ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 35, Number 3, 2005

EXTREMAL PROBLEMS OF INTERPOLATION THEORY

J. WILLIAM HELTON AND L.A. SAKHNOVICH

ABSTRACT. We consider problems where one seeks $m \times m$ matrix valued H^∞ functions $w(\xi)$ which satisfy interpolation constraints and a bound

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

where the $m \times m$ positive semi-definite matrix $\rho_{\min}$ is minimal (no smaller than) any other matrix ρ producing such a bound. That is, if

$$w^*(\xi)w(\xi) \le \rho, \quad |\xi| < 1, \tag{0.2}$$

and if $\rho_{\min} - \rho$ is positive semi-definite, then $\rho_{\min} = \rho$. This is an example of what we shall call a "minimal interpolation problem." Such problems are studied extensively in the book [**13**, Chapter 7]. When the bounding matrices ρ are restricted to be scalar multiples of the identity, then the problem where we extremize over them is just the classical matrix valued interpolation problem containing those of Schur and Nevalinna-Pick (which in typical cases has highly nonunique solutions). Our minimal interpolation forces tighter conditions.

In this paper we actually study a framework more general than that of Nevanlinna-Pick and Schur, and in this general context we show under some assumptions that our minimal interpolation problem, with $\rho_{\min}$ defined formally by a minimal rank condition in Definition 3.3, has a unique solution $\rho_{\min}$ and $w_{\min}(\xi)$. It is important both from applied and theoretical view points that the solution $w_{\min}(\xi)$ turns out to be a rational matrix function, indeed for the matrix Nevanlinna-Pick and Schur problems we obtain an explicit formulas generalizing those known classically.

Also in this paper we compare minimal interpolation problems to superoptimal interpolation problem, cf. [**14**] and [**11**], and see that they have very different answers. Whether one chooses super-optimal criteria or our minimal criteria in a particular situation depends on which issues are important in that situation.

The case $m = 1$ was investigated by many people with a formulation close to the one we use being found in Akhiezer

Received by the editors on October 11, 2001, and in revised form on July 28, 2003.
This work was partially supported by NSF, ONR, DARPA and the Ford Motor Company.

[**1**]. Interpolation with matrix valued analytic functions has found great application in control theory, cf. the books [**2, 3, 6, 7, 15**].

1. Outline. The main consequences for analytic function theory of the general results of this paper are presented in Sections 4 and 5. Theorem 4.1 and the corollaries which follow it thoroughly describe minimal solutions to a class of matrix valued Nevannlinna-Pick interpolation problems. Also these corollaries connect the definition of minimal interpolation given in the abstract, see inequality (0.1), with the more general minimal rank Definition 3.3.

Section 5 parallels Section 4 with Theorem 5.1 and its corollaries solving a class of matrix valued Schur problems as a consequence of the theory of a general interpolation problem.

The general interpolation problem and some consequences of it appear in Section 3. It is a problem, about matrices with consequences for analytic function theory. This matrix, or more generally operator theoretic approach, comes from the book [**13**]. There are different matrix theoretic approaches to analytic function theory, which correspond to state space linear systems theory, cf. [**2, 6, 15**]; linear systems theory. Also there is the approach in [**3**]. While it would be interesting to know the connection between these ways of converting between linear algebra and analytic function theory, this has not been done. Possibly state space methods might be effective on our minimal interpolation problems, however, this has never to our knowledge been tried.

In summary, this paper begins with some background on matrix inequalities, Section 2, moves to the general interpolation problem, Section 3, and then that goes to Nevanlinna-Pick interpolation and Schur interpolation applications, Sections 4 and 5. Finally in Section 6 we compare minimal interpolation to super-optimal interpolation.

2. Background on matrix equations. In the solution of extremal problem (0.1) an important role is played by the matrix nonlinear equation

$$X = R + C^*X^{-1}C, \quad R > 0 \tag{2.1}$$

where matrices X, R, C are $N \times N$ matrices. When studying equation

(2.1) we apply the method of successive approximations. We put

$$X_0 = R, \quad X_{n+1} = R + C^* X_n^{-1} C. \tag{2.2}$$

It follows from (2.2) that

$$X_n \geq X_0, \quad n \geq 0. \tag{2.3}$$

As the righthand side of (2.1) decreases with the growth of X, then in view of (2.2) and (2.3) the inequalities

$$X_n \leq X_1, \quad n \geq 1 \tag{2.4}$$

are true. Similarly we obtain that

$$X_n \geq X_2, \quad n \geq 2. \tag{2.5}$$

This leads to the following assertion (found in [**4, 5**]).

2.1 The case $R > 0$.

Proposition 2.1. *Suppose $R > 0$. Then we have*

(i) *the sequence $X_0, X_2, X_4, \dots$ monotonically increases and has a limit $\underline{X}$,*

(ii) *the sequence $X_1, X_3, X_5, \dots$ monotonically decreases and has a limit $\overline{X}$,*

(iii) *the inequality*

$$\underline{X} \leq \overline{X} \tag{2.6}$$

is true.

In the paper [**5**] the following assertion is proved.

Proposition 2.2. *Let $R > 0$. Then equation* (2.1) *has one and only one solution X such that $X > 0$. Here relations*

$$\underline{X} = \overline{X} = X \tag{2.7}$$

are fulfilled.

