PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 137, Number 1, January 2009, Pages 205–210
S 0002-9939(08)09623-8
Article electronically published on August 4, 2008

ANALYTIC APPROXIMATION OF MATRIX FUNCTIONS AND DUAL EXTREMAL FUNCTIONS

V. V. PELLER

(Communicated by Michael T. Lacey)

ABSTRACT. We study the question of the existence of a dual extremal function for a bounded matrix function on the unit circle in connection with the problem of approximation by analytic matrix functions. We characterize the class of matrix functions, for which a dual extremal function exists in terms of the existence of a maximizing vector of the corresponding Hankel operator and in terms of certain special factorizations that involve thematic matrix functions.

1. Introduction

In this paper we consider the problem of approximation of bounded matrix-valued functions on the unit circle $\mathbb{T}$ by bounded analytic matrix functions in the unit disk $\mathbb{D}$. In other words, for $\Phi \in L^\infty(\mathbb{M}_{m,n})$ (i.e., Φ is a bounded function that takes values in the space $\mathbb{M}_{m,n}$ of $m \times n$ matrices), we search for a matrix function $F \in H^\infty(\mathbb{M}_{m,n})$ (i.e., F is a bounded analytic function in $\mathbb{D}$ with values in $\mathbb{M}_{m,n}$) such that

$$\|\Phi - F\|_{L^\infty} = \operatorname{dist}_{L^\infty}\big(\Phi, H^\infty(\mathbb{M}_{m,n})\big). \tag{1.1}$$

Here for a function $G \in L^\infty(\mathbb{M}_{m,n})$,

$$\|G\|_{L^\infty} \stackrel{\text{def}}{=} \operatorname*{ess\,sup}_{\zeta\in\mathbb{T}} \|G(\zeta)\|_{\mathbb{M}_{m,n}},$$

where for a matrix $A \in \mathbb{M}_{m,n}$, the norm $\|A\|_{\mathbb{M}_{m,n}}$ is the norm of A as an operator from $\mathbb{C}^n$ to $\mathbb{C}^m$. It is well known (and follows easily from a compactness argument) that the distance on the right-hand side of (1.1) is attained. A matrix function Φ is called *badly approximable* if the zero matrix function is a best approximant to Φ or, in other words,

$$\|\Phi\|_{L^\infty} \le \|\Phi - F\|_{L^\infty} \quad \text{for every} \quad F \in H^\infty(\mathbb{M}_{m,n}).$$

Note that by the matrix version of Nehari's theorem, the right-hand side of (1.1) is the norm of the Hankel operator $H_\Phi : H^2(\mathbb{C}^n) \to H^2_-(\mathbb{C}^m)$ that is defined on the Hardy class H^2 of $\mathbb{C}^n$-valued functions by

$$H_\Phi f = \mathbb{P}_-(\Phi f), \tag{1.2}$$

Received by the editors December 18, 2007.
2000 *Mathematics Subject Classification.* Primary 47B35, 46E40, 30D55, 30E10.
Key words and phrases. Best approximation, badly approximable matrix functions, dual extremal function, Hankel operator, maximizing vector.
The author is partially supported by NSF grant DMS 0700995.

where $\mathbb{P}_-$ is the orthogonal projection from the vector space $L^2(\mathbb{C}^m)$ onto the subspace $H^2_-(\mathbb{C}^m) \stackrel{\text{def}}{=} L^2(\mathbb{C}^m) \ominus H^2(\mathbb{C}^m)$ (see, e.g., [P, Ch. 2, §2]).

Note that this problem is very important in applications in control theory; see e.g., [F] and [P, Ch. 11].

By the Hahn–Banach theorem,

$$\text{dist}_{L^\infty}\big(\Phi, H^\infty(\mathbb{M}_{m,n})\big) = \sup\left|\int_{\mathbb{T}} \text{trace}\big(\Phi(\zeta)\Psi(\zeta)\big)\, d\boldsymbol{m}(\zeta)\right|, \tag{1.3}$$

where the supremum is taken over all matrix functions $\Psi \in H^1_0(\mathbb{M}_{n,m})$ such that $\|\Psi\|_{L^1(\boldsymbol{S}_1)} = 1$. Here $H^1_0(\mathbb{M}_{n,m})$ is the subspace of the Hardy class $H^1(\mathbb{M}_{n,m})$ of $m \times n$ matrix functions vanishing at the origin and the norm $\|A\|_{\boldsymbol{S}_1}$ of a matrix A is its *trace norm*: $\|A\|_{\boldsymbol{S}_1} \stackrel{\text{def}}{=} \text{trace}(A^*A)^{1/2}$.

