Linear Algebra and its Applications 430 (2009) 762–781

Contents lists available at ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

Extremal classical interpolation problems (matrix case)

B. Fritzsche [a], B. Kirstein [a,*], L.A. Sakhnovich [b]

[a] Mathematisches Institut, Universität Leipzig, Johannisgasse 26, 04103 Leipzig, Germany
[b] 735 Crawford Avenue, Brooklyn, NY 11223, USA

A R T I C L E I N F O

Article history:
Received 27 May 2008
Accepted 12 September 2008
Available online 30 October 2008

Submitted by R.A. Brualdi

Keywords:
Extremal interpolation problem
Degenerate interpolation problem
Subspace of minimality
Condition of regularity

A B S T R A C T

In this paper, we try to find amongst the solutions $w(z)$ of the corresponding interpolation problem the solution which satisfies an additional extremal condition. We show that degenerate interpolation problems play an important role in the theory of extremal interpolation problems. At the end of the paper we accomplish a comparison of our approach with former known approaches.

0. Introduction

In this paper, we consider a particular matrix extremal interpolation problem. More precisely, we try to find amongst the solutions $w(z)$ of the corresponding interpolation problem the solution which satisfies the additional extremal condition

$$w^*(z)w(z) \leqslant \rho_{\min}^2, \quad |z| < 1, \tag{0.1}$$

where $\rho_{\min}$ is a positive Hermitian $m \times m$ matrix. What concerns the statement of the problem we follow the book [18, Chapter 7]. Degenerate interpolation problems play an important role in the theory of extremal interpolation problems. In the monograph [18, Chapter 5] of the third author the solution of a degenerate interpolation problem was constructed. In some special cases the solution could be expressed in a more simple form (see [12]). Using this more simple form it is proved that the corresponding solution is optimal in some sense (see [12]).

* Corresponding author.
E-mail addresses: fritzsche@math.uni-leipzig.de (B. Fritzsche), kirstein@math.uni-leipzig.de (B. Kirstein), Lev.Sakhnovich@verizon.net (L.A. Sakhnovich).

0024-3795/$ - see front matter
doi:10.1016/j.laa.2008.09.013

In this paper, it will be shown that the formula for the solution of a degenerate interpolation problem which was obtained in [18, Chapter 5] can be transformed into an essentially more simple and more efficient formula. In the case of the interpolation problems named after Nevanlinna–Pick and Schur, respectively, our formula coincides with the formulas obtained in the paper [12].

Using the obtained formula we formulate and solve the extremal interpolation problems for a considerably broader class than studied in the paper [12]. We mention that the results of the paper [12] had a conditional character since the conditions for the existence of the matrix $\rho_{\min}$ were not known at this time. In the meantime these existence conditions were found for all problems studied in this paper (see [10]).

The scalar versions of the considered extremal problems have a long history and were studied by several authors (see, e.g., [21,1,7,2,3,4]). Nevertheless some of our results will be even new for the scalar case. (This concerns particularly our investigations on the situation which we call the Jordan block case.)

At the end of this paper we accomplish a short comparison of our approach to interpolation problems with former known approaches (see [21,1,7,2,3,4]).

1. On an interpolation problem associated with an operator identity

It was shown in the monograph [18, Chapters 1 and 2] that an operator identity generates in a natural way a corresponding interpolation problem. In this paper, we will concentrate on operator identities which produce interpolation problems for bounded holomorphic matrix functions.

Let m and n be positive integers. We consider complex $n \times n$ matrices A and S and complex $n \times m$ matrices Ψ_1 and Ψ_2 which are connected by the operator identity

$$S - ASA^* = \Pi j \Pi^*, \tag{1.1}$$

where

$$\Pi := (\Psi_1, \Psi_2), \quad j := \begin{pmatrix} -I_m & 0 \\ 0 & I_m \end{pmatrix}. \tag{1.2}$$

Then the following interpolation problem is associated with the operator identity (1.1): Determine all pairs $[\tau(\varphi), \alpha]$ consisting of a monotonically increasing $m \times m$ matrix-valued function $\tau(\varphi)$ and a Hermitian $m \times m$ matrix α such that the integral representations:

$$S = \frac{1}{2} \int_{-\pi}^{\pi} (I_n - e^{i\varphi} A)^{-1} (\Psi_1 - \Psi_2)[d\tau(\varphi)] \cdot (\Psi_1^* - \Psi_2^*)(I_n - e^{-i\varphi} A^*)^{-1} \tag{1.3}$$

and

$$\begin{aligned} \Psi_1 + \Psi_2 = {} & \int_{-\pi}^{\pi} (I_n - e^{i\varphi} A)^{-1} (-\Psi_1 + \Psi_2) d\tau(\varphi) \\ & + [-\Psi_1 + \Psi_2] \left[i\alpha - \frac{1}{2} \int_{-\pi}^{\pi} d\tau(\varphi) \right] \end{aligned} \tag{1.4}$$

are satisfied.

With each pair $[\tau(\varphi), \alpha]$ satisfying (1.3) and (1.4) we associate the function

$$F(\zeta) := -i\alpha + \frac{1}{2} \int_{-\pi}^{\pi} \frac{e^{i\varphi} - \zeta}{e^{i\varphi} + \zeta} d\tau(\varphi), \quad |\zeta| < 1. \tag{1.5}$$

The pairs $[\tau(\varphi), \alpha]$ satisfying (1.3) and (1.4) are closely related to some matrix inequality associated with the operator identity (1.1). This will be explained now. Let

$$\widetilde{L}(\zeta) := \begin{pmatrix} S & \widetilde{B}(\zeta) \\ \widetilde{B}^*(\zeta) & \widetilde{C}(\zeta) \end{pmatrix}, \quad |\zeta| < 1, \tag{1.6}$$

$$\widetilde{B}(\zeta) = (I_n + \zeta A)^{-1} (\Psi_2 - \Psi_1 w(\zeta)) \tag{1.7}$$

and

$$\widetilde{C}(\zeta) := \frac{I_m - w^*(\zeta)w(\zeta)}{1-\bar{\zeta}\zeta}. \tag{1.8}$$

Then it has been proved in [9, Section 6.2] that a holomorphic $m \times m$ matrix-valued function $w(\zeta)$, $|\zeta| < 1$, satisfies the matrix inequality

$$\widetilde{L}(\zeta) \geqslant 0, \quad |\zeta| < 1 \tag{1.9}$$

if and only if there exists a pair $[\tau(\varphi), \alpha]$ satisfying (1.3) and (1.4) such that the function $F(\zeta)$ defined in (1.5) fulfills

$$w(\zeta) = (F(\zeta) + I_m)(F(\zeta) - I_m)^{-1}, \quad |\zeta| < 1. \tag{1.10}$$

We note that inequality (1.9) is an abstract form of Potapov's fundamental matrix inequality [16].

Now we are going to express the holomorphic $m \times m$ matrix functions $w(\zeta)$ satisfying (1.9) in terms of a particular operatorial calculus which was developed in [9, Chapter 11, 19]. Outgoing from the function $F(\zeta)$ given by (1.5) we set

$$G(\zeta) := [F(-\bar{\zeta})]^*, \quad |\zeta| < 1. \tag{1.11}$$

From (1.5) and (1.11) we conclude

$$G(\zeta) = \mathrm{i}\alpha + \frac{1}{2}\int_{-\pi}^{\pi} \frac{1+\mathrm{e}^{\mathrm{i}\varphi}\zeta}{1-\mathrm{e}^{\mathrm{i}\varphi}\zeta}\,\mathrm{d}\tau(\varphi). \tag{1.12}$$

We define now for each complex $n \times n$ matrix A without spectrum on the unit circle an operator $G(A)$ mapping the set of complex $n \times m$ matrices Ψ into itself by the formula

$$G(A)\Psi := \mathrm{i}\Psi\alpha + \frac{1}{2}\int_{-\pi}^{\pi} (I_n + \mathrm{e}^{\mathrm{i}\varphi}A)(I_n - \mathrm{e}^{\mathrm{i}\varphi}A)^{-1}\Psi\,\mathrm{d}\tau(\varphi). \tag{1.13}$$

Let A be a complex $n \times n$ matrix without spectrum on the unit circle which satisfies

$$I_n - AA^* \geqslant 0. \tag{1.14}$$

Then from [9, Proposition 11.2.3] it follows that the operator $I + G(A)$ is invertible. Thus, the operator $(I + G(A))^{-1}(G(A) - I)$ is well defined. This operator turns out to be closely related to the $m \times m$ matrix function $\tilde{w}(\zeta)$ defined by

$$\tilde{w}(\zeta) := (I_m + G(\zeta))^{-1}(G(\zeta) - I_m), \quad |\zeta| < 1. \tag{1.15}$$

Namely, if

$$\tilde{w}(\zeta) = \sum_{k=0}^{\infty} \zeta^k b_k, \quad |\zeta| < 1, \tag{1.16}$$

denotes the Taylor series of $\tilde{w}(\zeta)$ and Ψ is a complex $n \times m$ matrix, then

$$(I + G(A))^{-1}(G(A) - I)\Psi = \sum_{k=0}^{\infty} A^k \Psi b_k. \tag{1.17}$$

In view of (1.14)–(1.16) we define

$$\tilde{w}(A) := (I + G(A))^{-1}(G(A) - I). \tag{1.18}$$

If the function $w(\zeta)$ is defined by (1.10) then from (1.11) and (1.15) we conclude

$$\tilde{w}(\zeta) = [w(-\bar{\zeta})]^*, \quad |\zeta| < 1. \tag{1.19}$$

From [18, Section 6.2], we obtain now the following result.