2.2 The case $R = 0$. We shall consider separately the case when $R = 0$. Then equation (2.1) has the form

$$X = C^* X^{-1} C. \tag{2.8}$$

The necessary condition for the solvability of equation (2.8) is the inequality

$$\det C \neq 0. \tag{2.9}$$

Example 2.3. Let $C > 0$. We write equation (2.8) in the form

$$Y^2 = I_m, \tag{2.10}$$

where $Y = C^{-1/2} X C^{-1/2}$. Equation (2.10) has only one positive solution $Y = I_m$. It means that equation (2.8) has as its only positive solution

$$X = C. \tag{2.11}$$

Example 2.4. Let

$$N = 2, \quad C = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = J. \tag{2.12}$$

Then equation (2.8) can be written in the form

$$XJX = J. \tag{2.13}$$

It is known that the last equation is satisfied by the matrices

$$X_\varphi = \begin{bmatrix} \cosh\varphi & \sinh\varphi \\ \sin h\varphi & \cosh\varphi \end{bmatrix}. \tag{2.14}$$

The matrices X_φ are positive, which means that under conditions (2.12) equation (2.8) has an infinite set of positive solutions.

Conclusion. When $R = 0$ equation (2.1) can have no solution at all, when $\det C = 0$, can have only one positive solution, when $C > 0$, and can have an infinite number of positive solutions, $N = 2$, $C = J$.

3. General interpolation problems. Let Hilbert spaces H and $\mathcal{G}$ be given, $\dim \mathcal{G} = m < \infty$. Suppose we are given operators

$$A : H \longrightarrow H \qquad S : H \to H$$
$$\Phi_l : \mathcal{G} \longrightarrow H \quad \text{for } l = 1, 2,$$

satisfying the operator identity

$$S - ASA^* = \Phi_1\Phi_2^* + \Phi_2\Phi_1^*. \tag{3.1}$$

Let us state an interpolation problem associated with the operator identity (3.1).

The problem is to find a nondecreasing $m \times m$ matrix function $\tau(\varphi)$ such that

$$S = \int_{-\pi}^{\pi} (I - e^{i\varphi}A)^{-1}\Phi_2[d\tau(\varphi)]\Phi_2^*(I - e^{-i\varphi}A^*)^{-1} \tag{3.2}$$

$$\Phi_1 = \frac{1}{2}\int_{-\pi}^{\pi} (I + e^{i\varphi}A)(I - e^{i\varphi}A)^{-1}\Phi_2\, d\tau(\varphi) + i\Phi_2\alpha \tag{3.3}$$
$$\text{where} \quad (\alpha = \alpha^*).$$

The solution of various classical interpolation problems can be expressed in terms of the matrix function

$$F(\xi) = -i\alpha + \frac{1}{2}\int_{-\pi}^{\pi} \frac{e^{i\varphi} - \xi}{e^{i\varphi} + \xi}\, d\tau(\varphi), \quad |\xi| < 1, \tag{3.4}$$

which has positive semi-definite real part. Often F is of more direct interest than τ. Interpolation problems are also connected with the matrix function

$$w(\xi) = [F(\xi) + I_m]^{-1}[F(\xi) - I_m]. \tag{3.5}$$

It follows from (3.4) and (3.5) that

$$\|w(\xi)\| \leq 1, \quad |\xi| < 1. \tag{3.6}$$

The classical interpolation problems (Nevanlinna-Pick, Schur) are special cases. If the matrix A is diagonal we obtain the Nevanlinna-Pick problem. If A is a Jordan matrix we obtain the Schur problem. A number of other concrete problems are given both in the paper [**8**] and in the book [**13**]. We note that representation (3.3) can be formulated in terms of contour integral [**2**].

Formula (3.2) directly implies that the inequality

$$S \geq 0 \tag{3.7}$$

is a necessary condition for the interpolation problem to be solvable. The problem is called *nondegenerate* if the following stronger inequality

$$S \geq \delta I_H \quad \delta > 0$$

holds. Extremal cases of interpolation problems are all degenerate. Degenerate cases will be discussed below and after that extremal cases.

3.1 Degenerate interpolation problems. Let A and S be $n \times n$ matrices and let Φ_1, Φ_2 be $n \times m$ matrices. We assume that these matrices satisfy the operator identity (3.1). Further we shall assume that the following conditions hold.

1. rank $S = n - m$.
2. The matrices A and S have the following block forms

$$A = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}, \quad S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \tag{3.8}$$

and

$$S \geq 0, \quad S_{11} > 0 \tag{3.9}$$

(Here A_{22} and S_{22} are $m \times m$ matrices).

3.

$$[I_m, 0]\Phi_2 g \neq 0, \quad \text{if } g \neq 0 \tag{3.10}$$

4.

$$\text{rank}\,[M_1, M_2] = m, \tag{3.11}$$

where

$$M_k = [-X^*, I_m](I + \xi_0 A)^{-1}\Phi_k, \quad k = 1, 2, \tag{3.12}$$

$$X = S_{11}^{-1} S_{12}, \quad |\xi_0| = 1 \tag{3.13}$$

Condition 1 for $m \neq 0$ tells us that $S \not> \delta I$, so the problem is degenerate. The following assertion is proved in the book [**13**, Chapter 5].

Proposition 3.1. *Suppose conditions* (3.8)–(3.11) *are fulfilled and rank* $S = n - m$. *Then the corresponding interpolation problem* (3.2)–(3.5) *has one and only one solution* $w(\xi)$ *and this solution has rational entries.*

Remark 3.2. The method for constructing the solution $w(\xi)$ of the degenerate interpolation problem is given in [**13**, Chapter 5].

