This is to certify that the
dissertation entitled

THEMATIC INDICES AND SUPEROPTIMAL SINGULAR
VALUES OF MATRIX FUNCTIONS

presented by

ALBERTO A. CONDORI

has been accepted towards fulfillment
of the requirements for the

Ph.D. degree in Mathematics

V. Peller

Major Professor's Signature

May 7, 2009

Date

# THEMATIC INDICES AND SUPEROPTIMAL SINGULAR VALUES OF MATRIX FUNCTIONS

By

Alberto A. Condori

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Mathematics

2009

# ABSTRACT

## THEMATIC INDICES AND SUPEROPTIMAL SINGULAR VALUES OF MATRIX FUNCTIONS

**By**

**Alberto A. Condori**

In this dissertation, we discuss a number of results on superoptimal approximation by analytic and meromorphic matrix-valued functions on the unit circle. We first prove the existence of a monotone non-decreasing thematic factorization for admissible (e.g. continuous) very badly approximable matrix functions. Unlike the case of monotone non-increasing thematic factorizations, it is shown that thematic indices in a monotone non-decreasing thematic factorization are not uniquely determined. We then consider the problem of characterizing superoptimal singular values. An extremal problem is introduced and its connection with the sum of superoptimal singular values is explored by considering a new class of operators: *Hankel-type* operators on Hardy spaces of matrix functions. Lastly, we consider approximation by meromorphic matrix-valued functions; the so-called Nehari-Takagi problem. We provide a counterexample that shows that the index formula in connection with meromorphic approximation, which is well-known to hold in the case of scalar-valued functions, fails in the case of matrix-valued functions.

# DEDICATION

To Papa Alberto and “Don Manuel con el lago y la montaña.”

# ACKNOWLEDGMENT

First and foremost, I would like to express my gratitude to my dissertation adviser Professor Vladimir V. Peller. You have been a constant inspiration to me since my first day at Michigan State University. I thank you for seeing potential in me and taking me as your student. Without your support, patience and helpful remarks, this dissertation would have not been possible. To you, I am forever indebted.

I thank my fiancée and best friend, Cara. Without your continuous love and support, I would have not seen the sunshine in my darkest days at MSU; especially during my physical and metaphysical struggles. Thank you for believing in my mathematical ability in the times that I could not.

During the many times I felt like surrending, I have always remembered my dad telling me "los Condori nunca mueren." Dad, I thank you for teaching me to never give up. Mom, Sandro and Elthon, I thank you for always reminding me about living life, especially at the times when I was consumed by my mathematical and personal problems.

Lastly, I express my appreciation to Professors A. Volberg, M. Frazier, and V. Zeidan. You have been an inspiration to me as well. I thank you for believing in my potential and for stimulating conversations that led to both my mathematical and personal growth.

# TABLE OF CONTENTS

# Chapter 1

# Introduction

The problem of approximating a continuous function on the unit circle $\mathbb{T}$ by bounded analytic functions in the unit disk $\mathbb{D}$ with respect to the uniform norm has been studied for quite some time. A simple compactness argument reveals that *any* bounded measurable function $\varphi$ on $\mathbb{T}$ has a best uniform approximation $\mathcal{A}\varphi$ by bounded analytic functions, i.e.

$$\|\varphi - \mathcal{A}\varphi\|_\infty = \operatorname{dist}(\varphi, H^\infty) = \inf\{\|\varphi - f\|_\infty : f \in H^\infty\}.$$

The uniqueness of a best approximation for continuous functions was first proved by S. Khavinson in [Kh] and rediscovered later by several mathematicians.

Different authors have studied the error function $\varphi - \mathcal{A}\varphi$, or equivalently, functions $\psi$ for which the zero function is a best approximation. These functions $\psi$ are called *badly approximable.* For example, it was proved by Poreda in [Po] that a continuous function $\psi$ is badly approximable if and only if it has constant modulus and negative winding number.

