Math. Proc. Camb. Phil. Soc. (1996), **119**, 497
Printed in Great Britain

Superoptimal approximation by meromorphic functions

BY V. V. PELLER†
Department of Mathematics, Kansas State University, Manhattan, Kansas 66502, U.S.A.

AND N. J. YOUNG
Department of Mathematics and Statistics, Lancaster University, Lancaster, LA1 4YF

(*Received* 11 *July* 1994; *revised* 26 *September* 1994)

Abstract

Let G be a matrix function of type $m \times n$ and suppose that G is expressible as the sum of an H^∞ function and a continuous function on the unit circle. Suppose also that the $(k-1)$th singular value of the Hankel operator with symbol G is greater than the kth singular value. Then there is a unique superoptimal approximant to G in $H^\infty_{(k)}$: that is, there is a unique matrix function Q having at most k poles in the open unit disc which minimizes $s^\infty(G-Q)$ or, in other words, which minimizes the sequence

$$(s_0^\infty(G-Q), s_1^\infty(G-Q), s_2^\infty(G-Q), \ldots)$$

with respect to the lexicographic ordering, where

$$s_j^\infty(F) = \sup_{z\in\mathbb{T}} s_j(F(z))$$

and $s_j(.)$ denotes the jth singular value of a matrix. This result is due to the present authors [**PY1**] in the case $k = 0$ (when the hypothesis on the Hankel singular values is vacuous) and to S. Treil [**T2**] in general. In this paper we give a proof of uniqueness by a diagonalization argument, a high level algorithm for the computation of the superoptimal approximant and a recursive parametrization of the set of all optimal solutions of a matrix Nehari–Takagi problem.

Introduction

The celebrated results of Adamyan, Arov and Krein show how to approximate scalar functions on the circle by meromorphic functions in the open unit disc $\mathbb{D}$ with a prescribed number of poles. These results, besides being of mathematical interest, have made an impact on some engineering questions, notably the problem of constructing good low-order models of a given linear system (see the tutorial papers [**G2**, **Y3**]). Because most engineering systems have several inputs and outputs, they are described by matrix-valued functions, and so generalizations of the results of [**AAK**] to such functions are needed. There have been many papers carrying out such generalizations, starting with [**KL**]. There are often significant new complications in the matrix case: in particular, this is so for the question of uniqueness. Is the best

† V. V. Peller's research was supported by an NSF grant in modern analysis.

approximation to a given matrix function G on $\mathbb{T}$ by an element $Q \in H^{\infty}_{(k)}$ unique? Here $H^{\infty}_{(k)}$ is the set of meromorphic matrix-valued functions on $\mathbb{D}$, bounded on $\mathbb{T}$ and having at most k poles in $\mathbb{D}$ (a fuller definition is given below). For an affirmative answer one needs a more refined criterion of optimality than the simple minimization of the L^{∞} norm. A natural approach is to take account of all the singular values of the error function matrices $G(z) - Q(z)$: this leads to the notion of superoptimality. In [**PY1**] we carried out a thorough investigation of superoptimal approximations for the case $k = 0$ – that is, when G is to be approximated by a function Q bounded and analytic in $\mathbb{D}$ (the Nehari problem). The key result is that the superoptimal approximant is indeed unique as long as G is expressible as the sum of a continuous function on $\mathbb{T}$ and a function bounded and analytic in $\mathbb{D}$. Furthermore, we obtained detailed structural information about the superoptimal error $G - Q$. Briefly this error can be diagonalized by certain unitary-valued functions to give a diagonal matrix function whose diagonal entries each have constant modulus and negative winding number.

For applications to the problem of model reduction it is the case $k > 0$ which is important, and so the question naturally arises as to whether the uniqueness statement remains valid when $k > 0$. A simple example shows that it does not, but S. Treil [**T2**] showed that uniqueness does hold subject to a simple condition on the Hankel singular values of G. His method of proof is entirely different from that of [**PY1**]: he makes interesting and original use of matricial weighting functions, which amounts to a kind of implicit diagonalization. His method is efficient for numerous purposes, but it is not clear how it can be used for the construction of the superoptimal approximant. The method of [**PY1**] proceeds via explicit diagonalization and does yield an algorithm, as well as detailed structural information.

In this paper we extend our method to the case $k > 0$. We obtain a proof by diagonalization of Treil's uniqueness result and a high level algorithm for the construction of the superoptimal approximant. We also obtain a parametrization of all optimal solutions of the matrix Nehari–Takagi problem which is of independent interest.

The 'continuous time' analogues of the present results (relating to the approximation of functions on the imaginary axis by functions meromorphic in the right half plane) are given in [**PY2**].

1. *The superoptimal Nehari–Takagi problem*

The problem of finding $Q \in H^{\infty}_{(k)}$ which is closest to a given L^{∞} function G on the unit circle with respect to the L^{∞}-norm is sometimes called the Nehari–Takagi problem [**BGR**]. For matrix functions there is typically a high degree of non-uniqueness in the solution, and the question arises as to whether there is a natural way of selecting a 'very best' approximation. One idea is the superoptimal approximation, which is easily grasped with the aid of the following example. Let

$$g(z) = \frac{1}{z^2}.$$

It is easy to calculate with the aid of the familiar scalar theory of the Nehari–Takagi problem (see, for example, [**AAK**] or [**Y**, chapter 16]) that the best approximation to g with respect to the L^∞ norm by functions with at most a single pole in the open unit disc is the zero function. Now let

$$G(z) = \begin{bmatrix} g(z) & 0 \\ 0 & 0 \end{bmatrix}.$$

Then the distance from G to $H^\infty_{(1)}$ is 1, and this distance is attained at every element of the form

$$Q(z) = \begin{bmatrix} 0 & 0 \\ 0 & q(z) \end{bmatrix}$$

with any $q \in H^\infty_{(1)}$ of norm no greater than 1. Thus there is a whole L^∞ ball of optimal solutions to the present approximation problem. The most natural of these to pick is the one for which $q = 0$, and one might ask for a modification of the optimality condition which would select this solution. One way to achieve this aim is to ask for the Q among the optimal approximants which minimizes

$$s_1^\infty(G-Q) \overset{\text{def}}{=} \operatorname*{ess\,sup}_{z \in \mathbb{T}} s_1(G-Q)(z)$$

since

$$s_1^\infty(G-Q) = \|q\|_\infty$$

as long as the right-hand side is no greater than 1. This quantity is minimized uniquely when $q = 0$.

