}^{1/2}$$

and, hence,

$$\begin{aligned} H_n\{\kappa_n^R\}^{-1}\kappa_{n+1}^R H_{n+1}^* &= \{\gamma_{nn}^{(n)}\}^{-1/2}\{\gamma_{n+1,n+1}^{(n)}\}^{-1/2} \\ &= V_n(I_p - \beta_{n+1}\beta_{n+1}^*)^{1/2} V_{n+1}^* \quad \text{for } n = 0,1,\ldots, \end{aligned}$$

where (5.31) was used to obtain the last line. Thus,

$$V_n^* H_n\{\kappa_n^R\}^{-1}\kappa_{n+1}^R H_{n+1}^* V_{n+1} \succ 0,$$

However, as $V_0 = H_0 = I_p$ and $\{\kappa_0^R\}^{-1}\kappa_1^R \succ 0$ and

$$\{\kappa_0^R\}^{-1}\kappa_1^R H_1^* V_1 \succ 0,$$

Lemma B.1 impllies that $H_1 = V_1^*$. One can continue inductively and deduce that

$$H_n = V_n^* \quad \text{for } n = 0,1,\ldots.$$

Therefore,

$$\varphi_n^R(\lambda) = E_n^-(\lambda) V_n \quad \text{for } n = 0,1,\ldots.$$

Since $E_n^-(\lambda)V_n = F_n^-(\lambda)$ (see (5.26)), the first equality in row 1 of the table holds. The remaining equalities are verified in a similar manner.

2. Verification of the equalities for SP. In view of the identifications made in Step 1, the recursion appearing above formula (3.12) in [9] can be rewritten as

$$\begin{bmatrix} F_{n+1}^-(\lambda) & F_{n+1}^+(\lambda) \end{bmatrix} = \begin{bmatrix} \lambda F_n^-(\lambda) & F_n^+(\lambda) \end{bmatrix} H(-\alpha_n) \quad \text{for } n = 0,1,\ldots,$$

which, upon comparison with (5.4) implies that

$$H(-\alpha_n) = H(\beta_{n+1}) \quad \text{for } n = 0,1,\ldots.$$

Therefore, the equality $-\beta_{n+1} = \alpha_n$ for $n = 0,1,\ldots$ follows from (4.3). The remaining equalities in row 5 of the table follow from the identifications for the matrix orthogonal polynomials.

C. Baxter's inequality

In this section we present a matrix version of Baxter's inequality that is adapted from a paper of Findley [20].

Theorem C.1. *If Δ meets condition (D1), $h(\zeta) = \sum_{j=0}^{n} \zeta^j h_j$ is a $p \times p$ matrix polynomial,*

$$g_k = \begin{cases} \sum_{j=0}^{n} \Delta_{k-j} h_j & \text{for } k = 0, \dots, n \\ 0 & \text{for } k \le -1 \text{ and } k \ge n+1 \end{cases}$$

and $g(\zeta) = \sum_{j=0}^{n} \zeta^j g_j$, then for every choice of $\varepsilon \in (0,1)$, there exists a positive integer n_ε such that

$$\|h\|_{\mathcal{W}} \le \left\{ \frac{\left(\|Q^{-*}\|_{\mathcal{W}} + \|R^{-1}\|_{\mathcal{W}}\right)^2}{1-\varepsilon} + \|\Delta^{-1}\|_{\mathcal{W}} \right\} \|g\|_{\mathcal{W}} \quad \text{when } n \ge n_\varepsilon. \tag{C.1}$$

Proof. Let

$$(\Pi_a^b f)(\zeta) = \sum_{j=a}^{b} \zeta^j f_j \quad \text{for mvf's } f = \sum_{j=-\infty}^{\infty} \zeta^j f_j \text{ in } \mathcal{W}^{p\times p}$$

and set

$$p(\zeta) = \sum_{j=-\infty}^{-1} \zeta^j (\Delta h)_j \quad \text{and} \quad f(\zeta) = \sum_{j=n+1}^{\infty} \zeta^j (\Delta h)_j.$$

(To keep the notation in mind, think of p as the past and f as the future.) Then the following four identities are valid for every point $\zeta \in \mathbb{T}$:

$$\Delta(\zeta)h(\zeta) = p(\zeta) + g(\zeta) + f(\zeta), \tag{C.2}$$

$$Q(\zeta)h(\zeta) = Q(\zeta)^{-*}p(\zeta) + Q(\zeta)^{-*}g(\zeta) + Q(\zeta)^{-*}f(\zeta), \tag{C.3}$$

$$R(\zeta)^{*}h(\zeta) = R(\zeta)^{-1}p(\zeta) + R(\zeta)^{-1}g(\zeta) + R(\zeta)^{-1}f(\zeta), \tag{C.4}$$

and

$$h(\zeta) = \Delta(\zeta)^{-1}p(\zeta) + \Delta(\zeta)^{-1}g(\zeta) + \Delta(\zeta)^{-1}f(\zeta). \tag{C.5}$$

The identity (C.5) implies that

$$\|h\|_{\mathcal{W}} \le \|Q^{-1}\|_{\mathcal{W}}\|Q^{-*}p\|_{\mathcal{W}} + \|\Delta^{-1}\|_{\mathcal{W}}\|g\|_{\mathcal{W}} + \|R^{-*}\|_{\mathcal{W}}\|R^{-1}f\|_{\mathcal{W}}. \tag{C.6}$$

The rest of the proof is broken into steps.

1. Verify the inequality

$$\|\Pi_{-\infty}^{-1} Q^{-*} f\|_{\mathcal{W}} \le \|\Pi_{-\infty}^{-(n+2)} Q^{-*}\|_{\mathcal{W}} \|f\|_{\mathcal{W}}. \tag{C.7}$$

Let $Q(\zeta)^{-*} = L(\zeta) = \sum_{j=-\infty}^{0} \zeta^j L_j$ and $f(\zeta) = \sum_{j=n+1}^{\infty} \zeta^j f_j$. Then

$$\begin{aligned} \|\Pi_{-\infty}^{-1} L f\|_{\mathcal{W}} &= \|\Pi_{-\infty}^{-1}\{(\Pi_{-\infty}^{-(n+2)} L) f\}\|_{\mathcal{W}} \\ &\le \|\Pi_{-\infty}^{-(n+2)} L\|_{\mathcal{W}} \|f\|_{\mathcal{W}}, \end{aligned}$$

which is equivalent to (C.7).

2. Verify the inequality

$$\|\Pi_{n+1}^{\infty} R^{-1} p\|_{\mathcal{W}} \le \|\Pi_{n+2}^{\infty} R^{-1}\|_{\mathcal{W}} \|p\|_{\mathcal{W}}. \tag{C.8}$$

Let $R(\zeta)^{-1} = \sum_{j=0}^{\infty} \zeta^j M_j$ and $p(\zeta) = \sum_{j=-\infty}^{-1} \zeta^j p_j$. Then

$$\begin{aligned}\|\Pi_{n+1}^{\infty} R^{-1} p\|_{\mathcal{W}} &= \|\Pi_{n+1}^{\infty}\{(\Pi_{n+2}^{\infty} R^{-1}) p\}\|_{\mathcal{W}} \\ &\le \|\Pi_{n+2}^{\infty} R^{-1}\|_{\mathcal{W}} \|p\|_{\mathcal{W}},\end{aligned}$$

which is equivalent to (C.8).

3. Verify the inequality

$$\|Q^{-*} p\|_{\mathcal{W}} \le \|Q^{-*}\|_{\mathcal{W}} \|g\|_{\mathcal{W}} + \|\Pi_{-\infty}^{-(n+2)} Q^{-*}\|_{\mathcal{W}} \|R\|_{\mathcal{W}} \|R^{-1} f\|_{\mathcal{W}} \tag{C.9}$$

Since $\Pi_{-\infty}^{-1} Q^{-*} p = Q^{-*} p$, formula (C.3) implies that

$$Q^{-*} p + \Pi_{-\infty}^{-1} \{Q^{-*} g + Q^{-*} f\} = \Pi_{-\infty}^{-1} Q h = 0.$$

Therefore,

$$\begin{aligned}\|Q^{-*} p\|_{\mathcal{W}} &= \|\Pi_{-\infty}^{-1} \{Q^{-*} g + Q^{-*} f\}\|_{\mathcal{W}} \\ &\le \|Q^{-*}\|_{\mathcal{W}} \|g\|_{\mathcal{W}} + \|\Pi_{-\infty}^{-1} Q^{-*} f\|_{\mathcal{W}}.\end{aligned}$$

The inequality (C.9) now follows easily from the last inequality, (C.7) and the observation that

$$\|f\|_{\mathcal{W}} = \|R R^{-1} f\|_{\mathcal{W}} \le \|R\|_{\mathcal{W}} \|R^{-1} f\|_{\mathcal{W}}.$$

4. Verify the inequality

$$\|R^{-1} f\|_{\mathcal{W}} \le \|\Pi_{n+2}^{\infty} R^{-1}\|_{\mathcal{W}} \|Q^*\|_{\mathcal{W}} \|Q^{-*} p\|_{\mathcal{W}} + \|R^{-1}\|_{\mathcal{W}} \|g\|_{\mathcal{W}}. \tag{C.10}$$

Since $\Pi_{n+1}^{\infty}(R^{-1} f) = R^{-1} f$, formula (C.4) implies that

$$\Pi_{n+1}^{\infty} R^{-1} p + \Pi_{n+1}^{\infty} R^{-1} g + R^{-1} f = \Pi_{n+1}^{\infty} R^* h = 0.$$

Therefore,

$$\begin{aligned}\|R^{-1} f\|_{\mathcal{W}} &= \|\Pi_{n+1}^{\infty} \left\{R^{-1} p + R^{-1} g\right\}\|_{\mathcal{W}} \\ &\le \|\Pi_{n+1}^{\infty} R^{-1} p\|_{\mathcal{W}} + \|R^{-1}\|_{\mathcal{W}} \|g\|_{\mathcal{W}}.\end{aligned}$$

The inequality (C.10) now follows easily from the last inequality, (C.8) and the observation that

$$\|p\|_{\mathcal{W}} = \|Q^* Q^{-*} p\|_{\mathcal{W}} \le \|Q^*\|_{\mathcal{W}} \|Q^{-*} p\|_{\mathcal{W}}.$$