Proposition 1.1. *Let the relations (1.1) and (1.2) be fulfilled and let A be a complex $n \times n$ matrix without spectrum on the unit circle which satisfies (1.14). Let $w(\zeta)$ be a holomorphic $m \times m$ matrix function in the unit disc $|\zeta| < 1$ and let the matrix $\widetilde{L}(\zeta)$, $|\zeta| < 1$, be defined by formulas (1.6)–(1.9). Then*

$$\widetilde{L}(\zeta) \geqslant 0, \quad |\zeta| < 1,$$

if and only if

$$\Psi_1 = \tilde{w}(A)\Psi_2, \tag{1.20}$$

where

$$\tilde{w}(\zeta) = [w(-\bar{\zeta})]^*, \quad |\zeta| < 1.$$

Some interrelations of formula (1.20) with well-known results due to Sarason [20] are described in the book [18, Section 11.3] and in the paper [19].

2. On the degenerate case of an interpolation problem associated with an operator identity: the form of the solution

In this section, we will study the interpolation problem associated with the operator identity (1.1) in the degenerate situation. This will be explained now in more detail.

Let the positive integers m and n satisfy the inequality $n > m$ and let $N := n - m$. We suppose that the matrix S is nonnegative Hermitian and satisfies

$$\operatorname{rank} S = N. \tag{2.1}$$

We partition S into blocks via

$$S = \begin{array}{cc} \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} & \begin{array}{l} \}N \\ \}m \end{array} \\ \begin{array}{cc} \underbrace{\phantom{S_{11}}}_{N} & \underbrace{\phantom{S_{12}}}_{m} \end{array} & \end{array} \tag{2.2}$$

and suppose that the matrix S_{11} is positive Hermitian, i.e.

$$S_{11} > 0. \tag{2.3}$$

Moreover, we suppose that the matrix A has the lower block triangular form

$$A = \begin{array}{cc} \begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix} & \begin{array}{l} \}N \\ \}m \end{array} \\ \begin{array}{cc} \underbrace{\phantom{A_{11}}}_{N} & \underbrace{\phantom{A_{22}}}_{m} \end{array} & \end{array}. \tag{2.4}$$

From condition (2.3) and (2.4) it follows that

$$S\begin{pmatrix} -X \\ I_m \end{pmatrix} = 0, \tag{2.5}$$

where the $N \times m$ matrix X is defined by

$$X := S_{11}^{-1} S_{12}. \tag{2.6}$$

We introduce the $m \times m$ matrices

$$M_1 := [-X^*, I_m](I_n + \zeta_0 A)^{-1}\Psi_1 \tag{2.7}$$

and

$$M_2 := [-X^*, I_m](I_n + \zeta_0 A)^{-1}\Psi_2, \tag{2.8}$$

where ζ_0 is a fixed complex number satisfying $|\zeta_0| = 1$ and $\det(I_n + \zeta_0 A) \neq 0$. Following the book [18, p. 84] we introduce the matrix

$$\widetilde{\mathfrak{A}}(\zeta) := I_{2m} + (\zeta_0 - \zeta)\widetilde{\Pi}^*(\zeta I_N + \widetilde{A}^*)^{-1} S_{11}^{-1} (I_N + \zeta_0 \widetilde{A})^{-1} \widetilde{\Pi} j, \tag{2.9}$$

where

$$\widetilde{\Pi}:=[\widetilde{\Psi}_1,\widetilde{\Psi_2}], \tag{2.10}$$
$$\widetilde{\Psi}_k:=[I_N,0_{N\times m}]\Psi_k,\quad k\in\{1,2\} \tag{2.11}$$

and

$$\widetilde{A}=A_{11}. \tag{2.12}$$

We note that the matrices $\widetilde{\Psi}_k, k\in\{1,2\}$, have the size $N\times m$. In the sequel we will use the block partition

$$\tilde{\mathfrak{U}}(\zeta)=\begin{array}{cc} \begin{pmatrix}\tilde a(\zeta) & \tilde b(\zeta)\\ \tilde c(\zeta) & \tilde d(\zeta)\end{pmatrix} & \begin{array}{l}\}m\\ \}m\end{array} \\ \underbrace{\qquad}_{m}\ \underbrace{\qquad}_{m} & \end{array}. \tag{2.13}$$

As it was shown in [18, Section 5.3] the degenerate interpolation problem associated with the operator identity (1.1) has a unique solution $\widetilde{F}(\zeta)$ which is given by the formula

$$\widetilde{F}(\zeta)=[\tilde a(\zeta)M_1^*+\tilde b(\zeta)M_2^*][\tilde c(\zeta)M_1^*+\tilde b(\zeta)M_2^*]^{-1}. \tag{2.14}$$

Remark 2.1. The degenerate interpolation problem under consideration is defined by the relations (2.1)–(2.4). The connection between the function $\widetilde{F}(\zeta)$ and the interpolation problem will be explained later (see Sections 3–5).

Our aim is to reduce formula (2.14) to a simpler form. This is the content of the following result which is crucial for our further considerations.

Lemma 2.2. *The solution $\widetilde{F}(\zeta)$ can be written as*

$$\widetilde{F}(\zeta)=\Psi_1^*\begin{pmatrix}-(\zeta I_N+\widetilde{A}^*)^{-1}[X+A_{21}^*(\zeta I_m+A_{22}^*)^{-1}]\\(\zeta I_m+A_{22}^*)^{-1}\end{pmatrix}$$
$$\times\left[\Psi_2^*\begin{pmatrix}-(\zeta I_N+\widetilde{A}^*)^{-1}[X+A_{21}^*(\zeta I_m+A_{22}^*)^{-1}]\\(\zeta I_m+A_{22}^*)^{-1}\end{pmatrix}\right]^{-1}.$$

Before proving Lemma 2.2 we present its main application.

Hereby we consider the case that $n=km$ with some $k\in\{2,3,\ldots\}$. Let the $m\times m$ matrices $Y_1,Y_2,\ldots,Y_{k-1}$ be defined by the block partition

$$-X=\mathrm{col}\,(Y_1,\ldots,Y_{k-1}). \tag{2.15}$$

(Hereby, the symbol $\mathrm{col}(Y_1,\ldots,Y_{k-1})$ stands for the block column withs blocks $Y_1,\ldots,Y_{k-1}$.) Moreover, let

$$Y_k:=I_m \tag{2.16}$$

and

$$Y:=\mathrm{col}(Y_1,\ldots,Y_k). \tag{2.17}$$

Then the combination of (2.5), (2.15)–(2.17) yields

$$SY=0. \tag{2.18}$$

In view of (2.4) and (2.9) we have

$$(\zeta I_n+A^*)^{-1}=\begin{pmatrix}(\zeta I_N+\widetilde{A}^*)^{-1} & -(\zeta I_N+\widetilde{A}^*)^{-1}A_{21}^*(\zeta I_m+A_{22}^*)^{-1}\\ 0 & (\zeta I_m+A_{22}^*)^{-1}\end{pmatrix}. \tag{2.19}$$

Taking into account formulas (2.15)–(2.19) the application of Lemma 2.2 leads us to our main result.

Theorem 2.3. *Let $n = km$ with some $k \in \{2, 3, \ldots\}$ and let Y be defined by (2.15)–(2.17). Then the solution $\widetilde{F}(\zeta)$ has the form*

$$\widetilde{F}(\zeta) = [\Psi_1^*(\zeta I_n + A^*)^{-1} Y][\Psi_2^*(\zeta I_n + A^*)^{-1} Y]^{-1}. \tag{2.20}$$

Proof of Lemma 2.2. Successively applying (2.13), (2.9), (2.7) and (2.8) we obtain

$$\begin{aligned}
&\begin{pmatrix} \tilde{a}(\zeta) & \tilde{b}(\zeta) \\ \tilde{c}(\zeta) & \tilde{d}(\zeta) \end{pmatrix} \begin{pmatrix} M_1^* \\ M_2^* \end{pmatrix} \\
&\quad = \left\{ I_{2m} + (\zeta_0 - \zeta)\widetilde{\Pi}^*(\zeta I_N + \widetilde{A}^*)^{-1} S_{11}^{-1} (I_N + \zeta_0 \widetilde{A})^{-1} \widetilde{\Pi} j \right\} \\
&\qquad \times \begin{pmatrix} \Psi_1^* \\ \Psi_2^* \end{pmatrix} (I_n + \overline{\zeta_0} A^*)^{-1} \begin{pmatrix} -X \\ I_m \end{pmatrix}.
\end{aligned} \tag{2.21}$$

In view of (2.10), (2.11), and (1.2) we get

$$\widetilde{\Pi} = (I_N, 0_{N\times m})(\Psi_1, \Psi_2) = (I_N, 0_{N\times m})\Pi. \tag{2.22}$$