3.2 Extremal interpolation problems. Let the matrices A, S_k and Ψ_k, $k = 1, 2$, be of dimensions $mN \times mN$ and $mN \times m$ respectively with S_k positive semi-definite. We suppose that the matrices are connected by the relations

$$S_k - AS_kA^* = \Psi_k\Psi_k^*, \quad k = 1, 2. \tag{3.14}$$

Setting

$$S = S_2 - S_1.$$

We deduce from (3.14) the equality

$$S - ASA^* = \Psi_2\Psi_2^* - \Psi_1\Psi_1^*. \tag{3.15}$$

We introduce the block-diagonal matrix

$$R = \text{diag}\,\{\underbrace{\rho, \rho, \dots, \rho}_{N}\}$$

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

$$AR = RA. \tag{3.16}$$

This is justified, since it will be proved later that condition (3.16) is true in a number of concrete examples.

From equations (3.14) and (3.16) it follows that

$$S_\rho - AS_\rho A^* = \Psi_2\Psi_2^* - \Psi_{1,\rho}\Psi_{1,\rho}^*, \tag{3.17}$$

where we define

$$S_\rho = S_2 - R^{-1}S_1R^{-1} \tag{3.18}$$

$$\Psi_{1,\rho} = R^{-1}\Psi_1. \tag{3.19}$$

Thus we have constructed a set of operator identities (3.18), where the positive matrix ρ plays the role of a parameter. A set of interpolation problems, see [**13**, Chapter 6], corresponds to this set of operator identities. A necessary condition for the solvability of these problems is the inequality

$$RS_2R - S_1 \geq 0. \tag{3.20}$$

Now we turn to extremal interpolation.

Definition 3.3. We shall call the matrix $\rho = \rho_{\min} > 0$ a minimal solution of inequality (3.20) if the following two requirements are fulfilled

1. the inequality

$$R_{\min}S_2R_{\min} - S_1 \geq 0 \tag{3.21}$$

where

$$R_{\min} = \operatorname{diag}\{\rho_{\min}, \rho_{\min}, \dots, \rho_{\min}\}$$

is valid.

2. If $\rho > 0$ satisfies inequality (3.20), then

$$\operatorname{rank}(R_{\min}S_2R_{\min} - S_1) \leq \operatorname{rank}(RS_2R - S_1). \tag{3.22}$$

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

Remark 3.4. The existence of $\rho_{\min}$ follows directly from Definition 3.3.

We shall write the positive semi-definite matrices S_1, S_2 and R in the following block forms

$$\begin{aligned} S_k &= \begin{bmatrix} S_{11}^{(k)} & S_{12}^{(k)} \\ \\ S_{21}^{(k)} & S_{22}^{(k)} \end{bmatrix}, \quad k = 1, 2 \\ R &= \begin{bmatrix} R_1 & 0 \\ 0 & \rho \end{bmatrix}, \quad R_1 = \operatorname{diag}\{\underbrace{\rho, \dots, \rho}_{N-1}\}, \end{aligned} \tag{3.23}$$

where $S_{22}^{(k)}$ are blocks of size $m \times m$, $S_{11}^{(k)}$ has size $(N-1)m \times (N-1)m$ and $S_{12}^{(k)}$ has size $(N-1)m \times m$. The following proposition is proved in [**13**].

Proposition 3.5. *Suppose for all $\rho > 0$ satisfying inequality* (3.20) *the upper diagonal block of* (3.20) *is strictly positive, that is,*

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

holds. If $\rho = q > 0$ satisfies inequality (3.20) *and the relation*

$$qS_{22}^{(2)} q = S_{22}^{(1)} + C_1^* (Q_1 S_{11}^{(2)} Q_1 - S_{11}^{(1)})^{-1} C_1 \tag{3.24}$$

where

$$Q_1 = diag\{\underbrace{q, q, \dots, q}_{N-1}\}, \quad C_1 = Q_1 S_{12}^{(2)} q - S_{12}^{(1)}, \tag{3.25}$$

then

$$\rho_{\min} = q. \tag{3.26}$$

3.3 Solutions in a special case. Let us consider these equations in the special case where

$$S_2 = I. \tag{3.27}$$

In this case equation (3.24) has the form

$$q^2 = S_{22}^{(1)} + S_{12}^{(1)*}(Q^2 - S_{11}^{(1)})^{-1} S_{12}^{(1)} \tag{3.28}$$

with $Q = \operatorname{diag}\{q, \cdot = q\}$ a block $(N-1) \times (N-1)$ matrix.

We analyze solving this equation by setting

$$q_0^2 = S_{22}^{(1)} \tag{3.29}$$

$$q_{n+1}^2 = S_{22}^{(1)} + S_{12}^{(1)*}(Q_n^2 - S_{11}^{(1)})^{-1} S_{12}, \quad n \geq 0, \tag{3.30}$$

where

$$Q_n = \operatorname{diag}\{\underbrace{q_n, \dots, q_n}_{N-1}\}, \quad n \geq 0.$$

If we suppose

$$\operatorname{diag}\left\{S_{22}^{(1)}, \dots, S_{22}^{(1)}\right\} - S_{11}^{(1)} > 0, \tag{3.31}$$

then this is the same as $Q_0^2 - S_{11}^{(1)} > 0$, and we can apply the monotonicity technique of Section 2 to obtain.

Lemma 3.6. *Suppose relations* (3.27) *and* (3.31) *hold. Then we have the following consequences.*

(i) *The sequence* $q_0^2, q_1^2, \dots$ *monotonically increases and has the limit* $\underline{q}^2$.