5. Verify (C.1). Fix $\varepsilon \in (0,1)$ and choose n_ε so large that

$$\|\Pi_{-\infty}^{-(n+2)} Q^{-*}\|_{\mathcal{W}} \|R\|_{\mathcal{W}} \le \varepsilon \quad \text{and} \quad \|\Pi_{n+2}^{\infty} R^{-1}\|_{\mathcal{W}} \|Q^*\|_{\mathcal{W}} \le \varepsilon \quad \text{when } n \ge n_\varepsilon.$$

Then, by (C.9) and (C.10), the sum

$$\begin{aligned}&\|Q^{-*}p\|_{\mathcal{W}} + \|R^{-1}f\|_{\mathcal{W}} \\ &\qquad \le \varepsilon\left\{\|Q^{-*}p\|_{\mathcal{W}} + \|R^{-1}f\|_{\mathcal{W}}\right\} + \left\{\|Q^{-*}\|_{\mathcal{W}} + \|R^{-1}\|_{\mathcal{W}}\right\}\|g\|_{\mathcal{W}},\end{aligned}$$

since $\|Q^{-1}\|_{\mathcal{W}} = \|Q^{-*}\|_{\mathcal{W}}$ and $\|R^{-1}\|_{\mathcal{W}} = \|R^{-*}\|_{\mathcal{W}}$. Thus, by (C.6),

$$\begin{aligned}\|h\|_{\mathcal{W}} &\le \left\{\|Q^{-1}\|_{\mathcal{W}} + \|R^{-*}\|_{\mathcal{W}}\right\}\left\{\|Q^{-*}p\|_{\mathcal{W}} + \|R^{-1}f\|_{\mathcal{W}}\right\} + \|\Delta^{-1}\|_{\mathcal{W}}\|g\|_{\mathcal{W}} \\ &\le \left\{\frac{\left(\|Q^{-*}\|_{\mathcal{W}} + \|R^{-1}\|_{\mathcal{W}}\right)^2}{1-\varepsilon} + \|\Delta^{-1}\|_{\mathcal{W}}\right\}\|g\|_{\mathcal{W}}.\end{aligned}$$

□

D. Proofs for the Nehari problem

D.1. Preliminary observations.

Lemma D.1. *If $\Phi \in \mathfrak{W}^{p\times p}$, $\gamma_j = \Phi_{-j}$ for $j = 1,2,\ldots$ and $f(\zeta) = \sum_{k=0}^{\infty} \zeta^k f_k$ belongs to H_2^p, then*

$$(\widehat{\Gamma}_\Phi f)(\zeta) = \sum_{j=1}^{\infty} \zeta^{-j} \sum_{k=0}^{\infty} \gamma_{j+k} f_k = \zeta^{-1} \sum_{j=0}^{\infty} \zeta^{-j} (\Gamma_\Phi \mathbf{f})_j \quad \text{for } \zeta \in \mathbb{T}, \tag{D.1}$$

where $\mathbf{f}$ denotes the vector in ℓ_2^p with components f_k, $k = 0,1,\ldots$, and

$$\|\Gamma_\Phi\| = \|\widehat{\Gamma}_\Phi\|. \tag{D.2}$$

Proof. If $\Phi(\zeta) = \sum_{n=-\infty}^{\infty} \zeta^n \Phi_n$, then

$$\begin{aligned}(\widehat{\Gamma}_\Phi f)(\zeta) &= (\mathfrak{q}\Phi f)(\zeta) \\ &= \sum_{j=1}^{\infty} \zeta^{-j} \sum_{k=0}^{\infty} \Phi_{-j-k} f_k\end{aligned}$$

which agrees with the first formula in (D.1). The second formula in (D.1) follows from (13.12). Formula (D.2) follows from the Plancherel formula for Fourier series. □

Lemma D.2. *If $\Phi \in \mathfrak{W}^{p\times p}$, $\gamma_j = \Phi_{-j}$ for $j = 1,2,\ldots$ and $g(\zeta) = \sum_{k=1}^{\infty} \zeta^{-k} g_k$ belongs to $(H_2^p)^\perp$, then*

$$(\widehat{\Gamma}_\Phi^* g)(\zeta) = \sum_{j=0}^{\infty} \zeta^j \left(\sum_{k=1}^{\infty} \gamma_{j+k}^* g_k\right) = \sum_{j=0}^{\infty} \zeta^j (\Gamma_\Phi^* T^* \mathbf{g})_j \quad \text{for } \zeta \in \mathbb{T}, \tag{D.3}$$

where $\mathbf{g}$ denotes the vector in ℓ_2^p with components g_k, $k = 0,1,\ldots$.

Proof. If $g(\zeta) = \sum_{k=1}^{\infty} \zeta^{-k} g_k$ belongs to $(H_2^p)^\perp$, then

$$\begin{aligned}(\widehat{\Gamma}_\Phi^* g)(\zeta) &= (\mathfrak{p} M_{\Phi^*} g)(\zeta)\\ &= \mathfrak{p} \sum_{j=-\infty}^{\infty} \zeta^j \left(\sum_{k=-\infty}^{\infty} \Phi_{j-k}^* g_k \right)\\ &= \sum_{j=0}^{\infty} \zeta^j \left(\sum_{k=1}^{\infty} \gamma_{j+k}^* g_k \right).\end{aligned}$$

This justifies the first equality in (D.3); the second follows from (13.12). □

Lemma D.3. *If* $\Phi \in \mathfrak{W}^{p\times p}$, *then*

$$\widehat{\Gamma}_\Phi \zeta^k f = \mathfrak{q}\zeta^k \widehat{\Gamma}_\Phi f \quad \textit{for } f \in H_2^p \textit{ and } k = 0, 1, \dots \tag{D.4}$$

and

$$\widehat{\Gamma}_\Phi^* \zeta^{-k} g = \mathfrak{p}\zeta^{-k} \widehat{\Gamma}_\Phi^* g \quad \textit{for } g \in (H_2^p)^\perp \textit{ and } k = 0, 1, \dots. \tag{D.5}$$

Proof. If $f \in H_2^p$ and $k = 0, 1, \dots$, then clearly

$$\widehat{\Gamma}_\Phi \zeta^k f = \mathfrak{q}\zeta^k (\mathfrak{p} + \mathfrak{q})\Phi f = \mathfrak{q}\zeta^k \mathfrak{q}\Phi f = \mathfrak{q}\zeta^k \widehat{\Gamma}_\Phi f,$$

which justifies (D.4). The verification of (D.5) is similar. □

Let

$$\Theta(\zeta) = \begin{bmatrix} \theta_{11}(\zeta) & \theta_{12}(\zeta) \\ \theta_{21}(\zeta) & \theta_{22}(\zeta) \end{bmatrix} \stackrel{\text{def}}{=} \begin{bmatrix} \sum_{j=0}^{\infty} \zeta^{-j} a_j & \sum_{j=1}^{\infty} \zeta^{-j} b_j \\ \sum_{j=1}^{\infty} \zeta^{j} c_j & \sum_{j=0}^{\infty} \zeta^{j} d_j \end{bmatrix}. \tag{D.6}$$

With the help of Lemmas D.1 and D.2 it is readily checked that $\Theta(\lambda)$ is a solution of (13.6) in the Wiener algebra $\mathcal{W}^{2p\times 2p}$ if and only if the system of equations

$$\begin{bmatrix} I & -\Gamma_\Phi \\ -\Gamma_\Phi^* & I \end{bmatrix} \begin{bmatrix} \mathbf{a} & T^*\mathbf{b} \\ T^*\mathbf{c} & \mathbf{d} \end{bmatrix} = \begin{bmatrix} YN^{-1/2} & 0 \\ 0 & YM^{-1/2} \end{bmatrix} \tag{D.7}$$

for the vectors

$$\mathbf{a} = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} c_0 \\ c_1 \\ \vdots \end{bmatrix}, \quad \mathbf{d} = \begin{bmatrix} d_0 \\ d_1 \\ \vdots \end{bmatrix},$$

admits a solution with **a**, **b**, **c** and **d** in ℓ_1^p. Since the operator Γ_Φ is compact in ℓ_2^p, a theorem that seems to have originated with Krein (see e.g., the discussion in Gohberg and Zambickii [24], Lemma 7.1 in Adamjan, Arov and Krein [1], and the formulation in Theorem 3.1 in [18]) guarantees that Γ_Φ has the same nonzero spectrum in both ℓ_1^p and ℓ_2^p (as well as a host of other Banach spaces). Therefore, since $\|\Gamma_\Phi\| < 1$ as an operator from ℓ_2^p into itself and the columns of the right hand side of (D.7) belong to ℓ_1^p, the vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ and $\mathbf{d}$ belong to $\ell_1^{p\times p}$, as needed.

It is useful to note that (D.7) is equivalent to the four equations:

$$\mathbf{a} = \Gamma_\Phi T^*\mathbf{c} + YN^{-1/2} = (I - \Gamma_\Phi\Gamma_\Phi^*)^{-1}YN^{-1/2}; \tag{D.8}$$

$$\mathbf{c} = T\Gamma_\Phi^*\mathbf{a} \quad \text{(and hence } c_0 = 0_{p\times p}); \tag{D.9}$$

$$\mathbf{d} = \Gamma_\Phi^* T^*\mathbf{b} + YM^{-1/2} = (I - \Gamma_\Phi^*\Gamma_\Phi)^{-1}YM^{-1/2}; \tag{D.10}$$

$$\mathbf{b} = T\Gamma_\Phi\mathbf{d} \quad \text{(and hence } b_0 = 0_{p\times p}). \tag{D.11}$$

Thus, for example, Lemma D.2 implies that

$$\widehat{\Gamma}_\Phi^*\zeta^{-1}\theta_{11} = \sum_{j=0}^{\infty}\zeta^j\{\Gamma_\Phi^* T^* T\mathbf{a}\}_j.$$

Consequently,

$$\zeta^{-1}\theta_{21} = \widehat{\Gamma}_\Phi^*\zeta^{-1}\theta_{11} \iff \sum_{j=0}^{\infty}\zeta^j c_{j+1} = \sum_{j=0}^{\infty}\zeta^j\{\Gamma_\Phi^* T^* T\mathbf{a}\}_j$$
$$\iff T^*\mathbf{c} = \Gamma_\Phi^*\mathbf{a}.$$

The remaining identifications are verified in much the same way. Moreover,

$$G_\Psi^{(\beta)}(\mathbf{c}\beta + \mathbf{d})\xi = (\mathbf{a}\beta + \mathbf{b})\xi \tag{D.12}$$

and

$$(G_\Psi^{(\beta)})^*(\mathbf{a}\beta + \mathbf{b})\xi = (\mathbf{c}\beta + \mathbf{d})\xi. \tag{D.13}$$

□

D.2. Verification of items (1)–(4) in Theorem 13.1. The preceding discussion guarantees the existence of exactly one mvf solution Θ with blocks θ_{jk}, $j,k = 1,2$, of the form (13.4) to the equation (13.6). The rest of the proof is divided into a number of steps.