The combination of (2.21), (2.22) and (1.2) yields

$$\begin{aligned}
\begin{pmatrix} \tilde{a}(\zeta) & \tilde{b}(\zeta) \\ \tilde{c}(\zeta) & \tilde{d}(\zeta) \end{pmatrix} \begin{pmatrix} M_1^* \\ M_2^* \end{pmatrix} &= \Pi^*(I_n + \overline{\zeta_0} A^*)^{-1} \begin{pmatrix} -X \\ I_m \end{pmatrix} \\
&\quad + (\zeta_0 - \zeta)\Pi^* \begin{pmatrix} I_N \\ 0_{m\times N} \end{pmatrix} (\zeta I_N + \widetilde{A}^*)^{-1} S_{11}^{-1} (I_N + \zeta_0 \widetilde{A})^{-1} \\
&\quad \times (I_N, 0_{N\times m})\Pi j \Pi^*(I_n + \overline{\zeta_0} A^*)^{-1} \begin{pmatrix} -X \\ I_m \end{pmatrix}.
\end{aligned} \tag{2.23}$$

Now we consider some part of the second term at the right-hand side of (2.23). Using (1.1) we obtain

$$\begin{aligned}
\Pi j \Pi^*(I_n + \overline{\zeta_0} A^*)^{-1} \begin{pmatrix} -X \\ I_m \end{pmatrix} &= (S - ASA^*)(I_n + \overline{\zeta_0} A^*)^{-1} \begin{pmatrix} -X \\ I_m \end{pmatrix} \\
&= \left(S - \zeta_0 AS[(\overline{\zeta_0} A^* + I_n) - I_n]\right)(I_n + \overline{\zeta_0} A^*)^{-1} \begin{pmatrix} -X \\ I_m \end{pmatrix}.
\end{aligned}$$

According to (2.5) the relations

$$\begin{aligned}
\Pi j \Pi^*(I_n + \overline{\zeta_0} A^*)^{-1} \begin{pmatrix} -X \\ I_m \end{pmatrix} &= [S(I_n + \overline{\zeta_0} A^*)^{-1} - \zeta_0 AS + \zeta_0 AS(I_n + \overline{\zeta_0} A^*)^{-1}] \begin{pmatrix} -X \\ I_m \end{pmatrix} \\
&= (I_n + \zeta_0 A) S (I_n + \overline{\zeta_0} A^*)^{-1} \begin{pmatrix} -X \\ I_m \end{pmatrix}
\end{aligned} \tag{2.24}$$

are true. Taking into account successively (2.4), (2.2), (2.12), and (2.6) we get

$$S_{11}^{-1}(I_N + \zeta_0 \widetilde{A})^{-1}(I_N, 0_{N\times m})(I_n + \zeta_0 A)S = S_{11}^{-1}(S_{11}, S_{12}) = (I_N, X). \tag{2.25}$$

Because of (2.24) and (2.25) we infer

$$\begin{aligned}
&(\zeta_0 - \zeta)\Pi^* \begin{pmatrix} I_N \\ 0_{m\times N} \end{pmatrix} (\zeta I_N + \widetilde{A}^*)^{-1} S_{11}^{-1} (I_n + \zeta_0 \widetilde{A})^{-1} \\
&\quad \cdot (I_n, 0_{N\times m}) \Pi j \Pi^* (I_n + \overline{\zeta_0} A^*) \begin{pmatrix} -X \\ I_m \end{pmatrix} \\
&\quad = (\zeta_0 - \zeta)\Pi^* \begin{pmatrix} I_N \\ 0_{m\times N} \end{pmatrix} (\zeta I_n + \widetilde{A}^*)^{-1} (I_N, X) \cdot (I_n + \overline{\zeta_0} A^*)^{-1} \begin{pmatrix} -X \\ I_m \end{pmatrix}.
\end{aligned} \tag{2.26}$$

The combination of (2.23), (2.26) and (1.2) provides

$$\begin{pmatrix} \tilde{a}(\zeta) & \tilde{b}(\zeta) \\ \tilde{c}(\zeta) & \tilde{d}(\zeta) \end{pmatrix} \begin{pmatrix} M_1^* \\ M_2^* \end{pmatrix} = \begin{pmatrix} \Psi_1^* \\ \Psi_2^* \end{pmatrix} R(\zeta)(I_n + \overline{\zeta_0}A^*)^{-1} \begin{pmatrix} -X \\ I_m \end{pmatrix}, \tag{2.27}$$

where

$$R(\zeta) := \begin{pmatrix} I_N & 0 \\ 0 & I_m \end{pmatrix} + (\zeta_0 - \zeta) \begin{pmatrix} I_N \\ 0_{m\times N} \end{pmatrix} (\zeta I_N + \widetilde{A}^*)^{-1}(I_N, X). \tag{2.28}$$

Our next aim is to compute the matrix

$$R(\zeta)(I_n + \overline{\zeta_0}A^*)^{-1} \begin{pmatrix} -X \\ I_m \end{pmatrix}.$$

In view of (2.24) we have

$$R(\zeta) = \begin{pmatrix} I_N + (\zeta_0 - \zeta)(\zeta I_N + \widetilde{A}^*)^{-1} & (\zeta_0 - \zeta)(\zeta I_N + \widetilde{A}^*)^{-1}X \\ 0 & I_m \end{pmatrix}. \tag{2.29}$$

Taking into account (2.4), (2.12) and (2.29) we infer

$$\left(I_n + \overline{\zeta_0}A^*\right)^{-1} = \begin{pmatrix} (I_N + \overline{\zeta_0}\widetilde{A}^*)^{-1} & -\overline{\zeta_0}(I_N + \overline{\zeta_0}\widetilde{A}^*)^{-1}A_{21}^*(I_m + \overline{\zeta_0}A_{22}^*)^{-1} \\ 0 & (I_m + \overline{\zeta_0}A_{22}^*)^{-1} \end{pmatrix}. \tag{2.30}$$

Using (2.29) and (2.30) we obtain

$$R(\zeta)(I_n + \overline{\zeta_0}A^*)^{-1} = \begin{pmatrix} q_{11}(\zeta) & q_{12}(\zeta) \\ 0 & (I_m + \overline{\zeta_0}A_{22}^*)^{-1} \end{pmatrix}, \tag{2.31}$$

where

$$\begin{aligned} q_{11}(\zeta) &= \left[I_N + (\zeta_0 - \zeta)(\zeta I_N + \widetilde{A}^*)^{-1}\right](I_N + \overline{\zeta_0}\widetilde{A}^*)^{-1} \\ &= (\zeta_0 I_N + \widetilde{A}^*)(\zeta I_N + \widetilde{A}^*)^{-1}(I_N + \overline{\zeta_0}\widetilde{A}^*)^{-1} \\ &= \zeta_0(\zeta I_N + \widetilde{A}^*)^{-1} \end{aligned}$$

and

$$\begin{aligned} q_{12}(\zeta) &= \Big\{-[I_N + (\zeta_0 - \zeta)(\zeta I_N + \widetilde{A}^*)^{-1}]\overline{\zeta_0}(I_N + \overline{\zeta_0}\widetilde{A}^*)^{-1}A_{21}^* \\ &\quad + (\zeta_0 - \zeta)(\zeta I_N + \widetilde{A}^*)^{-1}X\Big\}(I_m + \overline{\zeta_0}A_{22}^*)^{-1} \\ &= \Big\{-(\zeta I_N + \widetilde{A}^*)^{-1}[(\zeta I_N + \widetilde{A}^*) + (\zeta_0 - \zeta)I_N] \\ &\quad \times \overline{\zeta_0}(I_N + \overline{\zeta_0}\widetilde{A}^*)^{-1}A_{21}^* + (\zeta_0 - \zeta)(\zeta I_N + \widetilde{A}^*)^{-1}X\Big\}(I_m + \overline{\zeta_0}A_{22}^*)^{-1}. \end{aligned}$$

Let us simplify the expression for $q_{12}(\zeta)$ in the following form:

$$\begin{aligned} q_{12}(\zeta) &= \Big\{-(\zeta I_N + \widetilde{A}^*)^{-1}\zeta_0(I_N + \overline{\zeta_0}\widetilde{A}^*)\overline{\zeta_0}(I_N + \overline{\zeta_0}\widetilde{A}^*)^{-1}A_{21}^* \\ &\quad + (\zeta_0 - \zeta)(\zeta I_N + \widetilde{A}^*)^{-1}X\Big\}(I_m + \overline{\zeta_0}A_{22}^*)^{-1} \\ &= \Big\{-(\zeta I_N + \widetilde{A}^*)^{-1}A_{21}^* + (\zeta_0 - \zeta)(\zeta I_N + \widetilde{A}^*)^{-1}X\Big\}(I_m + \overline{\zeta_0}A_{22}^*)^{-1} \\ &= (\zeta I_N + \widetilde{A}^*)^{-1}[-A_{21}^* + (\zeta_0 - \zeta)X](I_m + \overline{\zeta_0}A_{22}^*)^{-1}. \end{aligned}$$