However, it is well known that the supremum is not necessarily attained even for scalar matrix functions (see the remark following Theorem 3.1). If there exists a matrix function $\Psi \in H^1_0(\mathbb{M}_{n,m})$ such that

$$\|\Psi\|_{L^1(\boldsymbol{S}_1)} = 1 \quad \text{and} \quad \text{dist}_{L^\infty}\big(\Phi, H^\infty(\mathbb{M}_{m,n})\big) = \int_{\mathbb{T}} \text{trace}\big(\Phi(\zeta)\Psi(\zeta)\big)\, d\boldsymbol{m}(\zeta), \tag{1.4}$$

then Ψ is called a *dual extremal function* of Φ.

Note that the technique of dual extremal functions was used in [Kh] to study the problem of best analytic approximation in the scalar case.

In this paper we characterize the class of matrix functions Φ that have dual extremal functions. It turns out that this is equivalent to the fact that the Hankel operator H_Φ defined by (1.2) has a maximizing vector in $H^2(\mathbb{C}^n)$ which in turn is equivalent to the fact that the matrix function $\Phi - F$ (where F is a best approximant to Φ) admits a certain special factorization in terms of thematic matrix functions. The main result will be established in §3.

In §2 we state Sarason's factorization theorem [S], which will be used in §3, and we define the notion of a thematic matrix function.

2. Preliminaries

1. Sarason's Theorem. We are going to use the following result by D. Sarason:

Sarason's Theorem ([S])**.** *Let $\mathcal{H}$ be a separable Hilbert space and let Ψ be an analytic integrable $\mathcal{B}(\mathcal{H})$-valued function on $\mathbb{T}$. Then there exist analytic square integrable functions Q and R such that*

$$\Psi = QR, \quad R^*R = \big(\Psi^*\Psi\big)^{1/2}, \quad \textit{and} \quad Q^*Q = RR^* \quad \textit{a.e. on } \mathbb{T}. \tag{2.1}$$

Sarason's theorem implies the following fact:

Let Ψ be a matrix function in $H^1_0(\mathbb{M}_{n,n})$. Then there exist matrix functions $Q \in H^2(\mathbb{M}_{n,n})$ and $R \in H^2_0(\mathbb{M}_{n,n})$ such that

$$\Psi = QR \quad \text{and} \quad \|\Psi\|_{L^1(\boldsymbol{S}_1)} = \|Q\|_{L^2(\boldsymbol{S}_2)}\|R\|_{L^2(\boldsymbol{S}_2)}.$$

Here $H^2_0(\mathbb{M}_{n,n})$ is the Hardy class of $n \times n$ matrix functions vanishing at the origin. Recall that the *Hilbert–Schmidt norm* $\|A\|_{\boldsymbol{S}_2}$ of a matrix A is defined by $\|A\|_{\boldsymbol{S}_2} = \text{trace}\, A^*A$.

2. Thematic matrix functions. The notion of a thematic matrix function was introduced in [PY]. It turned out that it is very useful in the study of best approximation by analytic matrix functions (see [P, Ch. 14]).

Recall that a bounded analytic matrix function Θ is called an *inner function* if $\Theta^*(\zeta)^*\Theta(\zeta) = I$ for almost all $\zeta \in \mathbb{T}$, where I is the identity matrix. A matrix function $F \in H^\infty(m,n)$ is called *outer* if the operator of multiplication by F on $H^2(\mathbb{C}^n)$ has dense range in $H^2(\mathbb{C}^m)$. Finally, we say that a bounded analytic matrix function G is called *co-outer* if the transposed matrix function G^{t} is outer.

An $n \times n$ matrix function V is called a *thematic matrix function* if it has the form

$$V = \begin{pmatrix} \boldsymbol{v} & \overline{\Theta} \end{pmatrix},$$

where $\boldsymbol{v}$ is a column function, both functions $\boldsymbol{v}$ and Θ are inner and co-outer bounded analytic functions such that V takes unitary values on $\mathbb{T}$, i.e.,

$$V^*(\zeta)V(\zeta) = I, \quad \text{for almost all} \quad \zeta \in \mathbb{T}.$$

Note that a bounded analytic column function is co-outer if and only if its entries are coprime; i.e., they do not have a common nonconstant inner factor.

3. The main result

It is easy to see that a matrix function $\Phi \in L^\infty(\mathbb{M}_{m,n})$ has a dual extremal function if and only if $\Phi - F$ has a dual extremal function for any $F \in H^\infty(\mathbb{M}_{m,n})$. Moreover, if Ψ is a dual extremal function of Φ, then Ψ is also a dual extremal function for $\Phi - F$ for any $F \in H^\infty(\mathbb{M}_{m,n})$. Thus to characterize the class of matrix functions that possess extremal functions, it suffices to consider badly approximable matrix functions.