1. Verification of (1) *of Theorem 13.1*. This follows from the formulas for the the blocks in (13.6):

$$\zeta^{-1}\theta_{11} = \widehat{\Gamma}_\Phi\zeta^{-1}\theta_{21} + \zeta^{-1}N^{-1/2} = (I - \widehat{\Gamma}_\Phi\widehat{\Gamma}_\Phi^*)^{-1}\zeta^{-1}N^{-1/2}, \tag{D.14}$$

$$\theta_{12} = \widehat{\Gamma}_\Phi\theta_{22}, \tag{D.15}$$

$$\zeta^{-1}\theta_{21} = \widehat{\Gamma}_\Phi^*\zeta^{-1}\theta_{11}, \tag{D.16}$$

and

$$\theta_{22} = \widehat{\Gamma}_\Phi^*\theta_{12} + M^{-1/2} = (I - \widehat{\Gamma}_\Phi^*\widehat{\Gamma}_\Phi)^{-1}M^{-1/2}. \tag{D.17}$$

Thus, for example, in view of formula (13.5),

$$\theta_{22}(0) = [(I - \widehat{\Gamma}_\Phi^*\widehat{\Gamma}_\Phi)^{-1}M^{-1/2}, I_p]_{\mathrm{st}}$$
$$= [(I - \widehat{\Gamma}_\Phi^*\widehat{\Gamma}_\Phi)^{-1}I_p, I_p]_{\mathrm{st}}M^{-1/2} = M^{1/2}.$$

The verification of the formula $\theta_{11}^{\#}(0) = N^{1/2}$ is similar; the verification of the remaining two formulas is easy.

2. Verification of (2) *of Theorem 13.1.* In view of (D.16),

$$\begin{aligned}\langle \zeta^k \zeta^{-1}\theta_{21}u, \zeta^{-1}\theta_{21}v\rangle_{\mathrm{st}} &= \langle \zeta^k \zeta^{-1}\theta_{21}u, \widehat{\Gamma}_\Phi^* \zeta^{-1}\theta_{11}v\rangle_{\mathrm{st}} \\ &= \langle \widehat{\Gamma}_\Phi \zeta^k \zeta^{-1}\theta_{21}u, \zeta^{-1}\theta_{11}v\rangle_{\mathrm{st}} \\ &= \langle \mathfrak{q}\zeta^k \widehat{\Gamma}_\Phi \zeta^{-1}\theta_{21}u, \zeta^{-1}\theta_{11}v\rangle_{\mathrm{st}} \\ &= \langle \zeta^k (\zeta^{-1}\theta_{11} - \zeta^{-1}N^{-1/2})u, \zeta^{-1}\theta_{11}v\rangle_{\mathrm{st}}\end{aligned}$$

for $k = 0, 1, \ldots$ and $u, v \in \mathbb{C}^p$. Therefore,

$$\begin{aligned}\langle \zeta^k(\theta_{11}^*\theta_{11} - \theta_{21}^*\theta_{21})u, v\rangle_{\mathrm{st}} &= \langle \zeta^k N^{-1/2}u, \theta_{11}v\rangle_{\mathrm{st}} \\ &= \begin{cases}\langle N^{-1/2}u, N^{1/2}v\rangle_{\mathrm{st}} & \text{if } k = 0 \\ 0_{p\times p} & \text{if } k = 1, 2, \ldots\end{cases} \\ &= \begin{cases} I_p & \text{if } k = 0 \\ 0_{p\times p} & \text{if } k = 1, 2, \ldots .\end{cases}\end{aligned}$$

Since

$$\theta_{11}(\zeta)^*\theta_{11}(\zeta) - \theta_{21}(\zeta)^*\theta_{21}(\zeta) = \{\theta_{11}(\zeta)^*\theta_{11}(\zeta) - \theta_{21}(\zeta)^*\theta_{21}(\zeta)\}^*,$$

the Fourier coefficients

$$(\theta_{11}^*\theta_{11} - \theta_{21}^*\theta_{21})_k = 0 \quad \text{also for } k = -1, -2, \ldots .$$

Thus,

$$\theta_{11}(\zeta)^*\theta_{11}(\zeta) - \theta_{21}(\zeta)^*\theta_{21}(\zeta) = I_p \quad \text{for every point } \zeta \in \mathbb{T}. \tag{D.18}$$

This justifies the 11 block of the first asserted identity in (13.7). The remaining identities:

$$\theta_{11}(\zeta)^*\theta_{12}(\zeta) - \theta_{21}(\zeta)^*\theta_{22}(\zeta) = 0 \quad \text{for } \zeta \in \mathbb{T} \tag{D.19}$$

and

$$\theta_{12}(\zeta)^*\theta_{12}(\zeta) - \theta_{22}(\zeta)^*\theta_{22}(\zeta) = -I_p \quad \text{for } \zeta \in \mathbb{T} \tag{D.20}$$

are verified in much the same way. The second asserted identity in (13.7) is immediate from the first.

3. Verification of

$$(\theta_{21}\mathcal{E} + \theta_{22})^{-1} \in \mathcal{W}^{p\times p} \quad \textit{and} \quad (\theta_{11} + \theta_{12}\mathcal{E})^{-1} \in \mathcal{W}^{p\times p}$$

for $\mathcal{E} \in \mathcal{W}^{p\times p}$ *and* $\|\mathcal{E}(\zeta)\| \leq 1$ *when* $\zeta \in \mathbb{T}$.

If $\xi^*(\theta_{21}(\zeta)\mathcal{E}(\zeta) + \theta_{22}(\zeta)) = 0$ for some vector $\xi \in \mathbb{C}^p$ and some point $\zeta \in \mathbb{T}$, then it follows from the 22 block of the formula $\Theta(\zeta)j_p\Theta(\zeta)^* = j_p$ on $\mathbb{T}$:

$$\theta_{22}(\zeta)\theta_{22}(\zeta)^* = \theta_{21}(\zeta)\theta_{21}(\zeta)^* + I_p \quad \text{for } \zeta \in \mathbb{T}, \tag{D.21}$$

implies that

$$\begin{aligned}\xi^*\xi &= \xi^*\{\theta_{22}(\zeta)\theta_{22}(\zeta)^* - \theta_{21}(\zeta)\theta_{21}(\zeta)^*\}\xi \\ &= \xi^*\theta_{21}(\zeta)\{\mathcal{E}(\zeta)\mathcal{E}(\zeta)^* - I_p\}\theta_{21}(\zeta)^*\xi \le 0.\end{aligned}$$

Therefore, $\xi = 0$, and hence $(\theta_{21}\mathcal{E} + \theta_{22})$ is invertible on $\mathbb{T}$; $(\theta_{21}\mathcal{E} + \theta_{22})^{-1} \in \mathcal{W}^{p\times p}$ follows by item (1) of Theorem 5.1.

The proof of the second assertion is easily modelled on the proof of the first starting from the 11 block of the formula

$$\Theta(\zeta) j_p \Theta(\zeta)^* = j_p \quad \text{for } \zeta \in \mathbb{T}.$$

4. Verification of (3) *of Theorem 13.1.* Let $\Psi(\zeta) = (T_\Theta[\tau])(\zeta)$. Since $(\theta_{21}\tau + \theta_{22})^{-1} \in \mathcal{W}^{p\times p}$ by Step 3, Ψ belongs to $\mathcal{W}^{p\times p}$. By a straightforward calculation,

$$\begin{aligned}I_p - \Psi(\zeta)^*\Psi(\zeta) &= \{\theta_{21}(\zeta)\tau + \theta_{22}(\zeta)\}^{-*}(I_p - \tau^*\tau)\{\theta_{21}(\zeta)\beta + \theta_{22}(\zeta)\}^{-1} \\ &= 0_{p\times p} \quad \text{for } \zeta \in \mathbb{T},\end{aligned}$$

since $\tau(\zeta)\tau(\zeta)^* = I_p$.

5. Verification of (4) *of Theorem 13.1.* Let $X(\zeta) = \theta_{22}(\zeta)^{-1}\theta_{21}(\zeta)$. In view of (D.21),

$$X(\zeta)X(\zeta)^* = I_p - \theta_{22}(\zeta)^{-1}\theta_{22}(\zeta)^{-*} \quad \text{for } \zeta \in \mathbb{T}.$$

Since $\|X(\zeta)\|$ is continuous on $\mathbb{T}$ and $\|X(\zeta)\| < 1$ for each point $\zeta \in \mathbb{T}$, there exists $0 \le \varepsilon < 1$ such that

$$\|X(\zeta)^{-1}\| = \|\theta_{22}(\zeta)^{-1}\theta_{21}(\zeta)\| \le \varepsilon < 1 \quad \text{for } \zeta \in \mathbb{T}.$$

This completes the proof of the second assertion in (6); the proof of the first is similar. $\square$

D.3. The one step extension. In this subsection we shall show that if $\Phi \in \mathfrak{W}^{p\times p}$, then:

(1) The Hankel operator $\widehat{G}_{\Psi^\circ} = \tilde{\mathfrak{q}}\Psi^\circ|_{H_2^p}$ based on the mvf

$$\Psi^\circ(\zeta) = \gamma_0 + \sum_{j=-\infty}^{-1} \zeta^j\Phi_j = \gamma_0 + \sum_{j=1}^{\infty} \zeta^{-j}\gamma_j$$

is contractive if and only if γ_0 is in the matrix ball

$$\{N^{-1/2}KM^{-1/2} + C_0 : \ K \in \mathbb{C}^{p\times p} \text{ and } K^*K \preceq I_p\} \tag{D.22}$$

with center

$$\begin{aligned}C_0 &= -Y^*\Gamma_1 T^*(I - \Gamma_1^*\Gamma_1)^{-1}YM^{-1} &\text{(D.23)}\\ &= -Y^*\Gamma_1 T^*\Gamma_1^*(I - \Gamma_1 TT^*\Gamma_1^*)^{-1}\Gamma_1 Y, &\text{(D.24)}\end{aligned}$$

where Γ_1 is defined below in (D.25).