Thus, we deduced the following formulas:

$$q_{11}(\zeta) = \zeta_0(\zeta I_N + \widetilde{A}^*)^{-1} \tag{2.32}$$

and

$$q_{12}(\zeta) = (\zeta I_N + \widetilde{A}^*)^{-1}[-A_{21}^* + (\zeta_0 - \zeta)X](I_m + \overline{\zeta_0}A_{22}^*)^{-1}. \tag{2.33}$$

In view of (2.31) we have

$$R(\zeta)(I_n+\overline{\zeta_0}A^*)^{-1}\begin{pmatrix}-X\\ I_m\end{pmatrix}=\begin{pmatrix}-q_{11}(\zeta)X+q_{12}(\zeta)\\ (I_m+\overline{\zeta_0}A_{22}^*)^{-1}\end{pmatrix}. \tag{2.34}$$

Using (2.32) and (2.33) we infer

$$\begin{aligned}&-q_{11}(\zeta)X+q_{12}(\zeta)\\ &\quad=-\zeta_0(\zeta I_n+\widetilde{A}^*)^{-1}X+(\zeta I_N+\widetilde{A}^*)^{-1}[-A_{21}^*+(\zeta_0-\zeta)X]\cdot(I_m+\overline{\zeta_0}A_{22}^*)^{-1}\\ &\quad=(\zeta I_N+\widetilde{A}^*)^{-1}[-\zeta_0X(I_m+\overline{\zeta_0}A_{22}^*)-A_{21}^*+(\zeta_0-\zeta)X](I_m+\overline{\zeta_0}A_{22}^*)^{-1}.\end{aligned}$$

The right-hand side of the last equality can be transformed in the following way:

$$\begin{aligned}-q_{11}(\zeta)X+q_{12}(\zeta)=&-(\zeta I_N+\widetilde{A}^*)^{-1}[X+A_{21}^*(\zeta I_m+A_{22})^{-1}](\zeta I_m+A_{22})\\ &\times(I_m+\overline{\zeta_0}A_{22}^*)^{-1}.\end{aligned} \tag{2.35}$$

The combination of (2.14) and (2.35) shows that the solution $\widetilde{F}(z)$ can be written in the form:

$$\begin{aligned}\widetilde{F}(\zeta)=&\,\Psi_1^*\begin{pmatrix}-(\zeta I_N+\widetilde{A}^*)^{-1}[X+A_{21}^*(\zeta I_m+A_{22}^*)^{-1}]\\ (\zeta I_m+A_{22}^*)^{-1}\end{pmatrix}\\ &\times\left[\Psi_2^*\begin{pmatrix}-(\zeta I_N+\widetilde{A}^*)^{-1}[X+A_{21}^*(\zeta I_m+A_{22}^*)^{-1}]\\ (\zeta I_m+A_{22}^*)^{-1}\end{pmatrix}\right]^{-1}.\end{aligned}$$

The proof is complete. $\square$

3. On extremal matrix interpolation problems

In this section, we state some basic facts on extremal matrix interpolation problems. Most of the material is taken from the monograph [18, Chapter 7] and the authors' paper [10].

Let the matrices A, S_k and $\Psi_k, k=1,2$, have the sizes $mL\times mL$ and $mL\times m$, respectively, where S_k is nonnegative Hermitian. We suppose that these matrices are connected by the relations

$$S_k-AS_kA^*=\Psi_k\Psi_k^*,\quad k=1,2. \tag{3.1}$$

Setting

$$S:=S_2-S_1. \tag{3.2}$$

We deduce from (3.1) and (3.2) the equality

$$S-ASA^*=\Psi_2\Psi_2^*-\Psi_1\Psi_1^*. \tag{3.3}$$

We introduce the block-diagonal matrix

$$R:=\operatorname{diag}\underbrace{(\rho,\ldots,\rho)}_{L},$$

where ρ is a positive Hermitian matrix of size $m\times m$. In addition we shall assume the equality

$$AR=RA. \tag{3.4}$$

This is justified, since it was shown in [12] that condition (3.4) is true in a number of concrete examples.

From Eqs. (3.1) and (3.4) it follows that

$$S_\rho-AS_\rho A^*=\Psi_2\Psi_2^*-\Psi_{1,\rho}\Psi_{1,\rho}^*, \tag{3.5}$$

where

$$S_\rho:=S_2-R^{-1}S_1R^{-1}, \tag{3.6}$$

$$\Psi_{1,\rho}:=R^{-1}\Psi_1. \tag{3.7}$$

Thus we have constructed a set of operator identities (3.6), where the positive Hermitian matrix ρ plays the role of a parameter. A set of interpolation problems, see [18, Chapter 6] corresponds to this set of operator identities. A necessary condition for the solvability of these problems is the inequality

$$RS_2R - S_1 \geqslant 0. \tag{3.8}$$

Now we turn to extremal interpolation.

Definition 3.1. We shall call the matrix $\rho = \rho_{\min} > 0$ a minimal solution of inequality (3.8) if the following two requirements are fulfilled:

1. The inequality

$$R_{\min}S_2R_{\min} - S_1 \geqslant 0 \tag{3.9}$$

holds where

$$R_{\min} = \operatorname{diag} \underbrace{(\rho_{\min}, \ldots, \rho_{\min})}_{L}$$

is valid.

2. If $\rho > 0$ satisfies inequality (3.8), then

$$\operatorname{rank}(R_{\min}S_2R_{\min} - S_1) \leqslant \operatorname{rank}(RS_2R - S_1). \tag{3.10}$$

(In other words, $R_{\min}$ minimizes the rank of $RS_2R - S_1 \geqslant 0$.)

Remark 3.2. The existence of $\rho_{\min}$ follows directly from Definition 3.1.

We shall write the nonnegative Hermitian matrices S_1, S_2 and R in the following block forms:

$$S_k = \begin{pmatrix} S_{11}^{(k)} & S_{12}^{(k)} \\ S_{21}^{(k)} & S_{22}^{(k)} \end{pmatrix}, \quad k = 1, 2, \tag{3.11}$$

$$R = \begin{pmatrix} R_1 & 0 \\ 0 & \rho \end{pmatrix}, \quad R_1 = \operatorname{diag} \underbrace{(\rho, \ldots, \rho)}_{L-1}, \tag{3.12}$$

where $S_{22}^{(k)}$ are blocks of size $m \times m$, $S_{11}^{(k)}$ has the size $(L-1)m \times (L-1)m$ and $S_{12}^{(k)}$ has the size $(L-1)m \times m$. The following result is proved in [18, Proposition 7.1.1].

Proposition 3.3. *Suppose that for all $\rho > 0$ satisfying inequality (3.8) the upper diagonal block is positive Hermitian, i.e., that*

$$R_1S_{11}^{(2)}R_1 - S_{11}^{(1)} > 0$$

holds. If $\rho = q > 0$ satisfies inequality (3.8) and the relation

$$qS_{22}^{(2)}q = S_{22}^{(1)} + C_1^*(Q_1S_{11}^{(2)}Q_1 - S_{11}^{(1)})^{-1}C_1, \tag{3.13}$$

where

$$Q_1 := \operatorname{diag} \underbrace{(q, q, \ldots, q)}_{L-1}, \quad C_1 := Q_1S_{12}^{(2)}q - S_{12}^{(1)}, \tag{3.14}$$

then

$$\rho_{\min} = q.$$

Now we investigate the question of existence of the corresponding matrix q. For this reason, we consider the equation

$$qS_{22}^{(2)}q = S_{22}^{(1)} + S_{12}^{*}(Q_1 S_{11}^{(2)} Q_1 - S_{11}^{(1)})^{-1} S_{12}, \tag{3.15}$$

where

$$S_{12} := Q_1 S_{12}^{(2)} q - S_{12}^{(1)}. \tag{3.16}$$

We make the following two assumptions.

Condition 1. The matrix S_2 has the block structure

$$S_2 = (C_{jk})_{j,k=1}^{n},$$

where all $m \times m$ blocks C_{jk} have the shape

$$C_{jk} = \alpha_{jk} I_m$$

with some complex number α_{jk}.

Condition 2. The matrix S_2 is positive Hermitian ($S_2 > 0$).