Let us now formalize the above notions. We write $M_{m,n}$ for the space of complex $m \times n$ matrices with the usual operator norm (the largest singular value). We denote by L^∞ the space of essentially bounded matrix functions on the unit circle with essential supremum norm (when we need to emphasize that the functions are of type $m \times n$ we write $L^\infty(M_{m,n})$). Thus, for $G \in L^\infty(M_{m,n})$,

$$\|G\|_\infty \overset{\text{def}}{=} \operatorname*{ess\,sup}_{z \in \mathbb{T}} \|G(z)\|.$$

H^∞ (or $H^\infty(M_{m,n})$) denotes the space of bounded analytic matrix functions in the open unit disc, with supremum norm. By Fatou's theorem, $H^\infty(M_{m,n})$ can be isometrically identified with a subspace of $L^\infty(M_{m,n})$. We denote by $H^\infty_{(k)}(M_{m,n})$ or simply $H^\infty_{(k)}$ the subset of $L^\infty(M_{m,n})$ consisting of functions essentially bounded on the unit circle, meromorphic on the open unit disc and having at most k poles there. These poles are counted as follows. A finite Blaschke–Potapov product of degree k is a function of the form

$$\Phi(z) = U_0 \begin{bmatrix} \frac{z-a_1}{1-\bar{a}_1 z} & 0 \\ 0 & I_{m-1} \end{bmatrix} U_1 \begin{bmatrix} \frac{z-a_2}{1-\bar{a}_2 z} & 0 \\ 0 & I_{m-1} \end{bmatrix} \cdots \begin{bmatrix} \frac{z-a_k}{1-\bar{a}_k z} & 0 \\ 0 & I_{m-1} \end{bmatrix} U_k,$$

where $m \in \mathbb{N}$, $|a_j| < 1$, $1 \leqslant j \leqslant k$, and U_j is a constant $m \times m$ unitary matrix, $0 \leqslant j \leqslant k$. We say that $G \in L^\infty$ of type $m \times n$ has at most k poles in the open unit disc if G is expressible on $\mathbb{T}$ in the form

$$G = \Phi^{-1} F$$

for some $F \in H^\infty(M_{m,n})$ and some Blaschke–Potapov product Φ of degree k. We denote by C the space of continuous matrix functions on the unit circle. The singular values of a matrix A are denoted by

$$s_0(A) \geqslant s_1(A) \geqslant \dots$$

and for $G \in L^\infty$ we write

$$s_j^\infty(G) \stackrel{\text{def}}{=} \operatorname*{ess\,sup}_{z \in \mathbb{T}} s_j(G(z)),$$

$$s^\infty(G) \stackrel{\text{def}}{=} (s_0^\infty(G), s_1^\infty(G), s_2^\infty(G), \dots).$$

Of course the first term of this sequence is $\|G\|_\infty$, and if G is of type $m \times n$ then at most the first $\min(m, n)$ terms can be non-zero. We shall say that Q is a superoptimal approximation to G in $H^\infty_{(k)}$ if Q is an element of $H^\infty_{(k)}$ at which the sequence $s^\infty(G-Q)$ attains its minimum with respect to the lexicographic ordering.

$H^\infty + C$ is the space of matrix functions on the unit circle expressible in the form $F+G$ with $F \in H^\infty, G \in C$. It was shown in [**PY1**] that if $G \in H^\infty + C$ then there is a unique superoptimal approximation to G in $H^\infty_{(0)}$, and it is natural to hope that the same will be true for $H^\infty_{(k)}$ with positive k. It was claimed in [**Y2**] that this conclusion does hold in the case of rational G, but the following example shows it is not true.

Let

$$G(z) = \begin{bmatrix} \frac{1}{z} & 0 \\ 0 & \frac{1}{z} \end{bmatrix}.$$

Then the lexicographic minimum of $s^\infty(G-Q)$ over $Q \in H^\infty_{(1)}$ is $(1, 0, 0, \dots)$, and this minimum is attained for either of the $H^\infty_{(1)}$ functions

$$Q_1(z) = \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{z} \end{bmatrix}, \quad Q_2(z) = \begin{bmatrix} \frac{1}{z} & 0 \\ 0 & 0 \end{bmatrix}$$

(or indeed for any function P/z with P a rank one orthogonal projection on $\mathbb{C}^2$). However the conclusion does follow if $s_k(H_G) < s_{k-1}(H_G)$, as was shown by Treil [**T2**]. We give an alternative proof using the diagonalization method of [**PY1**].

The starting point of most studies of the Nehari–Takagi problem is the theorem of Adamyan, Arov and Krein which relates best $H^\infty_{(k)}$ approximations to singular values and vectors of Hankel operators. Let $H^2(\mathbb{C}^n)$ denote the Hardy space of $\mathbb{C}^n$-valued functions on the circle (see [**He**]). Corresponding to the given G we define the Hankel operator

$$H_G : H^2(\mathbb{C}^n) \to (H^2(\mathbb{C}^m))^\perp$$

by

$$H_G x = P_-(Gx),$$

where

$$P_- : L^2(\mathbb{C}^m) \to (H^2(\mathbb{C}^m))^\perp$$

is the orthogonal projection operator. For $G \in H^\infty + C$, H_G is a compact operator. There is a matrix version of the theorem of Adamyan, Arov and Krein (see [**T1**]) which tells us that

$$\inf_{Q \in H^\infty_{(k)}} \|G - Q\|_\infty = s_k(H_G), \tag{1}$$

the kth singular value or s-number of H_G. The following simple fact is folklore.

LEMMA 1. *If $G \in H^\infty_{(k)}$ then H_G has rank at most k.*

Proof. We have $G = \Phi^* F$ for some Blaschke–Potapov product Φ of degree k and $F \in H^\infty$. For $x \in (H^2)^\perp$,

$$H_G^* x = P_+ G^* x = P_+ F^* \Phi x$$

from which it is clear that $\ker H_G^* \supset \Phi^*(H^2)^\perp$ and hence that the co-kernel of H_G^* is contained in the k-dimensional space $(H^2)^\perp \ominus \Phi^*(H^2)^\perp$. Thus H_G^* has rank $\leqslant k$.