(2) If K is unitary, then $\|\widehat{G}_{\Psi^\circ}\| = 1$ and

$$\dim\ker(I - \widehat{G}^*_{\Psi^\circ}\widehat{G}_{\Psi^\circ}) = p.$$

The construction is based on the one step extension method of Adamjan, Arov and Krein [2] and is adapted from [2] with some variations based on the analysis in [18].

It is convenient to work in the (discrete) time domain. Towards this end, let

$$\Gamma_j = \begin{bmatrix} \gamma_j & \gamma_{j+1} & \gamma_{j+2} & \cdots \\ \gamma_{j+1} & \gamma_{j+2} & \gamma_{j+3} & \cdots \\ \gamma_{j+2} & \gamma_{j+3} & \gamma_{j+4} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \quad \text{for } j = 1, 2 \tag{D.25}$$

denote the Hankel operators on ℓ_2^p based on the Fourier coefficients $\{\Phi_{-j}\}_{j=1}^\infty$ and let

$$G = \begin{bmatrix} A & \Gamma_1 \end{bmatrix} = \begin{bmatrix} \gamma_0 & C \\ B & \Gamma_2 \end{bmatrix} \tag{D.26}$$

with

$$A = \begin{bmatrix} \gamma_0 \\ B \end{bmatrix}, \quad B = \Gamma_1 Y \quad \text{and} \quad C = Y^*\Gamma_1. \tag{D.27}$$

Thus,

$$\|\Gamma_2\| \preceq \|\Gamma_1\| = \|\Gamma_\Phi\| < 1,$$

and, in terms of this notation, the matrices M and N in (13.5) can be expressed as

$$N = Y^*(I - \Gamma_1\Gamma_1^*)^{-1}Y \quad \text{and} \quad M = Y^*(I - \Gamma_1^*\Gamma_1)^{-1}Y. \tag{D.28}$$

It is readily checked that

$$\begin{aligned} GG^* \preceq I &\iff AA^* \preceq I - \Gamma_1\Gamma_1^* \\ &\iff (I - \Gamma_1\Gamma_1^*)^{-1/2}AA^*(I - \Gamma_1\Gamma_1^*)^{-1/2} \preceq I \\ &\iff A^*(I - \Gamma_1\Gamma_1^*)^{-1}A \preceq I. \end{aligned}$$

Thus, upon expressing $(I - \Gamma_1\Gamma_1^*)^{-1}$ in block form as

$$(I - \Gamma_1\Gamma_1^*)^{-1} = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix}, \quad \text{with } Z_{11} \in \mathbb{C}^{p\times p}, \tag{D.29}$$

it is easily seen that $GG^* \preceq I$ if and only if

$$\begin{bmatrix} \gamma_0^* & B^* \end{bmatrix} \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} \begin{bmatrix} \gamma_0 \\ B \end{bmatrix} \preceq I. \tag{D.30}$$

The rest of the discussion is broken into steps.

1. $$Z_{11} = Y^*(I - \Gamma_1\Gamma_1^*)^{-1}Y = \{I - C(I - \Gamma_2^*\Gamma_2)^{-1}C^*\}^{-1} = N$$

and

$$Z_{12} = Z_{11}C\Gamma_2^*(I - \Gamma_2\Gamma_2^*)^{-1}.$$

Since $\Gamma_1 = \begin{bmatrix} C \\ \Gamma_2 \end{bmatrix}$,

$$\begin{aligned} I - \Gamma_1\Gamma_1^* &= \begin{bmatrix} I - CC^* & -C\Gamma_2^* \\ -\Gamma_2 C^* & I - \Gamma_2\Gamma_2^* \end{bmatrix} \\ &= \begin{bmatrix} I & -C\Gamma_2^* W^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} X & 0 \\ 0 & W \end{bmatrix} \begin{bmatrix} I & 0 \\ -W^{-1}\Gamma_2 C^* & I \end{bmatrix}, \end{aligned}$$

with

$$\begin{aligned} X &= I - CC^* - C\Gamma_2^*(I - \Gamma_2\Gamma_2^*)^{-1}\Gamma_2 C^* \\ &= I - C(I - \Gamma_2^*\Gamma_2)^{-1}C^* \end{aligned}$$

and $$W = I - \Gamma_2\Gamma_2^*.$$

Therefore, X is invertible,

$$\begin{aligned} \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} &= (I - \Gamma_1\Gamma_1^*)^{-1} \\ &= \begin{bmatrix} I & 0 \\ W^{-1}\Gamma_2 C^* & I \end{bmatrix} \begin{bmatrix} X^{-1} & 0 \\ 0 & W^{-1} \end{bmatrix} \begin{bmatrix} I & C\Gamma_2^* W^{-1} \\ 0 & I \end{bmatrix}, \end{aligned}$$

and hence

$$Z_{11} = Y^*(I - \Gamma_1\Gamma_1^*)^{-1}Y = X^{-1}$$

is invertible, and

$$Z_{12} = X^{-1}C\Gamma_2^* W^{-1}.$$

2. $Z_{22} - Z_{21}Z_{11}^{-1}Z_{12} = (I - \Gamma_2\Gamma_2^*)^{-1}$. By Schur complements,

$$\begin{aligned} (Z_{22} - Z_{21}Z_{11}^{-1}Z_{12})^{-1} &= \begin{bmatrix} 0 & I \end{bmatrix} \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ I \end{bmatrix} \\ &= \begin{bmatrix} 0 & I \end{bmatrix} (I - \Gamma_1\Gamma_1^*) \begin{bmatrix} 0 \\ I \end{bmatrix} \\ &= \begin{bmatrix} 0 & I \end{bmatrix} (I - \begin{bmatrix} C \\ \Gamma_2 \end{bmatrix} \begin{bmatrix} C^* & \Gamma_2^* \end{bmatrix}) \begin{bmatrix} 0 \\ I \end{bmatrix} \\ &= I - \Gamma_2\Gamma_2^*. \end{aligned}$$

3. $GG^* \preceq I$ *if and only if*

$$(\gamma_0^* + B^*Z_{21}Z_{11}^{-1})Z_{11}(\gamma_0 + Z_{11}^{-1}Z_{12}B) \preceq I - B^*(Z_{22} - Z_{21}Z_{11}^{-1}Z_{12})B. \quad \text{(D.31)}$$

This follows easily from (D.30).

4. $\{I - B^*(I - \Gamma_2\Gamma_2^*)^{-1}B\}^{-1} = Y^*(I - \Gamma_1^*\Gamma_1)^{-1}Y = M$. Since $\Gamma_1 = \begin{bmatrix} B & \Gamma_2 \end{bmatrix}$,

$$I - \Gamma_1^*\Gamma_1 = \begin{bmatrix} I - B^*B & -B^*\Gamma_2 \\ -\Gamma_2^*B & I - \Gamma_2^*\Gamma_2 \end{bmatrix}.$$

Therefore, as

$$I - B^*B - B^*\Gamma_2(I - \Gamma_2^*\Gamma_2)^{-1}\Gamma_2^*B = I - B^*(I - \Gamma_2\Gamma_2^*)^{-1}B,$$

$$I - \Gamma_1^*\Gamma_1 = \begin{bmatrix} I & -B^*\Gamma_2(I - \Gamma_2^*\Gamma_2)^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} I - B^*(I - \Gamma_2\Gamma_2^*)^{-1}B & 0 \\ 0 & I - \Gamma_2^*\Gamma_2 \end{bmatrix} \times \begin{bmatrix} I & 0 \\ -(I - \Gamma_2^*\Gamma_2)^{-1}\Gamma_2^*B & I \end{bmatrix}$$

and the advertised formula drops out by computing the 11 block of $(I - \Gamma_1^*\Gamma_1)^{-1}$.

5. The inequality $GG^* \preceq I$ *holds if and only if*

$$\gamma_0 \in \{N^{-1/2}KM^{-1/2} + C_0 : K \in \mathbb{C}^{p\times p} \text{ and } K^*K \preceq I_p\},$$

where

$$C_0 = -Y^*\Gamma_1T^*\Gamma_1^*(I - \Gamma_1TT^*\Gamma_1^*)^{-1}\Gamma_1Y.$$

In view of the formulas in Steps 1 and 4, the constraint (D.31) can be reexpressed as

$$(\gamma_0^* + B^*Z_{21}Z_{11}^{-1})N(\gamma_0 + Z_{11}^{-1}Z_{12}B) \preceq M^{-1}.$$

But this holds if and only if

$$N^{1/2}(\gamma_0 + Z_{11}^{-1}Z_{12}B)M^{1/2} = K$$

is a contraction, i.e., if and only if

$$\begin{aligned}\gamma_0 &= N^{-1/2}KM^{-1/2} - Z_{11}^{-1}Z_{12}B \\ &= N^{-1/2}KM^{-1/2} - C\Gamma_2^*(I - \Gamma_2\Gamma_2^*)^{-1}B \\ &= N^{-1/2}KM^{-1/2} + C_0,\end{aligned}$$

with

$$C_0 = -C\Gamma_2^*(I - \Gamma_2\Gamma_2^*)^{-1}B.$$

But this is the same as the formula for C_0 in (D.24), since $\Gamma_2 = \Gamma_1T$.

