View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Diposit Digital de Documents de la UAB

ANALYTIC APPROXIMATION OF MATRIX FUNCTIONS IN L^p

L. BARATCHART, F.L. NAZAROV, AND V.V. PELLER

Abstract. We consider the problem of approximation of matrix functions of class L^p on the unit circle by matrix functions analytic in the unit disk in the norm of L^p, $2 \le p < \infty$. For an $m \times n$ matrix function Φ in L^p, we consider the Hankel operator $H_\Phi \colon H^q(\mathbb{C}^n) \to H^2_-(\mathbb{C}^m)$, $1/p + 1/q = 1/2$. It turns out that the space of $m \times n$ matrix functions in L^p splits into two subclasses: the set of respectable matrix functions and the set of weird matrix functions. If Φ is respectable, then its distance to the set of analytic matrix functions is equal to the norm of H_Φ. For weird matrix functions, to obtain the distance formula, we consider Hankel operators defined on spaces of matrix functions. We also describe the set of p-badly approximable matrix functions in terms of special factorizations and give a parametrization formula for all best analytic approximants in the norm of L^p. Finally, we introduce the notion of p-superoptimal approximation and prove the uniqueness of a p-superoptimal approximant for rational matrix functions.

1. Introduction

The classical problem of analytic approximation of functions on the unit circle $\mathbb{T}$ is for a given function $\varphi \in L^\infty$, to find a best H^∞ approximant to φ, i.e., to find a bounded analytic function ψ in the unit disk $\mathbb{D}$ such that

$$\|\varphi - \psi\|_{L^\infty(\mathbb{T})} = \operatorname{dist}_{L^\infty}(\varphi, H^\infty).$$

A standard compactness argument shows that such a best approximant always exists, though it is not necessarily unique in general. However, under certain mild assumptions the best approximation is indeed unique. For example, this happens if φ is continuous which was proved for the first time in [Kha]. We refer the reader to [Pe1] for a comprehensive study of the problem of best uniform approximation by analytic functions.

The second author is partially supported by NSF grant DMS 0501067, the third author is partially supported by NSF grant DMS 0700995.

It turns out that this approximation problem is closely related to Hankel operators on the Hardy class H^2. For a function $\varphi \in L^\infty$ the *Hankel operator*

$$H_\varphi \colon H^2 \to H^2_- \stackrel{\text{def}}{=} L^2 \ominus H^2$$

is defined by

$$H_\varphi f = \mathbb{P}_- \varphi f, \quad f \in H^2,$$

where $\mathbb{P}_-$ is the orthogonal projection from L^2 onto H^2_-. It was proved by Nehari (see [Pe1], Ch. 1, § 1) that

$$\|H_\varphi\| = \operatorname{dist}_{L^\infty}(\varphi, H^\infty).$$

Moreover, it turns out that the Hankel operators provide a powerful tool to constructively study the problem of best uniform analytic approximation, see Chapters 1, 5, and 7 of [Pe1]. The problem of uniform approximation by analytic functions is also called the *Nehari problem*.

The Nehari problem is very important in applications in control theory (see [F] and [Pe1]) and is also a useful tool in identification, see [Pa] and [BLPT]. Moreover, for the needs of control theory it is important to consider not only the scalar case, but also the case of matrix-valued functions.

Let Φ be a bounded function with values in the space $\mathbb{M}_{m,n}$ of $m \times n$ matrices (notationally, $\Phi \in L^\infty(\mathbb{M}_{m,n})$). The problem of best analytic approximation is to find a bounded analytic matrix function Q of size $m \times n$ such that

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

where $H^\infty(\mathbb{M}_{m,n})$ is the space of bounded analytic $m \times n$ matrix functions and for a matrix function $\Psi \in L^\infty(\mathbb{M}_{m,n})$ we use the notation

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

where for a matrix A in $\mathbb{M}_{m,n}$ we denote by $\|A\|_{\mathbb{M}_{m,n}}$ the operator norm of A as an operator from $\mathbb{C}^n$ to $\mathbb{C}^m$.

As in the scalar case, the following distance formula holds:

$$\operatorname{dist}_{L^\infty}\big(\Phi, H^\infty(\mathbb{M}_{m,n})\big) = \|H_\Phi\|, \quad \Phi \in L^\infty(\mathbb{M}_{m,n}),$$

where the Hankel operator $H_\Phi \colon H^2(\mathbb{C}^n) \to H^2_-(\mathbb{C}^m) \stackrel{\text{def}}{=} L^2(\mathbb{C}^m) \ominus H^2(\mathbb{C}^m)$ is defined by

$$H_\Phi f = \mathbb{P}_- \Phi f, \quad f \in H^2(\mathbb{C}^n),$$

and $\mathbb{P}_-$ is the orthogonal projection onto $H^2_-(\mathbb{C}^m)$ (see, e.g., [Pe1], Ch. 2).

However, unlike the scalar case, even if Φ is a polynomial matrix function, generically Φ has infinitely many best approximants. To choose among all best approximants the "very best approximant", it is natural to consider

the notion of superoptimal approximation. We refer the reader to § 2 of this paper for the definition of superoptimal approximation.

In this paper we are going to consider the problem of analytic approximation in the L^p norm, $2 \le p < \infty$.

Let φ be a scalar function in L^p. The problem of best analytic approximation is to find a function ψ in the Hardy class H^p such that

$$\|\varphi - \psi\|_{L^p} = \operatorname{dist}_{L^p}(\varphi, H^p).$$

If $1 < p < \infty$, then the space L^p is uniformly convex which implies that every function $\varphi \in L^p$ has a unique best analytic approximant ψ in the L^p. The function ψ is said to be the *p-best analytic approximant to* φ.

In [BS] Hankel operators have been used to study the problem of best analytic and meromorphic approximation in L^p for $2 \le p < \infty$ (see also [Pr] for a dual approach). For $\varphi \in L^p$ the Hankel operator

$$H_\varphi \colon H^q \to H^2_-$$

is defined by

$$H_\varphi f = \mathbb{P}_- \varphi f, \quad f \in H^q,$$

where the exponent q satisfies the equality

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{2}. \tag{1.1}$$

Throughout this paper we always assume that $2 \le p < \infty$ *and* q *satisfies* (1.1). *In the proofs of the results we assume that* $2 < p < \infty$, *it is an elementary exercise to adjust the proofs for* $p = 2$.

As in the case of uniform analytic approximation, the following formula holds

$$\|H_\varphi\|_{H^q \to H^2_-} = \operatorname{dist}_{L^p}(\varphi, H^p)$$

In § 2 of this paper we discuss in more detail the problem of best analytic approximation by scalar analytic functions in L^p.

In this paper we deal with the problem of approximation in L^p by analytic matrix functions: given a function Φ in $L^p(\mathbb{M}_{m,n})$ (i.e., all entries of Φ belong to L^p), we search for a best analytic approximant $Q \in H^p(\mathbb{M}_{m,n})$, i.e.,

$$\|\Phi - Q\|_{L^p} = \operatorname{dist}_{L^p}\big(\Phi, H^p(\mathbb{M}_{m,n})\big),$$

where for a matrix function $\Psi \in L^p(\mathbb{M}_{m,n})$,

$$\|\Psi\|_{L^p} \stackrel{\text{def}}{=} \|\Psi\|_{L^p(\mathbb{M}_{m,n})} = \left(\int_{\mathbb{T}} \|\Psi(\zeta)\|^p_{\mathbb{M}_{m,n}} \, d\boldsymbol{m}(\zeta)\right)^{1/p}.$$

If we consider the Hankel operator

$$H_\Phi \colon H^q(\mathbb{C}^n) \to H^2_-(\mathbb{C}^m)$$

defined by

$$H_\Phi f = \mathbb{P}_- \Phi f, \quad f \in H^q(\mathbb{C}^n),$$

it is easy to verify that

$$\|H_\Phi\| \le \operatorname{dist}_{L^p}\big(\Phi, H^p(\mathbb{M}_{m,n})\big)$$

(see Lemma 4.3). It will be shown in § 4 that if Φ has a p-best analytic approximant Q such that for ζ in a subset of $\mathbb{T}$ of positive measure, the space of maximizing vectors of $(\Phi - Q)(\zeta)$ is one-dimensional, then

$$\|H_\Phi\| = \operatorname{dist}_{L^p}\big(\Phi, H^p(\mathbb{M}_{m,n})\big). \tag{1.2}$$

Clearly, generically for an $m \times n$ matrix A, the maximizing vectors of A span a one-dimensional subspace.

This makes it plausible that for a dense subset of matrix functions Φ in $L^p(\mathbb{M}_{m,n})$ the distance formula (1.2) holds which would imply that (1.2) holds for all matrix functions Φ in $L^p(\mathbb{M}_{m,n})$.

Surprisingly, this is false!

In § 3 of this paper we obtain certain factorization theorems for analytic matrix functions that will be used to study Hankel operators. The main tool used in § 3 is Sarason's factorization theorem [Sa].

In § 4 we study the class of matrix functions $\Phi \in L^\infty(\mathbb{M}_{m,n})$, for which the distance formula (1.2) holds. Such matrix functions are called *respectable*. We obtain several characterizations of the class of respectable matrix functions.

The main result of § 5 is a construction of a 2×2 matrix function Φ, for which (1.2) is false. Such matrix functions are called *weird*.

Thus the space $L^p(\mathbb{M}_{m,n})$ splits in two subsets: the set of respectable matrix functions and the set of weird matrix functions. To compute the distance from a respectable matrix function Φ to the set of analytic matrix functions, we can use the distance formula (1.2). However, to compute $\operatorname{dist}_{L^p}\big(\Phi, H^p(\mathbb{M}_{m,n})\big)$ for weird matrix functions Φ, we have to search for another formula. Note that in a sense both the set of respectable matrix functions and the set of weird matrix functions are massive subsets of $L^p(\mathbb{M}_{m,n})$; see the discussion at the end of § 5.

It turns out, however, that the distance $\operatorname{dist}_{L^p}\big(\Phi, H^p(\mathbb{M}_{m,n})\big)$ from Φ to the set of analytic matrix functions can be obtained for all matrix functions in L^p as the norm of a Hankel operator if we consider Hankel operators acting on spaces of matrix functions rather than vector functions. Indeed, If we consider the Hankel operator $\boldsymbol{H}_\Phi$ defined on the space $H^q(\boldsymbol{S}_2^n)$ of $n \times n$ matrix functions with the norm

$$\|F\|_{L^q(\boldsymbol{S}_2^n)} = \left(\int_{\mathbb{T}} \|F(\zeta)\|_{\boldsymbol{S}_2^n}^q \, d\boldsymbol{m}(\zeta)\right)^{1/q},$$

Then the norm of the Hankel operator

$$\boldsymbol{H}_\Phi \colon H^q(\boldsymbol{S}_2^n) \to H^2_-(\boldsymbol{S}_2^n)$$

is equal to $\operatorname{dist}_{L^p}\big(\Phi, H^p(\mathbb{M}_{m,n})\big)$. Here for an $n \times k$ matrix A the norm $\|A\|_{\boldsymbol{S}_2^{n,k}}$ is the Hilbert–Schmidt norm of A and $\|A\|_{\boldsymbol{S}_2^n} \stackrel{\text{def}}{=} \|A\|_{\boldsymbol{S}_2^{n,n}}$. This will be proved in § 6. We also consider in § 6 Hankel operators acting on spaces of $n \times k$ matrix functions and we introduce in § 6 the class of $n \times n$ matrix functions in L^p of order k, $1 \le k \le n$.

In § 7 we obtain a description of the set of p-badly approximable matrix functions. A matrix function $\Phi \in L^p(\mathbb{M}_{m,n})$ is called *p-badly approximable* if

$$\|\Phi\|_{L^p} = \operatorname{dist}_{L^p}\big(\Phi, H^p(\mathbb{M}_{m,n})\big).$$

To obtain such a description, we use special factorizations that involve balanced matrix functions (see [Pe1], Ch. 14, § 1).

We also obtain in § 7 a parametrization formula for all p-best approximants.

In the last section we define the notion of p-superoptimal approximation and prove for rational matrix functions the uniqueness of a p-superoptimal approximant.

In § 2 we collect necessary information. In § 2.1 we present results on analytic approximation in L^p of scalar functions. In § 2.2 we define the notion of superoptimal approximation and state some uniqueness results and properties of superoptimal approximants. Finally, in § 2.3 we define the notion of balanced matrix functions and state factorization formulas for badly approximable matrix function.

Note that it suffices to study the problem of analytic approximation only for square matrix functions. Indeed, if a matrix function Φ is not square, we can add to Φ zero columns or zero rows to make it square. For the sake of simplicity, *beginning § 6, we state all the results only for square matrix functions.*

Notation and terminology. Throughout the paper we are going to use the following notation and terminology:

if X and Y are normed spaces and $T\colon X \to Y$ is a bounded linear operator, a vector $x \in X$ is called a *maximizing vector of T* if

$$x \ne \boldsymbol{0} \quad \text{and} \quad \|Tx\|_Y = \|T\| \cdot \|x\|_X;$$

if both X and Y are Hilbert spaces and T is a bounded linear operator from X to Y, then, by definition, the *space of maximizing vectors of T* is

$$\{x \in X : \ \|Tx\|_Y = \|T\| \cdot \|x\|_X\}$$

(it is well known that the space of maximizing vectors is a closed subspace of X that consists of the maximizing vectors and the zero vector);

$\mathbb{M}_{m,n}$ is the space of $m \times n$ matrices;

$\mathbb{M}_n \stackrel{\text{def}}{=} \mathbb{M}_{n,n}$;

if X is a normed space of functions on $\mathbb{T}$, then $X(\mathbb{M}_{m,n})$ means the space of $m \times n$ matrix functions whose entries belong to X. If this does not lead to a confusion, we say that $\Phi \in X$ for an $m \times n$ matrix function Φ if $\Phi \in X(\mathbb{M}_{m,n})$;

if $X = L^s$, $1 \le s \le \infty$, and $\Phi \in X(\mathbb{M}_{m,n})$, then

$$\|\Phi\|_X \stackrel{\text{def}}{=} \|\Phi\|_{X(\mathbb{M}_{m,n})} \stackrel{\text{def}}{=} \|\rho\|_X, \quad \text{where} \quad \rho(\zeta) \stackrel{\text{def}}{=} \|\Phi(\zeta)\|_{\mathbb{M}_{m,n}}, \quad \zeta \in \mathbb{T};$$

for $1 \le p \le \infty$, the space $H^p(\mathbb{M}_{m,n})$ is the subspace of $L^p(\mathbb{M}_{m,n})$ that consists of matrix functions with entries in H^p. By definition,

$$H^p_0(\mathbb{M}_{m,n}) = \{F \in L^p(\mathbb{M}_{m,n}) : \ F(0) = 0\};$$

for an operator A on Hilbert space (or for a matrix A), the *singular values* $s_j(A)$ are defined by

$$s_j(A) = \inf\{\|A - K\| : \ \operatorname{rank} K \le j\};$$

the Schatten–von Neumann class $\boldsymbol{S}_r$, $1 \le r < \infty$, consists of operators A on Hilbert space with finite norm

$$\|A\|_{\boldsymbol{S}_r} = \left(\sum_{j \ge 0} s_j^r(A) \right)^{1/r}; \tag{1.3}$$

for $r \in [1, \infty)$, we denote by $\boldsymbol{S}_r^{m,n}$ the space of $m \times n$ matrices A equipped with the Schatten–von Neumann norm (1.3);

$\boldsymbol{S}_r^n \stackrel{\text{def}}{=} \boldsymbol{S}_r^{n,n}$;

if $X = L^s$, $1 \le s \le \infty$, then $X(\boldsymbol{S}_r^{m,n})$ is the space of $m \times n$ matrix functions with entries in X equipped with the norm

$$\|\Phi\|_{X(\boldsymbol{S}_r^{m,n})} \stackrel{\text{def}}{=} \|\rho\|_X, \quad \text{where} \quad \rho(\zeta) \stackrel{\text{def}}{=} \|\Phi(\zeta)\|_{\boldsymbol{S}_r^{m,n}}, \quad \zeta \in \mathbb{T}.$$

Acknowledgements

We would like to thank Gilles Pisier and Ilya Spitkovskii for helpful remarks. We are also grateful to Alberto Condori who has thoroughly read the manuscript and suggested several corrections.

2. Preliminaries

2.1. **Analytic approximation by scalar functions in L^p, $2 \le p < \infty$.** The problem of analytic approximation in L^p was studied by many mathematicians, see, e.g., [Sh] and [Ka]. As we have already mentioned in the introduction, in [BS] to study the problem of best analytic approximation in L^p, Hankel operators from H^q to H^2_- were used, where the exponent q satisfies (1.1) (see also [Pr] in which a similar approach is used). The approach of [BS] and [Pr] is based on the analog of Nehari's theorem:

$$\|H_\varphi\|_{H^q \to H^2_-} = \operatorname{dist}_{L^p}(\varphi, H^p), \quad \varphi \in L^p. \tag{2.1}$$

Moreover, it can be shown that if $\varphi \in L^2$, then the Hankel operator H_φ defined on the set of analytic polynomials by the formula

$$H_\varphi f = \mathbb{P}_- \varphi f$$

extends to a bounded operator from H^q to H^2_- if and only if $\mathbb{P}_-\varphi \in L^p$. This can be proved in exactly the same way as in the case of classical Hankel operators from H^2 to H^2_- (see, e.g., [Pe1], Ch. 1, § 1). In particular, this implies that all bounded Hankel operators from H^q to H^2_- are compact, since the trigonometric polynomials are dense in L^p and H_φ has finite rank if φ is a trigonometric polynomial.

A scalar function $\varphi \in L^p$ is called *p-badly approximable* if

$$\|\varphi - \psi\|_{L^p} \ge \|\varphi\|_{L^p}$$

for any $\psi \in H^p$.

The following result describes the class of all p-badly approximable functions.