LEMMA 2. *Let ϕ be an $n \times 1$ inner function. There exists a co-outer $n \times (n-1)$ function ϕ_c such that*

$$\Phi \overset{\text{def}}{=} [\phi \quad \bar{\phi}_c]$$

is unitary-valued, and all minors on the first column of Φ belong to H^∞.

Here a function $\phi \in H^\infty$ is *inner* if $\phi(z)$ is an isometry (i.e. $\phi(z)^*\phi(z)$ is the identity) for almost all $z \in \mathbb{T}$. A function $F \in H^\infty(M_{m,n})$ is *outer* if $FH^2(\mathbb{C}^n)$ is dense in $H^2(\mathbb{C}^m)$. F is co-outer if F^T, the transpose of F, is outer. A *minor* of Φ is the determinant of a square submatrix of Φ; if the submatrix corresponds to rows $i_1, \ldots, i_p$ and columns, $j_1, \ldots, j_p$ where $j_1 = 1$ then the minor is said to be on the first column of Φ.

For the proof see [**PY**, theorem 1·1]. At first sight the class of such Φ looks less natural than the class of inner functions. However, it seems to be essential to use such functions to obtain superoptimality results in the desired generality. There are moreover unexpected other benefits, notably a simplification of the diagonalization process.

THEOREM 1. *Let $G \in H^\infty + C$ be an $m \times n$ matrix function and let k be a non-negative integer. If either $k = 0$ or $s_k(H_G) < s_{k-1}(H_G)$ then there is a unique $Q \in H^\infty_{(k)}(M_{m,n})$ such that $s^\infty(G-Q)$ is a minimum with respect to the lexicographic ordering. For this Q the singular values*

$$s_j(G(z) - Q(z))$$

are constant a.e. on the unit circle for each $j \geqslant 0$.

Proof. Consider any $Q \in H^\infty_{(k)}$ which is at minimal distance from G. From (1),

$$\|G-Q\|_\infty = s_k \overset{\text{def}}{=} s_k(H_G).$$

If $s_k = 0$ then the rank of H_G is at most k and hence $G \in H^\infty_{(k)}$. Thus the only Q which minimizes $\|G-Q\|_\infty$ is G itself, and so the theorem holds in this case. Now assume $s_k > 0$. Let $v \in H^2$ be a unit singular vector of H_G corresponding to s_k. We claim that

$$(G-Q)\,v = H_G v.$$

In the case $k = 0$ this was established in [**AAK**]. Consider the case $k > 0$. Since H_Q has rank at most k, by Lemma 1, it follows from the definition of singular value that

$$\|H_{G-Q}\| = \|H_G - H_Q\| \geqslant s_k.$$

On the other hand,

$$\|H_{G-Q}\| \leqslant \|G-Q\|_\infty = s_k,$$

and so
$$\|H_{G-Q}\| = s_k.$$
We can thus apply the following elementary observation (see [**P**, theorem 6·14]).

LEMMA 3. *Let T be a bounded linear operator on Hilbert space and let k be a positive integer. Suppose that $s_{k-1}(T) > s_k(T)$ and that R is an operator of rank at most k such that $\|T-R\| = s_k(T)$. Then every singular vector of T corresponding to $s_k(T)$ is a maximizing vector of $T-R$ and is in* $\ker R$.

It follows that v is a maximizing vector of H_{G-Q}, and so
$$s_k = \|H_{G-Q}\| = \|H_{G-Q}v\| = \|P_-(G-Q)v\| \leqslant \|(G-Q)v\| \leqslant \|G-Q\|_\infty \|v\| = s_k.$$
Equality holds throughout, and so
$$\|P_-(G-Q)v\| = \|(G-Q)v\|,$$
whence
$$(G-Q)v = P_-(G-Q)v = H_{G-Q}v = H_G v,$$
and our claim is established.

We wish to use v to block-diagonalize $G-Q$. Let $\mathscr{E}$ denote the class of all $G-Q$, where $Q \in H^\infty_{(k)}$ attains the infimum in (1). Write
$$t_0 = s_k(H_G), \quad w = t_0^{-1}H_G v \in (H^2)^\perp.$$
Then, $v, \overline{zw} \in H^2$ and we may perform inner-outer factorizations
$$v = v_i h, \quad \overline{zw} = w_i h_1,$$
where v_i, w_i are inner column matrices and h, h_1 are scalar outer functions. The Adamyan–Arov–Krein theorem gives us the further information that
$$\|w(z)\|_{\mathbb{C}^m} = t_0 \|v(z)\|_{\mathbb{C}^n} \quad \text{a.e.},$$
whence we may take $h_1 = h$. We thus have
$$Ev_i h = t_0 \overline{zw_i} \bar{h}$$
for every $E \in \mathscr{E}$. Let the scalar unimodular function u_0 be defined by
$$u_0 = \overline{zh}/h. \tag{2}$$
Then
$$Ev_i = t_0 \bar{w}_i u_0 \tag{3}$$
for all $E \in \mathscr{E}$.

If $n = 1$ then v_i is a non-zero scalar H^2 function and we may divide through to obtain the uniqueness result for the $m \times 1$ case. By considering G^T we may deduce uniqueness and constancy of the singular values for the $1 \times n$ case. Thus the result is true if $\min(m, n) = 1$. Now consider $\min(m, n) > 1$ and suppose the result holds for any lesser value of $\min(m, n)$. By Lemma 2 there exist co-outer functions α, β such that
$$V \stackrel{\text{def}}{=} [v_i \quad \bar{\alpha}], \quad W^T \stackrel{\text{def}}{=} [w_i \quad \bar{\beta}]$$
are unitary-valued and have the property that all minors on the first column belong to H^∞. Equation (3) can be expressed
$$EV[1 \quad 0 \ldots 0]^T = W^*[t_0 u_0 \quad 0 \ldots 0]^T$$

for all $E \in \mathscr{E}$, whence

$$WEV = \begin{bmatrix} t_0 u_0 & * \\ 0 & * \end{bmatrix}.$$

In view of the fact that $\|E\|_\infty = t_0$ and u_0 has unit modulus a.e. the $(1,2)$ block must be identically zero, and so, for every $E \in \mathscr{E}$ there exists $F \in L^\infty$ such that

$$WEV = \begin{bmatrix} t_0 u_0 & 0 \\ 0 & F \end{bmatrix}, \tag{4}$$

F being of type $(m-1) \times (n-1)$. We now have to characterize the class of all F which can occur for some $E \in \mathscr{E}$.