6. If γ_0 is in the matrix ball (D.22) *and G is defined by* (D.26)*, then*

$$\dim \ker (I - G^*G) = \dim \ker(I_p - K^*K).$$

If $\xi \in \ell_2^p \cap \ker(I - G^*G)$ and $G\xi = \eta$, then ξ and η must satisfy the following system of equations:

$$\gamma_0\xi_0 + CT^*\xi = \eta_0 \tag{D.32}$$
$$B\xi_0 + \Gamma_2 T^*\xi = T^*\eta \tag{D.33}$$
$$\gamma_0^*\eta_0 + B^*T^*\eta = \xi_0 \tag{D.34}$$
$$C^*\eta_0 + \Gamma_2^*T^*\eta = T^*\xi. \tag{D.35}$$

Equations (D.33) and (D.35) imply that

$$T^*\eta = (I - \Gamma_2\Gamma_2^*)^{-1}\{B\xi_0 + \Gamma_2 C^*\eta_0\}$$

and

$$T^*\xi = (I - \Gamma_2^*\Gamma_2)^{-1}\{\Gamma_2^* B\xi_0 + C^*\eta_0\}.$$

But, upon inserting the last two formulas into (D.32) and (D.34) it follows that

$$\begin{aligned}\{\gamma_0 + C(I - \Gamma_2^*\Gamma_2)^{-1}\Gamma_2^* B\}\xi_0 &= \{I - C(I - \Gamma_2^*\Gamma_2)^{-1}C^*\}\eta_0 \\ &= N^{-1}\eta_0,\end{aligned}$$

by the formulas in Step 1 and, similarly, with the aid of the formulas in Step 4,

$$\{\gamma_0^* + B^*(I - \Gamma_2\Gamma_2^*)^{-1}\Gamma_2 C^*\}\eta_0 = M^{-1}\xi_0.$$

The last two displayed formulas reduce to

$$KM^{-1/2}\xi_0 = N^{-1/2}\eta_0 \quad \text{and} \quad K^*N^{-1/2}\eta_0 = M^{-1/2}\xi_0$$

when γ_0 belongs to the matrix ball specified in Step 5. But this in turn leads easily to the conclusion:

$$(I - G^*G)\xi = 0 \quad \Longrightarrow \quad \text{that the components of } \xi = \begin{bmatrix} Y^*\xi \\ T^*\xi \end{bmatrix} = \begin{bmatrix} \xi_0 \\ T^*\xi \end{bmatrix}$$

meet the constraints

$$(I_p - K^*K)M^{-1/2}\xi_0 = 0 \tag{D.36}$$

and

$$T^*\xi = (I - \Gamma_2^*\Gamma_2)^{-1}\{\Gamma_2^* B + C^*N^{1/2}KM^{-1/2}\}\xi_0. \tag{D.37}$$

A lengthy but straightforward calculation serves to establish the converse: Thus,

$$(I - G^*G)\xi = 0 \quad \Longleftrightarrow \quad \text{(D.36) and (D.37) hold.}$$

Therefore, the assertion in Step 6 holds.

D.4. The proof of Theorem 13.2. Let $\widehat{G}^{(\beta)}_{\Phi^\circ}$ denote the Hankel operator based on the mvf $\Phi^\circ \in \mathcal{W}_-^{p\times p}$ with Fourier coefficients $\{\Phi_k\}_{k=-\infty}^{0}$, where Φ_0 is given by (13.8). Then,

$$(\widehat{G}^{(\beta)}_{\Phi^\circ} f)(\zeta) = (\Phi_0 + \sum_{j=1}^{\infty} \zeta^{-j}\gamma_j) f_0 + \zeta(\widehat{\Gamma}_\Phi R_0 f)(\zeta), \tag{D.38}$$

where in (D.38)

$$(R_0 f)(\lambda) = \frac{f(\lambda) - f(0)}{\lambda},$$

and, in view of assertion (2) in Subsection C.3,

$$\|\widehat{G}^{(\beta)}_{\Phi^\circ}\| = 1.$$

Moreover, with the help of (D.38), it is readily checked that

$$\widehat{G}^{(\beta)}_{\Phi^\circ}(\theta_{21}\beta + \theta_{22}) = \theta_{11}\beta + \theta_{12}. \tag{D.39}$$

Thus,

$$\begin{aligned}
&\langle (I - (\widehat{G}^{(\beta)}_{\Phi^\circ})^*\widehat{G}^{(\beta)}_{\Phi^\circ})(\theta_{21}\beta + \theta_{22}), (\theta_{21}\beta + \theta_{22})\rangle_{\rm st} \\
&\quad = \langle \theta_{21}\beta + \theta_{22}, \theta_{21}\beta + \theta_{22}\rangle_{\rm st} - \langle \widehat{G}^{(\beta)}_{\Phi^\circ}(\theta_{21}\beta + \theta_{22}), \widehat{G}^{(\beta)}_{\Phi^\circ}(\theta_{21}\beta + \theta_{22})\rangle_{\rm st} \\
&\quad = \langle \theta_{21}\beta + \theta_{22}, \theta_{21}\beta + \theta_{22}\rangle_{\rm st} - \langle \theta_{11}\beta + \theta_{12}, \theta_{11}\beta + \theta_{21}\rangle_{\rm st} = 0,
\end{aligned}$$

since $(\theta_{11}\beta + \theta_{12})(\theta_{21}\beta + \theta_{22})^{-1}$ is unitary on $\mathbb{T}$.

1. Verification of formulas (13.19), (13.20) *and* (13.22). Since $I - (\widehat{G}^{(\beta)}_{\Phi^\circ})^*\widehat{G}^{(\beta)}_{\Phi^\circ}$ is positive semidefinite, the preceding set of displayed formulas imply that

$$0_{p\times p} = (I - (\widehat{G}^{(\beta)}_{\Phi^\circ})^*\widehat{G}^{(\beta)}_{\Phi^\circ})(\theta_{21}\beta + \theta_{22}) = \theta_{21}\beta + \theta_{22} - (\widehat{G}^{(\beta)}_{\Phi^\circ})^*(\theta_{11}\beta + \theta_{12}),$$

which justifies (13.20); (13.19) is verified in (D.39). Suppose next that $(I - (\widehat{G}^{(\beta)}_{\Phi^\circ})^*\widehat{G}^{(\beta)}_{\Phi^\circ})f = 0$ and set

$$g = f - (\theta_{21}\beta + \theta_{22})u$$

with $u = \theta_{22}(0)^{-1}f(0)$. Then, since $g(0) = 0$ and $g \in \ker\{I - (\widehat{G}^{(\beta)}_{\Phi^\circ})^*\widehat{G}^{(\beta)}_{\Phi^\circ}\}$, formula (13.27) implies that

$$\|g\|^2 = \|\widehat{G}^{(\beta)}_{\Phi^\circ} g\|^2 = \|\widehat{\Gamma}_\Phi\|^2 \le \|\widehat{\Gamma}_\Phi\|^2\|g\|^2$$

and hence, as $\|\widehat{\Gamma}_\Phi\| < 1$, that $g = 0$. Thus, $\ker\{I - (\widehat{G}^{(\beta)}_{\Phi^\circ})^*\widehat{G}^{(\beta)}_{\Phi^\circ}\}$ is spanned by the p columns of the mvf $\theta_{21}\beta + \theta_{22}$.

2. If $\beta \in \mathbb{C}^{p\times p}$ and $\beta\beta^* = I_p$, then $(\theta_{21}\beta + \theta_{22})^{-1} \in \mathcal{W}_+^{p\times p}$. It suffices to show that the mvf $F(\lambda) = \theta_{21}(\lambda)\beta + \theta_{22}(\lambda)$ is invertible in $\overline{\mathbb{D}}$. Theorem 13.1 guarantees that $F(\lambda)$ is invertible for $\lambda \in \mathbb{T}$. If there exists a point $\omega \in \mathbb{D}$ and a vector $u \in \mathbb{C}^p$ such that $F(\omega)u = 0$, let

$$b_\omega(\lambda) = \frac{\lambda - \omega}{1 - \overline{\omega}\lambda} \quad \text{for } \lambda \in \overline{\mathbb{D}}.$$

Then $b_\omega^{-1}Fu \in \mathcal{W}_+^{p\times 1}$ and

$$\begin{aligned}
\|\widehat{G}_{\Phi^\circ}^{(\beta)}Fu\| &= \|\widetilde{\mathfrak{q}}b_\omega\Phi b_\omega^{-1}Fu\| = \|\widetilde{\mathfrak{q}}b_\omega\widehat{G}_{\Phi^\circ}^{(\beta)}b_\omega^{-1}Fu\| \\
&\le \|\widehat{G}_{\Phi^\circ}^{(\beta)}b_\omega^{-1}Fu\| \le \|b_\omega^{-1}Fu\| \\
&= \|Fu\| = \|(\widehat{G}_{\Phi^\circ}^{(\beta)})^*\widehat{G}_{\Phi^\circ}^{(\beta)}Fu\| \\
&\le \|\widehat{G}_{\Phi^\circ}^{(\beta)}Fu\|.
\end{aligned}$$

Therefore,

$$\|\widehat{G}_{\Phi^\circ}^{(\beta)}b_\omega^{-1}Fu\|^2 = \|b_\omega^{-1}Fu\|^2,$$

i.e.,

$$b_\omega^{-1}Fu \ \text{ is in the kernel of the operator } I - (\widehat{G}_{\Phi^\circ}^{(\beta)})^*\widehat{G}_{\Phi^\circ}^{(\beta)}.$$

Thus, in view of (13.22), there exists a vector $v \in \mathbb{C}^p$ such that

$$b_\omega(\lambda)^{-1}F(\lambda)u = F(\lambda)v \quad \text{for } \lambda \in \overline{\mathbb{D}}, \tag{D.40}$$

which is only possible when $u = v = 0$.