Theorem 3.1. *Let Φ be a nonzero badly approximable function in $L^\infty(\mathbb{M}_{m,n})$ with $m \ge 2$ and $n \ge 2$. The following are equivalent:*

(i) *the Hankel operator H_Φ has a maximizing vector;*

(ii) *Φ has a dual extremal function $\Psi \in H^1_0(\mathbb{M}_{n,m})$;*

(iii) *Φ has a dual extremal function $\Psi \in H^1_0(\mathbb{M}_{n,m})$ such that $\operatorname{rank}\Psi(\zeta) = 1$ almost everywhere on $\mathbb{T}$;*

(iv) *Φ admits a factorization*

$$\Phi = W^* \begin{pmatrix} tu & \mathbf{0} \\ \mathbf{0} & \Phi_\# \end{pmatrix} V^*, \tag{3.1}$$

where $t = \|\Phi\|_{L^\infty(\mathbb{M}_{m,n})}$, V and W^{t} are thematic matrix functions, u is a scalar function of the form $u = \bar{z}\bar{\vartheta}\bar{h}/h$ for an inner function ϑ and an outer function h in H^2, and $\Phi_\#$ is an $(n-1)\times(m-1)$ matrix function such that $\|\Phi_\#(\zeta)\| \le t$ for almost all $\zeta \in \mathbb{T}$.

Note that the proof of the implication (i)$\Rightarrow$(iv) is contained in [PY]; see also [P, Ch. 14, Th. 2.2]. However, we give here the proof of this implication for completeness.

Proof. (ii)$\Rightarrow$(i). By adding zero columns or zero rows if necessary, we may reduce the general case to the case $m = n$. Let Ψ be a matrix function in $H^1_0(\mathbb{M}_{n,n})$

that satisfies (1.4). By Sarason's theorem, there exist functions $Q \in H^2(\mathbb{M}_{n,n})$ and $R \in H^2_0(\mathbb{M}_{n,n})$ such that

$$\Psi = QR \quad \text{and} \quad 1 = \|\Psi\|_{L^1(\boldsymbol{S}_1)} = \|Q\|_{L^2(\boldsymbol{S}_2)}\|R\|_{L^2(\boldsymbol{S}_2)}.$$

Let $e_1, \cdots, e_n$ be the standard orthonormal basis in $\mathbb{C}^n$. We have

$$\begin{aligned}\int_{\mathbb{T}} \operatorname{trace}\big(\Phi(\zeta)\Psi(\zeta)\big)\, d\boldsymbol{m}(\zeta) &= \int_{\mathbb{T}} \operatorname{trace}\big(\Phi(\zeta)Q(\zeta)R(\zeta)\big)\, d\boldsymbol{m}(\zeta)\\ &= \int_{\mathbb{T}} \operatorname{trace}\big(R(\zeta)\Phi(\zeta)Q(\zeta)\big)\, d\boldsymbol{m}(\zeta)\\ &= \sum_{j=1}^{k} \int_{\mathbb{T}} \big(\Phi(\zeta)Q(\zeta)e_j, R^*(\zeta)e_j\big)\, d\boldsymbol{m}(\zeta)\\ &= \sum_{j=1}^{k} (H_\Phi Q e_j, R^* e_j)\end{aligned}$$

(we consider here Qe_j and R^*e_j as vector functions). By the Cauchy–Bunyakovsky–Schwarz inequality, we have

$$\begin{aligned}\left|\int_{\mathbb{T}} \operatorname{trace}\big(\Phi(\zeta)\Psi(\zeta)\big)\, d\boldsymbol{m}(\zeta)\right| &\le \sum_{j=1}^{n} |(H_\Phi Q e_j, R^* e_j)|\\ &\le \left(\sum_{j=1}^{n} \|H_\Phi Q e_j\|^2_{L^2(\mathbb{C}^n)}\right)^{1/2} \left(\sum_{j=1}^{n} \|R^* e_j\|^2_{L^2(\mathbb{C}^n)}\right)^{1/2}\\ &\le \|H_\Phi\| \left(\sum_{j=1}^{n} \|Q e_j\|^2_{L^2(\mathbb{C}^n)}\right)^{1/2} \left(\sum_{j=1}^{n} \|R^* e_j\|^2_{L^2(\mathbb{C}^n)}\right)^{1/2}.\end{aligned}$$