(ii) *The sequence* $q_1^2, q_3^2, \dots$ *monotonically decreases and has the limit* $\overline{q}^2$.

(iii) *The inequality*

$$\underline{q}^2 \leq \overline{q}^2$$

is true.

(iv) *If*

$$\underline{q}^2 = \overline{q}^2,$$

then the relation

$$\rho_{\min}^2 = \overline{q}^2$$

is true.

Proof. From (3.30) and (3.31) we have the relations

$$q_n^2 \geq q_0^2, \quad Q_n^2 \geq Q_0^2, \quad n \geq 0 \tag{3.32}$$

As the right side of (3.28) decreases with the growth of q^2, then in view of (3.30) and (3.32) the inequalities

$$q_n^2 \leq q_1^2, \quad Q_n^2 \leq Q_1^2, \quad n \geq 1 \tag{3.31}$$

are true. Similarly we obtain that

$$q_n^2 \geq q_2^2, \quad Q_n^2 \geq Q_2^2, \quad n \geq 2.$$

In this way we deduce parts (i), (ii) and (iii) of the lemma. Part (iv) follows from Proposition 3.5. □

Ran and Reurings proved the following important result [**12**].

Proposition 3.7. *Suppose relations* (3.27) *and* (3.31) *hold. Then equation* (3.14) *has one and only one positive solution*

$$q^2 = \rho_{\min}^2 = \overline{q}^2 = \underline{q}^2.$$

4. Extremal Nevanlinna-Pick problem.

4.1 The problem. Let the $m \times m$ matrices $w_1, w_2, \ldots, w_n$ and the points $z_1, z_2, \ldots, z_n$, $|z_k| < 1$, be given. We seek an $m \times m$ matrix valued function $w(z)$ which is holomorphic in the circle $|z| < 1$ such that

$$w(z_k) = w_k \tag{4.1}$$

and

$$w^*(z)w(z) \leq \rho_{\min}^2, \quad |z| < 1. \tag{4.2}$$

Here $\rho_{\min}$ will be defined by a minimal rank condition which turns out to be stronger than the minimality defined in (0.1) and (0.2).

4.2 An operator reformulation. The matrices A and S in the case of the Nevanlinna-Pick problem have the form, see [**1**, Chapter 7],

$$A = \operatorname{diag}\{\overline{z}_1 I_m, \overline{z}_2 I_m, \dots, \overline{z}_n I_m\} \tag{4.3}$$

$$S = S_2 - R^{-1} S_1 R^{-1}, \quad R = \operatorname{diag}\{\underbrace{\rho, \dots, \rho}_{n}\} \tag{4.4}$$

where

$$S_2 = \left\{ \frac{w_k^* w_l}{1 - \overline{z}_k z_l} \right\}_{kl=1}^{n}, \quad S_1 = \left\{ \frac{I_m}{1 - \overline{z}_k z_l} \right\}_{k,l=1}^{n}. \tag{4.5}$$

The matrices Φ_1, Φ_2 are defined by formulas

$$\Phi_1 = \frac{\Psi_1 + \Psi_2}{\sqrt{2}}, \quad \Phi_2 = \frac{\Psi_2 + \Psi_1}{\sqrt{2}} \tag{4.6}$$

where

$$\Psi_1 = R^{-1} col[w_1^*, w_2^*, \dots, w_n^*], \quad \Psi_2 = col[I_m, I_m, \dots, I_m]. \tag{4.7}$$

We seek a minimal rank solution in the sense of Definition 3.3. Note that inequality (3.21) implies that ρ satisfies inequality (0.2).

4.3 A solution. To obtain the solution of the extremal Nevanlinna-Pick problem we shall use both the results of the general theory [**13**] and the ideas of Akhiezer [**1**] concerning the scalar case. Let us consider the following set of equations

$$\sum_{k=1}^{n} \frac{w_j^* w_k - \rho^2}{1 - \overline{z}_j z_k} \quad Y_k = 0, \tag{4.8}$$

where Y_k are $m \times m$ matrices. Equation (4.8) can be written in the form

$$(S_1 - R S_2 R) Y = 0, \tag{4.9}$$

where

$$Y = \operatorname{col}[Y_1, Y_2, \dots, Y_n]. \tag{4.10}$$

We suppose that

$$\rho = \rho_{\min} \quad \text{and} \quad \operatorname{rank}(S_1 - R_{\min} S_2 R_{\min}) \leq (n-1)m. \tag{4.11}$$

In this case system (4.9) has a solution Y satisfying

$$\operatorname{rank} Y = m. \tag{4.12}$$

Theorem 4.1. *Let the $m \times m$ matrix function $\Psi(z)$ be holomorphic in the unit circle $|z| < 1$ and satisfy the conditions*

$$\Psi(z_k) = w_k, \quad 1 \leq k \leq n \tag{4.13}$$

$$\Psi^*(z)\Psi(z) \leq \rho_1^2, \quad |z| < 1 \tag{4.14}$$

where ρ_1 is a positive $m \times m$ matrix. Then we have the following inequality

$$\sum_{k,l=1}^{n} \frac{Y_k^* \rho_{\min}^2 Y_l}{1 - \overline{z}_k z_l} \leq \sum_{k,l=1}^{n} \frac{Y_k^* \rho_1^2 Y_l}{1 - \overline{z}_k z_l} \tag{4.15}$$

or equivalently

$$Y^* T^* R_{\min}^2 T Y \leq Y^* T^* R_{\rho_1}^2 T Y$$

where R and S_1 are defined by (4.4) and (4.5). Moreover, if $\rho_{\min} \neq \rho_1$, then there exists an $m \times 1$ vector $h \neq 0$ such that

$$h^* \sum_{k,l=1}^{n} \frac{Y_k \rho_{\min}^2 Y_l}{1 - \overline{z}_k z_l} \, h < h^* \sum_{k,l=1}^{n} \frac{Y_k^* \rho_1^2 Y_l}{1 - \overline{z}_k z_l} \, h.$$