3. If $\beta \in \mathbb{C}^{p\times p}$ and $\beta\beta^* = I_p$, then $(T_\Theta[\beta] - \Phi) \in \mathcal{W}_+^{p\times p}$. In view of (D.14) and (D.15),

$$\begin{aligned}
\mathfrak{q}\{\theta_{11}\beta + \theta_{12} - \Phi(\theta_{21}\beta + \theta_{22})\} &= \mathfrak{q}\zeta\mathfrak{q}\zeta^{-1}\theta_{11}\beta - \mathfrak{q}\zeta\widehat{\Gamma}_\Phi\zeta^{-1}\theta_{21}\beta \\
&= \mathfrak{q}N^{-1/2} = 0.
\end{aligned}$$

Therefore,

$$(\theta_{11}\beta + \theta_{12}) - \Phi(\theta_{21}\beta + \theta_{22}) \in \mathcal{W}_+^{p\times p}$$

and hence, in view of Step 2,

$$(\theta_{11}\beta + \theta_{12})(\theta_{21}\beta + \theta_{22})^{-1} - \Phi \in \mathcal{W}_+^{p\times p}.$$

4. If $\beta \in \mathbb{C}^{p\times p}$ and $\beta\beta^* = I_p$, then $(\theta_{11}\beta + \theta_{12})^{-1} \in \mathcal{W}_-^{p\times p}$. The formula

$$\begin{bmatrix} I_p & \mathcal{E} \end{bmatrix} \Theta(\zeta)^* j_p \Theta(\zeta) \begin{bmatrix} \mathcal{E} \\ I_p \end{bmatrix} = 0_{p\times p}$$

leads easily to a second formula for the linear fractional transformation $T_\Theta[\mathcal{E}]$ for every mvf $\mathcal{E} \in \mathcal{W}^{p\times p} \cap \mathcal{S}^{p\times p}$:

$$T_\Theta[\mathcal{E}] = T_\Theta^\ell[\mathcal{E}] \stackrel{\text{def}}{=} (\theta_{11}^* + \mathcal{E}\theta_{12}^*)^{-1}(\theta_{21}^* + \mathcal{E}\theta_{22}^*). \tag{D.41}$$

Thus, if σ and τ are contractive mvf's in $\mathcal{W}^{p\times p}$, then

$$\begin{aligned} T_\Theta[\sigma] - T_\Theta[\tau] &= T_\Theta^\ell[\sigma] - T_\Theta[\tau] \\ &= (\theta_{11}^* + \sigma\theta_{12}^*)^{-1}(\sigma-\tau)(\theta_{21}\tau+\theta_{22})^{-1}. \end{aligned}$$

Therefore, if $\alpha, \beta \in \mathbb{C}^{p\times p}$ are unitary, then

$$(\theta_{11}^* + \alpha\theta_{12}^*)^{-1}(\alpha-\beta) = (T_\Theta[\alpha] - T_\Theta[\beta])(\theta_{21}\beta+\theta_{22})$$

and hence as the right hand side of the last formula belongs to $\mathcal{W}_+^{p\times p}$ thanks to Step 3, so does the left hand side. The stated result follows easily by choosing unitary matrices α and β for which $\alpha-\beta$ is invertible (e.g., $\beta=-\alpha$).

D.5. Verification of the inclusion $\{T_\Theta[\mathcal{E}] : \mathcal{E} \in \mathcal{S}^{p\times p} \cap \mathcal{W}_+^{p\times p}\} \subseteq \mathcal{N}(\Phi)$ in the setting of Theorem 13.3 and (5) and (6) in Theorem 13.1. The verification is divided into steps.

1. $\theta_{22}^{-1} \in \mathcal{W}_+^{p\times p}$. To this point, we know that $\theta_{22} \in \mathcal{W}_+^{p\times p}$, $\theta_{22}^{-1} \in \mathcal{W}_+^{p\times p}$ and if $\beta \in \mathbb{C}^{p\times p}$ with $\beta\beta^* = I_p$, then $(\theta_{21}\beta+\theta_{22})^{-1} \in \mathcal{W}_+^{p\times p}$. Thus, if $X(\zeta) \stackrel{\text{def}}{=} \theta_{22}(\zeta)^{-1}\theta_{21}(\zeta)\beta$, then

$$\{\theta_{22}(X+I_p)\}^{-1} = (I_p+X)^{-1}\theta_{22}^{-1} \in \mathcal{W}_+^{p\times p}, \tag{D.42}$$

and hence, as $\theta_{22} \in \mathcal{W}_+^{p\times p}$,

$$(I_p+X)^{-1} \in \mathcal{W}_+^{p\times p}.$$

Consequently,

$$G = \{I_p - X\}\{I_p+X\}^{-1} \in \mathcal{W}_+^{p\times p}$$

and

$$\begin{aligned} G &= \{I_p - X\}\{I_p+X\}^{-1} \\ &= \{2I_p - (I_p+X)\}\{I_p+X\}^{-1} \\ &= 2\{I_p+X\}^{-1} - I_p \end{aligned}$$

also belongs to $\mathcal{W}_+^{p\times p}$, since $\|X(\zeta)\| \le \varepsilon < 1$ for $\zeta \in \mathbb{T}$ by item (6) of Theorem 13.1. It is easily seen that

$$G(\zeta)+G(\zeta)^* = 2\{I_p+X(\zeta)^*\}^{-1}\{I_p - X(\zeta)^*X(\zeta)\}\{I_p+X(\zeta)\}^{-1} \succ 0$$

for $\zeta \in \mathbb{T}$, and hence $G(\lambda)+G(\lambda) \succ 0$ for $\lambda \in \overline{\mathbb{D}}$. Thus, the mvf

$$S(\lambda) = \{I_p - G(\lambda)\}\{I_p + G(\lambda)\}^{-1}$$

belongs to $\mathcal{S}^{p\times p} \cap \mathcal{W}_+^{p\times p}$ and $X(\lambda) = S(\lambda)$ for $\lambda \in \overline{\mathbb{D}}$. Consequently,

$$(I_p+X) \in \mathcal{W}_+^{p\times p} \quad \text{and} \quad \theta_{22}^{-1} \in \mathcal{W}_+^{p\times p}$$

by (D.42).

2. $\theta_{11}^{-1} \in \mathcal{W}_-^{p\times p}$. The proof is completed in much the same way as the proof of Step 1.

3. The mvf $T_\Theta[0_{p\times p}] = \theta_{12}\theta_{22}^{-1}$ *belongs to* $\mathcal{N}(\Phi)$. Since $\theta_{22}^{-1} \in \mathcal{W}_+^{p\times p}$, it suffices to check that $\Phi\theta_{22} - \theta_{12} \in \mathcal{W}_+^{p\times p}$. But this follows easily from (D.15).

4. $(\theta_{11} - \theta_{12}\theta_{22}^{-1}\theta_{21}) \in \mathcal{W}_+^{p\times p}$. In view of the identity

$$\theta_{11} - \theta_{12}\theta_{22}^{-1}\theta_{21} = \theta_{11} - (\theta_{12}\theta_{22}^{-1} - \Phi)\theta_{21} - \Phi\theta_{21}$$

and Step 1, it suffices to show that $\theta_{11} - \Phi\theta_{21} \in \mathcal{W}_+^{p\times p}$. But this follows from (D.14) and the observation that

$$\begin{aligned}\mathfrak{q}(\theta_{11} - \Phi\theta_{21}) &= \mathfrak{q}\zeta\{\mathfrak{q}\zeta^{-1}\theta_{11} - \widehat{\Gamma}_\Phi\zeta^{-1}\theta_{21}\} \\ &= \mathfrak{q}\zeta\mathfrak{q}\zeta^{-1}N^{-1/2} = \mathfrak{q}N^{-1/2} = 0.\end{aligned}$$

5. $(\theta_{11} - \theta_{12}\theta_{22}^{-1}\theta_{21})^{-1} \in \mathcal{W}_+^{p\times p}$. In view of (13.7),

$$\begin{aligned}\theta_{11}^*(\theta_{11} - \theta_{12}\theta_{22}^{-1}\theta_{21}) &= \theta_{11}^*\theta_{11} - \theta_{11}^*\theta_{12}\theta_{22}^{-1}\theta_{21} \\ &= \theta_{11}^*\theta_{11} - \theta_{21}^*\theta_{22}\theta_{22}^{-1}\theta_{21} = I_p.\end{aligned}$$

Therefore,

$$(\theta_{11} - \theta_{12}\theta_{22}^{-1}\theta_{21})^{-1} = \theta_{11}^* \in \mathcal{W}_+^{p\times p}.$$

6. $(\theta_{21}\mathcal{E} + \theta_{22})^{-1} \in \mathcal{W}_+^{p\times p}$ *and* $(\theta_{11} + \theta_{12}\mathcal{E}^*)^{-1} \in \mathcal{W}_-^{p\times p}$. To verify the first assertion, it suffices to show that the mvf $\theta_{21}\mathcal{E} + \theta_{22}$ is invertible in $\mathbb{D}$. Suppose to the contrary that

$$\{\theta_{21}(\omega)\mathcal{E}(\omega) + \theta_{22}(\omega)\}\eta = 0 \quad \text{for some } \omega \in \mathbb{D} \text{ and } \eta \in \mathbb{C}^p.$$

Then

$$\|\eta\| = \|\theta_{22}(\omega)^{-1}\theta_{21}(\omega)\mathcal{E}(\omega)\| \leq \varepsilon\|\eta\|,$$

thanks to item (4) of Theorem 13.1 and the maximum modulus principle applied to the mvf $\theta_{22}^{-1}\theta_{21}$ which is holomorphic on $\mathbb{D}$ and belongs to $\mathcal{W}_+^{p\times p}$ thanks to Step 1. Therefore, $\eta = 0$.

The second assertion is verified in much the same way, with the help of item (4) of Theorem 13.1 and Step 2.

7. $T_\Theta[\mathcal{E}] \in \mathcal{N}(\Phi)$. It suffices to show that $T_\Theta[\mathcal{E}] - \Phi \in \mathcal{W}_+^{p\times p}$. In view of Step 1 in D.5 and the identity

$$T_\Theta[\mathcal{E}] - \Phi = T_\Theta[\mathcal{E}] - T_\Theta[0_{p\times p}] + T_\Theta[0_{p\times p}] - \Phi$$

this reduces to showing that $T_\Theta[\mathcal{E}] - T_\Theta[0_{p\times p}] \in \mathcal{W}_+^{p\times p}$. But this is immediate from Steps 4 and 6 of D.5, since

$$T_\Theta[\mathcal{E}] - T_\Theta[0_{p\times p}] = (\theta_{11} - \theta_{12}\theta_{22}^{-1}\theta_{21})\mathcal{E}(\theta_{21}\mathcal{E} + \theta_{22})^{-1}.$$

D.6. A preliminary bound.