Theorem 2.1. *Let φ be a nonzero function in L^p. Then φ is p-badly approximable if and only if there exists an inner function ϑ and an outer function h in H^2 such that*

$$\varphi = \bar z \bar\vartheta \frac{\bar h}{h^{2/q}} = \bar z \bar\vartheta \frac{\bar h}{h^{\frac{p-2}{p}}}. \tag{2.2}$$

Proof. Suppose that φ is p-badly approximable. Let $f \in H^q$ be a maximizing vector of H_φ. Such a vector exists, since H_φ is compact. We have

$$\|H_\varphi f\|_{L^2} = \|\mathbb{P}_-\varphi f\|_{L^2} \le \|\varphi f\|_{L^2}$$

$$\le \|\varphi\|_{L^p}\|f\|_{L^q} = \|H_\varphi\| \cdot \|f\|_{L^q} = \|H_\varphi f\|_{L^2},$$

since f is a maximizing vector. Thus all inequalities in the above chain are equalities. The fact that the first inequality turns into equality means that $\varphi f \in H^2_-$. The second inequality turns into equality if and only if

$|\varphi|^p = c|f|^q$ for some $c > 0$. We can multiply f by a constant after which c becomes equal to 1. Let h be an outer function in H^2 such that $|h| = |f|^{q/2}$. Then f admits a factorization $f = \vartheta_1 h^{2/q}$, where ϑ_1 is an inner function.

Put $g = \bar{z}\overline{H_\varphi f} \in H^2$. We have $|g|^2 = |\varphi f|^2 = |h|^2$. Let $g = \vartheta_2 h$, where ϑ_2 is an inner function. Then

$$\varphi = \frac{\bar{z}\bar{g}}{f} = \bar{z}\bar{\vartheta}_1\bar{\vartheta}_2\frac{\bar{h}}{h^{2/q}}.$$

It remains to put $\vartheta = \vartheta_1\vartheta_2$.

Suppose now that φ is of the form (2.2). Put $f = h^{2/q}$. We have

$$\|H_\varphi\| \cdot \|f\|_{L^q} \ge \|H_\varphi f\|_{L^2} = \|h\|_{L^2} = \|\varphi\|_{L^p}\|f\|_{L^q} \ge \|H_\varphi\| \cdot \|f\|_{L^q}.$$

Thus $\|\varphi\|_{L^p} = \|H_\varphi\|$, and so φ is p-badly approximable. ■

Remark. Note that in the case $p = \infty$ the situation is slightly different. A bounded Hankel operators from H^2 to H^2_- is not necessarily compact and does not necessarily have a maximizing vector. A badly approximable function φ has the form

$$\varphi = c\bar{z}\bar{\vartheta}\frac{\bar{h}}{h},$$

where $c \in \mathbb{C}$, ϑ is an inner function, and h is an outer function in H^2, if and only if the Hankel operator $H_\varphi\colon H^2 \to H^2_-$ has a maximizing vector, see [Pe1], Ch. 1, § 1.

In the case $p = 2$, Theorem 2.1 means that the 2-badly approximable functions are precisely the functions in H^2_- and a function $f \in H^\infty$ is a maximizing vector of the Hankel operator $H_\varphi\colon H^\infty \to H^2_-$ with a 2-badly approximable symbol φ if and only if $f = c\vartheta$, where c is a nonzero complex number and ϑ is an inner divisor of $\bar{z}\bar{\varphi}$.

Corollary 2.2. *Let ω be a nonnegative function in L^p. The following are equivalent:*

(i) *there exists a p-badly approximable function $\varphi \in L^p$ such that $|\varphi| = \omega$;*

(ii) $\log\omega \in L^1$.

Proof. The implication (i)⇒(ii) is an immediate consequence of Theorem 2.1 and the fact that the logarithm of the modulus of any outer function is in L^1.

Conversely, suppose that $\log\omega \in L^1$. Let h be an outer function such that $|h| = \omega^{p/2}$. Clearly, $h \in H^2$. Let $\varphi = \bar{z}\frac{\bar{h}}{h^{2/q}}$. By Theorem 2.1, φ is badly approximable. We have

$$|\varphi| = |h|^{1-2/q} = |h|^{2/p} = \omega. \quad ■$$

Corollary 2.3. *Let ω be a nonnegative function in L^p such that $\log\omega \in L^1$ and let h be an outer function such that $|h| = \omega^{p/2}$. Then the set of p-badly approximable functions with modulus ω coincides with*

$$\left\{\bar z\bar\vartheta\frac{\bar h}{h^{2/q}} :\ \vartheta \ \text{ is an inner function}\right\}.$$

Proof. The result follows immediately from Theorem 2.1. ■

2.2. **Superoptimal approximation.** As we have already mentioned in the introduction, even for polynomial matrix functions Φ there can be many best analytic approximants in the L^∞ norm. For instance, if $\Phi = \begin{pmatrix}\bar z & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}$ and $F = \begin{pmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & f\end{pmatrix}$, where f is a scalar function in the unit ball of H^∞, then F is a best approximant to Φ.

To introduce the notion of superoptimal approximation, recall the notion of singular values of matrices. For a matrix A the jth singular value of A is defined by

$$s_j(A) \stackrel{\text{def}}{=} \inf\{\|A-K\| :\ \operatorname{rank} K \le j\},\quad j\ge 0.$$

Clearly, $s_0(A) = \|A\|$.

Definition. Given a matrix function $\Phi\in L^\infty(\mathbb{M}_{m,n})$ we define inductively the sets $\boldsymbol{\Omega}_j$, $0\le j\le \min\{m,n\}-1$, by

$$\boldsymbol{\Omega}_0 = \left\{Q\in H^\infty(\mathbb{M}_{m,n}) :\ Q \text{ minimizes } \ t_0 \stackrel{\text{def}}{=} \operatorname*{ess\,sup}_{\zeta\in\mathbb{T}}\|\Phi(\zeta)-Q(\zeta)\|\right\};$$

$$\boldsymbol{\Omega}_j = \left\{Q\in\Omega_{j-1} :\ Q \text{ minimizes } \ t_j \stackrel{\text{def}}{=} \operatorname*{ess\,sup}_{\zeta\in\mathbb{T}} s_j\big(\Phi(\zeta)-Q(\zeta)\big)\right\},\quad j>0.$$

Functions in $\bigcap\limits_{k\ge0}\boldsymbol{\Omega}_k = \boldsymbol{\Omega}_{\min\{m,n\}-1}$ are called *superoptimal approximants* to Φ by bounded analytic matrix functions. The numbers $t_j = t_j(\Phi)$ are called the *superoptimal singular values* of Φ. Note that the matrix functions in $\boldsymbol{\Omega}_0$ are just the best approximants by analytic matrix functions.

In other words, a superoptimal approximant minimizes the essential suprema of the singular values of $(\Phi-Q)(\zeta)$ lexicographically.

It was proved in [PY] that if $\Phi\in(H^\infty+C)(\mathbb{M}_{m,n})$ (i.e., each entry of Φ is a sum of a a continuous function and an H^∞ function), then Φ has a unique superoptimal approximant. Moreover, if Q is the unique superoptimal approximant to Φ, then

$$s_j\big((\Phi-Q)(\zeta)\big) = t_j,\quad \zeta\in\mathbb{T}.$$

Later in [PT1] the same results were obtained under a less restrictive assumption on Φ. We refer the reader to [Pe1], Ch. 14 for a detailed presentation of the theory of superoptimal approximation.

2.3. **Balanced matrix functions and factorizations of badly approximable matrix functions.** A matrix function Φ in $L^\infty(\mathbb{M}_{m,n})$ is called *badly approximable* if

$$\|\Phi\|_{L^\infty} \le \|\Phi - Q\|_{L^\infty}$$

for any $Q \in H^\infty(\mathbb{M}_{m,n})$.

A matrix function Φ is called *very badly approximable* if the zero matrix function is a superoptimal approximant to Φ.

In [PY] and [AP] the set of badly approximable matrix functions of class $(H^\infty + C)(\mathbb{M}_{m,n})$ was described in terms of certain special factorizations (see also [PT2] in which a geometric description of very badly approximable matrix functions was obtained). Such factorizations involve certain special unitary-valued matrix functions (balanced matrix functions), see [Pe1], Ch. 14, § 1. To define balanced matrix functions, we have to introduce several notions.

A matrix function $\Theta \in H^\infty(\mathbb{M}_{n,k})$ is called *inner* if on the unit circle $\Theta^*\Theta = \boldsymbol{I}_k$, where $\boldsymbol{I}_k$ is the matrix function identically equal to the identity matrix I_k.

A matrix function $F \in H^2(\mathbb{M}_{m,n})$ is called *outer* if the set

$$\{Ff : \ f \text{ is a polynomial in } H^2(\mathbb{C}^n)\}$$

is dense in $H^2(\mathbb{C}^m)$.

Finally, a matrix function $F \in H^2(\mathbb{M}_{m,n})$ is called *co-outer* if the transposed function F^{t} is outer.

It is well known (see, e.g., [N] or [SF]) that if Ψ is a matrix function of class H^2, then Φ admits an *inner-outer factorization*

$$\Phi = \Theta F,$$

where Θ is an inner matrix function and F is an outer matrix function.

Let $k < n$ and let Υ be an $n \times k$ inner and co-outer matrix function. It is well known (see [Pe1], Ch. 14, § 1 and [H], Ch. 9) that there exists an inner and co-outer matrix function Θ of size $n \times (n-k)$ such that the matrix function

$$\mathcal{V} = \begin{pmatrix} \Upsilon & \overline{\Theta} \end{pmatrix} \tag{2.3}$$

takes unitary values or, in other words, is *unitary-valued*. Matrix functions of the form (2.3) are called *balanced matrix functions*. If we want to specify that the analytic part of $\mathcal{V}$ has k columns, we say that $\mathcal{V}$ is a *k-balanced matrix function*. In the case $k = 1$, k-balanced matrix functions are also

called *thematic matrix functions*. If $k = n$ by a k-balanced matrix function, we mean a matrix function of the form $\tau \boldsymbol{I}_n$, where τ is a complex number of modulus 1.

Balanced matrix functions have many interesting properties, see [Pe1], Ch. 14, § 1. They have been used to obtain a description of badly approximable matrix functions, to parametrize the set of best analytic approximants, to characterize very badly approximable matrix functions, to prove the uniqueness of superoptimal approximants, and to construct the superoptimal approximant (see [PY], [PT1], [AP], and [Pe1], Ch. 14).

In particular, it was shown in [PY] (see also [Pe1], Ch. 14, § 2) that if Φ is a matrix function in $L^\infty(\mathbb{M}_{m,n})$ such that the Hankel operator $H_\Phi \colon H^2(\mathbb{C}^n) \to H^2_-(\mathbb{C}^m)$ has a maximizing vector, then Φ is badly approximable if and only if Φ admits a factorization

$$\Phi = \|H_\Phi\| W^* \begin{pmatrix} \bar{z}\bar{\vartheta}\bar{h}/h & \mathbf{0} \\ \mathbf{0} & \Phi_\# \end{pmatrix} V^*,$$

where V and W^{t} are thematic matrix functions, ϑ is a scalar inner function, h is a scalar outer function in H^2, and $\Phi_\#$ is a matrix function of size $(m-1) \times (n-1)$ such that $\|\Phi_\#(\zeta)\|_{\mathbb{M}_{m-1,n-1}} \le 1$ almost everywhere on $\mathbb{T}$.

Another characterization of badly approximable functions was obtained in [AP] (see also [Pe1], Ch. 14, § 15). Let $\Phi \in (H^\infty + C)(\mathbb{M}_{m,n})$ and let k be the number of superoptimal singular values $t_j(\Phi)$ equal to $t_0(\Phi)$ (in other words, k is the multiplicity of the superoptimal singular value $t_0(\Phi)$). Then Φ is badly approximable if and only if

$$\Phi = \|H_\Phi\| \mathcal{W}^* \begin{pmatrix} U & \mathbf{0} \\ \mathbf{0} & \Phi_\# \end{pmatrix} \mathcal{V}^*,$$

where U is a $k \times k$ very badly approximable unitary-valued function of class $H^\infty + C$, $\mathcal{V}$ and $\mathcal{W}$ are k-balanced matrix functions, and $\Phi_\#$ is a matrix function in $(H^\infty + C)(\mathbb{M}_{m-k,n-k})$ such that $\|\Phi_\#(\zeta)\|_{\mathbb{M}_{m-k,n-k}} \le 1$ almost everywhere on $\mathbb{T}$ and $\|H_{\Phi_\#}\| < 1$. Actually, the condition $\Phi \in (H^\infty + C)(\mathbb{M}_{m,n})$ can be relaxed (see [AP] and [Pe1], Ch. 14, § 15).

3. **Factorization of analytic matrix functions**

In this section we obtain several factorization theorems for analytic matrix functions that will be used to study Hankel operators.

We are going to use the following result by D. Sarason that is an analog of Riesz factorization:

Sarason's Theorem [Sa]. *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 $\mathcal{Q}$ and $\mathcal{R}$ such that

$$\Psi = \mathcal{Q}\mathcal{R}, \quad \mathcal{R}^*\mathcal{R} = \left(\Psi^*\Psi\right)^{1/2}, \quad \textit{and} \quad \mathcal{Q}^*\mathcal{Q} = \mathcal{R}\mathcal{R}^* \quad \textit{a.e. on } \mathbb{T}. \tag{3.1}$$

The following theorem can be deduced easily from Sarason's theorem. Recall that $2 \le p < \infty$ and q satisfies (1.1); as usual, p' is the dual exponent: $1/p + 1/p' = 1$.

Theorem 3.1. *Let $\Psi \in H^{p'}(\mathbb{M}_n)$. Then there exist matrix functions $F \in H^q(\mathbb{M}_n)$ and $G \in H^2(\mathbb{M}_n)$ such that*

$$\Psi = FG \quad \textit{and} \quad \|\Psi\|_{L^{p'}(\boldsymbol{S}_1)} = \|F\|_{L^q(\boldsymbol{S}_2)}\|G\|_{L^2(\boldsymbol{S}_2)}.$$

Proof of Theorem 3.1. Clearly, we may assume that Ψ is a nonzero function. Suppose that $\mathcal{Q}$ and $\mathcal{R}$ satisfy the requirements of Sarason's theorem. Let h be a scalar outer function such that

$$|h(\zeta)| = \|\Psi(\zeta)\|_{\boldsymbol{S}_1^n}^{1/2-p'/2}, \quad \zeta \in \mathbb{T}. \tag{3.2}$$

Put

$$F = h\mathcal{Q} \quad \text{and} \quad G = \frac{1}{h}\mathcal{R}.$$

By Sarason's Theorem,

$$\begin{aligned}
\|F\|_{L^q(\boldsymbol{S}_2^n)}^q &= \int_{\mathbb{T}} |h(\zeta)|^q \|\mathcal{Q}(\zeta)\|_{\boldsymbol{S}_2^n}^q \, d\boldsymbol{m}(\zeta) \\
&= \int_{\mathbb{T}} \|\Psi(\zeta)\|_{\boldsymbol{S}_1^n}^{(1/2-p'/2)q+q/2} \, d\boldsymbol{m}(\zeta) \\
&= \int_{\mathbb{T}} \|\Psi(\zeta)\|_{\boldsymbol{S}_1^n}^{p'} \, d\boldsymbol{m}(\zeta).
\end{aligned}$$

Similarly,

$$\begin{aligned}
\|G\|_{L^2(\boldsymbol{S}_2^n)}^2 &= \int_{\mathbb{T}} |h(\zeta)|^{-2} \|\mathcal{R}(\zeta)\|_{\boldsymbol{S}_2^n}^2 \, d\boldsymbol{m}(\zeta) \\
&= \int_{\mathbb{T}} \|\Psi(\zeta)\|_{\boldsymbol{S}_1^n}^{p'-1+1} \, d\boldsymbol{m}(\zeta) \\
&= \int_{\mathbb{T}} \|\Psi(\zeta)\|_{\boldsymbol{S}_1^n}^{p'} \, d\boldsymbol{m}(\zeta).
\end{aligned}$$

It follows that

$$\|F\|_{L^q(\boldsymbol{S}_2^n)}\|G\|_{L^2(\boldsymbol{S}_2^n)} = \|\Psi\|_{L^{p'}(\boldsymbol{S}_1^n)}^{p'/q+p'/2} = \|\Psi\|_{L^{p'}(\boldsymbol{S}_1^n)}. \quad \blacksquare$$

We need a version of Theorem 3.1 in the case $\Psi(\zeta)$ has rank k for $\zeta \in \mathbb{T}$. The following result can be deduced from Sarason's theorem.

Theorem 3.2. *Let $1 \le k \le n$ and let Ψ be a function in $H^1(\mathbb{M}_n)$ such that*

$$\operatorname{rank}\Psi(\zeta) = k \quad \text{on a subset of } \mathbb{T} \text{ of positive measure} \tag{3.3}$$

Then there exist matrix functions $\mathcal{F} \in H^2(\mathbb{M}_{n,k})$ and $\mathcal{G} \in H^2(\mathbb{M}_{k,n})$ such that

$$\Psi = \mathcal{F}\mathcal{G} \quad \text{and} \quad \|\Psi(\zeta)\|_{\boldsymbol{S}_1^n} = \|\mathcal{F}(\zeta)\|_{\boldsymbol{S}_2^{n,k}} \|\mathcal{G}(\zeta)\|_{\boldsymbol{S}_2^{k,n}}, \quad \zeta \in \mathbb{T}. \tag{3.4}$$

Proof. Clearly, each minor of Ψ belongs to the Hardy class H^s for some $s > 0$. It follows now from the uniqueness theorem for Hardy classes that condition (3.3) is equivalent to the fact that $\operatorname{rank}\Psi(\zeta) = k$ almost everywhere on $\mathbb{T}$.

Let $\mathcal{Q}$ and $\mathcal{R}$ be $n \times n$ matrix functions satisfying the requirements of Sarason's theorem. Then

$$\|\Psi(\zeta)\|_{\boldsymbol{S}_1^n} = \|\mathcal{Q}(\zeta)\|_{\boldsymbol{S}_2^n} \|\mathcal{R}(\zeta)\|_{\boldsymbol{S}_2^n}. \tag{3.5}$$

We need the following elementary lemma whose proof is given here for completeness.

Lemma 3.3. *If A and B are operators on Hilbert space,* $\operatorname{rank} AB = k$, *and* $\|AB\|_{\boldsymbol{S}_1} = \|A\|_{\boldsymbol{S}_2} \|B\|_{\boldsymbol{S}_2}$, *then* $\operatorname{rank} A = \operatorname{rank} B = k$.