Let $\hat{\mathscr{E}}$ denote the set of all $G-Q$ with $Q \in H^\infty_{(k)}$ such that $W(G-Q)V$ is of the form (4). Then $\hat{\mathscr{E}} \supset \mathscr{E}$, and, for $Q \in H^\infty_{(k)}$, the error $E \overset{\text{def}}{=} G-Q$ belongs to $\hat{\mathscr{E}}$ if and only if

$$Ev = t_0 w, \quad w^* E = t_0 v^*.$$

Fix some $E_1 \in \hat{\mathscr{E}}$, say $E_1 = G - Q_1$. Then

$$Q_1 v = Gv - t_0 w, \quad w^* Q_1 = w^* G - t_0 v^*.$$

For any $E = G - Q \in \mathscr{E}$ we have

$$WEV = WE_1 V + W(Q_1 V - QV). \tag{5}$$

Suppose

$$WE_1 V = \begin{bmatrix} t_0 u_0 & 0 \\ 0 & F_1 \end{bmatrix}. \tag{6}$$

Since $WEV, WE_1 V$ have the same first column (cf. (4)), so do $Q_1 V$ and QV, say

$$Q_1 V = [C \quad H_1], \quad QV = [C \quad H]. \tag{7}$$

Then (5) gives

$$\begin{bmatrix} t_0 u_0 & 0 \\ 0 & F \end{bmatrix} = \begin{bmatrix} t_0 u_0 & 0 \\ 0 & F_1 \end{bmatrix} + [0 \quad W(H_1 - H)].$$

The functions H which can appear in (7) are precisely those $H \in L^\infty(M_{m,n-1})$ satisfying

(i) $[C \quad H] \in H^\infty_{(k)}(M_{m,n}) V$;

(ii) $W(H_1 - H) \in \begin{bmatrix} 0 \\ L^\infty(M_{m-1,n-1}) \end{bmatrix}$ and $\begin{bmatrix} 0 \\ F_1 \end{bmatrix} + W(H_1 - H)$ has L^∞ norm at most t_0.

We wish to find H for which

(iii) $s^\infty(WEV)$ is minimized.

Now

$$s^\infty(WEV) = s^\infty\left(\begin{bmatrix} t_0 u_0 & 0 \\ 0 & F_1 \end{bmatrix} + W(H_1 - H)\right)$$

$$= \left(t_0, s^\infty\left(\begin{bmatrix} 0 \\ F_1 \end{bmatrix} + W(H_1 - H)\right)\right)$$

and since $W^T = [w_i \quad \bar{\beta}]$, (iii) is equivalent (in the presence of (ii)) to the statement that $s^\infty(F_1 + \beta^*(H_1 - H))$ is minimized.

We now parametrize the H satisfying (i).

LEMMA 4. *Let*

$$V = [v_i \quad \bar{\alpha}]$$

be unitary-valued of type $n \times n$ *with* v_i, α *inner and* α *co-outer. Let* $C \in H^\infty_{(k)}(M_{m,n})\, v_i$. *Then there exist* $K \in H^\infty(M_{m,n})$ *and a Blaschke–Potapov product* Φ *of type* $m \times m$ *and degree* $l \leqslant k$ *such that* $[C \quad H] \in H^\infty_{(k)}(M_{m,n})\, V$ *if and only if*

$$H \in \Phi^*(K\bar{\alpha} + H^\infty_{(k-l)}(M_{m,n-1})). \tag{8}$$

Proof. Pick a minimal inner $m \times m$ Φ such that $\Phi C \in H^\infty v_i$. Let l be the degree of Φ. Then $l \leqslant k$. Pick any $K \in H^\infty(M_{m,1})$ such that $\Phi C = K v_i$. Then the parametrization (8) holds with these functions.

Suppose H is of the form (8): $H = \Phi^*(K\bar{\alpha} + \Theta^* h)$ for some Blaschke–Potapov product Θ of degree $\leqslant k-l$ and some $h \in H^\infty$. Then

$$\begin{aligned}\Theta\Phi[C \quad H] &= [\Theta K v_i \quad \Theta K\bar{\alpha} + h] \\ &= (\Theta K + h\alpha^T)[v_i \quad \bar{\alpha}] \in H^\infty V.\end{aligned} \tag{9}$$

Hence $[C \quad H] \in H^\infty_{(k)}\, V$.

Conversely, suppose $[C \quad H] \in H^\infty_{(k)}\, V$. There exists a Blaschke–Potapov product X of degree k and $f \in H^\infty$ such that

$$[C \quad H] = X^* f V. \tag{10}$$

Consideration of the first column of this equation gives $XC = f v_i$. By choice of Φ, $X = \Theta\Phi$ for some inner Θ, which must have degree at most $k-l$. We have

$$(f - \Theta K)\, v_i = 0,$$

and hence, by choice of α, $f - \Theta K = Z\alpha^T$ for some Z in H^∞. Take the second block column in (10) to get

$$f\bar{\alpha} = XH = \Theta\Phi H,$$

whence
$$\Theta\Phi H = (\Theta K + Z\alpha^T)\,\bar{\alpha} = \Theta K\bar{\alpha} + Z.$$

Thus $H = \Phi^*(K\bar{\alpha} + \Theta^* Z)$ and so H is of the form (8).