Lemma D.4. *If $P \in \mathbb{C}^{p\times p}$ and $P + P^* \succeq \rho I_p$ for some $\rho > 0$, then P is invertible and*

$$\|P^{-1}\| \leq 2\rho^{-1}. \tag{D.43}$$

If $A \in \mathbb{C}^{p\times p}$ and $\|A\| < 1$, then $I_p + A$ is invertible and

$$(I_p + A)^{-1} + (I_p + A^*)^{-1} \succeq I_p. \tag{D.44}$$

Proof. If $P + P^* \succeq \rho I_p$ and $\langle Pv, v\rangle = \mu + i\nu$, with $\mu, \nu \in \mathbb{R}$, then

$$2\mu = \langle (P + P^*)v, v\rangle \geq \rho\langle v, v\rangle = \rho\|v\|^2.$$

Thus, as

$$|\mu| \leq \sqrt{\mu^2 + \nu^2} = |\langle Pv, v\rangle| \leq \|Pv\|\|v\|,$$

it is easily seen that

$$\|Pv\| \geq \rho 2^{-1}\|v\| \quad \text{for } v \in \mathbb{C}^p, \tag{D.45}$$

i.e., P is invertible and (D.43) follows by setting $u = Pv$ in (D.45).

Next, if $\|A\| < 1$, then $I_p + A$ is invertible. If

$$\begin{aligned} V &= (I_p - A)(I_p + A)^{-1} \\ &= \{2I_p - (I_p + A)\}\{I_p + A\}^{-1} \\ &= 2(I_p + A)^{-1} - I_p, \end{aligned}$$

then $V + V^* \succeq 0$. Finally, since $(I_p + A)^{-1} = (V + I_p)/2$,

$$(I_p + A)^{-1} + (I_p + A^*)^{-1} = I_p + (V + V^*)/2 \succeq I_p,$$

i.e., (D.44) holds. □

D.7. Verification of (13.25) and the inclusion $\mathcal{N}(\Phi) \subseteq \{T_\Theta[\mathcal{E}] : \mathcal{E} \in \mathcal{W}_+^{p\times p} \cap \mathcal{S}^{p\times p}\}$ in the setting of Theorem 13.3. If $\Psi \in \mathcal{N}(\Phi)$ and

$$\mathcal{E}(\zeta) = (T_{\Theta^{-1}}[\Psi])(\zeta) \quad \text{for } \zeta \in \mathbb{T},$$

then $\mathcal{E} \in \mathcal{W}^{p\times p}$ and $\|\mathcal{E}(\zeta)\| \leq 1$ for $\zeta \in \mathbb{T}$ and $\Theta^{-1}(\zeta)$ exists due to (13.7). It remains to show that $\mathcal{E} \in \mathcal{W}_+^{p\times p} \cap \mathcal{S}^{p\times p}$. The proof is divided into steps.

1. $\mathcal{E}(I_p + \theta_{22}^{-1}\theta_{21}\mathcal{E})^{-1} \in \mathcal{W}_+^{p\times p}$. Step 3 of D.5 guarantees that

$$T_\Theta[0_{p\times p}] = \theta_{12}\theta_{22}^{-1} \in \mathcal{N}(\Phi)$$

and, hence, $\theta_{12}\theta_{22}^{-1} - \Phi \in \mathcal{W}_+^{p\times p}$. Thus,

$$\Psi - \theta_{12}\theta_{22}^{-1} = (\Psi - \Phi) - (\theta_{12}\theta_{22}^{-1} - \Phi) \in \mathcal{W}_+^{p\times p}.$$

Assertion 1 is an easy consequence of the formula

$$\begin{aligned}\Psi - \theta_{12}\theta_{22}^{-1} &= T_\Theta[\mathcal{E}] - \theta_{12}\theta_{22}^{-1}\\ &= (\theta_{11} - \theta_{12}\theta_{22}^{-1}\theta_{21})\mathcal{E}(\theta_{21}\mathcal{E} + \theta_{22})^{-1},\end{aligned}$$

since

$$\theta_{22}^{-1} \quad \text{and} \quad (\theta_{11} - \theta_{12}\theta_{22}^{-1}\theta_{21})^{-1}$$

belong to $\mathcal{W}_+^{p\times p}$ by Steps 1 and 5 of Subsection D.5, respectively.

2. $(I_p + \theta_{22}^{-1}\theta_{21}\mathcal{E}) \in \mathcal{W}_+^{p\times p}$. Since $\theta_{22}^{-1}\theta_{21} \in \mathcal{W}_+^{p\times p}$, Step 1 and the identity

$$I_p - \theta_{22}^{-1}\theta_{21}\mathcal{E}(I_p + \theta_{22}^{-1}\theta_{21}\mathcal{E})^{-1} = (I_p + \theta_{22}^{-1}\theta_{21}\mathcal{E})^{-1}$$

imply that the mvf

$$F = (I_p + \theta_{22}^{-1}\theta_{21}\mathcal{E})^{-1} \in \mathcal{W}_+^{p\times p}.$$

Since

$$\|\theta_{22}(\zeta)^{-1}\theta_{21}(\zeta)\mathcal{E}(\zeta)\| \le \|\theta_{22}(\zeta)^{-1}\theta_{21}(\zeta)\| \le \varepsilon < 1$$

by item (4) of Theorem 13.1, Lemma D.4 implies that

$$F(\zeta) + F(\zeta)^* \succeq I_p \quad \text{for } \zeta \in \mathbb{T}$$

and, hence, by the Poisson formula

$$\begin{aligned}F(re^{i\theta}) + F(re^{i\theta})^* &= \frac{1}{2\pi}\int_0^{2\pi} \frac{1-r^2}{|e^{i\theta} - re^{i\theta}|^2}\{F(e^{i\theta}) + F(e^{i\theta})^*\}d\theta\\ &\succeq \frac{1}{2\pi}\int_0^{2\pi} \frac{1-r^2}{|e^{i\theta} - re^{i\theta}|^2} I_p d\theta\\ &= I_p,\end{aligned}$$

for $0 \le r < 1$. Thus, by another application of Lemma D.4, $F(\lambda)$ is invertible for every point $\lambda \in \overline{\mathbb{D}}$. Thus,

$$F^{-1} = I_p + \theta_{22}^{-1}\theta_{21}\mathcal{E} \in \mathcal{W}_+^{p\times p}.$$

3. $\mathcal{E} \in \mathcal{W}_+^{p\times p} \cap \mathcal{S}^{p\times p}$. Steps 1 and 2 clearly imply that $\mathcal{E} \in \mathcal{W}_+^{p\times p}$. Since $\|\mathcal{E}(\zeta)\| \le 1$ for $\zeta \in \mathbb{T}$ the maximum modulus principle yields that $\|\mathcal{E}(\lambda)\| \le 1$ for $\lambda \in \overline{\mathbb{D}}$ and hence $\mathcal{E} \in \mathcal{W}_+^{p\times p} \cap \mathcal{S}^{p\times p}$.

4. Verification of (13.25) *in Theorem 13.3.* Suppose first that $X \in \mathcal{N}(\Phi) \cap \mathcal{X}^{p\times p}$. Then, in view of (13.24),

$$X = T_\Theta[\mathcal{E}] \quad \text{for some } \mathcal{E} \in \mathcal{S}_{\text{in}}^{p\times p} \cap \mathcal{W}^{p\times p}.$$

Thus,

$$X_- X_+ = Y_- Y_+$$

with

$$Y_- = \theta_{11} + \theta_{12}\mathcal{E}^* \quad \text{and} \quad Y_+ = \mathcal{E}(\theta_{21}\mathcal{E} + \theta_{22})^{-1}.$$

Therefore, since $X_-^{\pm 1} \in \mathcal{W}_-^{p\times p}$, $X_+^{\pm 1} \in \mathcal{W}_+^{p\times p}$, $Y_-^{\pm 1} \in \mathcal{W}_-^{p\times p}$ and $Y_+ \in \mathcal{W}_+^{p\times p}$,

$$Y_-^{-1}X_- = Y_+X_+^{-1} \quad \text{belongs to } \mathcal{W}_-^{p\times p} \cap \mathcal{W}_+^{p\times p},$$

i.e.,

$$Y_-^{-1}X_- = Y_+X_+^{-1} = K \in \mathbb{C}^{p\times p}.$$

But this implies that

$$X_-(\zeta) = Y_-(\zeta)K \quad \text{and} \quad Y_+(\zeta) = KX_+(\zeta) \quad \text{for all points } \zeta \in \mathbb{T}$$

and hence, as $X_-(\zeta)$ and $Y_-(\zeta)$ are invertible, that K is an invertible matrix and, consequently, Y_+^{-1} also belongs to $\mathcal{W}_+^{p\times p}$. Therefore,

$$\mathcal{E}^{\pm 1} \in \mathcal{W}_+^{p\times p} \cap \mathcal{S}_{\text{in}}^{p\times p}.$$

But this means that both $\mathcal{E}$ and $\mathcal{E}^*$ belong to $\mathcal{W}_+^{p\times p}$ and hence that $\mathcal{E} \in \mathbb{C}^{p\times p}$ and is unitary.