Let us first complete the proof of Theorem 3.2.

By Lemma 3.3, (3.5) implies that

$$\operatorname{rank}\mathcal{Q}(\zeta) = k \quad \text{and} \quad \operatorname{rank}\mathcal{R}(\zeta) = k \quad \text{for almost all} \quad \zeta \in \mathbb{T}. \tag{3.6}$$

Consider the inner-outer factorization of $\mathcal{R}$:

$$\mathcal{R} = \Upsilon\mathcal{G},$$

where Υ is an inner matrix function and $\mathcal{G}$ is an outer matrix function. It follows from (3.6) that Υ has size $n \times k$ and $\mathcal{G}$ has size $k \times n$. We can define now the function $\mathcal{F}$ by $\mathcal{F} = \mathcal{Q}\Upsilon$. Since Υ takes isometric values almost everywhere on $\mathbb{T}$, it follows that

$$\|\mathcal{F}(\zeta)\|_{\boldsymbol{S}_2^{n,k}} = \|\mathcal{Q}(\zeta)\|_{\boldsymbol{S}_2^n} \quad \text{and} \quad \|\mathcal{G}(\zeta)\|_{\boldsymbol{S}_2^{k,n}} = \|\mathcal{R}(\zeta)\|_{\boldsymbol{S}_2^n},$$

and so (3.4) holds. ■

Proof of Lemma 3.3. Clearly, if $\operatorname{rank} A < k$ or $\operatorname{rank} B < k$, then $\operatorname{rank} AB < k$. Suppose now that the conclusion of the lemma is false. Without loss of generality we may assume that $\operatorname{rank} A > k$. Let P be the orthogonal projection onto $\operatorname{Range} AB$. Then $AB = PAB$. Clearly, $\|A\|_{\boldsymbol{S}_2}^2 = \|PA\|_{\boldsymbol{S}_2}^2 + \|(I-P)A\|_{\boldsymbol{S}_2}^2$. Since $\operatorname{rank} P = k$ and $\operatorname{rank} A > k$, it follows that $\|PA\|_{\boldsymbol{S}_2} < \|A\|_{\boldsymbol{S}_2}$. Thus

$$\|AB\|_{\boldsymbol{S}_1} = \|PAB\|_{\boldsymbol{S}_1} \le \|PA\|_{\boldsymbol{S}_2}\|B\|_{\boldsymbol{S}_2} < \|A\|_{\boldsymbol{S}_2}\|B\|_{\boldsymbol{S}_2} = \|AB\|_{\boldsymbol{S}_1}$$

and we get a contradiction. ■

We need the following consequence of Theorem 3.2.

Theorem 3.4. *Let $\Psi \in H^{p'}(\mathbb{M}_n)$ such that*

$$\operatorname{rank}\Psi(\zeta) = k, \quad \zeta \in \mathbb{T}.$$

Then there exist matrix functions $F \in H^q(\mathbb{M}_{n,k})$ and $G \in H^2(\mathbb{M}_{k,n})$ such that

$$\Psi = FG \quad \textit{and} \quad \|\Psi\|_{L^{p'}(\boldsymbol{S}_1^n)} = \|F\|_{L^q(\boldsymbol{S}_2^{n,k})}\|G\|_{L^2(\boldsymbol{S}_2^{k,n})}.$$

Proof. As in the proof of Theorem 3.1, we put

$$F = h\mathcal{F}, \quad G = \frac{1}{h}\mathcal{G},$$

where $\mathcal{F}$ and $\mathcal{G}$ are matrix functions satisfying the requirements of Theorem 3.2 and h is a scalar outer function satisfying (3.2). The fact that F and G satisfy the conclusions of the theorem is exactly the same as in the proof of Theorem 3.2. ■

The case of matrix functions of rank 1 is of special interest. We treat this case separately, without using Sarason's theorem.

Lemma 3.5. *Let $\Psi \in H^1(\mathbb{M}_{m,n})$ such that*

$$\operatorname{rank}\Psi(\zeta) = 1 \quad \textit{on a subset of } \mathbb{T} \textit{ of positive measure} \tag{3.7}$$

Then there exist vector functions $u \in H^2(\mathbb{C}^m)$, and $v \in H^2(\mathbb{C}^n)$, such that

$$\Psi(\zeta) = u(\zeta)v^{\mathrm{t}}(\zeta), \quad \zeta \in \mathbb{T}. \tag{3.8}$$

and

$$\|u(\zeta)\|_{\mathbb{C}^m} = \|v(\zeta)\|_{\mathbb{C}^n} = \|\Psi(\zeta)\|^{1/2}_{\mathbb{M}_{m,n}}. \tag{3.9}$$

Proof. Condition (3.7) means that each 2×2 minor of Ψ vanishes on a set of positive measure. Since $\Psi \in H^1(\mathbb{M}_{m,n})$, it follows that all 2×2 minors of Ψ are identically equal to zero. Thus condition (3.7) implies that $\operatorname{rank}\Psi(\zeta) = 1$ almost everywhere on $\mathbb{T}$.

Let h be an outer function such that

$$|h(\zeta)|^2 = \|\Psi(\zeta)\|_{L^1(\mathbb{M}_{m,n})}, \quad \zeta \in \mathbb{T},$$

and let $G = h^{-1}\Psi$. Clearly, $G \in H^2(\mathbb{M}_{m,n})$. Consider the columns of G. Let $\mathcal{L}$ be the invariant subspace of multiplication by z on $H^2(\mathbb{C}^m)$ spanned by the columns of G. By the Beurling–Lax theorem (see [N]), there exists an inner function Υ of size $m \times k$ such that $\mathcal{L} = \Upsilon H^2(\mathbb{C}^k)$. Since $\operatorname{rank} G(\zeta) = 1$ almost everywhere, it follows that $k = 1$. Then there exist

functions $v_1, v_2, \cdots, v_n$ such that the columns of the matrix function G are $v_1\Upsilon, v_2\Upsilon, \cdots, v_n\Upsilon$. Let

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}.$$

Clearly, $v \in H^2(\mathbb{C}^n)$ and $G = \Upsilon v^{\mathrm{t}}$. It remains to put $u = h\Upsilon$ and observe that $u \in H^2(\mathbb{C}^m)$ and both (3.8) and (3.9) hold. ■

Theorem 3.6. *Let Ψ be a rank one matrix function in $H^{p'}(\mathbb{M}_{m,n})$. Then there exist column functions $f \in H^q(\mathbb{C}^m)$ and $g \in H^2(\mathbb{C}^n)$ such that*

$$\Psi = fg^{\mathrm{t}} \quad \textit{and} \quad \|\Psi\|_{L^{p'}(\mathbb{M}_{m,n})} = \|f\|_{L^q(\mathbb{C}^m)}\|g\|_{L^2(\mathbb{C}^n)}. \tag{3.10}$$

Proof. Let u and v be the column functions satisfying (3.8) and (3.9). Let h be a scalar outer function satisfying (3.2). Put

$$f = hu \quad \text{and} \quad g = \frac{1}{h}v.$$

It is easy to verify that $f \in H^q(\mathbb{C}^m)$, $g \in H^2(\mathbb{C}^n)$, and the equalities in (3.10) hold. ■

4. **Respectable matrix functions**

The main result of this section is Theorem 4.4, which gives us several characterizations of the set of matrix functions $\Phi \in L^p(\mathbb{M}_{m,n})$, for which $\operatorname{dist}_{L^p}\big(\Phi, H^p(\mathbb{M}_{m,n})\big)$ is equal to the norm of the Hankel operator $H_\Phi\colon H^q(\mathbb{C}^n) \to H^2_-(\mathbb{C}^m)$. The description of this class of matrix functions (such matrix functions will be called *respectable*) makes it very natural to hope that all matrix functions in $L^p(\mathbb{M}_{m,n})$ are respectable. However, it will be shown in § 5 that this is not true.

Definition. A matrix function $\Phi \in L^p(\mathbb{M}_{m,n}) \setminus H^p(\mathbb{M}_{m,n})$ is called *regularly approximable* if there exists a best approximant $Q \in H^p(\mathbb{M}_{m,n})$ such that the space of maximizing vectors of $(\Phi - Q)(\zeta)$ is one-dimensional on a subset of $\mathbb{T}$ of positive measure.

It follows from the Hahn–Banach theorem that for $\Phi \in L^p(\mathbb{M}_{m,n})$,

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

where the supremum is taken over all $\Psi \in H^{p'}_0(\mathbb{M}_{n,m})$ (i.e., $\Psi \in H^{p'}(\mathbb{M}_{n,m})$ and $\Psi(0) = 0$) such that $\|\Psi\|_{L^{p'}(\boldsymbol{S}_1^{n,m})} \le 1$.

Since the space $L^p(\mathbb{M}_{m,n})$ is reflexive, it follows that for a matrix function $\Phi \in L^p(\mathbb{M}_{m,n}) \setminus H^p(\mathbb{M}_{m,n})$ there exists a matrix function $\Psi \in H_0^{p'}(\mathbb{M}_{n,m})$ such that

(4.1)

$$\|\Psi\|_{L^{p'}(\boldsymbol{S}_1^{n,m})} = 1 \text{ and } \int_{\mathbb{T}} \operatorname{trace}\big(\Phi(\zeta)\Psi(\zeta)\big)\, d\boldsymbol{m}(\zeta) = \operatorname{dist}_{L^p}\big(\Phi, H^p(\mathbb{M}_{m,n})\big).$$

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

Recall that for a matrix function $\Phi \in L^p(\mathbb{M}_{m,n})$, we consider the Hankel operator $H_\Phi\colon H^q(\mathbb{C}^n) \to H^2_-(\mathbb{C}^m)$ defined by

$$H_\Phi f = \mathbb{P}_- \Phi f, \quad f \in H^q(\mathbb{C}^n),$$

where $1/p + 1/q = 1/2$.

As we have mentioned in § 2, for Hankel operators with scalar symbols, formula (2.1) holds. Thus it is easy to see that the norm of the Hankel operator $H_\Phi\colon H^q(\mathbb{C}^n) \to H^2_-(\mathbb{C}^m)$ is equivalent to the distance in L^p from Φ to $H^p(\mathbb{M}_{m,n})$. Since in the case of scalar symbols all bounded Hankel operators from H^q to H^2_- are compact, we can obtain the following result.

Lemma 4.1. *For an arbitrary matrix function Φ in $L^p(\mathbb{M}_{m,n})$, the Hankel operator $H_\Phi\colon H^q(\mathbb{C}^n) \to H^2_-(\mathbb{C}^m)$ is compact.*

Corollary 4.2. *Let $\Phi \in L^p(\mathbb{M}_{m,n})$. Then H_Φ has a maximizing vector in $H^q(\mathbb{C}^n)$.*

The following lemma gives us an upper estimate for the norm of H_Φ.

Lemma 4.3. *Let $\Phi \in L^p(\mathbb{M}_{m,n})$. Then*

$$\|H_\Phi\| \le \operatorname{dist}_{L^p}\big(\Phi, H^p(\mathbb{M}_{m,n})\big).$$

Proof. Since $H_{\Phi - Q} = H_\Phi$ for an arbitrary Q in $H^p(\mathbb{M}_{m,n})$, it suffices to prove the inequality

$$\|H_\Phi\| \le \|\Phi\|_{L^p(\mathbb{M}_{m,n})}, \quad \Phi \in L^p(\mathbb{M}_{m,n}).$$

Suppose that $f \in H^q(\mathbb{C}^n)$ and $g \in H^2_-(\mathbb{C}^m)$. We have by Hölder's inequality,

$$\begin{aligned}
|(H_\Phi f, g)| = |(\Phi f, g)| &\le \int_{\mathbb{T}} |\Phi f g^*|\, d\boldsymbol{m} \\
&\le \left(\int_{\mathbb{T}} \|\Phi(\zeta)\|^p_{\mathbb{M}_{m,n}}\, d\boldsymbol{m}(\zeta)\right)^{1/p} \left(\int_{\mathbb{T}} \|f(\zeta)\|^q_{\mathbb{C}^n}\right)^{1/q} \left(\int_{\mathbb{T}} \|g(\zeta)\|^2_{\mathbb{C}^m}\right)^{1/2} \\
&= \|\Phi\|_{L^p(\mathbb{M}_{m,n})} \|f\|_{L^q(\mathbb{C}^n)} \|g\|_{L^2(\mathbb{C}^n)}. \quad \blacksquare
\end{aligned}$$

The following theorem gives us several characterizations of the class of matrix functions Φ, for which $\|H_\Phi\| = \operatorname{dist}_{L^p}\big(\Phi, H^p(\mathbb{M}_{m,n})\big)$.

Theorem 4.4. *Let $\Phi \in L^p(\mathbb{M}_{m,n}) \setminus H^p(\mathbb{M}_{m,n})$. The following are equivalent:*

(i) $\|H_\Phi\| = \operatorname{dist}_{L^p}\big(\Phi, H^p(\mathbb{M}_{m,n})\big)$;

(ii) Φ *belongs to the closure of the set of regularly approximable functions in* L^p;

(iii) Φ *has a dual extremal function* Ψ *such that* $\operatorname{rank}\Psi(\zeta) = 1$ *on a set of positive measure;*

(iv) Φ *has a dual extremal function* Ψ *such that* $\operatorname{rank}\Psi(\zeta) = 1$, $\zeta \in \mathbb{T}$;

(v) *if Q is a best approximant to Φ, then $\Phi - Q$ admits a factorization*

$$\Phi - Q = W^* \begin{pmatrix} \bar{z}\bar{\vartheta}\bar{h}/h^{2/q} & \mathbf{0} \\ \mathbf{0} & \Phi_\# \end{pmatrix} V^*, \tag{4.2}$$

where V and W^{t} are thematic matrix functions, ϑ is a scalar inner function, h is a scalar outer function in H^2, and $\Phi_\#$ is an $(m-1)\times(n-1)$ matrix function such that $\|\Phi_\#(\zeta)\|_{\mathbb{M}_{m-1,n-1}} \le |h(\zeta)|^{2/p}$, $\zeta \in \mathbb{T}$.

Note that in (4.2) the outer function h must satisfy the equality

$$|h(\zeta)|^{2/p} = \|(\Phi - Q)(\zeta)\|_{\mathbb{M}_{m,n}}, \quad \zeta \in \mathbb{T}.$$

Remark. Since the set of matrices, for which the space of maximizing vectors is one-dimensional is dense in the space of matrices, this suggests a hope that the set of regularly approximable $m \times n$ matrix functions is dense in $L^p(\mathbb{M}_{m,n})$. If this were true, then the distance formula $\|H_\Phi\| = \operatorname{dist}_{L^p}\big(\Phi, H^p(\mathbb{M}_{m,n})\big)$ would hold for an arbitrary matrix functions in $L^p(\mathbb{M}_{m,n})$. *Surprisingly, we will show in § 5 that this is not the case.*

Definition. Matrix functions $\Phi \in L^p(\mathbb{M}_{m,n}) \setminus H^p(\mathbb{M}_{m,n})$ satisfying one of the conditions (i)–(v) in the statement of Theorem 4.4 are called *respectable matrix functions.* If a matrix function $\Phi \in L^p(\mathbb{M}_{m,n}) \setminus H^p(\mathbb{M}_{m,n})$ is not respectable, it is called a *weird function.*

It follows immediately from Theorem 4.4 that the set of respectable functions is closed in L^p, while the set of weird functions is open.

Proof of Theorem 4.4. We start with the proof of the implication (iv)⇒(i). Let Ψ be a dual extremal function such that $\operatorname{rank}\Psi(\zeta) = 1$, $\zeta \in \mathbb{T}$. Then Ψ satisfies (4.1).