Proof. It follows from (4.13) that the matrix function

$$\varphi(z) = \Psi(z) \left(\sum_{k=1}^{n} \frac{Y_k}{z - z_k} \right) \tag{4.16}$$

has the form

$$\varphi(z) = \sum_{k=1}^{n} \frac{w_k Y_k}{z - z_k} + \Psi_1(z), \tag{4.17}$$

where

$$\Psi_1(z) = C_0 + C_1 z + \cdots \tag{4.18}$$

is a regular function on the circle $|z| < 1$. Using (4.16) we obtain

$$\frac{1}{2\pi}\int_0^{2\pi} \varphi^*(z)\varphi(z)\, d\theta \le \frac{1}{2\pi}\int_0^{2\pi} \sum_{k=1}^{n} \frac{Y_k^*}{\overline{z} - \overline{z}_k}\ \rho_1^2 \sum_{l=1}^{n} \frac{Y_k}{z - z_l}\, d\theta, \tag{4.19}$$

where $z = r_0 e^{i\theta}$, $\max_k |z_k| < r_0 < 1$. From (4.19) we deduce that

$$\frac{1}{2\pi}\int_0^{2\pi} \varphi^*(z)\varphi(z)\, d\theta \le \sum_{k,l=1}^{n} \frac{Y_k^* \rho_1^2 Y_l}{r_0^2 - \overline{z}_k z_l}. \tag{4.20}$$

In view of (4.17) and (4.18) we have

$$\frac{1}{2\pi}\int_0^{2\pi} \varphi^*(z)\varphi(z)\, d\theta = \sum_{k,l=1}^{n} \frac{Y_k^* w_k^* w_l Y_l}{r_0^2 - \overline{z}_k z_l} + \sum_{j=0}^{\infty} C_j^* C_j r_0^{2j}. \tag{4.21}$$

Using (4.8) we obtain the equality

$$\sum_{k,l=1}^{n} \frac{Y_k^* w_k^* w_l Y_l}{1 - \overline{z}_k z_l} = \sum_{k,l=1}^{n} \frac{Y_k^* \rho_{\min}^2 Y_l}{1 - \overline{z}_k z_l}. \tag{4.22}$$

The first inequality in (4.15) follows directly from (4.20)–(4.22). To prove the second inequality in (4.15) write the first inequality in the form

$$Y^* R_{\min} S_1 R_{\min} Y \le Y^* R_{\rho_1} S_1 R_{\rho_1} Y$$

where R_{ρ_1} and S_1 are defined by (4.4) and (4.5) and Y by (4.10). Represent S_1 in the form $S_1 = T^*T$ where T is a block triangular operator. Since the operators R and S_1 commute, the operators T and R also commute. Hence the inequality immediately above can be written as the second inequality in (4.15).

The last inequality follows from (4.17) and the fact that $\Psi_1(z) \neq 0$. The theorem is proved. □

Remark 4.2. Theorem 4.1 and its corollaries below show that $\rho_{\min}$ has a minimal property of a type different than just that of minimal rank. In particular Corollary 3.3 shows that $\rho_{\min}$ is minimal in the sense described in inequality (0.1) of the abstract.

From Theorem 4.1 we deduce the following assertions.

Corollary 4.3. *A solution of the Nevanlinna-Pick problem* (4.1), (4.2) *when* $\rho_1^2 \leq \rho_{\min}^2$ *and* $\rho_1^2 \neq \rho_{\min}^2$ *does not exist.*

Proof. This follows directly from the second inequality in (4.15).

Corollary 4.4. *If the solution of the Nevanlinna-Pick problem* (4.1), (4.2) *for* $\rho_1^2 = \rho_{\min}^2$ *exists, then this solution,* $\phi_{\min}(z)$*, has the form*

$$\varphi_{\min}(z) = \left(\sum_{k=1}^{n} \frac{w_k Y_k}{z - z_k}\right)\left(\sum_{k=1}^{n} \frac{Y_k}{z - z_k}\right)^{-1} \tag{4.23}$$

Corollary 4.5. *If the matrix function* $\varphi_{\min}(z)$ *defined by* (4.23) *is holomorphic in the unit circle* $|z| < 1$*, then* $\varphi_{\min}(z)$ *is the solution of the Nevanlinna-Pick problem* (4.1), (4.2)*, when* $\rho_1^2 = \rho_{\min}^2$.

Corollary 4.6. *Let* S, Φ_1, Φ_2 *and* $R = R_{\min}$ *be defined by formulas* (4.4)–(4.6) *satisfy the conditions of Proposition 3.1. Then the corresponding Nevanlinna-Pick problem* (4.1), (4.2) ($\rho_1^2 = \rho_{\min}^2$) *has one and only one solution* $\varphi_{\min}(z)$ *and this solution has form* (4.23).

5. Schur extremal problem.

5.1 The problem. The $m \times m$ matrices $a_0, a_1, \dots, a_p$ are given. We wish to describe the set of $m \times m$ matrix functions $w(z)$, holomorphic in the circle $|z| < 1$, satisfying

$$w(z) = a_0 + a_1 z + \cdots + a_p z^p + \cdots \tag{5.1}$$

and

$$w^*(z) w(z) \leq \rho_{\min}^2, \quad |z| < 1. \tag{5.2}$$

Here $\rho_{\min}$ will be defined by a minimal rank condition which turns out to be stronger than minimality in the sense of (0.1) and (0.2).