Let us extract from this proof a formula for future use. For any optimal Q and $[C \quad H]$ given by (7), if H is parametrized as in the lemma by

$$H = \Phi^*(K\bar{\alpha} + h),$$

where l is the degree of Φ and $h \in H^\infty_{(k-l)}$, then (9) gives

$$\Phi[C \quad H] = (K + h\alpha^T)\, V,$$

and so, by virtue of (7),

$$Q = \Phi^*(K + h\alpha^T). \tag{11}$$

Lemma 4 enables us to write H_1, H in (ii) above in the form

$$H_1 = \Phi^*(K\bar{\alpha} + h_1), \quad H = \Phi^*(K\bar{\alpha} + h)$$

with $h_1, h \in H^\infty_{(k-l)}$. We can thus express conditions (i)–(iii) as

(i′) $H = \Phi^*(K\bar{\alpha} + h), \quad h \in H^\infty_{(k-l)}(M_{m,n-1}),$

(ii′) $w_i^T \Phi^*(h_1 - h) = 0$ and $F_1 + \beta^*\Phi^*(h_1 - h)$ has L^∞ norm at most t_0,

(iii′) $s^\infty(F_1 + \beta^*\Phi^*(h_1 - h))$ is minimized,

where h_1 is a fixed element of $H^\infty_{(k-l)}(M_{m,n-1})$. Note that the norm condition in (ii′) automatically holds when (iii′) is satisfied. We next parametrize the h in $H^\infty_{(k-l)}(M_{m,n-1})$ which satisfy (ii′).

LEMMA 5. *Let $h_1 \in H^\infty_{(p)}(M_{m,n-1})$ and let $[x \quad \bar{\gamma}]$ be unitary-valued of type $m \times m$ with x, γ inner, x of type $m \times 1$ and γ co-outer. There exist $A \in H^\infty(M_{m,n-1})$ and an $(n-1)$-square Blaschke–Potapov product Ψ of degree $r \leqslant p$ such that, for $h \in H^\infty_{(p)}$, h satisfies $x^T(h - h_1) = 0$ if and only if*

$$h \in (A + \gamma H^\infty_{(p-r)})\Psi^*. \tag{12}$$

Proof. Let Ψ be a minimal inner function such that

$$x^T h_1 \Psi \in x^T H^\infty(M_{m,n-1})$$

and let r be the degree of Ψ. Since $h_1 \in H^\infty_{(p)}, r \leqslant p$. Let

$$x^T h_1 = x^T A \Psi^* \quad \text{where} \quad A \in H^\infty.$$

Then the parametrization (12) holds with these functions.

Suppose that $h = (A + \gamma Z)\Psi^*$ for some $Z \in H^\infty_{(p-r)}$. Then

$$x^T(h_1 - h) = -x^T \gamma Z \Psi^* = 0.$$

Conversely, suppose $x^T(h - h_1) = 0$. Pick an $n-1$ square inner function Ω of degree at most p such that $h\Omega \in H^\infty$. Then $x^T h_1 \Omega \in H^\infty$ and hence $\Omega = \Psi X$ for some inner X of degree at most $p - r$. Since

$$x^T(h - A\Psi^*) = 0,$$

we have $h\Psi - A = \gamma Z$ for some $Z \in L^\infty(M_{m-1,n-1})$. Then

$$h\Omega = h\Psi X = AX + \gamma Z X \in H^\infty.$$

Thus $\gamma Z X \in H^\infty$, and since γ is co-outer it follows that $ZX \in H^\infty$ and so that $Z \in H^\infty_{(p-r)}$.

Now multiply both sides of condition (ii′) by the scalar inner function $\det \Phi$ to see that it is equivalent to $x^T(h_1 - h) = 0$ where $x = (\operatorname{adj} \Phi)^T w_i$. To apply Lemma 5 we need an inner co-outer function γ of type $m \times (m-1)$ such that $[x \quad \bar{\gamma}]$ is unitary-valued. Now

$$\operatorname{adj} \Phi^T [w_i \quad \bar{\beta}] = [x \quad \overline{(\operatorname{adj} \Phi^*)\beta}] = [x \quad (\det \Phi)\overline{\Phi\beta}] \tag{13}$$

is unitary-valued, and so therefore is $[x \quad \overline{\Phi\beta}]$. Let

$$(\Phi\beta)^T = \Upsilon^T \gamma^T$$

be the inner–outer factorization of $(\Phi\beta)^T$, so that Υ is $(m-1)$-square and γ is inner and co-outer of type $m \times (m-1)$. Then $[x \quad \bar{\gamma}]$ is unitary-valued and $\Phi\beta = \gamma\Upsilon$. Thus

$$\Upsilon = \gamma^* \Phi \beta. \tag{14}$$

By Lemma 5 there exist $A \in H^\infty$ and a finite Blaschke–Potapov product Ψ of degree $r \leqslant k - l$ such that the general h satisfying (ii′) can be written as

$$h = (A + \gamma Z)\Psi^*, \tag{15}$$

where $Z \in H^\infty_{(k-l-r)}(M_{m-1,n-1})$. In particular

$$h_1 = (A + \gamma Z_1)\Psi^*$$

for some $Z_1 \in H^\infty_{(k-l-r)}(M_{m-1,n-1})$. The objective (iii′) becomes to minimize

$$\begin{aligned} s^\infty(F_1 + \beta^*(H_1 - H)) &= s^\infty(F_1 + \beta^*\Phi^*(h_1 - h)) \\ &= s^\infty(F_1 + \beta^*\Phi^*\gamma(Z_1 - Z)\Psi^*) \\ &= s^\infty(F_1 + \Upsilon^*(Z_1 - Z)\Psi^*) \\ &= s^\infty(\Upsilon F_1 \Psi + Z_1 - Z), \end{aligned}$$

as Z varies over $H^\infty_{(k-l-r)}$: there is a one–one correspondence between superoptimal approximants $Q \in H^\infty_{(k)}$ to G and $Z \in H^\infty_{(k-l-r)}$ to

$$G_1 \stackrel{\text{def}}{=} \Upsilon F_1 \Psi + Z_1.$$

We are faced with a superoptimal model reduction problem for matrix functions of type $(m-1) \times (n-1)$. We shall show that this problem satisfies the hypotheses of the theorem so that we can invoke the inductive hypothesis.