Since $\|H_\Phi\|$ is always less than or equal to $\operatorname{dist}_{L^p}\big(\Phi, H^p(\mathbb{M}_{m,n})\big)$, we have to show that

$$\|H_\Phi\| \ge \operatorname{dist}_{L^p}\big(\Phi, H^p(\mathbb{M}_{m,n})\big).$$

By Theorem 3.6, there exist functions $f \in H^2(\mathbb{C}^n)$ and $g \in H^2_0(\mathbb{C}^m)$ such that

$$\Psi = fg^{\mathrm{t}} \quad \text{and} \quad 1 = \|\Psi\|_{L^{p'}(\mathbb{M}_{m,n})} = \|f\|_{L^q(\mathbb{C}^m)}\|g\|_{L^2(\mathbb{C}^n)}.$$

Without loss of generality we may assume that $\|f\|_{L^q(\mathbb{C}^n)} = \|g\|_{L^2(\mathbb{C}^m)} = 1$. We have

$$\begin{aligned}\|H_\Phi\| \ge |(H_\Phi f, g^*)| &= \left|\int_{\mathbb{T}} \operatorname{trace}\left((H_\Phi f)g^{\mathrm{t}}\right) d\boldsymbol{m}\right| \\ &= \left|\int_{\mathbb{T}} \operatorname{trace}(\Phi f g^{\mathrm{t}})\, d\boldsymbol{m}\right| = \int_{\mathbb{T}} \operatorname{trace}\left(\Phi\Psi\right) d\boldsymbol{m} \\ &= \operatorname{dist}_{L^p}\left(\Phi, H^p(\mathbb{M}_{m,n})\right).\end{aligned}$$

Next, let us show that (i)⇒(v). Let $f \in H^q(\mathbb{C}^n)$ be a maximizing vector of H_Φ and let $Q \in H^p(\mathbb{M}_{m,n})$ be a best approximant to Φ. We have

$$\begin{aligned}\|H_\Phi f\|_{L^2(\mathbb{C}^m)} &= \|H_{\Phi-Q} f\|_{L^2(\mathbb{C}^m)} = \|\mathbb{P}_-(\Phi - Q)f\|_{L^2(\mathbb{C}^m)} \\ &\le \|(\Phi - Q)f\|_{L^2(\mathbb{C}^m)} \le \|(\Phi - Q)\|_{L^p(\mathbb{M}_{m,n})}\|f\|_{H^q(\mathbb{C}^n)} \\ &= \|H_\Phi\| \cdot \|f\|_{H^q(\mathbb{C}^n)} = \|H_\Phi f\|_{L^2(\mathbb{C}^m)}.\end{aligned}$$

Hence, both inequalities are equalities. The fact that the first inequality turns into equality means that $(\Phi - Q)f \in H^2_-(\mathbb{C}^m)$. The fact that the second inequality turns into equality means that $f(\zeta)$ is a maximizing vector of $(\Phi - Q)(\zeta)$ for almost all $\zeta \in \mathbb{T}$ and

$$\int_{\mathbb{T}} \left(\|(\Phi - Q)(\zeta)\|_{\mathbb{M}_{m,n}}\|f(\zeta)\|_{\mathbb{C}^n}\right)^2 d\boldsymbol{m}(\zeta) =$$

$$\left(\int_{\mathbb{T}} \|(\Phi - Q)(\zeta)\|^p_{\mathbb{M}_{m,n}} d\boldsymbol{m}(\zeta)\right)^{2/p} \left(\int_{\mathbb{T}} \|f(\zeta)\|^q_{\mathbb{C}^n} d\boldsymbol{m}(\zeta)\right)^{2/q},$$

i.e., the corresponding Hölder inequality turns into equality, which implies that $\|(\Phi - Q)(\zeta)\|_{\mathbb{M}_{m,n}} = c\|f(\zeta)\|^{q/p}_{\mathbb{C}^n}$ for some constant c. Since

$$\|(H_\Phi f)(\zeta)\|_{\mathbb{C}^m} = \|(\Phi - Q)(\zeta)\|_{\mathbb{M}_{m,n}}\|f(\zeta)\|_{\mathbb{C}^n},$$

it follows that $\|(H_\Phi f)(\zeta)\|_{\mathbb{C}^m} = c\|f(\zeta)\|^{q/2}_{\mathbb{C}^n}$. Multiplying the maximizing vector f by a suitable constant, one can always make the constant c equal to 1, and so we may assume that

$$\|(H_\Phi f)(\zeta)\|_{\mathbb{C}^m} = \|f(\zeta)\|^{q/2}_{\mathbb{C}^n}.$$

Let h be a scalar outer function such that

$$|h(\zeta)| = \|(H_\Phi f)(\zeta)\|_{\mathbb{C}^m}, \quad \zeta \in \mathbb{T},$$

and so

$$\|(\Phi - Q)(\zeta)\|_{\mathbb{M}_{m,n}} = |h(\zeta)|^{2/p}, \quad \zeta \in \mathbb{T}.$$

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

$$\|g(\zeta)\|_{\mathbb{C}^m} = |h(\zeta)| \quad \text{and} \quad \|f(\zeta)\|_{\mathbb{C}^n} = |h(\zeta)|^{2/q}, \quad \zeta \in \mathbb{T}.$$

The vector function f admits a factorization $f = \vartheta_1 h^{2/q}\boldsymbol{v}$, where ϑ_1 is a scalar inner function and $\boldsymbol{v}$ is an $n \times 1$ inner and co-outer function, while the vector function g admits a factorization $g = \vartheta_2 h\boldsymbol{w}$, where ϑ_2 is a scalar inner function and $\boldsymbol{w}$ is an $m \times 1$ inner and co-outer function.

Let now

$$V = \begin{pmatrix} \boldsymbol{v} & \overline{\Theta} \end{pmatrix} \quad \text{and} \quad W^{\mathrm{t}} = \begin{pmatrix} \boldsymbol{w} & \overline{\Xi} \end{pmatrix} \tag{4.3}$$

be thematic matrix functions (see § 2.3).

Consider the matrix function $W(\Phi - Q)V$. Its upper left entry is equal to

$$\xi = \boldsymbol{w}^{\mathrm{t}}(\Phi - Q)\boldsymbol{v} = \bar{\vartheta}_2 h^{-1} g^{\mathrm{t}}(\Phi - Q)\bar{\vartheta}_1 h^{-2/q} f = \bar{\vartheta}_1\bar{\vartheta}_2 h^{-2/q-1} g^{\mathrm{t}} H_\Phi f$$

$$= \bar{z}\bar{\vartheta} h^{-2/q-1} g^{\mathrm{t}}\bar{g} = \bar{z}\bar{\vartheta} h^{-2/q-1}|h|^2 = \bar{z}\bar{\vartheta}\frac{\bar{h}}{h^{2/q}} = \bar{z}\bar{\vartheta}\frac{\bar{h}}{h^{\frac{p-2}{p}}},$$

where $\vartheta = \vartheta_1\vartheta_2$.

We have $|\xi(\zeta)| = \|(\Phi - Q)(\zeta)\|_{\mathbb{M}_{m,n}}$. Since both V and W are unitary-valued, it is easy to see that $\Phi - Q$ has the form (4.2).

To prove the implication (v)⇒(ii), we need the following lemma.

Lemma 4.5. *Suppose that Φ is a matrix function that admits a factorization*

$$\Phi = W^* \begin{pmatrix} \bar{z}\bar{\vartheta}\bar{h}/h^{2/q} & \mathbf{0} \\ \mathbf{0} & \Phi_\# \end{pmatrix} V^*, \tag{4.4}$$

where ϑ, h, $\Phi_\#$, V, and W are as in the statement of Theorem 4.4. *Then Φ is p-badly approximable.*

Proof. As we have already observed, for an arbitrary matrix function Φ in $L^p(\mathbb{M}_{m,n})$ the following inequalities hold:

$$\|H_\Phi\| \le \operatorname{dist}_{L^p}\big(\Phi, H^p(\mathbb{M}_{m,n})\big) \le \|\Phi\|_{H^p(\mathbb{M}_{m,n})}.$$

It suffices to prove that if Φ is as in (4.4), then $\|H_\Phi\| \ge \|\Phi\|_{H^p(\mathbb{M}_{m,n})}$. Consider the matrix functions V and W:

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

Let $f = h^{2/q}\boldsymbol{v}$. It is easy to verify that

$$H_\Phi f = \Phi f = \bar{z}\bar{h}\bar{\vartheta}\overline{\boldsymbol{w}} \quad \text{and} \quad \|H_\Phi f\|_{L^2(\mathbb{C}^m)} = \|\Phi\|_{L^p(\mathbb{M}_{m,n})}\|f\|_{L^q(\mathbb{C}^n)}$$

which implies that $\|H_\Phi\| \ge \|\Phi\|_{L^p(\mathbb{M}_{m,n})}$. ■

(v)⇒(ii). Let $R = \Phi - Q$. For $\varepsilon > 0$ we consider the function R_ε defined by

$$R_\varepsilon = W^* \begin{pmatrix} (1+\varepsilon)\bar{z}\bar{\vartheta}\bar{h}/h^{2/q} & \mathbf{0} \\ \mathbf{0} & \Phi_\# \end{pmatrix} V^*.$$

By Lemma 4.5, R and R_ε are p-badly approximable matrix functions. We define the function Φ_ε by $\Phi_\varepsilon = R_\varepsilon + Q$.

Since $Q \in H^p(\mathbb{M}_{m,n})$ and R_ε is p-badly approximable, it follows that Q is a p-best approximant to Φ_ε. Clearly, for $\zeta \in \mathbb{T}$, the space of maximizing vectors of $R_\varepsilon(\zeta)$ is one-dimensional, and so Φ_ε is a regularly approximable matrix function. The result follows from the obvious fact that

$$\|\Phi_\varepsilon - \Phi\|_{L^p} \to 0 \quad \text{as} \quad \varepsilon \to \infty.$$

To show that (iii)⇒(iv), we observe that (iii) implies that each 2×2 minor of Ψ vanishes on a set of positive measure. By the uniqueness theorem for the Hardy classes, it follows that all 2×2 minors of Ψ are zero almost everywhere on $\mathbb{T}$ which proves (iv).

Let us prove now that (ii)⇒(i). Clearly, it suffices to show that if Φ is regularly approximable, then $\|H_\Phi\| = \operatorname{dist}_{L^p}\big(\Phi, H^p(\mathbb{M}_{m,n})\big)$. Let Q be a matrix function in $H^p(\mathbb{M}_{m,n})$ such that the space of maximizing vectors of $(\Phi - Q)(\zeta)$ is one-dimensional on a subset of $\mathbb{T}$ of positive measure. Let Ψ be a dual extremal function of Φ. It follows easily from (4.1) that

$$\operatorname{trace}\big((\Phi - Q)(\zeta)\Psi(\zeta)\big) = \|(\Phi - Q)(\zeta)\|_{\mathbb{M}_{m,n}}\|\Psi(\zeta)\|_{\boldsymbol{S}_1^{n,m}}, \quad \zeta \in \mathbb{T}. \tag{4.5}$$

We need the following elementary lemma.

Lemma 4.6. *Let $A \in \mathbb{M}_{m,n}$ and $B \in \mathbb{M}_{n,m}$ be matrices satisfying*

$$|\operatorname{trace}(AB)| = \|A\|_{\mathbb{M}_{m,n}}\|B\|_{\boldsymbol{S}_1^{n,m}}.$$

Assume that the space of maximizing vectors A is one-dimensional. Then B has rank 1.

Proof. Without loss of generality we may assume that $m = n$. By considering the polar decomposition of B, we may assume that B is positive, i.e., $(Bx, x) \ge 0$ for every vector x. Let $e_1, \cdots, e_n$ be an orthonormal basis

of eigenvectors of B and let $Be_j = \lambda_j e_j$. We have

$$|\operatorname{trace}(AB)| = |\operatorname{trace}(BA)| = \left|\sum_{j=1}^{n}(BAe_j, e_j)\right| = \left|\sum_{j=1}^{n}(Ae_j, Be_j)\right|$$

$$= \left|\sum_{j=1}^{n}\lambda_j(Ae_j, e_j)\right| \le \sum_{j=1}^{n}\lambda_j\|Ae_j\| \le \|A\|\sum_{j=1}^{n}\lambda_j.$$

On the other hand,

$$\|A\|\cdot\|B\|_{\boldsymbol{S}_1^n} = \|A\|\cdot\sum_{j=1}^{n}(Be_j, e_j) = \|A\|\sum_{j=1}^{n}\lambda_j.$$

It follows that if $\|Ae_j\| < \|A\|$, then $\lambda_j = 0$. By the hypotheses there can be only one j, for which $\|Ae_j\| = \|A\|$, which proves the result. ■

It follows from (4.5) and from Lemma 4.6 that Φ satisfies (iii). Since we have already proved that (iii)⇒(iv) and (iv)⇒(i), it follows that Φ satisfies (i).

The fact that (iv)⇒(iii) is obvious. It remains to prove that (v)⇒(iv). Suppose that $\Phi - Q$ is factorized as in (4.2). Without loss of generality we may assume that $\|\Phi - Q\|_{L^p} = 1$ Define the matrix function Ψ by

$$\Psi = z\vartheta h^{1+2/q}\begin{pmatrix}\boldsymbol{v} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\boldsymbol{w}^{\mathrm{t}} \\ \mathbf{0}\end{pmatrix},$$

where $\boldsymbol{v}$ and $\boldsymbol{w}$ are as in (4.3).

Clearly, $\operatorname{rank}\Psi(\zeta) = 1$, $\zeta \in \mathbb{T}$. We have

$$\|\Psi\|^{p'}_{L^{p'}(\boldsymbol{S}_1^{n,m})} = \int_{\mathbb{T}}|h(\zeta)|^{p'(1+2/q)}\,d\boldsymbol{m} = \|h\|^2_{L^2} = 1$$

and

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

$$= \int_{\mathbb{T}}\operatorname{trace}\begin{pmatrix}|h|^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0}\end{pmatrix}d\boldsymbol{m} = \|h\|^2_{L^2} = \|\Phi - Q\|_{L^p} = 1.$$

This completes the proof. ■

Remark. Note that in the case of analytic matrix approximation in the L^∞ norm it is not true that for an arbitrary matrix function $\Phi \in L^\infty(\mathbb{M}_{m,n})$ there exists a dual extremal function in $H^1_0(\boldsymbol{S}_1^{n,m})$. Moreover,

it was shown in [Pe2] that a dual extremal function exists if and only if the Hankel operator $H_\Phi\colon H^2(\mathbb{C}^n) \to H^2_-(\mathbb{C}^m)$ has a maximizing vector.

However, in the case $p = \infty$, if a dual extremal function exists, then there exists a dual extremal function Ψ such that $\operatorname{rank}\Psi(\zeta) = 1$ almost everywhere on $\mathbb{T}$, see [Pe2].

5. **Weird matrix functions**

The main result of this section is a construction of a weird matrix function of size 2×2.

Lemma 5.1. *There exists a bounded 2×2 matrix function B such that $B^* = B$, $\operatorname{trace} B(\zeta) = 1$, $\zeta \in \mathbb{T}$, the eigenvalues of $B(\zeta)$, $\zeta \in \mathbb{T}$, are positive and separated away from zero and there is no constant self-adjoint matrix C such that*

$$\operatorname{rank} C = 1 \quad \textit{and} \quad \operatorname{trace} B(\zeta)C = 1, \quad \zeta \in \mathbb{T}.$$

Proof. Let α be a real bounded scalar functions, β a complex scalar bounded function such that the functions α, β, $\bar\beta$, and $\mathbf{1}$ are linearly independent, and the function $\alpha(1-\alpha) - |\beta|^2$ is positive and separated away from zero. Put

$$B = \begin{pmatrix} \alpha & \beta \\ \bar\beta & 1-\alpha \end{pmatrix}.$$

Clearly, $B^* = B$, the eigenvalues of $B(\zeta)$, $\zeta \in \mathbb{T}$, are positive and separated away from zero, and $\operatorname{trace} B(\zeta) = 1$, $\zeta \in \mathbb{T}$. Suppose that C is a self-adjoint constant matrix such that $\operatorname{rank} C = 1$, and $\operatorname{trace} B(\zeta)C = 1$, $\zeta \in \mathbb{T}$. Then C has the form

$$C = \begin{pmatrix} a & b \\ \bar b & a^{-1}|b|^2 \end{pmatrix},$$

where a is a nonzero real number and b is a complex number. We have

$$\operatorname{trace} B(\zeta)C = a\alpha(\zeta) + b\bar\beta(\zeta) + \bar b\beta(\zeta) + a^{-1}|b|^2(1-\alpha(\zeta)) = 1, \quad \zeta \in \mathbb{T}.$$

Thus

$$(a - a^{-1}|b|^2)\alpha(\zeta) + b\bar\beta(\zeta) + \bar b\beta(\zeta) + a^{-1}|b|^2 - 1 = 0.$$

Since the functions α, β, $\bar\beta$, and $\mathbf{1}$ are linearly independent, this equality is impossible. ■

Consider the Wiener–Masani factorization of B (see [WM]):

$$B = \Psi^*\Psi, \tag{5.1}$$

where Ψ is an invertible bounded analytic function in $\mathbb{D}$. Put

$$A = \Psi\Psi^*$$

and consider the Wiener–Masani factorization of A^2:

$$A^2 = QQ^*.$$

Let U be the matrix function defined by

$$U = \bar{z}Q^{-1}A. \tag{5.2}$$

Then U is a unitary-valued function on $\mathbb{T}$:

$$U^*U = A(Q^*)^{-1}Q^{-1}A = I.$$

Clearly,

$$AU^{-1} = zQ \in H_0^\infty(\mathbb{M}_{2,2}).$$

Let us show that U is p-badly approximable.

Lemma 5.2. *Let A be a self-adjoint 2×2 matrix function such that*

$$\operatorname{trace} A(\zeta) = 1, \quad \zeta \in \mathbb{T},$$

and the eigenvalues of $A(\zeta)$ are positive and separated away from zero. Suppose that U is a unitary-valued matrix function on $\mathbb{T}$ such that $AU^{-1} \in H_0^\infty(\mathbb{M}_{2,2})$. Then U is a p-badly approximable matrix function.

Proof. Let $F \in H^p(\mathbb{M}_{2,2})$. For $\zeta \in \mathbb{T}$, we have

$$\left|\operatorname{trace}\big((U-F)AU^*)(\zeta)\big)\right| \le \|(U-F)AU^*)(\zeta)\|_{\boldsymbol{S}_1} \le \|(U-F)(\zeta)\|_{\mathbb{M}_{2,2}}.$$

Thus by Hölder's inequality,

$$\begin{aligned}
\|U-F\|_{L^p(\mathbb{M}_{2,2})} &= \left(\int_{\mathbb{T}} \|U-F\|^p \, d\boldsymbol{m}\right)^{1/p} \\
&\ge \left(\int_{\mathbb{T}} \left|\operatorname{trace}\big((U-F)AU^*\big)\right|^p \, d\boldsymbol{m}\right)^{1/p} \\
&\ge \int_{\mathbb{T}} \left|\operatorname{trace}\big((U-F)AU^*\big)\right| \, d\boldsymbol{m} \\
&\ge \left|\int_{\mathbb{T}} \operatorname{trace}\big((U-F)AU^*\big)\, d\boldsymbol{m}\right| \\
&= \left|\int_{\mathbb{T}} \operatorname{trace}(UAU^*)\, d\boldsymbol{m} - \int_{\mathbb{T}} \operatorname{trace}(FAU^*)\, d\boldsymbol{m}\right| \\
&= \left|\int_{\mathbb{T}} \operatorname{trace}(UAU^*)\, d\boldsymbol{m}\right| = \left|\int_{\mathbb{T}} \operatorname{trace} A \, d\boldsymbol{m}\right| = 1.
\end{aligned}$$

Note that

$$\int_{\mathbb{T}} \operatorname{trace}(FAU^*)\, d\boldsymbol{m} = 0,$$

since $FAU^* \in H^p_0(\mathbb{M}_{2,2})$. Thus U is p-badly approximable. ■

To prove that the matrix function U defined by (5.2) is weird, we need the following lemma.

Lemma 5.3. *Let A be a bounded positive definite matrix function on $\mathbb{T}$ whose inverse is also bounded and let $A = \Psi\Psi^*$, where Ψ is an invertible matrix function in H^∞. A matrix function F in H^∞ satisfies the equation*

$$AF^* = FA \tag{5.3}$$

if and only if

$$F = \Psi C \Psi^{-1}. \tag{5.4}$$

where C is a constant self-adjoint matrix.