5.2 Operator reformulation. It is well known that in this case

$$S_2 = I, \quad S_1 = C_p C_p^* \tag{5.3}$$

where

$$C_p = \begin{bmatrix} a_0 & 0 & \dots & 0 \\ a_1 & a_0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ a_p & a_{p-1} & \dots & a_0 \end{bmatrix}, \tag{5.4}$$

Moreover, the matrices A and S in the case of the Schur problem have the form

$$A = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ I_m & 0 & \dots & 0 & 0 \\ 0 & I_m & \dots & 0 & 0 \\ 0 & 0 & \dots & I_m & 0 \end{bmatrix}, \tag{5.5}$$

$$S = S_2 - R^{-1} S_1 R^{-1}, \tag{5.6}$$

where

$$S_2 = I, \quad S_1 = C_p C_p^*, \qquad \text{see [\textbf{1}, Chapter 7]}. \tag{5.7}$$

The matrices Φ_1, Φ_2 are defined by formulas

$$\Phi_1 = \frac{\Psi_1 + \Psi_2}{\sqrt{2}}, \quad \Phi_2 = \frac{\Psi_2 - \Psi_1}{\sqrt{2}} \tag{5.8}$$

where

$$\Psi_1 = R^{-1} col[a_0, a_1, \dots, a_p] \tag{5.9}$$

$$\Psi_2 = \mathrm{col}\,[I_m, 0, \dots, 0]. \tag{5.10}$$

Using the notation in (3.23) and (3.29), we have

$$S_{22}^{(1)} = a_p a_p^* + a_{p-1} a_{p-1}^* + \cdots + a_0 a_0^* \tag{5.11}$$

$$S_{11}^{(1)} = C_{p-1} C_{p-1}^* \tag{5.12}$$

$$[Q_0]^2 = \mathrm{diag}\,\{\underbrace{S_{22}^{(1)}, S_{22}^{(1)}, \dots, S_{22}^{(1)}}_{p}\}. \tag{5.13}$$

It follows from (5.7) that the conditions of Lemma 3.6 and Proposition 3.7 are satisfied if

$$[Q_0]^2 > C_{p-1}C_{p-1}^*. \tag{5.14}$$

We seek a minimal rank solution in the sense of Definition 3.3. Note that inequality (3.12) implies that $\rho_{\min}$ satisfies inequality (5.2).

5.3 A solution. To obtain the solution of the extremal Schur problem we shall use both the results of the general theory [**13**] and the ideas of Akhiezer [**1**] concerning the scalar case. A necessary condition for the solvability of the Schur extremal problem is the inequality, see the book [**13**]

$$R_{\min}^2 - C_pC_p^* \geq 0, \tag{5.15}$$

which can be written in the following equivalent form

$$\begin{bmatrix} I & -R_{\min}^{-1}C_p \\ -C_p^*R_{\min}^{-1} & I \end{bmatrix} \geq 0. \tag{5.16}$$

Let us introduce matrices

$$X = \operatorname{col}[X_0, X_1, \ldots, X_p]$$
$$Y = \operatorname{col}[Y_0, Y_1, \ldots, Y_p],$$

where X_k and Y_k are $m \times m$ matrices. We consider the equation

$$\begin{bmatrix} I & -R_{\min}^{-1}C_p \\ -C_p^*R_{\min}^{-1} & I \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} = 0,$$

i.e.,

$$R_{\min}X = C_pY, \quad C_p^*R_{\min}^{-1}X = Y. \tag{5.17}$$

Theorem 5.1. *Let the $m \times m$ matrix function $\Phi(z)$ be holomorphic in the unit circle $|z| < 1$ and satisfy the conditions*

$$\Phi(z) = a_0 + a_1z + \cdots + a_pz^p + \cdots \tag{5.18}$$

$$\Phi^*(z)\Phi(z) \leq \rho_1^2, \quad |z| < 1 \tag{5.19}$$

where ρ_1 is a positive $m \times m$ matrix. Then we have the following inequality

$$\sum_{k=0}^{p} X_k^* \rho_{\min}^2 X_k \leq \sum_{k=0}^{p} Y_k^* \rho_1^2 Y_k \tag{5.20}$$

or equivalently

$$X^* R_{\min}^2 X \leq Y^* R_{\rho_1}^2 Y.$$

Moreover, if $\rho_{\min} \neq \rho_1$, then there exists an $m \times 1$ vector $h \neq 0$ such that

$$h^* \biggl(\sum_{k=0}^{p} X_k^* \rho_{\min}^2 X_k \biggr) h < h^* \biggl(\sum_{k=0}^{p} Y_k^* \rho_1^2 Y_k \biggr) h.$$

Proof. We introduce the matrix function

$$\varphi(z) = \Phi(z)(Y_0 + Y_1 z + \cdots + Y_p z^p). \tag{5.21}$$

It follows from (5.4) and (5.18), (5.21) that

$$\varphi(z) = \rho_{\min}(X_0 + X_1 z + \cdots + X_p z^p) + \cdots . \tag{5.22}$$