In view of Condition 1 we obtain

$$\begin{aligned} S_{12} &= Q_1^2 S_{12}^{(2)} - S_{12}^{(1)} \\ &= Q_1^2 S_{11}^{(2)} (S_{11}^{(2)})^{-1} S_{12}^{(2)} - S_{12}^{(1)} \\ &= [Q_1^2 S_{11}^{(2)} - S_{11}^{(1)}](S_{11}^{(2)})^{-1} S_{12}^{(2)} + S_{11}^{(1)} (S_{11}^{(2)})^{-1} S_{12}^{(2)} - S_{21}^{(1)} \end{aligned}$$

and

$$Q_1 S_{11}^{(2)} Q_1 = Q_1^2 S_{11}^{(2)}.$$

We introduce the following notations:

$$E := Q_1^2 S_{11}^{(2)} - S_{11}^{(1)}, \quad B := (S_{11}^{(2)})^{-1} S_{12}^{(2)} \tag{3.17}$$

and

$$C := S_{11}^{(1)} (S_{11}^{(2)})^{-1} S_{12}^{(2)} - S_{12}^{(1)}. \tag{3.18}$$

Then obviously $E^* = E$ and Eq. (3.15) can be written in the form:

$$\alpha_{nn} q^2 = S_{22}^{(1)} + (B^* E + C^*) E^{-1} (EB + C)$$

or

$$\alpha_{nn} q^2 = S_{22}^{(1)} + B^* E B + B^* C + C^* B + C^* E^{-1} C. \tag{3.19}$$

Using (3.17) and (3.19) we infer

$$q^2 T = U + C^* E^{-1} C, \tag{3.20}$$

where

$$T := \alpha_{nn} I_m - (S_{12}^{(2)})^* (S_{11}^{(2)})^{-1} S_{12}^{(2)} \tag{3.21}$$

and

$$\begin{aligned} U &:= B^* C + C^* B - B^* S_{11}^{(1)} B + S_{22}^{(1)} \\ &= (S_{12}^{(2)})^* \left(S_{11}^{(2)}\right)^{-1} S_{11}^{(1)} (S_{11}^{(2)})^{-1} S_{12}^{(2)} - (S_{12}^{(2)})^* (S_{11}^{(2)})^{-1} S_{12}^{(1)} \\ &\quad - (S_{12}^{(1)})^* (S_{11}^{(2)})^{-1} S_{12}^{(2)} + S_{22}^{(1)}. \end{aligned} \tag{3.22}$$

In view of Condition 2 the relation

$$T > 0 \tag{3.23}$$

is true.

According to Condition 1 and (3.21) the matrix T has scalar type, i.e.

$$T = \beta I_m. \tag{3.24}$$

From (3.23) and (3.24) it follows that

$$\beta > 0.$$

Hence Eq. (3.20) takes the form

$$q^2 = \frac{1}{\beta}(U + C^* E^{-1} C). \tag{3.25}$$

If we compare Eqs. (3.15) and (3.25) we see that S_{12} depends on q, but C does not depend on q. Taking into account this fact we can apply Theorem 3.3 of the paper [17] to Eq. (3.25). Moreover, we observe that the matrix E can be represented in the form

$$E = DQ_1^2 D - S_{11}^{(1)},$$

where

$$D := \sqrt{S_{11}^{(2)}} > 0.$$

Now we rewrite Eq. (3.25) in the form

$$q^2 = \frac{1}{\beta}[U + C_1^*(Q_1^2 - D^{-1} S_{11}^{(1)} D^{-1}) C_1], \tag{3.26}$$

where

$$C_1 := D^{-1} C.$$

We introduce the notation

$$\widetilde{U} := \mathrm{diag}\,\underbrace{(U, U, \ldots, U)}_{L-1}.$$

Definition 3.4. We call an interpolation problem regular if the condition

$$\frac{1}{\beta}\widetilde{U} > D^{-1} S_{11}^{(1)} D^{-1}$$

is satisfied.

In the paper [10], we obtained the following result.

Theorem 3.5. *Let the Conditions 1 and 2 be fulfilled and let the interpolation problem be regular. If $S_{11}^{(1)} \geqslant 0$ then Eq. (3.15) has a unique solution q such that $q > 0$ and $Q_1 S_{11}^{(2)} Q_1 > S_{11}^{(1)}$.*

Corollary 3.6. *Under the assumptions of Theorem 3.5 the relation $\rho_{\min}^2 = q^2$ holds.*

Remark 3.7. Under the assumption that $\rho_{\min}$ is known an explicit representation of the solution of the corresponding extremal interpolation problem is given in monograph [18, Chapter 7]. In the special case of the extremal interpolation problems named after Schur and Nevanlinna–Pick, respectively, this solution is written in a simpler form in the paper [12]. In the general case the solution was written in a simpler form in Section 2 of this paper.

We have shown that there is one and only one positive Hermitian solution of Eq. (3.15) which satisfies condition (3.8). In the case $L = 2$ in [9] it was proved that Eq. (3.15) has one and only one positive Hermitian solution.

4. On extremal interpolation for bounded holomorphic matrix functions

First we assume that the matrix S is defined by (3.2) where the matrices S_1 and S_2 satisfy the operator identities (3.1).

Let the conditions of Theorem 2.3 be fulfilled. Then the matrix function $w_{\min}(\zeta):=\widetilde{F}(-\zeta)$ can be written in the form (see (2.20))

$$w_{\min}(\zeta) = P_1(\zeta)[P_2(\zeta)]^{-1}, \tag{4.1}$$

where

$$P_s(\zeta) = \Psi_s^*(-\zeta I_{mL} + A^*)^{-1}Y, \quad s \in \{1,2\} \tag{4.2}$$

and

$$Y = \mathrm{col}(Y_1, \dots, Y_L). \tag{4.3}$$

Now we can formulate the first result of this section.

Theorem 4.1. *Let the $m \times m$ matrix function $w(\zeta)$ be holomorphic in the unit disk $|\zeta| < 1$ and satisfy the following conditions*:

1. *The $m \times m$ matrix function*

$$\varphi(\zeta):=w(\zeta)P_2(\zeta) \tag{4.4}$$

can be represented in the form

$$\varphi(\zeta) = P_1(\zeta) + Q(\zeta), \tag{4.5}$$

where

$Q(\zeta) = c_0 + c_1\zeta + c_2\zeta^2 + \cdots$

is holomorphic matrix function in the disk $|\zeta| < 1$.

2. *The inequality*

$$w^*(\zeta)w(\zeta) \leqslant \rho_1^2, \quad |\zeta| < 1 \tag{4.6}$$

is fulfilled where ρ_1 is some positive Hermitian $m \times m$ matrix.

Let T be a complex $mL \times mL$ matrix which is a lower block triangular matrix with $m \times m$ blocks and satisfies

$$T^*T = S_2. \tag{4.7}$$

Then we have the inequality

$$Y^*T^*TY \leqslant Y^*T^*R_{\rho_1}^2TY, \tag{4.8}$$

where

$$R_{\rho_1} = \mathrm{diag}(\underbrace{\rho_1, \dots, \rho_1}_{L}). \tag{4.9}$$

Moreover, if $\rho_1 \neq I_m$ then there exists some vector $h \in \mathbb{C}^m$ such that

$$h^*Y^*T^*TYh < h^*Y^*T^*R_{\rho_1}^2 TYh. \tag{4.10}$$

Proof. Let the number $r_0 \in (0,1)$ be chosen such that each point α belonging to the spectrum of the matrix A satisfies the inequality $|\alpha| < r_0$. From (4.4) and (4.6) we obtain then

$$\frac{1}{2\pi}\int_0^{2\pi} \varphi^*(r_0 e^{i\theta})\varphi(r_0 e^{i\theta})d\theta \leqslant \frac{1}{2\pi}\int_0^{2\pi} P_2^*(r_0 e^{i\theta})\rho_1^2 P_2(r_0 e^{i\theta})d\theta. \tag{4.11}$$

From (4.11) and (4.2) we deduce

$$\frac{1}{2\pi}\int_0^{2\pi} \varphi^*(r_0 e^{i\theta})\varphi(r_0 e^{i\theta})d\theta \leqslant Y^*\left(\sum_{j=0}^{\infty} A^j \Psi_2 \rho_1^2 \Psi_2^*(A^j)^* r_0^{-(2j+2)}\right) Y. \tag{4.12}$$

Since the $m \times m$ blocks of Ψ_2 and A commute with ρ_1 the identity

$$\sum_{j=0}^{\infty} A^j \Psi_2 \rho_1^2 \Psi_2^*(A_j)^* = R_{\rho_1}\left(\sum_{j=0}^{\infty} A^j \Psi_2 \Psi_2^*(A^j)^*\right) R_{\rho_1} \tag{4.13}$$

is true. In view of (3.1) we infer from [6, Theorem A3.4, part (a)] the identity

$$\sum_{j=0}^{\infty} A^j \Psi_l \Psi_l^*(A^j)^* = S_l, \quad l \in \{1,2\}. \tag{4.14}$$

Combining (4.12)–(4.14) we get

$$\lim_{r_0 \to 1-0} \frac{1}{2\pi}\int_0^{2\pi} \varphi^*(r_0 e^{i\theta})\varphi(r_0 e^{i\theta})d\theta \leqslant Y^* R_{\rho_1} S_2 R_{\rho_1} Y. \tag{4.15}$$

In view of (4.2) and (4.5) the equality

$$\frac{1}{2\pi}\int_0^{2\pi} \varphi^*(r_0 e^{i\theta})\varphi(r_0 e^{i\theta})d\theta = Y^*\left(\sum_{j=0}^{\infty} A^j \Psi_1 \Psi_1^*(A^j)^* r_0^{-(2j+2)}\right) Y + \sum_{j=0}^{\infty} c_j^* c_j r_0^{2j} \tag{4.16}$$

is valid. From (4.16) and (4.14) it follows

$$\lim_{r_0 \to 1-0} \frac{1}{2\pi}\int_0^{2\pi} \varphi^*(r_0 e^{i\theta})\varphi(r_0 e^{i\theta})d\theta \geqslant Y^*\left(\sum_{j=0}^{\infty} A^j \Psi_1 \Psi_1^*(A^j)^*\right) Y = Y^* S_1 Y,$$

whereas inequality (4.15) implies

$$\lim_{r_0 \to 1-0} \frac{1}{2\pi}\int_0^{2\pi} \varphi^*(r_0 e^{i\theta})\varphi(r_0 e^{i\theta})d\theta \leqslant Y^* R_{\rho_1} S_2 R_{\rho_1} Y.$$