Proof. Put

$$C = \Psi^{-1} F \Psi.$$

Then C is an H^∞ matrix function and (5.4) holds. By (5.3), we have

$$AF^* = A(\Psi^*)^{-1} C^* \Psi^* = FA = \Psi C \Psi^{-1} A.$$

Since $A = \Psi\Psi^*$, we obtain

$$\Psi\Psi^*(\Psi^*)^{-1} C^* \Psi^* = \Psi C \Psi^{-1} \Psi\Psi^*$$

which implies $C = C^*$. Since C is an H^∞ matrix function, it must be constant.

Clearly, if C is a constant self-adjoint matrix and F is defined by (5.4), then F satisfies equation (5.3). ■

Theorem 5.4. *The matrix function U defined by* (5.2) *is a weird p-badly approximable function.*

Proof. Assume that U is respectable. By Lemma 5.2, U is p-badly approximable. Then $\|H_U\| = \|U\|_{L^p(\mathbb{M}_{2,2})} = 1$.

Let $f \in H^q(\mathbb{C}^2)$ be a maximizing vector of H_U of norm 1. We have

$$1 = \|H_U f\|_{L^2(\mathbb{C}^2)} = \|\mathbb{P}_- U f\|_{L^2(\mathbb{C}^2)}$$

$$\le \|Uf\|_{L^2(\mathbb{C}^2)} = \|f\|_{L^2(\mathbb{C}^2)} \le \|f\|_{L^q(\mathbb{C}^2)} = 1.$$

Thus all inequalities in this chain of inequalities are equalities. The equality $\|f\|_{L^2(\mathbb{C}^2)} = \|f\|_{L^q(\mathbb{C}^2)}$ means that $\|f(\zeta)\|_{\mathbb{C}^2} = 1$, $\zeta \in \mathbb{T}$, while the equality $\|\mathbb{P}_- Uf\|_{L^2(\mathbb{C}^2)} = \|Uf\|_{L^2(\mathbb{C}^2)}$ means that $Uf \in H^2_-(\mathbb{C}^2)$, and so $Uff^* \in H^\infty_-(\mathbb{M}_{2,2})$ or, in other words, $ff^*U^{-1} \in H^\infty_0(\mathbb{M}_{2,2})$. Put

$$F = ff^* A^{-1}.$$

Then F satisfies (5.3). Hence, by Lemma 5.3, F has the form $F = \Psi C \Psi^{-1}$, where C is a constant self-adjoint matrix. Since F has rank one on $\mathbb{T}$, it follows that $\operatorname{rank} C = 1$. Clearly,

$$ff^* = FA = \Psi C \Psi^{-1} \Psi \Psi^* = \Psi C \Psi^*.$$

Let $B = \Psi^* \Psi$ be the matrix function obtained in Lemma 5.1. By (5.1), we have

$$\operatorname{trace} BC = \operatorname{trace} \Psi^* \Psi C = \operatorname{trace} \Psi C \Psi^* = \operatorname{trace} f f^* = 1.$$

This contradicts Lemma 5.1. ■

Remark. The results of Sections 4 and 5 show that the class $L^p(\mathbb{M}_{m,n})$ splits into two subsets. The first subset consists of respectable matrix functions and for respectable matrix functions Φ the distance $\operatorname{dist}_{L^p}\big(\Phi, H^p(\mathbb{M}_{m,n})\big)$ can be computed by formula (1.2). The second subset consists of weird matrix functions and for weird matrix functions Φ we have to find another formula to compute the distance $\operatorname{dist}_{L^p}\big(\Phi, H^p(\mathbb{M}_{m,n})\big)$. Such a formula will be obtained in the next section.

Let us explain that in a sense both the set of respectable matrix functions and the set of weird matrix functions are massive subsets of $L^p(\mathbb{M}_{m,n})$. First of all, the set of weird matrix functions is open and nonempty, as we have just seen.

Secondly, if $\Phi \in L^p(\mathbb{M}_{m,n})$ and Q is an arbitrary function in $H^\infty(\mathbb{M}_{m,n})$, then Φ is respectable if and only if $\Phi - Q$ is. Thus to characterize the set of respectable matrix functions, we can restrict ourselves to the case of p-badly approximable respectable matrix functions. It is easy to see that the set of respectable badly approximable matrix functions has nonempty interior in the set of p-badly approximable matrix functions. Indeed, it is easy to verify that the p badly approximable matrix function

$$\Phi = \begin{pmatrix} \bar{z} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \tag{5.5}$$

belongs to the interior of the set of respectable p-badly approximable functions.

However, we do not know whether the set of respectable matrix functions has nonempty interior in the space $L^p(\mathbb{M}_{m,n})$. In particular, we do not know whether the matrix function Φ defined in (5.5) belongs to the interior of the set of respectable matrix functions has nonempty interior in the space $L^p(\mathbb{M}_{m,n})$.

6. **Hankel operators on spaces of matrix-valued functions**

We have already mentioned in the introduction that the problem of analytic approximation of matrix functions can be reduced to the case of square matrix functions and beginning this section we assume that $\Phi \in L^p(\mathbb{M}_n)$.

For $\Phi \in L^p(\mathbb{M}_n)$, we consider the Hankel operator $\boldsymbol{H}_\Phi$ defined on the space $H^q(\boldsymbol{S}_2^n)$ to the space $H^2_-(\boldsymbol{S}_2^n)$ defined by

$$\boldsymbol{H}_\Phi F = \mathbb{P}_- \Phi F,$$

where $\mathbb{P}_-$ is an orthogonal projection from the space $L^2(\boldsymbol{S}_2^n)$ onto the subspace $H^2_-(\boldsymbol{S}_2^n) \stackrel{\text{def}}{=} L^2(\boldsymbol{S}_2^n) \ominus H^2(\boldsymbol{S}_2^n)$.

Theorem 6.1. *Let $\Phi \in L^p(\mathbb{M}_n)$. Then*

$$\|\boldsymbol{H}_\Phi\|_{H^q(\boldsymbol{S}_2^n)\to H^2_-(\boldsymbol{S}_2^n)} = \operatorname{dist}_{L^p}\big(\Phi, H^p(\mathbb{M}_n)\big).$$

Proof. Suppose that $\Phi \in L^p(\mathbb{M}_n)$. Then for $F \in H^q(\boldsymbol{S}_2^n)$ and $Q \in H^p(\mathbb{M}_n)$, we have

$$\|\boldsymbol{H}_\Phi F\|_{L^2(\boldsymbol{S}_2^n)} = \|\mathbb{P}_-\big((\Phi - Q)F\big)\|_{L^2(\boldsymbol{S}_2^n)} \le \|(\Phi - Q)F\|_{L^2(\boldsymbol{S}_2^n)}$$

$$\le \|\Phi - Q\|_{L^p(\mathbb{M}_n)} \|F\|_{H^q(\boldsymbol{S}_2^n)}$$

by Hölder's inequality. Thus $\|\boldsymbol{H}_\Phi\| \le \operatorname{dist}_{L^p}\big(\Phi, H^p(\mathbb{M}_n)\big)$.

To prove the opposite inequality, we are going to use Theorem 3.1 that has been deduced from Sarason's theorem. Let Ψ be a dual extremal function of Φ, i.e., Ψ belongs to $H^{p'}_0(\mathbb{M}_n)$ and satisfies (4.1). By Theorem 3.1, there exist matrix functions $F \in H^q(\mathbb{M}_n)$ and $G \in H^2_0(\mathbb{M}_n)$ such that

$$\Psi = FG \quad \text{and} \quad \|F\|_{L^q(\boldsymbol{S}_2^n)} \|G\|_{L^2(\boldsymbol{S}_2^n)} = 1.$$

Without loss of generality we may assume that $\|F\|_{L^q(\boldsymbol{S}_2^n)} = 1$ and $\|G\|_{L^2(\boldsymbol{S}_2^n)} = 1$. We have

$$\|\boldsymbol{H}_\Phi\| \ge |(\boldsymbol{H}_\Phi F, G^*)_{L^2(\boldsymbol{S}_2^n)}| = \left| \int_{\mathbb{T}} \operatorname{trace}\big((\boldsymbol{H}_\Phi F)G\big)\, d\boldsymbol{m} \right|$$

$$= \left| \int_{\mathbb{T}} \operatorname{trace}(\Phi F G)\, d\boldsymbol{m} \right| = \int_{\mathbb{T}} \operatorname{trace}\big(\Phi\Psi\big)\, d\boldsymbol{m}$$

$$= \operatorname{dist}_{L^p}\big(\Phi, H^p(\mathbb{M}_{m,n})\big)$$

by (4.1). ■

It follows immediately from Theorem 6.1 that

$$\|H_\Phi\|_{H^q(\mathbb{C}^n)\to H^2_-(\mathbb{C}^n)} \le \|\boldsymbol{H}_\Phi\|_{H^q(\boldsymbol{S}_2^n)\to H^2_-(\boldsymbol{S}_2^n)}.$$

Note that this inequality can also be obtained easily from the definitions of H_Φ and $\boldsymbol{H}_\Phi$.

Theorem 6.2. *Let $\Phi \in L^p(\mathbb{M}_n)$ and let Q be a best approximant to Φ in $H^p(\mathbb{M}_n)$. Then the following assertions hold:*

(i) *if $F \in H^q(\mathbb{M}_n)$ is a maximizing vector of $\boldsymbol{H}_\Phi$, then $(\Phi - Q)F \in H^2_-(\boldsymbol{S}_2^n)$;*

(ii) *the function*

$$\zeta \mapsto \|(\Phi - Q)(\zeta)\|_{\mathbb{M}_n} \tag{6.1}$$

does not depend on the choice of a best approximant Q;

(iii) *if $\boldsymbol{H}_\Phi$ has a maximizing vector F such that $\operatorname{rank} F(\zeta) = n$ on a subset of $\mathbb{T}$ of positive measure, then Φ has a unique best approximant in $H^p(\mathbb{M}_n)$;*

(iv) *if F_1 and F_2 are maximizing vectors of $\boldsymbol{H}_\Phi$, then*

$$\|F_1(\zeta)\|_{\boldsymbol{S}_2^n} = c\|F_2(\zeta))\|_{\boldsymbol{S}_2^n}$$

for some positive constant c;

(v) *if Q is a best approximant to Φ in $H^p(\mathbb{M}_n)$ and F is a maximizing vector of $\boldsymbol{H}_\Phi$, then the matrix*

$$\frac{1}{\|(\Phi - Q)(\zeta)\|_{\mathbb{M}_n}}(\Phi - Q)(\zeta), \quad \zeta \in \mathbb{T},$$

is isometric on the range of $F(\zeta)$.

Proof. Let us fix a maximizing vector F of H_Φ. We have by Hölder's inequality,

$$\begin{aligned} \|\boldsymbol{H}_\Phi F\|_{L^2(\boldsymbol{S}_2^n)} = \|\mathbb{P}_-\big((\Phi - Q)F\big)\|_{L^2(\boldsymbol{S}_2^n)} &\le \|(\Phi - Q)F\|_{L^2(\boldsymbol{S}_2^n)} \\ &\le \|\Phi - Q\|_{L^p(\mathbb{M}_n)}\|F\|_{H^q(\boldsymbol{S}_2^n)} \\ &= \|\boldsymbol{H}_\Phi\|_{H^q(\boldsymbol{S}_2^n)\to H^2_-(\boldsymbol{S}_2^n)}\|F\|_{H^q(\boldsymbol{S}_2^n)}. \end{aligned} \tag{6.2}$$

Since $\|\boldsymbol{H}_\Phi F\|_{L^2(\boldsymbol{S}_2^n)} = \|\boldsymbol{H}_\Phi\| \cdot \|F\|_{H^q(\boldsymbol{S}_2^n)}$, it follows that both inequalities in (6.2) are equalities.

The fact that the first inequality in (6.2) turns into equality means that $(\Phi - Q)F \in H^2_-(\boldsymbol{S}_2^n)$, i.e.,

$$(\Phi - Q)F = \boldsymbol{H}_\Phi F \tag{6.3}$$

which proves (i).

To prove (iii), we observe that since $F \in H^q(\boldsymbol{S}_2^n)$, it follows that if $\operatorname{rank} F(\zeta) = n$ on a set of positive measure, then $\operatorname{rank} F(\zeta) = n$, $\zeta \in \mathbb{T}$, almost everywhere on $\mathbb{T}$. Hence,

$$\Phi - Q = (\boldsymbol{H}_\Phi F)F^{-1},$$

and so Q is uniquely determined by Φ.

The fact that the second inequality in (6.2) turns into equality means that there exists $c > 0$ such that

$$\|(\Phi - Q)(\zeta)\|_{\mathbb{M}_n}^p = c\|F(\zeta)\|_{\boldsymbol{S}_2^n}^q, \quad \zeta \in \mathbb{T}, \tag{6.4}$$

and

$$\|(\Phi - Q)(\zeta)F(\zeta)\|_{\boldsymbol{S}_2^n} = \|(\Phi - Q)(\zeta)\|_{\mathbb{M}_n}\|F(\zeta)\|_{\boldsymbol{S}_2^n}, \quad \zeta \in \mathbb{T}. \tag{6.5}$$

Clearly, (iv) follows immediately from (6.4).

If we normalize the maximizing vector F by the condition

$$\|F\|_{L^q(\boldsymbol{S}_2)}^q = \|\boldsymbol{H}_\Phi\|^p, \tag{6.6}$$

then integrating (6.4), we obtain

$$\|\boldsymbol{H}_\Phi\|^p = \|\Phi - Q\|_{L^p(\mathbb{M}_n)}^p = c\|F\|_{L^q(\boldsymbol{S}_2)}^q.$$

Hence, under condition (6.6),

$$\|(\Phi - Q)(\zeta)\|_{\mathbb{M}_n}^p = \|F(\zeta)\|_{\boldsymbol{S}_2^n}^q, \quad \zeta \in \mathbb{T}, \tag{6.7}$$

and so the function (6.1) is uniquely determined by Φ. This proves (ii).

It remains to observe that (v) follows from (6.5) and from the fact that for $n \times n$ matrices A and B the equality

$$\|AB\|_{\boldsymbol{S}_2^n} = \|A\| \cdot \|B\|_{\boldsymbol{S}_2^n}$$

holds if and only if the restriction of A to the range of B is a multiple of an isometry. ■

Definition. For a function $\Phi \in L^p(\mathbb{M}_n)$, the function (6.1) is called the *distance function* of Φ. We denote the distance function of Φ by d_Φ:

$$\mathrm{d}_\Phi(\zeta) = \|(\Phi - Q)(\zeta)\|_{\mathbb{M}_n}, \quad \zeta \in \mathbb{T}, \tag{6.8}$$

where Q is an arbitrary best approximant to Φ.

The following result describes the set of all nonzero distance functions of matrix functions in $L^p(\mathbb{M}_n)$.

Theorem 6.3. *Let $d \geq 0$ be a nonzero function in L^p. Then d is the distance function of a matrix function $\Phi \in L^p(\mathbb{M}_n)$ if and only if $\log d \in L^1$.*

Proof. If $\Phi \in L^p(\mathbb{M}_n) \setminus H^p(\mathbb{M}_n)$ and Q is a p-best approximant to Φ and $d(\zeta) = \|(\Phi - Q)(\zeta)\|_{\mathbb{M}_n}$, the fact that $\log d \in L^1$ follows immediately from (6.4).

The converse follows from Corollary 2.2 by considering matrix functions of the form

$$\begin{pmatrix} \varphi & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{pmatrix}. \quad \blacksquare$$

Theorem 6.4. *$\Phi \in L^p(\mathbb{M}_n) \setminus H^p(\mathbb{M}_n)$. Then Φ is respectable if and only if there exists a maximizing vector F of $\boldsymbol{H}_\Phi$ such that*

$$\operatorname{rank} F(\zeta) = 1, \quad \zeta \in \mathbb{T}. \tag{6.9}$$

Proof. Suppose that Φ is respectable. Consider the Hankel operator

$$H_\Phi \colon H^q(\mathbb{C}^n) \to H^2_-(\mathbb{C}^n).$$

Let $f \in H^q(\mathbb{C}^n)$ be a maximizing vector of H_Φ. Define the matrix function $F \in H^q(\mathbb{M}_n)$ by

$$F = \begin{pmatrix} f & \mathbf{0} & \cdots & \mathbf{0} \end{pmatrix}.$$

It is obvious that $\operatorname{rank} F(\zeta) = 1$ for $\zeta \in \mathbb{T}$. Clearly,

$$\|F(\zeta)\|_{\boldsymbol{S}_2} = \|f(\zeta)\|_{\mathbb{C}^n} \quad \text{and} \quad \|(\boldsymbol{H}_\Phi F)(\zeta)\|_{\boldsymbol{S}_2^n} = \|(H_\Phi f)(\zeta)\|_{\mathbb{C}^n},$$

and so F is a maximizing vector of $\boldsymbol{H}_\Phi$.

To prove the converse, we may assume that $\|H_\Phi\| = 1$. Suppose that F is a maximizing vector of $\boldsymbol{H}_\Phi$ of norm 1 that satisfies (6.9). Let Q be a best approximant to Φ in $H^p(\mathbb{M}_n)$. By Theorem 6.2, we have $(\Phi - Q)F \in H^2_-(\mathbb{M}_n)$. Put

$$G = \frac{1}{\|\boldsymbol{H}_\Phi\|}\Big((\Phi - Q)F\Big)^* \quad \text{and} \quad \Psi = FG \in H^{p'}_0(\mathbb{M}_n).$$

Clearly,

$$\operatorname{rank} \Psi(\zeta) = 1, \quad \zeta \in \mathbb{T}, \quad \text{and} \quad \|\Psi\|_{L^{p'}(\boldsymbol{S}_1^n)} \le 1.$$

Let us show that Ψ is a dual extremal function of Φ. Assuming that (6.6) holds, we have by (6.7),

$$\begin{aligned} \int_{\mathbb{T}} \operatorname{trace}(\Phi\Psi)\, d\boldsymbol{m} &= \int_{\mathbb{T}} \operatorname{trace}\big((\Phi - Q)FG\big)\, d\boldsymbol{m} \\ &= \int_{\mathbb{T}} \operatorname{trace}\Big(\mathbb{P}_-\big((\Phi - Q)F\big)G\Big)\, d\boldsymbol{m} \\ &= \int_{\mathbb{T}} \operatorname{trace}\big((\boldsymbol{H}_\Phi F)G\big)\, d\boldsymbol{m} = (\boldsymbol{H}_\Phi F, G^*)_{L^2(\boldsymbol{S}_2^n)} \\ &= \frac{1}{\|\boldsymbol{H}_\Phi\|}\big\|\boldsymbol{H}_\Phi F\big\|^2_{L^2(\boldsymbol{S}_2^n)} = \|\boldsymbol{H}_\Phi\| = 1. \end{aligned}$$

Thus Ψ is a dual extremal function of rank 1, and so by Theorem 4.4, Φ is respectable. ■

Note that the computation, in fact, shows that $\|\Psi\|_{L^{p'}(\boldsymbol{S}_1^n)} = 1$.