A new light into the best approximation problem was shed by Nehari [Ne]. He found the following formula for the distance from a bounded measurable function $\varphi$

to the Banach algebra $H^\infty$ of bounded analytic functions in $\mathbb{D}$:

$$\operatorname{dist}_{L^\infty}(\varphi, H^\infty) = \|H_\varphi\|. \tag{1.0.1}$$

As usual, the *Hankel operator* $H_\varphi : H^2 \to H^2_- = L^2 \ominus H^2$ with symbol $\varphi$ is defined by

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

where $\mathbb{P}_-$ denotes the orthogonal projection of $L^2$ onto $H^2_-$. (Throughout, $\mathcal{H} \ominus \mathcal{K}$ denotes the orthogonal complement of a subspace $\mathcal{K}$ of a Hilbert space $\mathcal{H}$.) Therefore, formula (1.0.1) motivated the consideration of Hankel operators in the study of the best approximation problem. In the years to follow, further evidence of the intimate connection between Hankel operators and the best approximation problem was revealed through many beautiful results. For instance, Adamyan, Arov and Krein found a more general condition that guarantees the uniqueness of the best approximation. In [AAK1], they showed that if $\varphi$ is admissible, then $\varphi$ has a unique best approximation in $H^\infty$. A function $\varphi \in L^\infty$ is said to be *admissible* if the essential norm $\|H_\varphi\|_e$ of the Hankel operator $H_\varphi$ is strictly less than its operator norm $\|H_\varphi\|$. As usual, $\|T\|_e$ denotes the *essential norm* of an operator $T : \mathcal{H} \to \mathcal{K}$ between Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$.

Even though the notion of winding number is not available for the class of admissible functions, the classification of badly approximable functions given by Poreda was also extended to this class of functions by using Hankel and Toeplitz operators. It is well-known now (e.g. see Chapter 7 in [Pe1]) that an admissible function $\psi$ is badly approximable if and only if $\psi$ has constant modulus, the Toeplitz operator $T_\psi$ is Fredholm, and $\operatorname{ind} T_\psi > 0$. Here, the Toeplitz operator $T_\psi : H^2 \to H^2$ with symbol $\psi \in L^\infty$ is defined by

$$T_\psi f = \mathbb{P}_+ \psi f, \quad \text{for } f \in H^2,$$

where $\mathbb{P}_+$ denotes the orthogonal projection of $L^2$ onto $H^2$.

Moreover, in [PK], Peller and Khruschëv also used Hankel operators to prove many general *hereditary* properties of the non-linear operator $\mathcal{A}$ of best approximation; loosely speaking, if $\varphi \in X$, then $\mathcal{A}\varphi \in X$ holds for many "large" classes of function spaces $X$ on $\mathbb{T}$.

The problem of best uniform approximation by *meromorphic* functions in $\mathbb{D}$ and its connection to Hankel (and Toeplitz) operators was also considered. In [AAK2], Adamyan, Arov and Krein showed that if $\varphi \in L^\infty$, then the $k$th singular value $s_k(H_\varphi)$ of the Hankel operator $H_\varphi$ is given by the formula

$$s_k(H_\varphi) = \operatorname{dist}_{L^\infty}(\varphi, H^\infty_{(k)}),$$

where $H^\infty_{(k)}$ denotes the collection of meromorphic functions in $\mathbb{D}$ bounded near $\mathbb{T}$ and having at most $k$ poles in $\mathbb{D}$ (counting multiplicities). Moreover, if $\varphi$ is $k$-admissible (i.e. $\|H_\varphi\|_{\mathrm{e}} < s_k(H_\varphi)$), then $\varphi$ has a unique meromorphic approximation $q \in H^\infty_{(k)}$, the function $u = s_k^{-1}(H_\varphi)(\varphi - q)$ has modulus equal to 1 a.e. on $\mathbb{T}$, the Toeplitz operator $T_u$ is Fredholm and

$$\operatorname{ind} T_u = 2k + \mu,$$

where $\mu$ denotes the multiplicity of the singular value $s_k(H_\varphi)$ of the Hankel operator $H_\varphi$. As usual, $\operatorname{ind} T$ denotes the index of Fredholm operator $T$, i.e. $\operatorname{ind} T \stackrel{\mathrm{def}}{=} \dim\ker T - \dim\ker T^*$. Moreover, for $n \geq 0$, the *singular value* $s_n(T)$ of a bounded operator $T : \mathcal{H} \to \mathcal{K}$ between two Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$ is defined by

$$s_n(T) = \inf\{\|T - R\| : R \text{ a bounded operator from } \mathcal{H} \text{ to } \mathcal{K}, \operatorname{rank} R \leq n\}.$$

Besides being of mathematical interest, the problem of best approximation by analytic and meromorphic functions is also important in applications. Since most engineering systems have several inputs and outputs, it is of interest to find analogous results in the case of matrix-valued functions. (For instance, see [F] and [Pe1].) Unfortunately, there are significant complications in the case of matrix-valued functions.

In this dissertation, we continue the study of the best approximation problem in the case of matrix-valued functions (for short, matrix functions). In order to appropriately discuss our results, we first introduce notation and recall several results.