Go back to equation (6) and take the 2×2 minor corresponding to rows 1, i and columns 1, j $(2 \leqslant i \leqslant m, 2 \leqslant j \leqslant n)$ of both sides. We obtain

$$\sum_{r,s,t,u} W_{1i,r,s}(E_1)_{rs,tu} V_{tu,1j} = t_0 u_0 (F_1)_{i-1,j-1}. \tag{16}$$

By choice of W, V, each of the 2×2 minors $W_{1i,rs}$ and $V_{tu,1j}$ are in H^∞, and since $E_1 \in H^\infty + C$, it follows that the 2×2 minor $(E_1)_{rs,tu} \in H^\infty + C$. Since $H^\infty + C$ is an algebra the left-hand side of (16) is in $H^\infty + C$, and so $u_0 F_1 \in H^\infty + C$ (we can suppose $t_0 \neq 0$). From (3) we have

$$t_0 u_0 = w_i^T E_1 v_i \in H^\infty + C,$$

while it is shown in [**PK**] that if u_0 has the form (2) and $u_0 \in H^\infty + C$ then also $\bar{u}_0 \in H^\infty + C$. Hence we have

$$F_1 = \bar{u}_0(u_0 F_1) \in H^\infty + C.$$

Since $\Upsilon, \Psi \in H^\infty$ and

$$Z_1 \in H^\infty_{(k-l-r)} \subset H^\infty + C$$

we conclude that

$$G_1 = \Upsilon F_1 \Psi + Z_1 \in H^\infty + C.$$

We are seeking $Z \in H^\infty_{(k-l-r)}$ to minimize $s^\infty(G_1 - Z)$. Uniqueness will follow from the inductive hypothesis provided either $k - l - r = 0$ or

$$s_{k-l-r-1}(H_{G_1}) > s_{k-l-r}(H_{G_1}).$$

Suppose these are both false, that is, that $l + r < k$ and

$$s_{k-l-r-1}(H_{G_1}) = s_{k-l-r}(H_{G_1}).$$

Then there exists $Z \in H^\infty_{(k-l-r-1)}$ which minimizes $\|G_1 - Z\|_\infty$ over $Z \in H^\infty_{(k-l-r)}$. This Z determines via (11) and (14) an element $Q \in H^\infty_{(k)}$ which is a best approximation to G. Now these formulae combine to give

$$Q = \Phi^*(K + (A + \gamma Z)\Psi^* \alpha^T).$$

Since Z has at most $k-l-r-1$ poles and Φ, Ψ are of degree l, r respectively, it follows that Q has at most $k-1$ poles. This implies

$$\operatorname{dist}(G, H^\infty_{(k)}) = \operatorname{dist}(G, H^\infty_{(k-1)}),$$

and hence, by the generalization to matrix functions of the Adamyan–Arov–Krein theorem,

$$s_{k-1}(H_G) = s_k(H_G).$$

This contradicts the hypothesis of the theorem. It follows that the singular values of H_{G_1} satisfy the requisite inequality for the inductive hypothesis to apply to G_1. There is thus a unique $Z \in H^\infty_{(k-l-r)}$ such that $s^\infty(G_1 - Z)$ attains its lexicographic minimum, and furthermore, for this Z, the singular values of $(G_1 - Z)(z)$ are constant a.e. on the unit circle. It follows that there is a unique $Q \in H^\infty_{(k)}$ which minimizes $s^\infty(G-Q)$, and since we have

$$W(G-Q)V = \begin{bmatrix} t_0 u_0 & 0 \\ 0 & \Upsilon^*(G_1 - Z)\Psi^* \end{bmatrix}, \tag{17}$$

it follows that the singular values of $(G-Q)(z)$ are constant a.e. on the unit circle. Theorem 1 is proved.

2. *Construction of superoptimal approximants*

The question arises as to whether the above proof shows how to construct a superoptimal approximation Q to a given G. In a sense it does: in principle, if we can follow the steps of the proof then we can compute the desired Q. In practice there may be difficulties in carrying out some of the steps. Nevertheless, we can give a high level algorithm for the calculation of superoptimal approximations, and in the next section we illustrate by an example that it can indeed be implemented, at least in simple cases.

Algorithm

Suppose given an $m \times n$ matrix function $G \in H^\infty + C$ and assume $s_{k-1}(H_G) > s_k(H_G)$ and $m \geqslant n$, where H_G is the Hankel operator with symbol G. Then the superoptimal approximation $Q \in H^\infty_{(k)}$ to G can be calculated as follows.

1. Find the kth singular value t_0 of H_G and a corresponding Schmidt pair $v \in H^2(\mathbb{C}^n)$, $w \in (H^2(\mathbb{C}^m))^\perp$. If $n = 1$ then the desired Q is $G - t_0 w/v$ and the algorithm terminates.

2. If $n > 1$ then perform the inner–outer factorizations

$$v = v_i h, \quad \overline{zw} = w_i h,$$

with h a scalar outer function.

3. Find inner co-outer functions α, β of types $n \times (n-1)$, $m \times (m-1)$ respectively such that

$$V \overset{\text{def}}{=} [v_i \quad \bar{\alpha}], \quad W^T \overset{\text{def}}{=} [w_i \quad \bar{\beta}]$$

are unitary-valued.

4. Find $Q_1 \in H^\infty_{(k)}(M_{m,n})$ such that

$$Q_1 v = Gv - t_0 v, \quad w^* Q_1 = w^* G - t_0 v^*.$$

5. Let Φ be a minimal $m\times m$ inner function such that $\Phi Q_1 v_i \in H^\infty(M_{m,n})v_i$, and pick $K\in H^\infty(M_{m,n})$ such that $\Phi Q_1 v_i = Kv_i$. Let

$$h_1 = (\Phi Q_1 - K)\bar{\alpha}.$$

6. Let

$$(\Phi\beta)^T = \Upsilon^T\gamma^T$$

be the inner–outer factorization of $(\Phi\beta)^T$, so that Υ is $(m-1)$-square and γ is inner and co-outer of type $m\times(m-1)$.