Corollary 6.5. *Let Φ be a weird function in $L^p(\mathbb{M}_2) \setminus H^p(\mathbb{M}_2)$. Then Φ has a unique best approximant in $H^p(\mathbb{M}_2)$.*

Proof. By Theorem 6.2, if Φ has more than one best approximant, then each maximizing vector F of $\boldsymbol{H}_\Phi$ has rank 1 almost everywhere on $\mathbb{T}$. By Theorem 6.4, the function Φ is respectable. ■

We consider now for a function $\Phi \in L^p(\mathbb{M}_n)$, the family of Hankel operators $H_\Phi^{\{k\}} \colon H^q\big(\boldsymbol{S}_2^{n,k}\big) \to H^2_-\big(\boldsymbol{S}_2^{n,k}\big)$, $1 \le k \le n$, defined by

$$H_\Phi^{\{k\}} F = \mathbb{P}_- \Phi F, \quad F \in H^q(\boldsymbol{S}_2^{n,k}).$$

Clearly, $H_\Phi^{\{1\}} = H_\Phi$ and $H_\Phi^{\{n\}} = \boldsymbol{H}_\Phi$.

Theorem 6.6. *Let $\Phi \in L^p(\mathbb{M}_n) \setminus H^p(\mathbb{M}_n)$ and let $1 \le k \le n$. The following are equivalent:*

(i) *there exists a maximizing vector F of $\boldsymbol{H}_\Phi$ such that*

$$\operatorname{rank} F(\zeta) \le k, \quad \zeta \in \mathbb{T}; \tag{6.10}$$

(ii) *the following distance formula holds:*

$$\left\|H_\Phi^{\{k\}}\right\| = \operatorname{dist}_{L^p}\big(\Phi, H^p(\mathbb{M}_n)\big).$$

Note that a standard argument with analyticity properties of minors shows that $\operatorname{rank} F(\zeta)$ is constant for almost all ζ in $\mathbb{T}$.

Proof of Theorem 6.6. Suppose that (ii) holds. Let $G \in H^q(\boldsymbol{S}_2^{n,k})$ be a maximizing vector of $H_\Phi^{(k)}$ (observe that $H^{(k)}$ is compact). Consider the matrix function $F \in H^q(\boldsymbol{S}_2^n)$ obtained from G by adding $n-k$ zero columns. Clearly.

$$\|\boldsymbol{H}_\Phi F\|_{L^2(\boldsymbol{S}_2^n)} = \left\|H_\Phi^{\{k\}} G\right\|_{L^2(\boldsymbol{S}_2^{n,k})} = \operatorname{dist}_{L^p}\big(\Phi, H^p(\mathbb{M}_n)\big)\|F\|_{L^q(\boldsymbol{S}_2^n)}.$$

Thus F is a maximizing vector of $\boldsymbol{H}_\Phi$ that satisfies (6.10).

Suppose now that F is a maximizing vector of $\boldsymbol{H}_\Phi$ such that

$$\operatorname{rank} F(\zeta) = k, \quad \zeta \in \mathbb{T}.$$

Without loss of generality we may assume that $\|F\|_{L^q(\boldsymbol{S}_2^n)} = 1$. As in the proof of Theorem 6.4, consider a best approximant Q to Φ in $H^p(\mathbb{M}_n)$ and define the matrix functions G and Ψ as in that proof. Then

$$\operatorname{rank} \Psi(\zeta) = k, \quad \zeta \in \mathbb{T}, \quad \text{and} \quad \|\Psi\|_{L^{p'}(\mathbb{M}_n)} = 1.$$

The fact that Ψ is a dual extremal function of Φ can be verified as in the proof of Theorem 6.4.

By Theorem 3.4, Ψ admits a factorization $\Psi = F_\flat G_\flat$, where $F_\flat \in H^q(\mathbb{M}_{n,k})$, $G_\flat \in H^2_0(\mathbb{M}_{k,n})$ and $\|\Psi\|_{L^{p'}(\boldsymbol{S}_1^n)} = \|F_\flat\|_{L^q(\boldsymbol{S}_2^{n,k})}\|G_\flat\|_{L^2(\boldsymbol{S}_2^{k,n})}$.

We claim that $\|\boldsymbol{H}_\Phi\| = \left\|H_\Phi^{\{k\}}\right\|$ and $F_\flat$ is a maximizing vector of $H_\Phi^{\{k\}}$. This can be proved in the same way as in the proof of Theorem 6.1. Indeed, without loss of generality we may assume that $\|F_\flat\|_{L^q(\boldsymbol{S}_2^{n,k})} = 1$ and $\|G_\flat\|_{L^2(\boldsymbol{S}_2^{k,n})} = 1$. Then

$$\begin{aligned}\|\boldsymbol{H}_\Phi\| \ge \left\|H_\Phi^{\{k\}}\right\| \ge \left|\left(H_\Phi^{\{k\}}F_\flat, G_\flat^*\right)_{L^2(\boldsymbol{S}_2^n)}\right| &= \left|\int_{\mathbb{T}} \operatorname{trace}\left(\left(H_\Phi^{\{k\}}F_\flat\right)G_\flat\right) d\boldsymbol{m}\right| \\ = \left|\int_{\mathbb{T}} \operatorname{trace}\left(\mathbb{P}_-(\Phi F_\flat), G_\flat\right) d\boldsymbol{m}\right| &= \left|\int_{\mathbb{T}} \operatorname{trace}(\Phi F_\flat G_\flat)\, d\boldsymbol{m}\right| \\ = \int_{\mathbb{T}} \operatorname{trace}\left(\Phi\Psi\right) d\boldsymbol{m} &= \operatorname{dist}_{L^p}\left(\Phi, H^p(\mathbb{M}_n)\right) = \|\boldsymbol{H}_\Phi\|\end{aligned}$$

by (4.1) and Theorem 6.1. ■

Definition. A matrix function $\Phi \in L^p(\mathbb{M}_n)$ is said to *have order* k if k is the smallest number such that

$$\operatorname{dist}_{L^p}\left(\Phi, H^p(\mathbb{M}_n)\right) = \left\|H_\Phi^{\{k\}}\right\|.$$

Clearly, a matrix function Φ is respectable if and only if it has order 1.

The reasoning given in the proof of Theorem 6.6 allows us to obtain the following formulae for the order of a matrix function in L^p.

Theorem 6.7. *Let Φ be a matrix function in $L^p(\mathbb{M}_n)$. Then the following assertion hold:*

(i) *the order of Φ is the minimal number k, for which there exists a maximizing vector F of $\boldsymbol{H}_\Phi$ that satisfies* (6.10).

(ii) *the order of Φ is the minimal number k such that Φ has a dual extremal function Ψ satisfying*

$$\operatorname{rank}\Psi(\zeta) \le k, \quad \zeta \in \mathbb{T}.$$

Proof. It is easy to see that the proof of Theorem 6.7 is contained in the proof of Theorem 6.6. ■

In §7 we obtain one more formula for the order of Φ, see Theorem 7.2.

We can obtain now an analog of Theorem 6.2 for the Hankel operators $H_\Phi^{\{k\}}$.

Theorem 6.8. *Let Φ be a matrix function in $L^p(\mathbb{M}_n)$ such that*

$$\operatorname{dist}_{L^p}\left(\Phi, H^p(\mathbb{M}_n)\right) = \left\|H_\Phi^{\{k\}}\right\|.$$

Then the following assertions hold:

(i) *if $F \in H^q(\mathbb{M}_{n,k})$ is a maximizing vector of $H_\Phi^{\{k\}}$, then $(\Phi - Q)F \in H^2_-(\boldsymbol{S}_2^{n,k})$;*

(ii) *if F_1 and F_2 are maximizing vectors of $H_\Phi^{(k)}$, then*

$$\|F_1(\zeta)\|_{\boldsymbol{S}_2^n} = c\|F_2(\zeta))\|_{\boldsymbol{S}_2^n}$$

for some positive constant c;

(iii) *if Q is a best approximant to Φ in $H^p(\mathbb{M}_n)$ and F is a maximizing vector of $H_\Phi^{\{k\}}$, then the matrix*

$$\frac{1}{\|(\Phi - Q)(\zeta)\|}(\Phi - Q)(\zeta), \quad \zeta \in \mathbb{T},$$

is isometric on the range of $F(\zeta)$.

Theorem 6.8 can be proved in the same way as Theorem 6.2.

Remark. Note that in the case $p = \infty$ and $k = 1$, (ii) is very far from being true. Indeed, we can take two different scalar outer functions h_1 and h_2 in H^2 and consider the matrix function Φ defined by

$$\Phi = \begin{pmatrix} \bar{z}\frac{\bar{h}_1}{h_1} & \mathbf{0} \\ \mathbf{0} & \bar{z}\frac{\bar{h}_2}{h_2} \end{pmatrix}.$$

It is easy to see that Φ is badly approximable, $\|H_\Phi\|_{H^2(\mathbb{C}^2)\to H^2_-(\mathbb{C}^2)} = 1$, and the vector functions

$$f_1 = \begin{pmatrix} h_1 \\ \mathbf{0} \end{pmatrix} \quad \text{and} \quad f_2 = \begin{pmatrix} \mathbf{0} \\ h_1 \end{pmatrix}$$

are maximizing vectors of H_Φ, though the functions

$$\zeta \mapsto \|f_1(\zeta)\|_{\mathbb{C}^2} = |h_1(\zeta)| \quad \text{and} \quad \zeta \mapsto \|f_2(\zeta)\|_{\mathbb{C}^2} = |h_2(\zeta)|$$

do not have to be proportional.

7. *p*-badly approximable functions

In this section we characterize the set of all badly approximable functions in terms of certain special factorizations. Such factorizations allow us in this section to obtain a parametrization of all p-best approximants to a given matrix function in $L^p(\mathbb{M}_n)$ in the case when such a best approximant is not unique.

To describe the set of p-badly approximable matrix functions, we prove the following result that can be considered as an analog of the corresponding result for analytic approximation in the L^∞ norm, see [Pe1], Ch. 14, § 15.

Theorem 7.1. *Let $\Phi \in L^p(\mathbb{M}_n)$ and let Q be a best approximant to Φ in $H^p(\mathbb{M}_n)$. Then $\Phi - Q$ admits the following factorization*

$$\Phi - Q = \mathcal{W}^* \begin{pmatrix} \Delta & \mathbf{0} \\ \mathbf{0} & \Phi_\# \end{pmatrix} \mathcal{V}^*, \tag{7.1}$$

where $\mathcal{V}$ and $\mathcal{W}^{\mathrm{t}}$ are k-balanced matrix functions for some $k \le n$, Δ is a $k \times k$ p-badly approximable matrix function such that the matrix function $\mathrm{d}_\Phi^{-1}\Delta$ is unitary-valued, and $\Phi_\#$ is a matrix function such that

$$\|\Phi_\#(\zeta)\|_{\mathbb{M}_{n-k}} \le \|\Delta(\zeta)\|_{\mathbb{M}_k}, \quad \zeta \in \mathbb{T}.$$

Proof. Clearly, without loss of generality we may assume that $Q = \mathbf{0}$, i.e., Φ is a p-badly approximable matrix function.

Suppose that $\boldsymbol{H}_\Phi$ has a maximizing vector of rank k. In the proof of Theorem 6.6 we have shown that $\left\|H_\Phi^{\{k\}}\right\| = \operatorname{dist}_{L^p}\left(\Phi, H^p(\mathbb{M}_n)\right)$ and there exists a maximizing vector $F \in H^q(\mathbb{M}_{n,k})$ of $H_\Phi^{\{k\}}$ such that $\operatorname{rank} F(\zeta) = k$, $\zeta \in \mathbb{T}$. Consider the inner-outer factorization of F^{t}:

$$F^{\mathrm{t}} = \mathcal{O}_1^{\mathrm{t}} F_{\mathrm{co}}^{\mathrm{t}}.$$

Then

$$F = F_{\mathrm{co}} \mathcal{O}_1$$

It is easy to see that $\mathcal{O}_1$ is an inner matrix function of size $k \times k$ and F_{co} is a co-outer matrix function of size $n \times k$. It follows easily from (i) of Theorem 6.8 that F_{co} is a maximizing vector of $H_\Phi^{\{k\}}$. Without loss of generality we may thus assume that F is co-outer.

Let G be the function in $H^2(\mathbb{M}_{n,k})$ defined by

$$G(\zeta) = \overline{\zeta}\, \overline{\left(H_\Phi^{\{k\}} F\right)(\zeta)}.$$

By (i) and (iii) of Theorem 6.8, we know that $G = \bar{z}\overline{\Phi F}$ has rank k on $\mathbb{T}$.

Similarly, we can consider the inner-outer factorization of G^{t} and obtain a factorization

$$G = G_{\mathrm{co}} \mathcal{O}_2,$$

where $\mathcal{O}_2$ is an inner matrix function of size $k \times k$ and G_{co} is a co-outer matrix function.

Consider now the inner-outer factorization of F

$$F = \Upsilon F_{\mathrm{o}}.$$

Since $\operatorname{rank} F(\zeta) = k$ almost everywhere on $\mathbb{T}$, it is easy to see that Υ has size $n \times k$. Similarly, we can consider the inner-outer factorization of G_{co}:

$$G_{\mathrm{co}} = \Omega G_{\mathrm{o}}$$

and Ω has size $n \times k$.

We can consider now balanced completions $\mathcal{V}$ and $\mathcal{W}$ of Υ and Ω:

$$\mathcal{V} = \begin{pmatrix} \Upsilon & \overline{\Theta} \end{pmatrix} \quad \text{and} \quad \mathcal{W}^{\mathrm{t}} = \begin{pmatrix} \Omega & \overline{\Xi} \end{pmatrix}, \tag{7.2}$$

where Θ and Ξ are inner and co-outer matrix functions such that the matrix functions $\mathcal{V}$ and $\mathcal{W}$ defined by (7.2) are unitary-valued (see § 2).

Let

$$A = \mathcal{W}\Phi\mathcal{V}.$$

By Theorem 6.8, $H_{\Phi}^{\{k\}} F = \Phi F$, and so

$$\Phi F = \mathcal{W}^* A \mathcal{V}^* F = \mathcal{W}^* A \begin{pmatrix} \Upsilon^* \\ \Theta^{\mathrm{t}} \end{pmatrix} \Upsilon F_{\mathrm{o}} = \mathcal{W}^* A \begin{pmatrix} F_{\mathrm{o}} \\ \mathbf{0} \end{pmatrix} = \bar{z}\overline{G} = \bar{z}\overline{G_{\mathrm{co}}\mathcal{O}_2}.$$

Thus

$$A \begin{pmatrix} F_{\mathrm{o}} \\ \mathbf{0} \end{pmatrix} = \mathcal{W}\Phi F = \bar{z} \begin{pmatrix} \Omega^{\mathrm{t}} \\ \Xi^* \end{pmatrix} \overline{\Omega G_{\mathrm{o}}\mathcal{O}_2} = \begin{pmatrix} \bar{z}\overline{G_{\mathrm{o}}\mathcal{O}_2} \\ \mathbf{0} \end{pmatrix}.$$

Clearly, $\|A(\zeta)\|_{\mathbb{M}_n} = \|\Phi(\zeta)\|_{\mathbb{M}_n}$, $\zeta \in \mathbb{T}$, and by Theorem 6.2 (see (6.5)), $\begin{pmatrix} F_{\mathrm{o}}(\zeta) \\ 0 \end{pmatrix}$ is a maximizing vector of $A(\zeta)$ for almost all $\zeta \in \mathbb{T}$. Let

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

where A_{11} has size $k \times k$.

By Theorem 6.8, the matrices $\|A(\zeta)\|^{-1} A_{11}(\zeta)$ take unitary values almost everywhere on $\mathbb{T}$. It is easy to verify (see e.g., [Pe1], Lemma 15.5 of Ch. 14) that $A_{21} = \mathbf{0}$, $A_{12} = \mathbf{0}$, and $\|A_{22}(\zeta)\| \le \|A_{11}(\zeta)\|$, $\zeta \in \mathbb{T}$.

Clearly, $\|A(\zeta)\| = \mathrm{d}_{\Phi}(\zeta)$. Put $\Delta = A_{11}$. Then $(\mathrm{d}_{\Phi})^{-1}\Delta$ is a unitary-valued matrix function and

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

where $\Phi_{\#} \stackrel{\mathrm{def}}{=} A_{22}$. Obviously, $\|\Phi_{\#}(\zeta)\| \le \mathrm{d}_{\Phi}(\zeta)$, $\zeta \in \mathbb{T}$.

It is easy to see that $\boldsymbol{H}_{\Delta} F_{\mathrm{o}} = \bar{z}\overline{G_{\mathrm{o}}\mathcal{O}_2}$, and so $\|\boldsymbol{H}_{\Delta}\| = \|\boldsymbol{H}_{\Phi}\|$, which implies that Δ is a p-badly approximable matrix function. ■

Remark 1. Note that the matrix function Δ is determined by the choice of a maximizing vector and it does not depend on the choice of a p-best approximant Q. It is also clear that the k-balanced matrix functions $\mathcal{V}$ and $\mathcal{W}$ do not depend on the choice of Q either.

Remark 2. Clearly, we can always take k to be the order of Φ. However, the choice of k is not always unique. For example, if $\varphi \in L^p$ is a scalar p-badly approximable function and $f \in H^q$ is a maximizing vector of H_φ, then it is easy to see that $\Phi = \begin{pmatrix} \varphi & \mathbf{0} \\ \mathbf{0} & \varphi \end{pmatrix}$ is a respectable p-badly approximable matrix function and the matrix function $F = \begin{pmatrix} f & \mathbf{0} \\ \mathbf{0} & f \end{pmatrix}$ is a maximizing vector of $\boldsymbol{H}_\Phi = H_\Phi^{\{2\}}$. Thus the matrix function Φ admits factorizations of the form 7.1 with $k = 1$ and $k = 2$.