## 1.1 Best and superoptimal approximation in $H^\infty_{(k)}$

Throughout, $\mathbb{M}_{m,n}$ denotes the space of $m \times n$ matrices equipped with the operator norm $\|\cdot\|_{\mathbb{M}_{m,n}}$ (under the usual identification of elements in $\mathbb{M}_{m,n}$ and operators from $\mathbb{C}^n$ to $\mathbb{C}^m$.) In the case $m = n$, we use the notation $\mathbb{M}_n$ to denote $\mathbb{M}_{n,n}$. For $A \in \mathbb{M}_{m,n}$ and $j \geq 0$, we denote by $s_j(A)$ the $j$th-singular value of $A$, i.e. the distance (under the operator norm) from $A$ to the set of matrices of rank at most $j$.

For any space $X$ of scalar functions on $\mathbb{T}$, $X(\mathbb{M}_{m,n})$ denotes the space of $m \times n$ matrix-valued functions on $\mathbb{T}$ whose entries belong to $X$. We also use the notation $X(\mathbb{C}^n)$ for $X(\mathbb{M}_{n,1})$. In the case of the space of (essentially) bounded $m \times n$ matrix functions $L^\infty(\mathbb{M}_{m,n})$, we use the norm $\|\cdot\|_{L^\infty(\mathbb{M}_{m,n})}$ defined by

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

A matrix function $B \in H^\infty(\mathbb{M}_n)$ is called a finite *Blaschke-Potapov product* if it admits a factorization of the form

$$B = UB_1B_2 \dots B_m,$$

where $U$ is a unitary matrix and, for each $1 \leq j \leq m$,

$$B_j = \frac{z - \lambda_j}{1 - \bar{\lambda}_j z} P_j + (I - P_j)$$

for some $\lambda_j \in \mathbb{D}$ and orthogonal projection $P_j$ on $\mathbb{C}^n$. The *degree* of the Blaschke-Potapov product $B$ is defined to be

$$\deg B \stackrel{\text{def}}{=} \sum_{j=1}^{m} \operatorname{rank} P_j.$$

It turns out that every subspace $\mathcal{L}$ of finite codimension invariant under multiplication by $z$ on $H^2(\mathbb{C}^n)$ is of the form $BH^2(\mathbb{C}^n)$ for some Blaschke-Potapov product of finite degree $\operatorname{codim} \mathcal{L}$ (e.g. see Lemma 5.1 in Chapter 2 of [Pe1]).

Let $H^\infty_{(k)}(\mathbb{M}_{m,n})$ denote the collection of matrix functions $Q \in L^\infty(\mathbb{M}_{m,n})$ that admit a factorization of the form $Q = FB^*$ for some $F \in H^\infty(\mathbb{M}_{m,n})$ and Blaschke-Potapov product $B$ of degree at most $k$. Alternatively, the class $H^\infty_{(k)}(\mathbb{M}_{m,n})$ consists of matrix functions $\Psi \in L^\infty(\mathbb{M}_{m,n})$ which can be written in the form $\Psi = R + F$ for some $F \in H^\infty(\mathbb{M}_{m,n})$ and some rational $m \times n$ matrix function $R$ with poles in $\mathbb{D}$ and whose McMillan degree is at most $k$. Here we do not need the notion of McMillan degree and so we refer the interested reader to consult Chapter 2 in [Pe1] for more information.

**Definition 1.1.1.** Let $k \geq 0$. Given an $m \times n$ matrix-valued function $\Phi \in L^\infty(\mathbb{M}_{m,n})$, we say that $Q$ is a *best approximation in $H^\infty_{(k)}(\mathbb{M}_{m,n})$ to $\Phi$* if $Q$ belongs to $H^\infty_{(k)}(\mathbb{M}_{m,n})$ and

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

Note that, by a compactness argument, a matrix function $\Phi \in L^\infty(\mathbb{M}_{m,n})$ always has a best approximation in $H^\infty_{(k)}(\mathbb{M}_{m,n})$. In other words, the set

$$\Omega_0^{(k)}(\Phi) \stackrel{\text{def}}{=} \left\{ Q \in H^\infty_{(k)}(\mathbb{M}_{m,n}) : Q \text{ minimizes } \operatorname*{ess\,sup}_{\zeta \in \mathbb{T}} \|\Phi(\zeta) - Q(\zeta)\|_{\mathbb{M}_{m,n}} \right\}$$

is non-empty.