Conversely, if $\mathcal{E}$ is a unitary $p \times p$ matrix, then, (13.16) and (13.17) guarantee that

$$(\theta_{11}\mathcal{E} + \theta_{12})^{\pm 1} \in \mathcal{W}_-^{p\times p} \quad \text{and} \quad (\theta_{21}\mathcal{E} + \theta_{22})^{\pm 1} \in \mathcal{W}_+^{p\times p},$$

respectively. Thus,

$$T_\Theta[\mathcal{E}] = X_-X_+,$$

with $X_- = \theta_{11}\mathcal{E} + \theta_{12}$ and $X_+ = (\theta_{21}\mathcal{E} + \theta_{22})^{-1}$, implies that $T_\Theta[\mathcal{E}] \in \mathcal{X}^{p\times p}$. This completes the proof, since $T_\Theta[\mathcal{E}] \in \mathcal{N}(\Phi)$ by formula (13.24). □

D.8. Proof of Corollary 13.4. Since β is a unitary matrix, Theorem 13.3 guarantees that $\Psi \in \mathcal{N}(\Phi)$ and hence that $\Psi - \Phi \in \mathcal{W}_+^{p\times p}$. Therefore,

$$\Psi_{-k} = \Phi_{-k} \quad \text{for } k = 1, 2, \ldots.$$

It remains to evaluate Ψ_0. In view of formulas (D.6) and the identity

$$\theta_{11}(\zeta)\beta + \theta_{12}(\zeta) = \Psi(\zeta)\{\theta_{21}(\zeta)\beta + \theta_{22}(\zeta)\} \quad \text{for } \zeta \in \mathbb{T},$$

it is easily seen by matching the coefficients of ζ^0 that

$$\begin{aligned}
a_0\beta &= \Psi_0 d_0 + \sum_{k=1}^{\infty} \Psi_{-k}(c_k\beta + d_k) \\
&= \Psi_0 d_0 + \sum_{k=1}^{\infty} \gamma_k c_k \beta + \sum_{k=1}^{\infty} \gamma_k d_k \\
&= \Psi_0 d_0 + \sum_{k=1}^{\infty} \gamma_{k+1}(T^*\mathbf{c})_k\beta + \sum_{k=0}^{\infty} \gamma_{k+1}(T^*\mathbf{d})_k \\
&= \Psi_0 d_0 + \{\Gamma_\Phi(T^*\mathbf{c}\}_0\beta + \{\Gamma_\Phi T^*(d)\}_0.
\end{aligned}$$

Thus, as $a_0 = \{\Gamma_\Phi T^* \mathbf{c}\}_0 + N^{-1/2}$ and $d_0 = M^{1/2}$, it follows that

$$\Psi_0 = \left\{N^{-1/2}\beta - (\Gamma_\Phi T^* \mathbf{d})_0\right\} M^{-1/2},$$

which coincides with (13.26).

The verification of (13.27) is a straightforward calculation. □

References

[1] V. M. Adamjan, D. Z. Arov and M. G. Krein, Analytic properties of the Schmidt pairs of a Hankel operator and the generalized Schur–Takagi problem, *Mat. Sb. (N.S.)*, **86** (1971), no. 128, 34–75. Zbl 0243.47023 MR 0298453

[2] V. M. Adamjan, D. Z. Arov and M. G. Krein, Infinite Hankel block matrices and related extension problems, *Amer. Math. Soc. Transl.*, **111** (1978), no. 2, 321–353.

[3] D. Z. Arov and H. Dym, *J-contractive matrix valued functions and related topics*, Encyclopedia of Mathematics and its Applications, 116, Cambridge University Press, Cambridge, 2008. Zbl 1159.30001 MR 2474532

[4] W. Arveson, *A short course on spectral theory*, Graduate Texts in Mathematics, 209, Springer-Verlag, New York, 2002. Zbl 0997.47001 MR 1865513

[5] G. Baxter, A convergence equivalence related to polynomials orthogonal on the unit circle, *Trans. Amer. Math. Soc.*, **99** (1961), 471–487. Zbl 0116.35705 MR 0126126

[6] N. H. Bingham, Multivariate prediction and matrix Szegő theory, *Probab. Surv.*, **9** (2012), 325–339. Zbl 1285.60038 MR 2956574

[7] N. H. Bingham, Szegő's theorem and its probabilistic descendants, *Probab. Surv.*, **9** (2012), 287–324. Zbl 1285.60037 MR 2956573

[8] M. J. Cantero, L. Moral and L. Velázquez, Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle, *Linear Algebra Appl.*, **362** (2003), 29–56. Zbl 1022.42013 MR 1955452

[9] D. Damanik, A. Pushnitski and B. Simon, The analytic theory of matrix orthogonal polynomials, *Surv. Approx. Theory*, **4** (2008), 1–85. Zbl 1193.42097 MR 2379691

[10] Ph. Delsarte, Y. V. Genin and Y. G. Kamp, Orthogonal polynomial matrices on the unit circle, *IEEE Trans. Circuits and Systems*, **CAS-25** (1978), no. 3, 149–160. Zbl 0408.15018 MR 0481886

[11] Ph. Delsarte, Y. V. Genin and Y. G. Kamp, The Nevanlinna–Pick problem for matrix-valued functions, *SIAM J. Appl. Math.*, **36** (1979), no. 1, 47–61. Zbl 0421.41002 MR 0519182

[12] Ph. Delsarte, Y. V. Genin and Y. G. Kamp, Schur parametrization of positive definite block-Toeplitz systems, *SIAM J. Appl. Math.*, **36** (1979), no. 1, 34–46. Zbl 0417.42013 MR 0519181

[13] Ph. Delsarte, Y. V. Genin and Y. G. Kamp, Schur parametrization of positive definite block-Toeplitz systems, *SIAM J. Appl. Math.*, **36** (1979), no. 1, 34–46. Zbl 0417.42013 MR 0519181

[14] P. Dewilde and H. Dym, Lossless chain scattering matrices and optimum linear prediction: the vector case, *Internat. J. Circuit Theory Appl.*, **9** (1981), no. 2, 135–175. Zbl 0458.94056 MR 0612268

[15] H. Dym, Hermitian block Toeplitz matrices, orthogonal polynomials, reproducing kernel Pontryagin spaces, interpolation and extension, in *Orthogonal matrix-valued polynomials and applications (Tel Aviv, 1987–88)*, 79–135, Oper. Theory Adv. Appl., 34, Birkhäuser, Basel, 1988. Zbl 0658.47018 MR 1021062

[16] H. Dym, Tutorial on a Nehari problem and related reproducing kernel spaces, in *Mathematical methods in systems, optimization, and control*, 129–149, Oper. Theory Adv. Appl., 222, Birkhäuser-Springer Basel AG, Basel, 2012. Zbl 1308.47026 MR 2962780

[17] H. Dym, *Linear algebra in action*, Graduate Studies in Mathematics, 78, American Mathematical Society, Providence, RI, second edition, 2013. Zbl 1285.15001 MR 3154813

[18] H. Dym and I. Gohberg, Extension of matrix valued functions and block matrices, *Indiana Univ. Math. J.*, **31** (1982), no. 5, 733–765. Zbl 0496.47020 MR 0667792

[19] H. Dym and D. P. Kimsey, Trace formulas for a class of truncated block Toeplitz operators, *Linear Algebra Appl.*, **439** (2013), no. 10, 3070–3099. Zbl 1310.47042 MR 3116415

[20] D. F. Findley, Convergence of finite multistep predictors from incorrect models and its role in model selection, Dedicated to the memory of Professor Gottfried Köthe, *Note Mat.*, **11** (1991), 145–155. Zbl 0792.62085 MR 1258544

[21] J. S. Geronimo, Matrix orthogonal polynomials on the unit circle, *J. Math. Phys.*, **22** (1981), no. 7, 1359–1365. Zbl 0505.42019 MR 0626123

[22] J. S. Geronimo, Scattering theory and matrix orthogonal polynomials on the real line, *Circuits Systems Signal Process.*, **1** (1982), no. 3-4, 471–495. Zbl 0506.15010 MR 0730128

[23] I. Gohberg, S. Goldberg and M. A. Kaashoek, *Classes of linear operators. Vol. II*, Operator Theory: Advances and Applications, 63, Birkhäuser Verlag, Basel, 1993. Zbl 0789.47001 MR 1246332

[24] I. Gohberg and M. Zambickii, On the theory of linear operators in spaces with two norms, *Amer. Math. Soc. Transl.*, **85** (1969), no. 2, 145–163.

[25] L. Golinskii, A. Kheifets, F. Peherstorfer and P. Yuditskii, Scattering theory for CMV matrices: uniqueness, Helson–Szegő and strong Szegő theorems, *Integral Equations Operator Theory*, **69** (2011), no. 4, 479–508. Zbl 1235.47029 MR 2784575

[26] I. I. Hirschman, Jr, Matrix-valued Toeplitz operators, *Duke Math. J.*, **34** (1967), 403–415. Zbl 0182.46203 MR 0220002

[27] M. G. Krein, Integral equations on a half-line with kernel depending on the difference of the arguments, *Amer. Math. Soc. Transl. (2)*, **22** (1962), 163–286. Zbl 0119.09601

[28] M. G. Krein and F. È. Melik-Adamjan, On the theory of the S-matrices of canonical differential equations with an integrable potential, *Akad. Nauk Armjan. SSR Dokl.*, **46** (1968), no. 4, 150–155. Zbl 0188.39506 MR 409963

[29] J. Nuttall and S. R. Singh, Orthogonal polynomials and Padé approximants associated with a system of arcs, *J. Approximation Theory*, **21** (1977), no. 1, 1–42. Zbl 0355.30004 MR 0487173

[30] I. Schur, On power series which are bounded in the interior of the unit circle. I, II, Translated from the German, in *Schur methods in operator theory and signal processing*, 31–59, 61–88, Oper. Theory Adv. Appl., 18, Birkhäuser, Basel, 1986. I: Zbl 0618.30001; II: Zbl 0618.30002 MR 0902602

[31] B. Simon, *Orthogonal polynomials on the unit circle. Part 1*, American Mathematical Society Colloquium Publications, 54, American Mathematical Society, Providence, RI, 2005. Zbl 1082.42020 MR 2105088

[32] B. Simon, CMV matrices: five years after, *J. Comput. Appl. Math.*, **208** (2007), no. 1, 120–154. Zbl 1125.15027 MR 2347741

[33] H. Stahl, Orthogonal polynomials with respect to complex-valued measures, in *Orthogonal polynomials and their applications (Erice, 1990)*, 139–154, IMACS Ann. Comput. Appl. Math., 9, Baltzer, Basel, 1991. Zbl 0852.42009 MR 1270224

[34] G. Szegő, *Orthogonal polynomials*, fourth edition, American Mathematical Society, Colloquium Publications, Vol. XXIII, American Mathematical Society, Providence, R.I., 1975. Zbl 0305.42011 MR 0372517

[35] S. Treil and A. Volberg, Wavelets and the angle between past and future, *J. Funct. Anal.*, **143** (1997), no. 2, 269–308. Zbl 0876.42027 MR 1428818

[36] S. Verblunsky, On Positive Harmonic Functions: A Contribution to the Algebra of Fourier Series, *Proc. London Math. Soc.*,, **S2-38** (1935), no. 1, 125–157. Zbl 0010.16202 MR 1576309

Received 02 October, 2015; revised 20 March, 2016