Definition. We say that a matrix function $\Phi \in L^p(\mathbb{M}_n)$ *has gender* k if k is the maximal number such that $\boldsymbol{H}_\Phi$ has a maximizing vector of rank k. Clearly, in Theorem 7.1 we can take k to be the gender of Φ.

Factorizations of the form (7.1) allow us to obtain one more formula for the order of matrix functions in L^p.

Theorem 7.2. *Let Φ be a matrix function in $L^p(\mathbb{M}_n)$ and let Q is a p-best approximant to Φ. Then the order of Φ is the minimal number k such that $\Phi - Q$ admits a factorization as in* (7.1) *with k-balanced matrix functions $\mathcal{V}$ and $\mathcal{W}^{\mathrm{t}}$.*

Proof. The proof of Theorem 7.1 shows that if k is the order of Φ, then $\Phi - Q$ admits a factorization of the form (7.1) with k-balanced matrix functions $\mathcal{V}$ and $\mathcal{W}^{\mathrm{t}}$.

Suppose now that (7.1) holds with k-balanced matrix functions $\mathcal{V}$ and $\mathcal{W}^{\mathrm{t}}$. Suppose that $\mathcal{V}$ and $\mathcal{W}$ are given by (7.2).

Let $G \in H^q(\boldsymbol{S}_2^k)$ be a maximizing vector of $\boldsymbol{H}_\Delta$. Consider the matrix function $F \in H^q(\boldsymbol{S}_2^{n,k})$ defined by

$$F = \Upsilon G.$$

We have

$$\begin{aligned}(\Phi - Q)F &= \mathcal{W}^* \begin{pmatrix} \Delta & \mathbf{0} \\ \mathbf{0} & \Phi_\# \end{pmatrix} \begin{pmatrix} \Upsilon^* \\ \Theta^{\mathrm{t}} \end{pmatrix} \Upsilon G \\ &= \mathcal{W}^* \begin{pmatrix} \Delta & \mathbf{0} \\ \mathbf{0} & \Phi_\# \end{pmatrix} \begin{pmatrix} G \\ \mathbf{0} \end{pmatrix} = \begin{pmatrix} \overline{\Omega} & \Xi \end{pmatrix} \begin{pmatrix} \Delta G \\ \mathbf{0} \end{pmatrix}.\end{aligned}$$

Since G is a maximizing vector of $\boldsymbol{H}_\Delta$ and Δ is a p-badly approximable matrix function, it follows from Theorem 6.2 that $\Delta G = \boldsymbol{H}_\Delta G$, and so

$$(\Phi - Q)F = \overline{\Omega}\boldsymbol{H}_\Delta G.$$

It is easy to see that F is a maximizng vector of $H_\Phi^{\{k\}}$ and $\|\boldsymbol{H}_\Phi\| = \big\|H_\Phi^{\{k\}}\big\|$. This proves the result. ■

Remark. If $\Phi - Q$ admits a factorization as in (7.1) with k-balanced $\mathcal{V}$ and $\mathcal{W}$, where k is the order of Φ, and Λ is a dual extremal function of Δ, then a dual extremal function Ψ of Φ satisfying the condition

$$\operatorname{rank}\Psi(\zeta) = k, \quad \zeta \in \mathbb{T},$$

can be obtained by the following explicit formula:

$$\Psi = \Upsilon\Lambda\Omega^{\mathrm{t}}.$$

Indeed, we may assume without loss of generality that Φ is p-badly approximable and $Q = \mathbf{0}$. We have

$$\|\Psi\|_{L^{p'}(\boldsymbol{S}_1^n)} = \|\Lambda\|_{L^{p'}(\boldsymbol{S}_1^k)} = 1.$$

Clearly,

$$\begin{aligned}\operatorname{trace}\big(\Phi(\zeta)\Psi(\zeta)\big)) &= \operatorname{trace}\big(\Omega^{\mathrm{t}}(\zeta)\Phi(\zeta)\Upsilon(\zeta)\Lambda(\zeta)\big)\\ &= \operatorname{trace}\left(\begin{pmatrix} I_k & 0\end{pmatrix}\begin{pmatrix}\Delta(\zeta) & 0\\ 0 & \Phi_{\#}(\zeta)\end{pmatrix}\begin{pmatrix} I_k\\ 0\end{pmatrix}\Lambda(\zeta)\right)\\ &= \operatorname{trace}\Big(\Delta(\zeta)\Lambda(\zeta)\Big).\end{aligned}$$

Thus

$$\begin{aligned}\int_{\mathbb{T}}\operatorname{trace}(\Phi(\zeta)\Psi(\zeta))\,d\boldsymbol{m}(\zeta) &= \int_{\mathbb{T}}\operatorname{trace}\big(\Delta(\zeta)\Lambda(\zeta)\big)\,d\boldsymbol{m}(\zeta)\\ &= \|\Delta\|_{L^p(\mathbb{M}_k)} = \|\Phi\|_{L^p(\mathbb{M}_n)}. \quad \blacksquare\end{aligned}$$

Let us now describe all p-badly approximable matrix functions. Note that similar results hold in the case $p = \infty$ under certain restrictions on the function, see [Pe1], Ch. 14, § 15.

Theorem 7.3. *Let Φ be matrix function in $L^p(\mathbb{M}_n)$. Then Φ is p-badly approximable if and only if there exists $k \le n$ such that Φ admits a factorization*

$$\Phi = \mathcal{W}^*\begin{pmatrix}\Delta & \mathbf{0}\\ \mathbf{0} & \Phi_{\#}\end{pmatrix}\mathcal{V}^*, \tag{7.3}$$

where $\mathcal{V}$ and $\mathcal{W}^{\mathrm{t}}$ are k-balanced matrix functions, Δ is a $k\times k$ p-badly approximable matrix function such that the matrix function $\mathrm{d}_\Phi^{-1}\Delta$ is unitary-valued, and $\Phi_{\#}$ is a matrix function such that $\|\Phi_{\#}(\zeta)\|_{\mathbb{M}_{n-k}} \le \|\Delta(\zeta)\|_{\mathbb{M}_k}$ for almost all $\zeta \in \mathbb{T}$.

Proof. The fact that p-badly approximable matrix functions Φ admit factorizations of the form (7.3) follows immediately from Theorem 7.1.

Suppose now that Φ is given by (7.3). Consider the Hankel operator $\boldsymbol{H}_\Delta\colon H^q(\boldsymbol{S}_2^k) \to H^2_-(\boldsymbol{S}_2^k)$. Let $\mathcal{F} \in H^q(\boldsymbol{S}_2^k)$ be a maximizing vector of $\boldsymbol{H}_\Delta$. Since Δ is p badly approximable, it follows from Theorems 6.1 and 6.2 that

$$\boldsymbol{H}_\Delta \mathcal{F} = \Delta\mathcal{F} \quad \text{and} \quad \|\boldsymbol{H}_\Delta \mathcal{F}\|_{L^2(\boldsymbol{S}_2^k)} = \|\Phi\|_{L^p(\mathbb{M}_n)} \|\mathcal{F}\|_{L^q(\boldsymbol{S}_2^k)}.$$

Consider the matrix function $F = \Upsilon\mathcal{F}$, where Υ is as in (7.2). We have

$$\begin{aligned}
\Phi F &= \mathcal{W}^* \begin{pmatrix} \Delta & \mathbf{0} \\ \mathbf{0} & \Phi_\# \end{pmatrix} \begin{pmatrix} \Upsilon^* \\ \Theta^{\mathrm{t}} \end{pmatrix} \Upsilon\mathcal{F} \\
&= \mathcal{W}^* \begin{pmatrix} \Delta & \mathbf{0} \\ \mathbf{0} & \Phi_\# \end{pmatrix} \begin{pmatrix} \mathcal{F} \\ \mathbf{0} \end{pmatrix} \\
&= \begin{pmatrix} \overline{\Omega} & \Xi \end{pmatrix} \begin{pmatrix} \Delta\mathcal{F} \\ \mathbf{0} \end{pmatrix} = \overline{\Omega}\Delta\mathcal{F} \in H^2_-(\boldsymbol{S}_2^{n,k}).
\end{aligned}$$

Thus

$$\begin{aligned}
\left\| H_\Phi^{\{k\}} F \right\|_{L^2(\boldsymbol{S}_2^{n,k})} &= \|\Delta\mathcal{F}\|_{L^2(\boldsymbol{S}_2^k)} = \|\boldsymbol{H}_\Delta \mathcal{F}\|_{L^2(\boldsymbol{S}_2^k)} \\
&= \|\Phi\|_{L^p(\mathbb{M}_n)} \|\mathcal{F}\|_{L^2(\boldsymbol{S}_2^k)} = \|\Phi\|_{L^p(\mathbb{M}_n)} \|F\|_{L^2(\boldsymbol{S}_2^{n,k})}.
\end{aligned}$$

It follows that $\left\| H_\Phi^{\{k\}} \right\| = \|\Phi\|_{L^p(\mathbb{M}_n)}$, and so Φ is p-badly approximable. ■

The next result allows us to parametrize all best approximants in the case when there are more than one best approximant. A similar result also holds in the case of approximation in the norm of L^∞ under certain restrictions on Φ, see [Pe1], Ch. 14, § 15.

Theorem 7.4. *Let Φ and Q be as in Theorem* 7.1 *and let $\Phi - Q$ be factorizred as in* (7.1)*. A matrix function $R \in H^p(\mathbb{M}_n)$ is a p-best approximant to Φ if and only if there exists a matrix function $R_\# \in H^p(\mathbb{M}_{n-k})$ such that*

$$\Phi - R = \mathcal{W}^* \begin{pmatrix} \Delta & \mathbf{0} \\ \mathbf{0} & \Phi_\# - R_\# \end{pmatrix} \mathcal{V}^*, \tag{7.4}$$

and

$$\|(\Phi_\# - R_\#)(\zeta)\|_{\mathbb{M}_{n-k}} \le \|\Delta(\zeta)\|_{\mathbb{M}_k}, \quad \zeta \in \mathbb{T}. \tag{7.5}$$

We need the following lemma.

Lemma 7.5. *Let $\mathcal{V}$ and $\mathcal{W}^{\mathrm{t}}$ be k-balanced matrix functions of size $n \times n$. Then*

$$\mathcal{W} H^p(\mathbb{M}_n) \mathcal{V} \bigcap \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & L^p(\mathbb{M}_{n-k}) \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & H^p(\mathbb{M}_{n-k}) \end{pmatrix}.$$

For $p = \infty$ this is Theorem 1.8 of Ch. 14 of [Pe1]. The proof given in [Pe1] also works in our case.

Proof of Theorem 7.4. Suppose that R is a best approximant to Φ. Then by Theorem 7.1, $\Phi - R$ admits a factorization

$$\Phi - R = \mathcal{W}^* \begin{pmatrix} \Delta & \mathbf{0} \\ \mathbf{0} & \Phi_\flat \end{pmatrix} \mathcal{V}^*,$$

where Δ, $\mathcal{V}$, and $\mathcal{W}$ are as in (7.1) and $\Phi_\flat$ is a matrix function such that $\|\Phi_\flat(\zeta)\|_{\mathbb{M}_{n-k}} \le \|\Delta(\zeta)\|_{\mathbb{M}_k}$, $\zeta \in \mathbb{T}$. Then

$$R - Q = \mathcal{W}^* \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \Phi_\# - \Phi_\flat \end{pmatrix} \mathcal{V}^*.$$

By Lemma 7.5, $R_\# \stackrel{\text{def}}{=} \Phi_\# - \Phi_\flat \in H^p(\mathbb{M}_{n-k})$.

Conversely, suppose that $R_\#$ is a matrix function in $H^p(\mathbb{M}_{n-k})$ such that (7.5) holds. Then by Lemma 7.5, there exists $R \in H^p(\mathbb{M}_n)$ such that

$$R - Q = \mathcal{W}^* \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & R_\# \end{pmatrix} \mathcal{V}^*.$$

Then (7.4) holds. It follows easily from Theorem 7.3 that R is a p-best analytic approximant to Φ. ■

Theorem 7.4 says that to describe all p best approximants, we should describe all functions $R_\# \in H^p(\mathbb{M}_{n-k})$ such that (7.5) holds. By Theorem 6.3 there exists a scalar outer function $\varkappa$ in H^p such that $|\varkappa(\zeta)| = \|\Delta(\zeta)\|_{\mathbb{M}_k} = \mathrm{d}_\Phi(\zeta)$, $\zeta \in \mathbb{T}$. Clearly, a matrix function $R_\#$ in $H^p(\mathbb{M}_{n-k})$ satisfies (7.5) if and only if the matrix function $\varkappa^{-1}R_\#$ satisfies the inequality

$$\|\varkappa^{-1}\Phi_\# - \varkappa^{-1}R_\#\|_{L^\infty} \le 1.$$

In other words, this reduces the problem of the description of all p-best approximants to the problem of describing all matrix functions $\mathcal{Q}$ in $H^\infty(\mathbb{M}_{n-k.n-k})$ such that

$$\|\varkappa^{-1}\Phi_\# - \mathcal{Q}\|_{L^\infty} \le 1. \tag{7.6}$$

Note that the problem to describe all H^∞ matrix function $\mathcal{Q}$ satisfying (7.6) is the classical Nehari problem and in the case of nonuniqueness there is formula parametrizing all solutions. It was obtained by Adamyan, Arov, and Krein in [AAK1] and [AAK2] under certain assumptions and by Kheifets [Khe] in the most general case; see also Ch. 5 of [Pe1].

8. *p*-superoptimal approximation

In this section we introduce the notion of p-superoptimal approximation and prove that if Φ is a rational matrix function then Φ has a unique p-superoptimal approximant.

Definition. Let $\Phi \in L^p(\mathbb{M}_n) \setminus H^p(\mathbb{M}_n)$. For a function $Q \in H^p(\mathbb{M}_n)$, we define the numbers $\tau_j(\Phi, Q)$, $0 \le j \le n-1$, by

$$\tau_j(\Phi, Q) = \operatorname{ess\,sup}_{\zeta \in \mathbb{T}} \frac{s_j\big((\Phi - Q)(\zeta)\big)}{\mathrm{d}_\Phi(\zeta)}.$$

A function $Q \in H^p(\mathbb{M}_n)$ is called a *p-superoptimal approximant* to Φ if it minimizes lexicographically the sequence $\tau_j(\Phi, Q)$, $0 \le j \le n-1$.

If Q is a p-superoptimal approximant to Φ, we put

$$\tau_j(\Phi) \stackrel{\text{def}}{=} \tau_j(\Phi, Q), \quad 0 \le j \le n-1.$$

Clearly, if Q is a best analytic approximant to Φ in $L^p(\mathbb{M}_{m,n})$, then $\tau_0(\Phi, Q) = 1$. It is also clear that if F is a p-superoptimal approximant, then F is a best analytic approximant in $L^p(\mathbb{M}_{m,n})$.

It is easy to see that if Φ has gender k, then

$$\tau_j(\Phi) = 1 \quad \text{for} \quad j = 0, \cdots, k-1.$$

In this section we are going to work with rational matrix functions. When we say that a matrix function defined on the unit circle $\mathbb{T}$ is rational, we mean that it is a restriction of a rational matrix function to the unit circle. It is easy to see that if A is a rational matrix function, then its adjoint A^* is also a rational matrix function.

Suppose now that Φ is a rational matrix function of size $n \times n$ with no poles in $\mathbb{T}$ and k is the gender of Φ. As in the proof of Theorem 7.1, we consider a co-outer maximizing vector F of the Hankel operator $H_\Phi^{\{k\}}$, the matrix function G defined by $G = \overline{\bar{z} H_\Phi^{\{k\}} F}$, the factorizations

$$G = G_{\mathrm{co}} \mathcal{O}, \quad F = \Upsilon F_{\mathrm{o}}, \quad \text{and} \quad G_{\mathrm{co}} = \Omega G_{\mathrm{o}},$$

where $\mathcal{O}$ is an inner matrix function of size $k \times k$, Υ and Ω are inner and co-outer matrix functions of size $n \times k$, G_{co} is a co-outer matrix functions of size $n \times k$, and F_{o} and G_{o} are outer matrix functions of size $k \times k$. We also assume that the k-balanced matrix functions $\mathcal{V}$ and $\mathcal{W}^{\mathrm{t}}$ are given by (7.2), Q is a p-best approximant to Φ and $\Phi - Q$ is factorized as in (7.3). Finally, we assume that F is normalized so that (6.7) holds.

Lemma 8.1. *Let Φ be a rational matrix function in $L^p(\mathbb{M}_n)$. Then the matrix functions $\mathcal{O}$, F, G, $\mathcal{V}$, and $\mathcal{W}$ are also rational.*

Proof. If Φ is rational, it is easy to see that $H^{\{k\}}_{\Phi}A$ is rational for an arbitrary function $A \in H^q(\boldsymbol{S}^{n,k}_2)$. In particular, this is true for the function $H^{\{k\}}_{\Phi}F$, and so G is rational.

Let us show that $\mathcal{O}$ is rational. It is well known (see e.g., [Pe1], Ch. 2, § 5) that a square inner function $\mathcal{U}$ is rational if and only if the subspace

$$K_{\mathcal{U}} \stackrel{\text{def}}{=} H^2(\mathbb{C}^k) \ominus \mathcal{U}H^2(\mathbb{C}^k) = \mathcal{U}H^2_-(\mathbb{C}^k) \bigcap H^2(\mathbb{C}^k)$$

is finite-dimensional. Since G is rational, the Hankel operator

$$H_{\overline{G}}\colon H^2(\mathbb{C}^k) \to H^2_-(\mathbb{C}^k)$$

has finite rank (see e.g., [Pe1], Ch. 2, §,5). It is easy to see that for $f \in K_{\mathcal{O}^{\mathrm{t}}}$,

$$H_{\overline{G}}f = \overline{G}f.$$

Since $\operatorname{rank} G(\zeta) = k$ almost everywhere on $\mathbb{T}$, it follows that multiplication by $\overline{G}$ has trivial kernel. Thus $K_{\mathcal{O}^{\mathrm{t}}}$ is finite-dimensional, and so $\mathcal{O}^{\mathrm{t}}$ is rational. Thus $\mathcal{O}$ is rational, and so $G_{\mathrm{co}} = G\mathcal{O}^*$ is also rational.