Thus, we have

$$Y^* S_1 Y \leqslant Y^* R_{\rho_1} S_2 R_{\rho_1} Y. \tag{4.17}$$

In view of (3.2), (4.3) and (2.18) we get

$$(S_2 - S_1)Y = SY = 0.$$

Hence

$$Y^*S_2Y = Y^*S_1Y. \tag{4.18}$$

From (4.18) and (4.17) we infer

$$Y^*S_2Y \leqslant Y^*R_{\rho_1}S_2R_{\rho_1}Y. \tag{4.19}$$

From Condition 1 and the shape of the matrix R_{ρ_1} it follows

$$S_2R_{\rho_1} = R_{\rho_1}S_2.$$

From this and the choice of T it follows

$$TR_{\rho_1} = R_{\rho_1}T.$$

Hence, inequality (4.19) implies inequality (4.8). The inequality (4.10) follows from (4.5) and (4.16) and the fact that the function $Q(\zeta)$ does not identically vanish. The theorem is proved. $\square$

Remark 4.2. From Theorem 2.3 and the construction of the function $\widetilde{F}(\zeta)$ it is clear that

$$\widetilde{F}^*(\zeta)\widetilde{F}(\zeta) \leqslant I_m, \quad |\zeta| < 1.$$

Hence, we have

$$w^*_{\min}(\zeta)w_{\min}(\zeta) \leqslant I_m, \quad |\zeta| < 1. \tag{4.20}$$

Now we consider the case that the matrices S_1 and S_2 satisfy the operator identities (3.1) and, additionally, the condition

$$\operatorname{rank}(R_{\min}S_2R_{\min} - S_1) = (m-1)L. \tag{4.21}$$

We set

$$\widetilde{S}_2 := R_{\min}S_2R_{\min}, \quad \widetilde{\Psi}_1 := \Psi_1, \quad \widetilde{\Psi}_2 := R_{\min}\Psi_2. \tag{4.22}$$

In view of (4.22) and (3.1) we obtain the operator identity

$$\widetilde{S}_2 - A\widetilde{S}_2A^* = \widetilde{\Psi}_2\widetilde{\Psi}_2^*. \tag{4.23}$$

Setting

$$S := \widetilde{S}_2 - S_1 \tag{4.24}$$

and taking into account (4.21)–(4.23) we can reduce our problem to the previous case. The corresponding $Lm \times m$ matrix $\widetilde{Y}$ satisfies now the equation

$$\widetilde{Y}^*(\widetilde{S}_2 - S_1)\widetilde{Y} = 0. \tag{4.25}$$

The corresponding matrix function $\tilde{w}_{\min}(\zeta)$ can be written in the form

$$\tilde{w}_{\min}(\zeta) = \widetilde{P}_1(\zeta) \cdot [\widetilde{P}_2(\zeta)]^{-1}, \tag{4.26}$$

where

$$\widetilde{P}_s(\zeta) = \widetilde{\Psi}_s^*(-\zeta I_N + A^*)^{-1}\widetilde{Y}, \quad s \in \{1,2\}. \tag{4.27}$$

Using the identity

$$\Psi_2^*R_{\rho_{\min}} = \rho_{\min}\Psi_2^*,$$

we obtain

$$\widetilde{P}_2(\zeta) = \rho_{\min}\Psi_2^*(-\zeta I_N + A^*)^{-1}\widetilde{Y}. \tag{4.28}$$

From (4.26)–(4.28) we infer

$$\tilde{w}_{\min}(\zeta) = \Psi_1^*(-\zeta I_N + A^*)^{-1}\widetilde{Y} \cdot [\Psi_2^*(-\zeta I_N + A^*)^{-1}\widetilde{Y}]^{-1} \cdot (\rho_{\min})^{-1}. \tag{4.29}$$

In view of Remark 4.2 we have

$$[\tilde{w}_{\min}(\zeta)]^*\tilde{w}_{\min}(\zeta) \leqslant I_m, \quad |\zeta| < 1. \tag{4.30}$$

Now we consider an $m \times m$ matrix function $\tilde{w}(\zeta)$ which is holomorphic in the unit disk $|\zeta| < 1$ and satisfies the following conditions:

1. The $m \times m$ matrix function $\tilde{w}(\zeta)\widetilde{P}_2(\zeta) - \widetilde{P}_1(\zeta)$ is holomorphic in the unit disk $|\zeta| < 1$.
2. The inequality

   $$\tilde{w}^*(\zeta)\tilde{w}(\zeta) \leqslant \tilde{\rho}_1^2, \quad |\zeta| < 1$$

   is fulfilled where $\tilde{\rho}_1$ is some positive Hermitian $m \times m$ matrix.

Let the matrix T be chosen as in Theorem 4.1 and let

$$\widetilde{T} := TR_{\min}. \tag{4.31}$$

Then $\widetilde{T}$ is a complex $mL \times mL$ matrix which is a lower block triangular matrix with $m \times m$ blocks and taking into account (4.31), (4.7), and (4.22) we obtain

$$\widetilde{T}^*\widetilde{T} = R_{\min}T^*TR_{\min} = R_{\min}S_2R_{\min} = \widetilde{S}_2.$$

Then from Theorem 4.1 we obtain

$$\widetilde{Y}\widetilde{T}^*\widetilde{T}\widetilde{Y} \leqslant \widetilde{Y}^*\widetilde{T}^*R^2_{\tilde{\rho}_1}\widetilde{T}\widetilde{Y} \tag{4.32}$$

and the existence of some vector $\tilde{h} \in \mathbb{C}^m$ such that

$$\tilde{h}^*\widetilde{Y}^*\widetilde{T}^*\widetilde{T}\widetilde{Y}\tilde{h} < \tilde{h}^*\widetilde{Y}^*\widetilde{T}^*R^2_{\widetilde{\rho}_1}\widetilde{T}\widetilde{Y}\tilde{h}. \tag{4.33}$$

Thus setting

$$\rho_1 := \sqrt{\rho_{\min}\tilde{\rho}_1^2\rho_{\min}}, \tag{4.34}$$

we see that ρ_1 is a positive Hermitian $m \times m$ matrix and the formulas (4.32) and (4.33) can be rewritten as

$$\widetilde{Y}^*T^*R^2_{\min}T\widetilde{Y} \leqslant \widetilde{Y}^*T^*R^2_{\rho_1}T\widetilde{Y} \tag{4.35}$$

and

$$\tilde{h}\widetilde{Y}^*T^*R^2_{\min}T\widetilde{Y}\tilde{h} < h\widetilde{Y}T^*R^2_{\rho_1}T\widetilde{Y}\tilde{h}. \tag{4.36}$$

Let

$$W_{\min}(\zeta) := \tilde{w}_{\min}(\zeta) \cdot \rho_{\min} \tag{4.37}$$

and

$$Q_s(\zeta) := \Psi_s^*(-\zeta I_{mL} + A^*)^{-1}\widetilde{Y}, \quad s \in \{1, 2\}. \tag{4.38}$$

From (4.37), (4.29), and (4.38) we obtain the representation

$$W_{\min}(\zeta) = Q_1(\zeta)[Q_2(\zeta)]^{-1}. \tag{4.39}$$

From (4.37) and (4.30) we infer

$$[W_{\min}(\zeta)]^*W_{\min}(\zeta) \leqslant \rho^2_{\min}. \tag{4.40}$$

From (4.22), (4.27), and (4.38) it follows

$$\widetilde{P}_1(\zeta) = Q_1(\zeta), \tag{4.41}$$

whereas formulas (4.28) and (4.38) imply

$$\widetilde{P}_2(\zeta) = \rho_{\min}Q_2(\zeta). \tag{4.42}$$

Now we are able to formulate the main result of this section.

Theorem 4.3. *Let the $m \times m$ matrix function $W(\zeta)$ be holomorphic in the unit disk $|\zeta| < 1$ and satisfy the following conditions*:

1. *The $m \times m$ matrix function $W(\zeta)Q_2(\zeta) - Q_1(\zeta)$ is holomorphic in the unit disk $|\zeta| < 1$.*
2. *The inequality*

$$W^*(\zeta)W(\zeta) \leqslant \rho_1^2, \quad |\zeta| < 1$$

is fulfilled where ρ_1 is some positive Hermitian $m \times m$ matrix.