Then the inequality

$$\sum_{k=0}^{p} X_k^* \rho_{\min}^2 X_k r^{2k} \leq \frac{1}{2\pi} \int_0^{2\pi} \varphi^*(z)\varphi(z)\, d\theta \tag{5.23}$$

where $z = re^{i\theta}, 0 < r < 1$, holds. Using (5.19) and (5.21) we obtain

$$\frac{1}{2\pi} \int_0^{2\pi} \varphi^*(z)\varphi(z) d\theta \leq \sum_{k=0}^{p} Y_k^* \rho_1^2 Y_k r^{2k} \tag{5.24}$$

the first inequality in (5.20) follows directly from (5.23) and (5.24). The second inequality follows from relations (5.21)–(5.24) which imply

$$X^* R_{\min}^2 X \leq Y^* R_{\min}^2 Y.$$

The last inequality in the theorem follows from (5.22) and the fact that $\Psi_1(z) \neq 0$. □

Remark 5.2. The notion of $\rho_{\min}$ was introduced in terms of min × rank $(RS_2R - S_1)$. We have proved that $\rho_{\min}$ satisfies the minimal property (4.15) (Nevanlinna-Pick problem) and the minimal property (5.20) (Schur problem) too. In Corollary 5.4 below we see that $\rho_{\min}$ fits the definition in the abstract at inequality (0.1). In the scalar case, $m = 1$, we deduce from (4.15) and (5.20) the well-known result [**1**]

$$\rho_{\min}^2 \le \rho_1^2. \tag{5.25}$$

From Theorem 5.1 we deduce the following assertions.

Corollary 5.3. *The solution of the Schur problem* (5.18), (5.19) *when* $\rho_1^2 \le \rho_{\min}^2$ *and* $\rho_1^2 \ne \rho_{\min}^2$ *does not exist.*

Proof. This follows immediately from (5.20). □

Corollary 5.4. *If the solution of the Schur problem* (5.18), (5.19) *when* $\rho_1^2 = \rho_{\min}^2$ *exists, as it does in the presence of* (5.14) *which says* $\operatorname{diag}\left\{\sum_{j=0}^{p} a_j a_j^*\right\} \ge C_{p-1}C_{p-1}^*$, *then this solution* $\varphi_{\min}(z)$ *has the form*

$$\varphi_{\min}(z) = \rho_{\min}\left(\sum_{k=0}^{p} X_k z^k\right)\left(\sum_{k=0}^{p} Y_k z^k\right)^{-1}. \tag{5.26}$$

Corollary 5.5. *If the matrix function* $\varphi_{\min}(z)$ *defined by* (5.26) *is holomorphic in the unit circle* $|z| < 1$, *then* $\varphi_{\min}(z)$ *is the solution of the Schur problem* (5.18), (5.19) *when* $\rho_1^2 = \rho_{\min}^2$.

Corollary 5.6. *Let* S, Φ_1, Φ_2 *and* $R = R_{\min}$ *defined by formulas* (5.5)–(5.10) *satisfy the conditions of Proposition* 3.1. *Then the corresponding Schur problem* ($\rho_1 = \rho_{\min}^2$) *has one and only one solution* $\varphi_{\min}(z)$ *and this solution has form* (5.26).

5.4 Examples.

Example 5.7. Let $p = 1$ and the given coefficients a_0 and a_1 have the form

$$a_0 = -\alpha I_m, \quad a_1 = \sqrt{Q^2 - \alpha^2 I_m}\, U \,\sqrt{Q^2 - \alpha^2 I_m}\, Q^{-1} \tag{5.27}$$

where Q and U are $m \times m$ matrices such that

$$U^* U = I_m, \quad Q > \alpha I_m \quad \text{with} \quad \alpha > 0. \tag{5.28}$$

The following assertion is proved in [**13**, Chapter 7, p. 101].

Proposition 5.8. *In case* (5.27), (5.28) *we have that* $\rho_{\min}$ *is unique and is given by*

$$\rho_{\min} = Q. \tag{5.29}$$

It follows from (5.17) *and* (5.26) *that in case* (5.27) *we have*

$$\varphi_{\min}(z) = [a_0 + (a_1 + a_0) y_1 z]\ [I_m + y_1 z]^{-1}, \tag{5.30}$$

where

$$y_1 = (I_m - a_0^* Q^{-2} a_0)^{-1} a_0^* Q^{-2} a_1.$$

Example 5.9. Let $m = 2$ in Example 5.7. Then we get

$$U = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad Q = \begin{bmatrix} \beta_1 & 0 \\ 0 & \beta_2 \end{bmatrix}, \tag{5.31}$$

where $\beta_1 > \alpha$, $\beta_2 > \alpha$. In view of (5.27) and (5.31) we have

$$a_0 = -\alpha I_2, \quad a_1 = \begin{bmatrix} 0 & \gamma/\beta_2 \\ \gamma/\beta_1 & 0 \end{bmatrix}, \tag{5.32}$$

where $\gamma = [(\beta_1^2 - \alpha^2)(\beta_2^2 - \alpha^2)]^{1/2}$.

6. Comparison of minimal rank, optimal and superoptimal H^∞ interpolation. A very appealing type of H^∞ interpolation and approximation was formulated by Young [**14**] (see more recent work in the paper [**11**] of Peller and Young). It is called *superoptimal* H^∞

interpolation and now we describe the superoptimal H^∞ interpolation problem. Denote the singular values or s- numbers of a matrix by $s_0 \geq s_1 \geq \cdots \geq 0$ and define

$$S_j(w) := \sup s_j(w), \quad |\xi| \leq 1.$$

Suppose we are given interpolation constraints $\mathcal{I}$.