Clearly,

$$\begin{aligned}\left(\sum_{j=1}^{n} \|Q e_j\|^2_{L^2(\mathbb{C}^n)}\right)^{1/2} &= \left(\sum_{j=1}^{n} \int_{\mathbb{T}} \|Q(\zeta) e_j\|^2_{\mathbb{C}^n}\, d\boldsymbol{m}(\zeta)\right)^{1/2}\\ &= \left(\int_{\mathbb{T}} \|Q(\zeta)\|^2_{\boldsymbol{S}_2}\, d\boldsymbol{m}(\zeta)\right)^{1/2} = \|Q\|_{L^2(\boldsymbol{S}_2)}\end{aligned}$$

and

$$\left(\sum_{j=1}^{n} \|R^* e_j\|^2_{L^2(\mathbb{C}^n)}\right)^{1/2} = \|R\|_{L^2(\boldsymbol{S}_2)}.$$

Since Φ is badly approximable, we have $\|H_\Phi\| = \|\Phi\|_{L^\infty(\mathbb{M}_{n,n})}$.

It follows that

$$\begin{aligned}\|\Phi\|_{L^\infty(\mathbb{M}_{n,n})} &= \left|\int_{\mathbb{T}} \operatorname{trace}\big(\Phi(\zeta)\Psi(\zeta)\big)\,d\boldsymbol{m}(\zeta)\right| \\ &\le \left(\sum_{j=1}^n \|H_\Phi Q e_j\|^2_{L^2(\mathbb{C}^n)}\right)^{1/2} \left(\sum_{j=1}^n \|R^* e_j\|^2_{L^2(\mathbb{C}^n)}\right)^{1/2} \\ &\le \|H_\Phi\|\cdot\|Q\|_{L^2(\boldsymbol{S}_2)}\|R\|_{L^2(\boldsymbol{S}_2)} = \|\Phi\|_{L^\infty(\mathbb{M}_{n,n})}.\end{aligned}$$

Thus all inequalities are equalities and if $Qe_j \ne \mathbf{0}$, then Qe_j is a maximizing vector of H_Φ.

The implication (iii)⇒(ii) is trivial.

(iv)⇒(iii). Suppose that Φ is a function given by (3.1). By multiplying h by a constant if necessary, we may assume without loss of generality that $\|h\|_{L^2} = 1$. Let

$$V = \begin{pmatrix}\boldsymbol{v} & \overline{\Theta}\end{pmatrix} \quad \textit{and} \quad W^{\mathrm{t}} = \begin{pmatrix}\boldsymbol{w} & \Xi\end{pmatrix}.$$

Put

$$\Psi = z\vartheta h^2 \begin{pmatrix}\boldsymbol{v} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\boldsymbol{w}^{\mathrm{t}} \\ \mathbf{0}\end{pmatrix}.$$

Clearly,

$$\|\Psi\|_{L^1(\boldsymbol{S}_1)} = \|h^2\|_{L^1} = 1,$$

and it is easy to see that

$$\begin{aligned}\int_{\mathbb{T}} \operatorname{trace}\big(\Phi(\zeta)\Psi(\zeta)\big)\,d\boldsymbol{m}(\zeta) &= \int_{\mathbb{T}} z\vartheta h^2 \operatorname{trace}\left(\begin{pmatrix}\boldsymbol{w}^{\mathrm{t}} \\ \mathbf{0}\end{pmatrix}\Phi\begin{pmatrix}\boldsymbol{v} & \mathbf{0}\end{pmatrix}\right) d\boldsymbol{m} \\ &= \int_{\mathbb{T}} \operatorname{trace}\begin{pmatrix}|h|^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0}\end{pmatrix} d\boldsymbol{m} = 1.\end{aligned}$$

(i)⇒(iv). Let f be a maximizing vector of H_Φ. It is well known (see [P, Ch. 2. Th. 2.3]) that

$$\|\Phi(\zeta)\|_{\mathbb{M}_{m,n}} = \|\Phi\|_{L^\infty} = \|H_\Phi\|, \quad \|\Phi(\zeta)f(\zeta)\|_{\mathbb{C}^m} = \|H_\Phi\|\cdot\|f(\zeta)\|_{\mathbb{C}^n}, \qquad \zeta \in \mathbb{T},$$

and

$$\Phi f \in H^2_-(\mathbb{C}^m).$$

Put

$$g = \frac{1}{\|H_\Phi\|}\bar{z}\overline{\Phi f} = \frac{1}{\|H_\Phi\|}\bar{z}\overline{H_\Phi f} \in H^2(\mathbb{C}^m).$$

Then

$$\|f(\zeta)\|_{\mathbb{C}^n} = \|g(\zeta)\|_{\mathbb{C}^m}, \quad \zeta \in \mathbb{T}.$$