Let T be a complex $mL \times mL$ matrix which is a lower block triangular matrix with $m \times m$ blocks and satisfies

$$T^*T = S_2.$$

Then we have the inequality

$$\widetilde{Y}^*T^*R^2_{\min}T\widetilde{Y} \leqslant \widetilde{Y}T^*R^2_{\rho_1}T\widetilde{Y}.$$

Moreover, if $\rho_1 \neq \rho_{\min}$ there exists some vector $\tilde{h} \in \mathbb{C}^m$ such that

$$\tilde{h}^*\widetilde{Y}^*T^*R^2_{\min}T\widetilde{Y}\tilde{h} < \tilde{h}^*\widetilde{Y}^*T^*R^2_{\rho_1}T\widetilde{Y}\tilde{h}.$$

Proof. Let

$$\tilde{w}(\zeta) := w(\zeta) \cdot \rho_{\min}^{-1}. \tag{4.43}$$

From (4.34) and the choice of $w(\zeta)$ we see that $\tilde{w}(\zeta)$ is an $m \times m$ matrix function which is holomorphic in the unit disk $|\zeta| < 1$. In view of (4.41)–(4.43) we have

$$\tilde{w}(\zeta)\widetilde{P}_2(\zeta) - \widetilde{P}_1(\zeta) = W(\zeta)Q_2(\zeta) - Q_1(\zeta).$$

Thus, the function $\tilde{w}(\zeta)\widetilde{P}_2(\zeta) - \widetilde{P}_1(\zeta)$ is holomorphic in the disk $|\zeta| < 1$. Let

$$\tilde{\rho}_1 := \sqrt{(\rho_{\min})^{-1}\rho_1^2(\rho_{\min})^{-1}}. \tag{4.44}$$

Then $\tilde{\rho}_1$ is a positive Hermitian $m \times m$ matrix satisfying

$$\rho_1^2 = \sqrt{\rho_{\min}\, \tilde{\rho}_1^2\, \rho_{\min}}. \tag{4.45}$$

From (4.43) and (4.44) and the choice of $w(\zeta)$ we obtain

$$\tilde{w}^*(\zeta)\tilde{w}(\zeta) \leqslant \tilde{\rho}_1^2, \quad |\zeta| < 1.$$

Hence from (4.34)–(4.36) and (4.45) we obtain all assertions. The theorem is proved. □

Corollary 4.4. *The regularity Condition 1 in Theorem 4.3 cannot be fulfilled if $\rho_1^2 \leqslant \rho^2_{\min}$ and $\rho_1^2 \neq \rho^2_{\min}$.*

Remark 4.5. Theorem 4.3, Corollary 4.4 and the examples below show that the matrix $\rho_{\min}$ has a minimality property which is different from the property of having minimal rank.

Taking into account Theorem 4.3 and Corollary 4.4 we call the m-dimensional subspace spanned by the vectors $T\widetilde{Y}$ *subspace of minimality*. We think that in applied problems the subspace of minimality can have a physical sense.

Example 4.6. In the Schur problem we have $S_2 = I$. Thus, $T = I$ and the space of minimality is spanned by $\widetilde{Y}$.

5. Regularity and interpolation problems

In this section, we show that in a broad class of cases the regularity condition is equivalent to the corresponding interpolation conditions. According to Proposition 1.1 the interpolation problem can be formulated as follows. Determine all bounded $m \times m$ matrix functions $W(\zeta)$ which are holomorphic in the unit disk $|\zeta| < 1$ and satisfy the condition:

$$V(A)\Psi_2 = \Psi_1, \tag{5.1}$$

where

$$V(\zeta) := [W(-\bar{\zeta})]^*. \tag{5.2}$$

The matrix $V(A)$ can be computed by transforming the matrix A to Jordan block form. Let U be a nonsingular matrix which transforms the matrix A to its Jordan block normal form A_J i.e.

$$A = UA_JU^{-1},$$

where

$$A_J = \operatorname{diag}(A_1, \ldots, A_p)$$

and where for $s \in \{1, \ldots, p\}$ the matrix A_s has the size $(p_s + 1)m \times (p_s + 1)m$ with some $p_s \in \{0, 1, 2, \ldots\}$ and has the $m \times m$ block partition

$$A_s = \begin{pmatrix} z_sI_m & 0 & 0 & \cdots & 0 & 0 \\ I_m & z_sI_m & 0 & \cdots & 0 & 0 \\ 0 & I_m & z_sI_m & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & I_m & z_sI_m \end{pmatrix}.$$

If $f(\zeta)$ is a holomorphic function in the unit disk then using a well-known formula (see, e.g., [11, Section 5.1]) we obtain

$$f(A_s) = \begin{pmatrix} f(z_s)I_m & 0 & \cdots & 0 \\ \frac{f'(z_s)}{1!}I_m & f(z_s)I_m & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ \frac{f^{(p_s)}(z_s)}{p_s!}I_m & \frac{f^{(p_s-1)}(z_s)}{(p_s-1)!}I_m & \cdots & f(z_s)I_m \end{pmatrix}, \quad s \in \{1, \ldots, p\}. \tag{5.3}$$

Setting

$$\widetilde{\Psi}_1 := U^{-1}\Psi_1, \quad \widetilde{\Psi}_2 := U^{-1}\Psi_2,$$

we obtain

$$V(A_J)\widetilde{\Psi}_2 = \widetilde{\Psi}_1.$$

Then

$$V^{(j)}(z_s) = V_{\min}^{(j)}(z_s), \quad s \in \{1, \ldots, p\}, \; j \in \{0, \ldots, p_s\}.$$

This implies that the function $W(\zeta)$ satisfies the regularity Condition 1 in Theorem 4.3.

Using the function $W_{\min}(\zeta)$ defined in (4.37) we introduce

$$V_{\min}(\zeta) := [W_{\min}(\overline{-\zeta})]^*. \tag{5.4}$$

From (1.20) and (5.4) it follows that

$$V_{\min}(A)\Psi_2 = \Psi_1. \tag{5.5}$$

Example 5.1 (*Nevanlinna–Pick problem*). Let the complex $m \times m$ matrices $\eta_1, \eta_2, \ldots, \eta_k$ and the pairwise different points $z_1, z_2, \ldots, z_k$ belonging to the open unit disk $|\zeta| < 1$ be given. Then the corresponding

Nevanlinna–Pick problem is to determine all bounded holomorphic $m \times m$ matrix-valued functions $w(\zeta), |\zeta| < 1$, which satisfy the conditions

$$w(-\bar{z}_s) = \eta_s^*, \quad s \in \{1, \dots, k\}.$$

As it was shown in [18, Section 7.3] in this case the $km \times m$ matrices Ψ_1 and Ψ_2 and the $km \times km$ matrix A are given by

$$\Psi_1 := \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_k \end{pmatrix}, \quad \Psi_2 := \begin{pmatrix} I_m \\ \vdots \\ I_m \end{pmatrix}, \quad A := \begin{pmatrix} z_1 I_m & 0 & \cdots & 0 \\ 0 & z_2 I_m & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & z_k I_m \end{pmatrix}. \tag{5.6}$$

Proposition 5.2. *Let the matrices Ψ_1, Ψ_2 and A be given as in (5.6). Let $W(\zeta)$ be a bounded $m \times m$ matrix function which is holomorphic in the unit disk $|\zeta| < 1$. Then $W(\zeta)$ satisfies the regularity Condition 1 from Theorem 4.3 if and only if the relations (5.1) and (5.2) are true.*

Proof. The assertion follows by a straightforward computation using the residue theorem. $\square$

Example 5.3 (*Schur problem*). Let the complex $m \times m$ matrices $a_0, a_1, \dots, a_k$ be given. Then the corresponding Schur problem is to determine all bounded holomorphic $m \times m$ matrix-valued functions $w(\zeta), |\zeta| < 1$, which satisfy the conditions

$$\frac{w^{(j)}(0)}{j!} = a_j, \quad j \in \{0, \dots, k\}.$$

As it was shown in [18, Section 7.3] in this case the $(k+1)m \times m$ matrices Ψ_1 and Ψ_2 and the $(k+1)m \times (k+1)m$ matrix A are given by

$$\Psi_1 := \begin{pmatrix} a_0^* \\ \vdots \\ a_k^* \end{pmatrix}, \quad \Psi_2 := \begin{pmatrix} I_m \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad A := \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ I_m & 0 & \cdots & 0 & 0 \\ 0 & I_m & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & I_m & 0 \end{pmatrix}. \tag{5.7}$$

Then for $|\zeta| < 1$ we have

$$V(\zeta) - V_{\min}(\zeta) = c_{k+1} z^{k+1} + c_{k+2} \zeta^{k+2} + \cdots. \tag{5.8}$$

Taking into account (5.8) we obtain the following result:

Proposition 5.4. *Let the matrices Ψ_1, Ψ_2 and A be given by (5.7). Let $W(\zeta)$ be a bounded $m \times m$ matrix function which is holomorphic in the unit disk $|\zeta| < 1$. Then $W(\zeta)$ satisfies the regularity Condition 1 from Theorem 4.3 if and only if the relations (5.1) and (5.2) are true.*

Example 5.5 (*Jordan block case*). Let $p \in \mathbb{N}$ and let $z_1, z_2, \dots, z_p$ be pairwise different points from the open unit disk $|\zeta| < 1$. For $s \in \{1, \dots, p\}$ let $p_s \in \mathbb{N}_0$ and $a_{0,s} a_{1,s}, \dots, a_{p_s,s}$ be a sequence of complex $m \times m$ matrices. Then we want to determine all bounded holomorphic $m \times m$ matrix-valued functions $w(\zeta), |\zeta| < 1$, which satisfy the conditions

$$\frac{w^{(j)}(-\bar{z}_s)}{j!} = a_{j,s}, \quad s \in \{1, \dots, p\},\ j \in \{0, \dots, p\}.$$