As in the case of scalar-valued bounded functions, Hankel operators on Hardy spaces are a very useful tool in the study of best approximation by matrix functions in $H^\infty_{(k)}(\mathbb{M}_{m,n})$. For a matrix function $\Phi \in L^\infty(\mathbb{M}_{m,n})$, we define the *Hankel operator* $H_\Phi$ by

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

where $\mathbb{P}_-$ denotes the orthogonal projection of $L^2(\mathbb{C}^m)$ onto $H^2_-(\mathbb{C}^m) = L^2(\mathbb{C}^m) \ominus H^2(\mathbb{C}^m)$. By a matrix analog of a theorem of Adamyan, Arov and Krein (see [Tr1] or Section 3 of Chapter 4 in [Pe1]), it is known that

$$\operatorname{dist}_{L^\infty(\mathbb{M}_{m,n})}(\Phi, H^\infty_{(k)}(\mathbb{M}_{m,n})) = s_k(H_\Phi). \tag{1.1.1}$$

However, in contrast to the case of scalar-valued functions, a best approximation is rarely unique (even under the assumption $s_k(H_\Phi) > \|H_\Phi\|_{\mathrm{e}}$).

**Example 1.1.2.** Consider the problem of best approximation of the matrix function

$$\Phi = \begin{pmatrix} 1/z^2 & \mathbb{O} \\ \mathbb{O} & \mathbb{O} \end{pmatrix}$$

in $H^\infty_{(k)}(\mathbb{M}_2)$ for $k = 0$ and $k = 1$, respectively. In this case, it is easy to see that any matrix function of the form

$$F_k = \begin{pmatrix} \mathbb{O} & \mathbb{O} \\ \mathbb{O} & f_k \end{pmatrix},$$

is a best approximation in $H^\infty_{(k)}(\mathbb{M}_2)$ to $\Phi$, where $f_k$ is any (scalar-valued) function in $H^\infty_{(k)}$ such that $\|f_k\|_\infty \le 1$, for $k = 0, 1$.

Notice that in Example 1.1.2, for $k = 0$ and $k = 1$, it is likely to say that the "very best" approximant to $\Phi$, in the respective $H^\infty_{(k)}(\mathbb{M}_2)$ class, should be the zero matrix function. Therefore it is natural to impose additional conditions in order to distinguish a "very best" approximant among all best approximants to $\Phi$ in $H^\infty_{(k)}(\mathbb{M}_2)$ for each $k = 0, 1$. This idea led N.J. Young [Y] to the notion of "superoptimal" approximation in the case $k = 0$, i.e. the case of *analytic* approximation.

**Definition 1.1.3.** Let $k \ge 0$ and $\Phi \in L^\infty(\mathbb{M}_{m,n})$. For $j > 0$, define the sets

$$\Omega_j^{(k)}(\Phi) \stackrel{\text{def}}{=} \left\{ Q \in \Omega_{j-1}^{(k)}(\Phi) : Q \text{ minimizes } \operatorname*{ess\,sup}_{\zeta\in\mathbb{T}} s_j(\Phi(\zeta) - Q(\zeta)) \right\}.$$

We say that $Q$ is a *superoptimal approximation of* $\Phi$ in $H^\infty_{(k)}(\mathbb{M}_{m,n})$ if $Q$ belongs to $\bigcap_{j\ge 0} \Omega_j^{(k)}(\Phi) = \Omega^{(k)}_{\min\{m,n\}-1}(\Phi)$ and in this case we define the *superoptimal singular values of* $\Phi$ *in* $H^\infty_{(k)}(\mathbb{M}_{m,n})$ by

$$t_j^{(k)}(\Phi) = \operatorname*{ess\,sup}_{\zeta\in\mathbb{T}} s_j((\Phi - Q)(\zeta)) \text{ for } j \ge 0.$$

In the case $k = 0$, we also use the notations $\Omega_j(\Phi)$ and $t_j(\Phi)$ to denote $\Omega_j^{(0)}(\Phi)$ and $t_j^{(0)}(\Phi)$, respectively, for $j \ge 0$.

The uniqueness of a superoptimal analytic approximation $F$ for matrix functions $\Phi \in (H^\infty + C)(\mathbb{M}_{m,n})$ was first established by Peller and Young in [PY1]. (Recall that $H^\infty + C$ denotes the closed subalgebra of $L^\infty$ that consists of functions of the form $f + g$ with $f \in H^\infty$ and $g \in C$.) Their method was based on a diagonalization of the error term $\Phi - F$; the so-called *thematic factorization* (see Section 1.3).