7. Let Ψ be a minimal $(n-1)$-square inner function such that

$$w_i^T(\operatorname{adj}\Phi)h_1\Psi \in w_i^T(\operatorname{adj}\Phi)H^\infty(M_{m,n-1}),$$

and pick $A\in H^\infty(M_{m,n-1})$ such that

$$w_i^T(\operatorname{adj}\Phi)h_1\Psi = w_i^T(\operatorname{adj}\Phi)A.$$

Let

$$Z_1 = \gamma^*(h_1\Psi - A).$$

8. Let G_1 of type $(m-1)\times(n-1)$ be defined by

$$G_1 = \Upsilon\beta^*(G-Q_1)\bar{\alpha}\Psi + Z_1.$$

Then $G_1\in H^\infty + C$. Let Z be the (unique) superoptimal approximation to G_1 in $H^\infty_{(p)}(M_{m-1,n-1})$, where

$$p = k - \operatorname{degree}\Phi - \operatorname{degree}\Psi.$$

Then the desired superoptimal approximation $Q\in H^\infty_{(k)}(M_{m,n})$ to G is

$$Q = \Phi^*(K + (A+\gamma Z)\Psi^*\alpha^T). \tag{18}$$

End of algorithm

The justification of this algorithm simply consists in following the steps through the proof of Theorem 1. The algorithm gives a recursive procedure for calculating Q since the computation in step 8 is a superoptimal Nehari–Takagi problem of type $(m-1)\times(n-1)$, and at the nth pass a problem of type $(m-n+1)\times 1$ will be reached, so that the algorithm will terminate at step 1.

If the constant singular values of $G_1 - Z$ are $t_1, t_2, \ldots$, then the constant singular values of $G-Q$ are $t_0, t_1, \ldots$.

The formula (18) for Q is obtained by combining (11) with the parametrization (15).

If it happens that $\operatorname{degree}\Phi + \operatorname{degree}\Psi = k$ then $p=0$ in Step 7 and the approximation problem reduces to the superoptimal Nehari problem studied in [**PY1**].

In [**PY3**, Sec. 2, remark 4] it is shown how α and β can be calculated.

3. *An example*

We illustrate the algorithm of the preceding section by outlining a worked example. Let $G = B^{-1}A$ where

$$A(z) = \begin{bmatrix}\sqrt{3}+2z & 0\\ 0 & 1\end{bmatrix},\quad B(z) = \frac{1}{\sqrt{2}}\begin{bmatrix} z^2 & z\\ z & -1\end{bmatrix}.$$

In [**PY3**] we found the superoptimal approximation to G in $H^\infty_{(0)}$; here we calculate the superoptimal approximation in $H^\infty_{(1)}$ (a more difficult problem). It is a simple exercise to show that the singular values of the rank 2 Hankel operator H_G are $\surd 6$ and 1. A Schmidt pair corresponding to the singular value 1 is

$$v(z) = \begin{bmatrix} -1+\surd 3z \\ 1 \end{bmatrix}, \quad w(z) = \frac{\bar{z}}{\surd 2}\begin{bmatrix} -\surd 3\bar{z}+2 \\ -\surd 3 \end{bmatrix}.$$

The outer factor of v and $\bar{z}\bar{w}$ is

$$h(z) = c(1+\gamma z),$$

where

$$c = \sqrt{\left[\frac{5+\surd 13}{2}\right]}, \quad \gamma = -\frac{5-\surd 13}{2\surd 3}.$$

Inner co-outer functions α, β as in step 3 are

$$\alpha = \frac{1}{h}\begin{bmatrix} 1 \\ 1-\surd 3z \end{bmatrix}, \quad \beta = \frac{1}{h\surd 2}\begin{bmatrix} -\surd 3 \\ \surd 3z-2 \end{bmatrix}.$$

It may be verified that

$$Q_1(z) = \begin{bmatrix} 0 & \surd 6 \\ 2\surd 2 & \surd 2(2-\surd 3z) \end{bmatrix}$$

satisfies the requirements of step 4. This Q_1 belongs to H^∞, not merely to $H^\infty_{(1)}$ as required by the algorithm, and in consequence steps 5–7 become very simple – we take

$$\Phi = I, \quad K = Q_1, \quad h_1 = 0, \quad \Upsilon = I, \quad \gamma = \beta, \quad \Psi = I, \quad A = 0, \quad Z_1 = 0.$$

Then in step 8 we have

$$G_1 = \beta^*(G-Q_1)\bar{\alpha}, \quad p = 1.$$

On calculating the scalar function G_1 we find that a cancellation occurs, and G_1 has only a simple pole in $\mathbb{D}$, at $z = -\gamma$. Thus $G_1 \in H^\infty_{(1)}$, and so the optimal (= superoptimal) approximation Z to G_1 is G_1 itself. It follows from (17) that the superoptimal approximation $Q \in H^\infty_{(1)}$ to G satisfies

$$G-Q = W^*\begin{bmatrix} t_0 u_0 & 0 \\ 0 & 0 \end{bmatrix} V^* = \frac{t_0}{|h|^2} wv^*.$$

Substitution of the known values of t_0, w etc., gives

$$Q(z) = \frac{1}{\surd 2(\surd 3z^2-5z+\surd 3)}\begin{bmatrix} -9+2\surd 3z & -3+\surd 3z \\ -2\surd 3-7z+2\surd 3z^2 & -2\surd 3+5z-\surd 3z^2 \end{bmatrix}.$$

In this example the calculation was greatly simplified by the fact that we found a $Q_1 \in \hat{\mathscr{E}}$ which belongs to H^∞. It would be pleasant if this were always possible, but unfortunately it is not: consider $G = \operatorname{diag}\{g_1, g_2\}$ where g_1 has Hankel singular values $4, 2, 0, \ldots$, and g_2 has values $3, 1, 0, \ldots$. Then, for $k = 2$, we have $s_2(H_G) = 2$ and any $Q \in \hat{\mathscr{E}}$ has at least one pole in $\mathbb{D}$.

4. *Parametrization of solutions of a Nehari–Takagi problem*

Although we are primarily concerned with the superoptimal Nehari–Takagi problem, the calculations carried out in the proof of Theorem 1 yield a result of interest for the usual 'optimal' Nehari–Takagi problem. We give a recursive parametrization of all solutions of a matrix Nehari–Takagi problem. That is, we can describe all optimal solutions of an $m\times n$ matrix problem in terms of the set of all solutions of a (possibly suboptimal) Nehari–Takagi problem for $(m-1)\times(n-1)$ matrix functions.