In this case, we have

$$A := \operatorname{diag}(A_1, A_2, \dots, A_p), \tag{5.9}$$

where for $s \in \{1, \dots, p\}$ the $(p_s + 1)m \times (p_s + 1)m$ matrix A_s is given by

$$A_s := \begin{pmatrix} z_s I_m & 0 & 0 & \cdots & 0 & 0 \\ I_m & z_s I_m & 0 & \cdots & 0 & 0 \\ 0 & I_m & z_s I_m & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & I_m & z_s I_m \end{pmatrix}. \tag{5.10}$$

Moreover, let

$$\Psi_1 := \begin{pmatrix} \Psi_{11} \\ \Psi_{12} \\ \vdots \\ \Psi_{1p} \end{pmatrix}, \quad \Psi_2 := \begin{pmatrix} \Psi_{21} \\ \Psi_{22} \\ \vdots \\ \Psi_{2p} \end{pmatrix}, \tag{5.11}$$

where for $s \in \{1, \ldots, p\}$ the $(p_s + 1)m \times m$ matrices Ψ_{1s} and Ψ_{2s} are defined by

$$\Psi_{1,s} := \begin{pmatrix} a_{0,s}^* \\ a_{1,s}^* \\ \vdots \\ a_{p_s,s}^* \end{pmatrix}, \quad \Psi_{2,s} := \begin{pmatrix} I_m \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{5.12}$$

and where the $m \times m$ matrix $a_{0,s}$ is nonsingular. Using (5.3), we obtain from the relations

$$V(A)\Psi_2 = \Psi_1 \tag{5.13}$$

and

$$V_{\min}(A)\Psi_2 = \Psi_1 \tag{5.14}$$

the expression

$$V(\zeta) - V_{\min}(\zeta) = c_s(\zeta - \zeta_s)^{p_s+1} + \cdots, \quad |\zeta| < 1, \ s \in \{1, \ldots, p\}. \tag{5.15}$$

From formulas (5.10)–(5.15) we see that

$$V(A)\Psi_2 = \Psi_1 \tag{5.16}$$

if and only if for all $s \in \{1, \ldots, p\}$ and all $j \in \{0, \ldots, p_s\}$ the relations

$$\frac{V^{(j)}(-\bar{z}_s)}{j!} = a_{j,s}^* \tag{5.17}$$

are true.

Proposition 5.6. *Let the matrices Ψ_1, Ψ_2 and A be defined by formulas (5.9)–(5.12). Let $W(\zeta)$ be a bounded $m \times m$ matrix function which is holomorphic in the unit disk $|\zeta| < 1$. Then $W(\zeta)$ satisfies the regularity Condition 1 from Theorem 4.3 if and only if the relations (5.4) and (5.5) are true.*

6. Comparison of different approaches and results to the extremal interpolation problems

1. The scalar extremal problems named after Schur and Nevanlinna–Pick were investigated in papers by Carathéodory and Fejér [8], Takagi [21], Akhiezer [1], Clark [7], Adamjan et al. [2,3,4] and others. The matrix versions of these problems were treated in the paper [12]. In all these papers the Schur problem and the Nevanlinna–Pick problem were handled separately. The approach developed in this paper is based on a unifying principle: The regularity of some matrix function is required (see Condition 1 in Theorem 4.3). This allows us to consider a whole class of matricial interpolation problems with the same method. This class contains a problem (the so-called Jordan case) which is even new for the scalar case.

Our method of construction of the extremal solutions is based on results from the book [18, Chapters 5–7 and 11], the paper [10] and Section 1 of this paper.

The method used to prove that the constructed solutions are extremal was prompted by Akhiezer's paper [1]. It should be mentioned that in the papers [1,4] even extremal problems for meromorphic functions having a finite number of poles in the unit disk $|\zeta| < 1$ were studied. Such types of results have not been obtained in the matrix case up to now.

2. As it was indicated in the Adamjan et al. paper [4] scalar versions of extremal problems are closely related to problems of best approximation.

Open problem: Investigate the connections between extremal problems and corresponding approximation problems in the matrix case.

In the important paper Adamjan et al. [5] the case of a matricial (even operatorial) function $w(\zeta)$ was considered, whereas ρ was chosen as a scalar.

3. In the paper [12] our extremal interpolation problem was compared with the superoptimal interpolation problem (see [22,15,14]). It was proved that these two interpolation problems have quite different answers.

Acknowledgements

This paper was written during the Leibniz Professorship of the third author at Leipzig University. He thanks the Leipzig University for kind hospitality and support.

Moreover, the authors are grateful to A.L. Sakhnovich, who carefully read the paper and gave a lot of useful remarks.

References

[1] N.I. Akhiezer, On a minimum problem in function theory and the number of roots of an algebraic equation inside the unit disc, Izv. Akad. Nauk SSR Otdel. Mat. i Est. Nauk, 9 (1930) 1169–1189 (in Russian, English Translation in: H. Dym, B. Fritzsche, V.E. Katsnelson, B. Kirstein (Eds.), Topics in Interpolation Theory, OT Series, vol. 95, Birkhä, Basel, Boston, Berlin, 1997, pp. 19–35).
[2] V.M. Adamjan, D.Z. Arov, M.G. Krein, On infinite Hankel matrices and generalized Carathé–Fejé and F. Riesz problems, Funktional. Anal. i Prilozhen. 2 (1) (1968) 1–19 (in Russian, English transl.: Functional Anal. Appl. 2 (1968), 1–18).
[3] V.M. Adamjan, D.Z. Arov, M.G. Krein, Infinite Hankel matrices and generalized problems of Carathé–Fejé and I. Schur, Funktional. Anal. i Prilozhen. 2 (4) (1968) 1–17 (in Russian, English transl.: Functional Anal. Appl. 2 (1968) 269–281).
[4] V.M. Adamjan, D.Z. Arov, M.G. Krein, Analytic properties of Schmidt pairs for a Hankel operator and the generalized Schur–Takagi problem, Mat. Sb. 86 (1971) 34–75 (in Russian, English transl.: Math. USSR Sb. 15 (1971) 31–73).
[5] V.M. Adamjan, D.Z. Arov, M.G. Krein, Infinite block Hankel matrices and related extension problems,Izv. Akad. Nauk. Armjan. SSR Ser. Mat. 6 (1971) 87–112 (in Russian, English transl.: Amer. Math. Soc. Transl. Series 2, vol. 111 (1978) 133–156).
[6] J.A. Ball, I. Gohberg, L. Rodman, Interpolation of Rational Matrix Functions, OT Series, Birkhä, Basel, Boston, Berlin, 1990.
[7] D.N. Clark, Hankel forms, Toeplitz forms and meromorphic functions, Trans. Amer. Math. Soc. 134 (1968) 109–116.
[8] C. Carathéodory, L. Fejér, Über den Zusammenhang der Extremen von harmonischen Funktionen mit ihren Koeffizienten und über den Picard–Landauschen Satz, Rend. Circ. Mat. Palermo 32 (1911) 218–239.
[9] A. Ferrante, B.C. Levy, Hermitian solutions of the equation $X = Q + NX^{-1}N^{*}$, Linear Algebra Appl. 247 (1996) 359–373.
[10] B. Fritzsche, B. Kirstein, L.A. Sakhnovich, On extremal problems of interpolation theory with unique solution, in press.
[11] F.R. Gantmacher, Matrizentheorie, VEB Deutscher Verlag der Wissenschaften, Berlin, 1986.
[12] J.W. Helton, L.A. Sakhnovich, Extremal problems of interpolation theory, Rocky Mountain J. Math. 35 (2005) 819–841.
[13] H. Kimura, State space approach to the classical interpolation problem and its applications, in: H. Nijmeijer, J.M. Schumacher (Eds.), Three Decades of Mathematical System Theory, Lecture Notes in Control and Inform. Sci., vol. 135, Springer, Berlin, 1989, pp. 243–275.
[14] V.V. Peller, Hankel Operators and Their Applications, Springer Monographs in Mathematics, Springer, New York, 2003.
[15] V.V. Peller, N.J. Young, Superoptimal analytic approximations of matrix functions, J. Funct. Anal. 120 (1994) 300–343.
[16] V.P. Potapov, Collected Papers of V.P. Potapov, Private Translation and Edition by T. Ando, Hokkaido University, Sapporo, Japan, 1982.
[17] A.C.M. Ran, M.C.B. Reurings, A nonlinear matrix equation connected to interpolation theory, Linear Algebra Appl. 379 (2004) 289–302.
[18] L.A. Sakhnovich, Interpolation Theory and its Applications, Kluwer, Dordrecht, 1997.
[19] L.A. Sakhnovich, On the Sarason scheme in the matrix case, Funktsional. Anal. i Prilozhen. 31 (4) (1997) 52–59. (in Russian, English transl.: Functional Anal. Appl. 31 (1997) 266–272).
[20] D. Sarason, Generalized interpolation in H^{∞}, Trans. Amer. Math. Soc. 137 (1967) 179–203.
[21] T. Takagi, On an algebraic problem related to an analytic theorem of Carathéodory and Fejér, Japan J. Math. Part I 1 (1924) 83–93., Part II, 2 (1925) 13–17.
[22] N.J. Young, The Nevanlinna–Pick problem for matrix-valued functions, J. Operator Theory 15 (1986) 239–265.