In [Tr2], Treil proved that a unique superoptimal *meromorphic* approximation $Q$ in $H^\infty_{(k)}(\mathbb{M}_{m,n})$ exists for matrix functions $\Phi \in (H^\infty + C)(\mathbb{M}_{m,n})$ such that $s_k(H_\Phi) < s_{k-1}(H_\Phi)$ by using geometric arguments and operator weights. Shorty after, Peller and Young also proved this result in [PY2] by using a diagonalization argument that also constructs (in principle) the superoptimal meromorphic approximant in $H^\infty_{(k)}(\mathbb{M}_{m,n})$. More generally, it is now known (see Section 17 of Chapter 14 in [Pel]) that if $\Phi \in L^\infty(\mathbb{M}_{m,n})$ is $k$-admissible, then $\Phi$ has a unique superoptimal approximation $Q$ in $H^\infty_{(k)}(\mathbb{M}_{m,n})$ and

$$s_j((\Phi - Q)(\zeta)) = t_j^{(k)}(\Phi) \text{ holds for a.e. } \zeta \in \mathbb{T},\ j \geq 0.$$

A matrix function $\Phi \in L^\infty(\mathbb{M}_{m,n})$ is *$k$-admissible* if $s_k(H_\Phi) < s_{k-1}(H_\Phi)$ and $\|H_\Phi\|_{\mathrm{e}}$ is strictly less than the smallest nonzero number in the set $\{t_j^{(k)}(\Phi)\}_{j\geq 0}$. (Note that the statement on the singular values of the Hankel operator is vacuous when $k = 0$ and the essential norm of the Hankel operator $H_\Phi$ equals zero for continuous matrix functions $\Phi$.) We also refer to 0-admissible matrix functions as *admissible*. Moreover, in the case of scalar-valued functions, to say that a function $\varphi$ is $k$-admissible simply means that $\|H_\varphi\|_{\mathrm{e}} < s_k(H_\varphi)$ and $s_k(H_\varphi) < s_{k-1}(H_\varphi)$. Thus the notion of $k$-admissibility for matrix functions (and so the uniqueness of a superoptimal approximation) is a natural extension of the notion (and results) for scalar functions mentioned at the beginning of this chapter.

## 1.2 Badly and very badly approximable matrix functions

A matrix function $G \in L^\infty(\mathbb{M}_{m,n})$ is called *badly approximable* if the zero matrix function is a best approximation in $H^\infty(\mathbb{M}_{m,n})$ of $G$. If, in addition, the zero matrix function is a superoptimal approximation in $H^\infty(\mathbb{M}_{m,n})$ of $G$, we say $G$ is *very badly approximable*. In particular, a matrix function is very badly approximable if and only if it is the difference of a bounded matrix function and its superoptimal approximant in $H^\infty(\mathbb{M}_{m,n})$.

For $\Phi \in L^\infty(\mathbb{M}_{m,n})$ and fixed $\ell \geq 0$, it is easy to observe from Definition 1.1.3 that if $Q \in \Omega_\ell^{(k)}(\Phi)$, then the zero matrix function belongs to $\Omega_\ell(\Phi - Q)$, $t_j(\Phi - Q) = t_j^{(k)}(\Phi)$ and $Q + F \in \Omega_j^{(k)}(\Phi)$ whenever $F \in \Omega_j(\Phi - Q)$ for $0 \leq j \leq \ell$. Therefore, if $\Phi$ has superoptimal approximant $Q$ in $H^\infty_{(k)}(\mathbb{M}_{m,n})$, then $\Phi - Q$ is very badly approximable.

## 1.3 Thematic factorizations

In [PY1], very badly approximable matrix functions in $(H^\infty + C)(\mathbb{M}_{m,n})$ were characterized algebraically in terms of thematic factorizations. It turns out that this same algebraic characterization remains valid for very badly approximable matrix functions which are only admissible. To appropriately discuss these factorizations, we first recall several definitions and refer the reader to [AP1] and [PT2] for other algebraic and geometric characterizations of admissible very badly approximable matrix functions.

Let $I_n$ denote the matrix function that equals the $n \times n$ identity matrix on $\mathbb{T}$. Recall that a matrix function $\Theta \in H^\infty(\mathbb{M}_{m,n})$ is called *inner* if $\Theta^*\Theta = I_n$ a.e. on $\mathbb{T}$. A matrix function $F \in H^\infty(\mathbb{M}_{m,n})$ is called *outer* if $FH^2(\mathbb{C}^n)$ is dense in $H^2(\mathbb{C}^m)$. Lastly, a matrix function $\Theta \in H^\infty(\mathbb{M}_{m,n})$ is called *co-outer* whenever the transposed function $\Theta^t$ is outer. In what follows, we shall make use of the following fact concerning co-outer matrix functions (see Chapter 14 of [Pe1] for a proof).