It follows that both f and g admit factorizations

$$f = \vartheta_1 h\boldsymbol{v}, \quad g = \vartheta_2 h\boldsymbol{w},$$

where ϑ_1 and ϑ_2 are scalar inner functions, h is a scalar outer function in H^2, and $\boldsymbol{v}$ and $\boldsymbol{w}$ are inner and co-outer column functions. Then $\boldsymbol{v}$ and $\boldsymbol{w}$ admit thematic completions; i.e., there are inner and co-outer matrix functions Θ and Ξ such that the matrix functions

$$\begin{pmatrix}\boldsymbol{v} & \overline{\Theta}\end{pmatrix} \quad \text{and} \quad \begin{pmatrix}\boldsymbol{w} & \Xi\end{pmatrix}$$

are thematic. Put

$$V = \begin{pmatrix} \boldsymbol{v} & \overline{\Theta} \end{pmatrix}, \quad W = \begin{pmatrix} \boldsymbol{w} & \overline{\Xi} \end{pmatrix}^{\mathrm{t}}, \quad \text{and} \quad u = \bar{z}\overline{\vartheta}_1\overline{\vartheta}_2\bar{h}/h.$$

Consider the matrix function $W\Phi V$. It is easy to see that its upper left entry is equal to

$$\boldsymbol{w}^{\mathrm{t}}\Phi\boldsymbol{v} = \frac{\overline{\vartheta}_2}{h}g^{\mathrm{t}}\Phi\frac{\overline{\vartheta}_1}{h}f = \|H_\Phi\|\bar{z}\frac{\overline{\vartheta}_1\overline{\vartheta}_2}{h^2}g^{\mathrm{t}}g = \|H_\Phi\|u = tu.$$

Since the norm of $(W\Phi V)(\zeta)$ is equal to t and its upper left entry $tu(\zeta)$ has modulus t almost everywhere, it is easy to see that the matrix function $W\Phi V$ has the form

$$W\Phi V = \begin{pmatrix} tu & \mathbf{0} \\ \mathbf{0} & \Phi_\# \end{pmatrix},$$

where $\Phi_\#$ is an $(m-1)\times(n-1)$ matrix function such that $\|\Phi_\#\|_{L^\infty} \le t$. It follows that

$$\Phi = W^* \begin{pmatrix} tu & \mathbf{0} \\ \mathbf{0} & \Phi_\# \end{pmatrix} V^*,$$

which completes the proof. □

Remark. In the case when Φ has size $m \times 1$, $m > 1$, Theorem 3.1 remains true if we replace the factorization in (3.1) with the factorization

$$\Phi = W^* \begin{pmatrix} tu \\ \mathbf{0} \end{pmatrix},$$

where W^{t} is a thematic matrix function and u has the form $u = \bar{z}\bar{\vartheta}\bar{h}/h$, where ϑ is a scalar inner function and h is a scalar outer function in H^2.

Similarly, the theorem can be stated in the case of size $1 \times n$, $n > 1$.

In the case of scalar functions, the result also holds if we replace (iv) with the condition that Φ admits a factorization in the form

$$\Phi = \bar{z}\bar{\vartheta}\bar{h}/h,$$

where ϑ is a scalar inner function and h is a scalar outer function in H^2.

Since it is well known that not all scalar badly approximable functions have constant modulus on $\mathbb{T}$ (see, e.g., [P, Ch. 1, §1]), there are scalar functions in L^∞ that have no dual extremal functions.

References

[F] B. A. Francis, *A Course in H^∞ Control Theory,* Lecture Notes in Control and Information Sciences, No. 88, Springer-Verlag, Berlin, 1987. MR0932459 (89i:93002)

[Kh] S. Havinson, On some extremal problems of the theory of analytic functions, *Uchen. Zapiski Mosk. Universiteta, Matem.* **148:4** (1951), 133–143. English Translation: Amer. Math. Soc. Translations (2) **32** (1963), 139–154. MR0049322 (14:155f)

[P] V.V. Peller, *Hankel operators and their applications,* Springer-Verlag, New York, 2003. MR1949210 (2004e:47040)

[PY] V.V. Peller and N.J. Young, Superoptimal analytic approximations of matrix functions, *J. Funct. Anal.* **120** (1994), 300-343. MR1266312 (94m:47030)

[S] D. Sarason, *Generalized interpolation in H^∞,* Trans. Amer. Math. Soc. **127** (1967), 179–203. MR0208383 (34:8193)

Department of Mathematics, Michigan State University, East Lansing, Michigan 48824