To prove that the matrix function G_{o} is rational, we observe that

$$G^*_{\mathrm{co}}G_{\mathrm{co}} = G^*_{\mathrm{o}}G_{\mathrm{o}},$$

and so $G^*_{\mathrm{o}}G_{\mathrm{o}}$ is a rational matrix function. The rationality of G_{o} follows now from the following well-known fact (see [Y]): if $\mathcal{Q}$ be a matrix outer function of class $H^2(\mathbb{M}_k)$, then $\mathcal{Q}$ is rational if and only if $\mathcal{Q}^*\mathcal{Q}$ is rational.

We have $\Omega = G_{\mathrm{co}}G_{\mathrm{o}}^{-1}$, and so Ω is rational. By Lemma 12.1 of Ch. 14 of [Pe1], the matrix function Ξ is rational, and so $\mathcal{W}$ is rational.

Let us show that $\mathcal{V}$ is a rational matrix function. Since Φ is rational, it follows that

$$\mathbb{P}_-\Phi = \mathbb{P}_-(\Phi - Q) = \mathbb{P}_-\mathcal{W}^* \begin{pmatrix} \Delta & \mathbf{0} \\ \mathbf{0} & \Phi_{\#} \end{pmatrix} \mathcal{V}^*$$

is a rational matrix function. Thus

$$\mathbb{P}_-\Omega^{\mathrm{t}}\mathcal{W}^* \begin{pmatrix} \Delta & \mathbf{0} \\ \mathbf{0} & \Phi_{\#} \end{pmatrix} \mathcal{V}^* = \mathbb{P}_-\Omega^{\mathrm{t}}\mathbb{P}_-\mathcal{W}^* \begin{pmatrix} \Delta & \mathbf{0} \\ \mathbf{0} & \Phi_{\#} \end{pmatrix} \mathcal{V}^*$$

is rational. We have

$$\Omega^{\mathrm{t}}\mathcal{W}^* \begin{pmatrix} \Delta & \mathbf{0} \\ \mathbf{0} & \Phi_{\#} \end{pmatrix} \mathcal{V}^* = \begin{pmatrix} \boldsymbol{I}_k & \mathbf{0} \end{pmatrix} \begin{pmatrix} \Delta & \mathbf{0} \\ \mathbf{0} & \Phi_{\#} \end{pmatrix} \mathcal{V}^* = \Delta\Upsilon^*,$$

and so $\mathbb{P}_-\Delta\Upsilon^*$ is a rational matrix function.

Let h be a scalar outer function such that

$$|h(\zeta)| = \|G(\zeta)\|_{\boldsymbol{S}^{n,k}_2}, \quad \zeta \in \mathbb{T}.$$

Then $h \in H^2$. It follows from (6.5) and (6.7) that

$$\|\Delta(\zeta)\|_{\mathbb{M}_k} = |h(\zeta)|^{2/p} \quad \text{and} \quad \|F(\zeta)\|_{\boldsymbol{S}^{n,k}_2} = |h(\zeta)|^{2/q}. \tag{8.1}$$

Since G is rational, the function $|h|^2$ is rational. It follows from the result from [Y] quoted above that the function h is also rational.

Let us show that F_{o} is a maximizing vector of $\boldsymbol{H}_\Delta$ and $H_\Phi^{\{k\}}F = \overline{\Omega}\Delta F_{\mathrm{o}}$. Since $F = \Upsilon F_{\mathrm{o}}$ is a maximizing vector of $H_\Phi^{\{k\}}$, we have by Theorem 6.8,

$$H_\Phi^{\{k\}}F = (\Phi - Q)F = \mathcal{W}^* \begin{pmatrix} \Delta & \mathbf{0} \\ \mathbf{0} & \Phi_\# \end{pmatrix} \mathcal{V}^* \Upsilon F_{\mathrm{o}}$$

$$= \begin{pmatrix} \overline{\Omega} & \Xi \end{pmatrix} \begin{pmatrix} \Delta & \mathbf{0} \\ \mathbf{0} & \Phi_\# \end{pmatrix} \begin{pmatrix} \Upsilon^* \\ \Theta^{\mathrm{t}} \end{pmatrix} \Upsilon F_{\mathrm{o}} = \overline{\Omega}\Delta F_{\mathrm{o}} \in H^2_-(\boldsymbol{S}_2^{n,k}).$$

Since the matrix function Ω is co-outer, it follows from Lemma 1.4 of Ch. 14 of [Pe1] that $\Delta F_{\mathrm{o}} \in H^2_-(\boldsymbol{S}_2^k)$. It is easy to see from (8.1) that F_{o} is a maximizing vector of $\boldsymbol{H}_\Delta$ and $H_\Phi^{\{k\}}F = \overline{\Omega}\Delta F_{\mathrm{o}}$. Thus

$$\Delta F_{\mathrm{o}} = \bar{z}\overline{G_{\mathrm{o}}\mathcal{O}}.$$

Consider now the inner-outer factorization of the matrix function $G_{\mathrm{o}}\mathcal{O}$:

$$G_{\mathrm{o}}\mathcal{O} = \mathfrak{O}\mathcal{G}_{\mathrm{o}}.$$

Clearly, both $\mathfrak{O}$ and $\mathcal{G}_{\mathrm{o}}$ are rational matrix functions. Then $\Delta F_{\mathrm{o}} = \bar{z}\overline{\mathfrak{O}\mathcal{G}_{\mathrm{o}}}$, and so

$$\Delta = \bar{z}\overline{\mathfrak{O}\mathcal{G}_{\mathrm{o}}}F_{\mathrm{o}}^{-1}$$

Put

$$U = zh^{-2/p}\mathfrak{O}^{\mathrm{t}}\Delta = \overline{\mathcal{G}}_{\mathrm{o}}(h^{2/p}F_{\mathrm{o}})^{-1}.$$

It is easy to see that U is unitary-valued. Put

$$\mathcal{Q} = h^{2/p}F_{\mathrm{o}} \quad \text{and} \quad \mathcal{Q}_\# = \mathcal{G}_{\mathrm{o}}^{\mathrm{t}}.$$

Since U is unitary-valued, it is easy to verify that

$$\mathcal{Q}^*\mathcal{Q} = \mathcal{Q}_\#\mathcal{Q}_\#^*.$$

Clearly, $\mathcal{Q}$ is outer. Since $\mathcal{Q}_\#\mathcal{Q}_\#^*$ is rational, it follows from the result of [Y] quoted above that $\mathcal{Q} = h^{2/p}F_{\mathrm{o}}$ is rational and U is rational.

We have $\Delta = \bar{z}h^{2/p}\overline{\mathfrak{O}}U$, and so

$$\mathbb{P}_-\Delta\Upsilon^* = \mathbb{P}_-\bar{z}h^{2/p}\overline{\mathfrak{O}}U\Upsilon^* = \mathbb{P}_-\bar{z}h^{2/p}\overline{\mathfrak{O}}\mathcal{Q}_\#^*\mathcal{Q}^{-1}\Upsilon^*$$

$$= \mathbb{P}_-h^{2/p}\big(\mathbb{P}_-\bar{z}\overline{\mathfrak{O}}\mathcal{Q}_\#^*\mathcal{Q}^{-1}\Upsilon^*\big)$$

is a rational matrix function. Put

$$\mathfrak{R} \stackrel{\mathrm{def}}{=} \mathbb{P}_-\bar{z}\overline{\mathfrak{O}}\mathcal{Q}_\#^*\mathcal{Q}^{-1}\Upsilon^*.$$

Let us show that $\mathfrak{R}$ is rational. We have

$$\mathbb{P}_-h\mathfrak{R} = \mathbb{P}_-h^{2/q}h^{2/p}\mathfrak{R} = \mathbb{P}_-h^{2/q}\mathbb{P}_-h^{2/p}\mathfrak{R},$$

and since $\mathbb{P}_- h^{2/p}\mathfrak{R}$ is rational and $h^{2/q} \in H^\infty$, it follows that $\mathbb{P}_- h\mathfrak{R}$ is rational. Since h is rational and $\mathfrak{R} = \mathbb{P}_-\mathfrak{R}$, it is easy to see that $\mathfrak{R}$ is rational.

Finally, since the matrix functions $\bar{z}\overline{\mathfrak{O}}\mathcal{Q}_\#^*\mathcal{Q}^{-1}$ and $\mathbb{P}_-\bar{z}\overline{\mathfrak{O}}\mathcal{Q}_\#^*\mathcal{Q}^{-1}\Upsilon^*$ are rational, it is easy to verify that Υ^* is rational. Again, it follows from Lemma 12.1 of Ch. 14 of [Pe1] that $\mathcal{V}$ is rational. ■

To prove the next theorem, we introduce the notation Λ_α, $0 < \alpha < 1$, for the class of Hölder functions of order α: a function φ on $\mathbb{T}$ is said to belong to the Hölder class Λ_α if

$$\sup_{\zeta\neq\tau} \frac{|\varphi(\zeta) - \varphi(\tau)|}{|\zeta - \tau|^\alpha} < \infty.$$

In the following theorem we keep all the notation as above.

Theorem 8.2. *If Φ is a rational matrix function, then $\Delta \in \Lambda_{2/p}(\mathbb{M}_k)$ and $h^{-2/p}\Phi_\# \in (H^\infty + C)(\mathbb{M}_{n-k})$.*

Proof. We have

$$\Delta = \bar{z}h^{2/p}\overline{\mathfrak{O}}U,$$

where $\overline{\mathfrak{O}}U$ is a rational function. If h has no zeros on $\mathbb{T}$, then Δ is infinitely differentiable. If h has zeros on $\mathbb{T}$, then, obviously, $h^{2/p} \in \Lambda_{2/p}$, which implies that $\Delta \in \Lambda_{2/p}(\mathbb{M}_k)$.

Next, since

$$\mathcal{W}(\Phi - Q)\mathcal{V} = \begin{pmatrix} \Delta & \mathbf{0} \\ \mathbf{0} & \Phi_\# \end{pmatrix},$$

it follows that $\Phi_\#$ is a sum of a rational matrix function and an H^p matrix function. Thus there exists a finite Blaschke product B such that $\Phi_\# = \overline{B}\Phi_\heartsuit$, where $\Phi_\heartsuit \in H^p(\mathbb{M}_{n-k})$. We also know that $\|\Phi_\heartsuit(\zeta)\|_{\mathbb{M}_{n-k}} \le |h(\zeta)|^{2/p}$. Since h is outer, it follows that $h^{-2/p}\Phi_\heartsuit \in H^\infty(\mathbb{M}_{n-k})$. Thus

$$h^{-2/p}\Phi_\# = \overline{B}h^{-2/p}\Phi_\heartsuit \in (H^\infty + C)(\mathbb{M}_{n-k}). \quad ■$$

Theorem 8.3. *Let Φ be an $n \times n$ rational matrix function. Then Φ has a unique p-superoptimal approximant Q. Moreover,*

$$\frac{s_j\big((\Phi - Q)(\zeta)\big)}{\mathrm{d}_\Phi(\zeta)} = \tau_j(\Phi), \quad 0 \le j \le n-1 \tag{8.2}$$

almost everywhere on $\mathbb{T}$.

Proof. Let R be a best analytic approximant to Φ in $L^p(\mathbb{M}_n)$. By Theorem 7.1, $\Phi - R$ admits a factorization of the form

$$\Phi - R = \mathcal{W}^* \begin{pmatrix} \Delta & \mathbf{0} \\ \mathbf{0} & \Phi_\# \end{pmatrix} \mathcal{V}^*,$$

where $\mathcal{V}$ and $\mathcal{W}^{\mathrm{t}}$ are k-balanced matrix function, k is the gender of Φ, Δ is a p-badly approximable $k \times k$ matrix function function such that the matrix function $\mathrm{d}_\Phi^{-1}\Delta$ is unitary-valued, and $\Phi_\#$ is a matrix function such that

$$\|\Phi_\#(\zeta)\|_{\mathbb{M}_{n-k}} \le \|\Delta(\zeta)\|_{\mathbb{M}_k}, \quad \zeta \in \mathbb{T}.$$

It follows from Theorem 7.4 that a matrix function $Q \in H^p(\mathbb{M}_n)$ is a p-superoptimal approximant to Φ if and only if

$$\Phi - Q = \mathcal{W}^* \begin{pmatrix} \Delta & \mathbf{0} \\ \mathbf{0} & \Phi_\# - Q_\# \end{pmatrix} \mathcal{V}^*,$$

where $Q_\# \in H^p(\mathbb{M}_{n-k})$ is a matrix function such that $h^{-2/p}Q_\#$ is a superoptimal approximant of $h^{-2/p}\Phi_\#$ in L^∞. Here h the scalar outer function as in the proof of Lemma 8.1, i.e., $|h^{2/p}| = \mathrm{d}_\Phi$. By Theorem 8.2, $h^{-2/p}\Phi_\# \in H^\infty + C$, and by Theorem 3.3 of Chapter 14 of [Pe1], $h^{-2/p}\Phi_\#$ has a unique superoptimal approximant in the L^∞ norm.

Formula (8.2) is an immediate consequence of Theorem 3.4 of Chapter 14 of [Pe1]. ■

The following example shows that there are matrix functions in L^p, for which there are different p-superoptimal approximants.

Example. Let φ be a scalar L^∞ function such that

$$\|\varphi\|_{L^\infty} = \mathrm{dist}_{L^\infty}(\varphi, H^\infty) = 1,$$

and such that there is a nonzero best approximant $f \in H^\infty$ in the norm of L^∞. It is well known that such functions φ exist (see, e.g., [Pe1], Ch. 1, § 1). Consider the matrix function $\Phi \in L^\infty(\mathbb{M}_2)$ defined by

$$\Phi = \begin{pmatrix} \bar{z} & \mathbf{0} \\ \mathbf{0} & \varphi \end{pmatrix}.$$

It is easy to see that both the zero function and the function $\begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & f \end{pmatrix}$ are p-superoptimal approximants for any $p \in (2, \infty)$.

References

[AAK1] V.M. Adamyan, D.Z. Arov, and M.G. Krein, Infinite Hankel matrices and generalized problems of Carathéodory–Fejér and I. Schur, *Funktsional. Anal. i Prilozhen.*, **2(2)** (1968), 1–17; English transl. in *Functional Anal Appl.* **2(2)** (1968).

[AAK2] V.M. Adamyan, D.Z. Arov, and M.G. Krein, Infinite Hankel block matrices and some related extension problems, *Izv. Akad. Nauk Armyan. SSR Ser. Mat.*, **6** (1971), 87–112 (Russian); English transl. in *Am. Math. Soc. Transl.*, Ser. 2, **111** (1978), 133–156.

[AP] R.B. Alexeev and V.V. Peller, Badly approximable matrix functions and canonical factorizations, *Indiana Univ. Math. J.*, **49** (2000), 1247–1285.

[BLPT] L. Baratchart, J. Leblond, J. Partington, and N. Torkhani, *Robust identification in the disk algebra from band-limited data,* IEEE Trans. Autom. Control, **42** (1997), 1318–1325.

[BS] L. Baratchart and F. Seyfert, *An L^p analog to AAK theory for $p \geq 2$*, J. Funct. Anal., **191** (2002), 52–122.

[F] B. A. Francis, *A Course in H^∞ Control Theory,* Lecture Notes in Control and Information Sciences No. 88, Springer Verlag, Berlin, 1986.

[H] H. Helson, *Lectures on Invariant subspaces*, Acad. Press, New Yourk, 1964.

[Ka] J.-P. Kahane, *Best approximation in $L^1(\mathrm{T})$*, Bull. Amer. Math. Soc., **80** (1974), 788–804.

[Kha] S. Khavinson, On some extremal problems of the theory of analytic functions, *Uchen. Zapiski Mosk. Universiteta, Matem.*, **144(4)** (1951), 133–143. English Translation: Amer. Math. Soc. Translations (2) **32** (1963), 139–154.

[Khe] A.Y. Kheifets, Parametrization of solutions of the Nehari problem and nonorthogonal dynamics. In: Operator theory and interpolation (Bloomington, IN, 1996), 213–233, *Oper. Theory Adv. Appl.*, **115**, Birkhäuser, Basel, 2000.

[N] N. K. Nikolskii, *Treatise on the Shift Operator,* Springer Verlag, Berlin 1986.

[Pa] J.R. Partington, *Robust identification in H^∞*, J. of Math. Anal. and Appl., **166** (1992), 428–441.

[Pe1] V.V. Peller, *Hankel operators and their applications,* Springer-Verlag, New York, 2003.

[Pe2] V.V. Peller, *Analytic approximation of matrix functions and dual extremal functions*, Preprint arXiv:0711.2771, 2007.

[PT1] V.V. Peller and S.R. Treil, Approximation by analytic matrix functions. The four block problem, *J. Funct. Anal.*, **148** (1997), 191-228.

[PT2] V.V. Peller and S.R. Treil, *Very badly approximable matrix functions*, Selecta Math., **11** (2005), 127–154.

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

[Pr] V.A. Prokhorov, *On L^p-generalization of a theorem of Adamyan, Arov, and Krein*, J. Approx. Theory **116** (2002), 380–396.

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

[Sh] H.S. Shapiro, *Regularity properties of the element of closest approximation*, Trans. Amer. Math. Soc., **181** (1973), 127–142.

[SF] B. Sz.-Nagy and C. Foias, *Harmonic analysis of operators on Hilbert space,* Akadémiai Kiadó, Budapest, 1970.

[WM] N. Wiener and P. Masani, The prediction theory of multivariate stochastic processes, I. The regularity conditions, *Acta Math.*, **98** (1957), 111-150.

[Y] D.C. Youla, *On the factorization of rational matrices*, IRE Trans. **IT-7** (1961), 172–189.

L. Baratchart
INRIA
BP 93
06902 Sophia-Antipolis Cedex, France

F.L. Nazarov
Department of Mathematics
University of Wisconsin
Madison, Wisconsin 53706, USA

V.V. Peller
Department of Mathematics
Michigan State University
East Lansing, Michigan 48824, USA