THEOREM 2. *Let k be a positive integer and let $G\in H^\infty + C$ be an $m\times n$ matrix function on the circle such that $s_k(H_G) < s_{k-1}(H_G)$. Suppose $m, n \geqslant 2$. There exist H^∞ functions A, K of types $m\times n$ and $m\times(n-1)$ respectively, finite Blaschke–Potapov products Φ, Ψ of types m-square and $(n-1)$-square and degrees l, r respectively where $l+r\leqslant k$, inner co-outer functions α, γ of types $n\times(n-1)$ and $m\times(m-1)$, and a function $G_1\in H^\infty+C$ of type $(m-1)\times(n-1)$ such that the set of all optimal solutions of the Nehari–Takagi problem for G with k poles, or in other words the set*

$$\{Q\in H^\infty_{(k)}(M_{m,n}) : \|G-Q\|_\infty = s_k(H_G)\},$$

is equal to

$$\{\Phi^*(K+(A+\gamma Z)\Psi^*\alpha^T) : Z\in H^\infty_{(k-l-r)}(M_{m-1,n-1}),\quad \|G_1 - Z\|_\infty \leqslant s_k(H_G)\}. \tag{19}$$

Indeed, K, A, Φ, Ψ, α, γ and G_1 may be chosen as in the Algorithm above.

Proof. This is just a repetition of part of the proof of Theorem 1. We retain the notation of that proof. We showed that all optimal Qs satisfy $(G-Q)v = H_G v$ and hence that

$$W(G-Q)V = \begin{bmatrix} t_0 u_0 & 0 \\ 0 & F \end{bmatrix} = \begin{bmatrix} t_0 u_0 & 0 \\ 0 & F_1 \end{bmatrix} + [0 \quad W(H_1 - H)],$$

where $H\in L^\infty(M_{m-1,n-1})$ satisfies conditions (i) and (ii) above. Lemma 4 gives the parametrization (8) of H satisfying (i), in terms of functions h satisfying conditions (i′) and (ii′). Lemma 5 then applies to give us the parametrization (15) of h in (ii′). Retracing our steps we find

$$H = \Phi^*(K\bar\alpha + h) = \Phi^*(K\bar\alpha + (A+\gamma Z)\Psi^*),$$

where Z varies over $H^\infty_{(k-l-r)}$ subject to $\|F_1 + W(H_1 - H)\|_\infty \leqslant t_0$, or equivalently

$$\|G_1 - Z\|_\infty = \|F_1 + \Upsilon^*(Z_1 - Z)\Psi^*\|_\infty \leqslant t_0,$$

so that the parameter Z varies as described. To express Q in terms of Z, we combine equations (11) and (15) to get the stated formula (19) for Q.

A different parametrization, in terms of a linear fractional map, was given by Ball and Helton in [**BH1**, theorem 3·9; corrected proof in **BH2**].

It should be mentioned that in the case of rational matrix functions G there is a well known parametrization of all optimal solutions of the matrix Nehari–Takagi problem in state space terms [**G1**].

REFERENCES

[**AAK**] V. M. ADAMYAN, D. Z. AROV and M. G. KREIN. Infinite Hankel block matrices and some related continuation problems. *Izv. Akad. Nauk Armyan. SSR Ser. Mat* **6** (1971), 87–112.

[**BGR**] J. A. BALL, I. GOHBERG and L. RODMAN. *Interpolation of rational matrix functions* (Birkhäuser Verlag, 1990).

[**BH1**] J. A. BALL and J. W. HELTON. A Beurling–Lax theorem for the Lie group $U(m, n)$ which contains most classical interpolation theory. *J. Operator Theory* **9** (1983), 107–142.

[BH2] J. A. BALL and J. W. HELTON. Interpolation problems of Pick–Nevanlinna and Loewner types for meromorphic matrix functions: parametrisation of the set of all solutions. *Integral Equations and Operator Theory* **9** (1986), 155–203.

[G1] K. GLOVER. All optimal Hankel-norm approximations of linear multivariable systems and their L^∞ error bounds. *Int. J. Control* **39** (1984), 1115–1193.

[G2] K. GLOVER. A tutorial on Hankel-norm approximation; in: *From data to model*, ed. J. C. Willems (Springer-Verlag, 1989).

[He] H. HELSON. *Lectures on invariant subspaces* (Academic Press, 1964).

[Ho] K. HOFFMAN. *Banach spaces of analytic functions* (Prentice Hall, 1962).

[KL] S. Y. KUNG and D. W. LIN. Optimal Hankel-norm model reduction: multivariable systems. *IEEE Trans. Automatic Control* **26** (1981), 832–852.

[P] J. R. PARTINGTON. *An introduction to Hankel operators* (Cambridge University Press, 1988).

[PK] V. V. PELLER and S. V. KHRUSCHEV. Hankel operators, best approximation and stationary Gaussian processes. *Russian Math. Surveys* **37** (1982), 53–124.

[PY1] V. V. PELLER and N. J. YOUNG. Superoptimal analytic approximations of matrix functions. *J. Functional Analysis* **120** (1994), 300–343.

[PY2] V. V. PELLER and N. J. YOUNG. *Superoptimal model reduction for continuous time systems.* Technical Report 94/94, Lancaster University Mathematics and Statistics Department.

[PY3] V. V. PELLER and N. J. YOUNG. Construction of superoptimal approximants (to be published in *Mathematics of Communications, Systems and Control*).

[T1] S. R. TREIL. The Adamyan–Arov–Krein theorem: a vector version. *Zap. Nauchn. Semin. LOMI* **141** (1985), 56–71 (Russian).

[T2] S. R. TREIL. On superoptimal approximation by analytic and meromorphic matrix-valued functions. *J. Functional Analysis* (to appear).

[Y1] N. J. YOUNG. The Nevanlinna–Pick problem for matrix-valued functions. *J. Operator Theory* **15** (1986), 239–265.

[Y2] N. J. YOUNG. Superoptimal Hankel norm approximations; in: *Modelling, robustness and sensitivity reduction in control systems*, ed. R. F. Curtain (Springer-Verlag, 1987).

[Y3] N. J. YOUNG. The Nehari problem and optimal Hankel norm approximation; in: *Analysis and optimization of systems: state and frequency domain approaches for infinite-dimensional systems,* edited by R. F. Curtain *et al.* Springer Lecture Notes in Control and Information Sciences No. **185** (1993), 199–221.