We seek a solution w meeting the interpolation constraints $\mathcal{I}$ which minimizes $S_0(w)$, say we obtain value S_0^*, then minimize $S_1(w)$ subject to the constraint that $S_0(w) = S_0^*$, we continue this procedure down the sequence $S_j(w)$ with $j = 0, 1, \ldots, m$. A function, denoted $w_{sopt}(\xi)$, obtained in this way is called a superoptimal solution of the interpolation problem $\mathcal{I}$. Since the first term of this sequence is $S_0 = \sup \|w(\xi)\|$, a superoptimal solution is also an optimal solution.

There are various correspondences one could imagine between superoptimal and minimal interpolation and we list them as questions.

(a) Is a minimal solution for $\mathcal{I}$ also superoptimal for $\mathcal{I}$?

(b) Is a superoptimal solution for $\mathcal{I}$ also minimal for $\mathcal{I}$, that is, is there a ρ_{min} for which it is minimal?

6.1 Examples showing that minimal and superoptimal solutions are different. The optimal condition has the form

$$\sigma_{opt} := \sup_{|\xi|<1} \|w_{\min}(\xi)\| \leq \sup_{|\xi|<1} \|w(\xi)\|, \tag{6.1}$$

where the $m \times m$ matrix function $w(\xi)$ satisfies the interpolation constraints defining the problem. In Example 5.9 equation (5.32) says that σ_{opt} is defined by the relation

$$\frac{(\beta_2^2 - \alpha^2)(\beta_1^2 - \alpha^2)}{\beta_2^2} = (\sigma_{opt}^2 - \alpha^2)^2 \sigma_{opt}^{-2}, \tag{6.2}$$

where $\beta_1 > \beta_2 > \alpha$. It follows from (6.2) that

$$\beta_1 > \sigma_{opt} > \beta_2, \tag{6.3}$$

that is, some eigenvalues of $\rho_{\min}$ are greater than σ_{opt} but some of them are smaller than ρ_{opt}. The superoptimal solution has singular values

$S_0^* = \sigma_{opt}$ and $S_1^* \leq \sigma_{opt}$, thus we have shown that the minimal solution is not superoptimal. This shows that the answer to question (a) is no.

On the other hand the optimal and superoptimal solutions do not satisfy the extremal relations (4.15) and (5.20) which are fulfilled for the minimal rank solutions. This shows that the answer to question (b) is no.

An important property of the minimal rank approach is the explicit and simple form of $w_{\min}(\xi)$.

The choice of whether to use superoptimal or minimal approaches depends on the concrete scientific or engineering application. We should like to quote here Young's words [**14**] about superoptimal (strong) approach: "On the assumption that God is a good engineer as well as a geometer, I am inclined to expect that the stronger minimization condition, seeming so mathematically right, will have physical significance." We think that these words are true for the minimal rank approach as well.

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*, in *Topics in Interpolation Theory*, Oper. Theory Adv. Appl., vol. 95, Birkhaüser, Basel, 1997, pp. 19–35.

2. J.A. Ball, I. Gohberg and L. Rodman, *Interpolation of rational matrix functions*, Birkhaüser, Basel, 1990.

3. H. Dym, *J Contractive matrix functions, reproducing kernel Hilber spaces and interpolation*, CBMS, No. 71, Amer. Math. Soc., Providence, 1989.

4. J.A. Engwerda, A.C.M. Ran and A.L. Rijvabeer, *Necessary and sufficient conditions for the existence of a positive definite solution of the matrix equation*, Linear Algebra Appl. **186** (1993), 255–275.

5. A. Ferrante and B.C. Levy, *Hermitian solutions of the equation $X = Q + NX^{-1}N^*$*, Linear Algebra Appl. **247** (1996), 359–373. (See the reduction process in the proof of Theorem 2.8.)

6. M. Green and D.J.N. Limebeer, *Linear robust control*, Prentice-Hall, Englewood Cliffs, NJ, 1995.

7. J.W. Helton, J. Ball, C. Johnson and C. Palmer, *Operator theory, analytic functions, matrices and electrical engineering*, CBMS Regional Conf. Ser. in Math., vol. 68, Amer. Math. Soc., Providence, 1987.

8. T.S. Ivanchenko and L.A. Sakhnovich, *Operator identities in the theory of interpolation problems*, Soviet J. Contemporary Math. Anal. **22** (1987), 84–94.

9. H. Kimura, *State space approach to the classical interpolation problem and its applications*, Lecture Notes in Control Inform. Sci., vol. 135, Springer, New York, 1989, pp. 243–275.

10. ———, *Chain scattering approach to H^∞-control*, Birkhaüser, Boston, 1997.

11. V.V. Peller and N.J. Young, *Superoptimal analytic approximations of matrix functions*, J. Funct. Anal. **120** (1994), 300–343.

12. A.C.M.Ran and M.C.B. Reurings, *On the nonlinear matrix equation $X + A^\star F(X)A + Q$ solution and perturbation theory*, Linear Algebra Appl. **346** (2002), 15–26.

13. L. Sakhnovich, *Interpolation theory and its applications*, Kluwer Acad. Publ., Dordrecht, 1997.

14. N.J. Young, *The Nevalinna-Pick problem for matrix-valued functions*, J. Operator Theory **15** (1986), 289–265.

15. K. Zhou, J. Doyle and K. Glover, *Robust and optimal control*, Prentice-Hall, Upper Saddle River, NJ, 1996.

Department of Mathematics, University of California, La Jolla, CA 92093-0112
E-mail address: helton@math.ucsd.edu

Courant Institute, 251 Mercer Street, New York, NY