**Fact 1.3.1.** *Suppose $\Theta$ is a co-outer matrix function in $H^\infty(\mathbb{M}_{m,n})$. If $\eta \in L^2(\mathbb{C}^n)$ is such that $\Theta\eta \in H^2(\mathbb{C}^m)$, then $\eta \in H^2(\mathbb{C}^n)$.*

Let $n \geq 2$ and $0 < r < n$. For an $n \times r$ inner and co-outer matrix function $\Upsilon$, it is known that there is an $n \times (n - r)$ inner and co-outer matrix function $\Theta$ such that

$$V = (\,\Upsilon\; \overline{\Theta}\,) \tag{1.3.1}$$

is a unitary-valued matrix function on $\mathbb{T}$. Functions of the form (1.3.1) are called *r-balanced.* We refer the reader to Chapter 14 in [Pe1] for a detailed presentation of many interesting properties of $r$-balanced matrix functions.

Our main interest lies with 1-balanced matrix functions, which are also referred to as *thematic.*

**Definition 1.3.2.** A *partial thematic factorization* of an $m \times n$ matrix function is a factorization of the form

$$W_0^* \cdot \ldots \cdot W_{r-1}^* \begin{pmatrix} t_0 u_0 & \mathbb{O} & \ldots & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & t_1 u_1 & \ldots & \mathbb{O} & \mathbb{O} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbb{O} & \mathbb{O} & \ldots & t_{r-1} u_{r-1} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \ldots & \mathbb{O} & \Psi \end{pmatrix} V_{r-1}^* \cdot \ldots \cdot V_0^* \tag{1.3.2}$$

where the numbers $t_0, t_1, \ldots, t_{r-1}$ satisfy

$$t_0 \geq t_1 \geq \ldots \geq t_{r-1} > 0; \tag{1.3.3}$$

the function $u_j$ is unimodular and such that the Toeplitz operator $T_{u_j}$ is Fredholm with positive index, for $0 \leq j \leq r-1$; the $n \times n$ matrix function $V_j$ and $m \times m$ matrix function $W_j$ have the form

$$V_j = \begin{pmatrix} I_j & \mathbb{O} \\ \mathbb{O} & \breve{V}_j \end{pmatrix} \text{ and } W_j = \begin{pmatrix} I_j & \mathbb{O} \\ \mathbb{O} & \breve{W}_j \end{pmatrix}, \tag{1.3.4}$$

for some thematic matrix functions $\breve{V}_j$ and $\breve{W}_j^t$, respectively, for $1 \leq j < r-1$; $V_0$ and $W_0^t$ are thematic matrix functions; and the matrix function $\Psi$ satisfies

$$\|\Psi\|_{L^\infty(\mathbb{M}_{m-r,n-r})} \leq t_{r-1} \text{ and } \|H_\Psi\| < t_{r-1}. \tag{1.3.5}$$

The positive integers $k_0, \ldots, k_{r-1}$ defined by

$$k_j = \operatorname{ind} T_{u_j}, \quad \text{for } 0 \le j \le r-1,$$

are called the *thematic indices* associated with the factorization in (1.3.2).

As usual, if $r = m$ or $r = n$, we use the convention that the corresponding row or column does not exist.

**Definition 1.3.3.** A *thematic factorization* of an $m \times n$ matrix function is a partial thematic factorization of the form (1.3.2) in which $\Psi$ is identically zero.

It can be shown that any admissible very badly approximable matrix function admits a thematic factorization. Conversely, any matrix function of the form (1.3.2) with $\Psi = \mathbb{O}$ is a very badly approximable matrix function whose $j$th-superoptimal singular value equals $t_j$ for $0 \le j \le r-1$. Actually, to deduce the latter, the assumption in (1.3.3) is essential as the following example illustrates.

**Example 1.3.4.** Consider the matrix function

$$G = W^* \begin{pmatrix} \bar{z} & \mathbb{O} \\ \mathbb{O} & 3\bar{z} \end{pmatrix} V^*, \quad \text{where } V = W^t = \frac{1}{\sqrt{2}} \begin{pmatrix} z & -1 \\ 1 & \bar{z} \end{pmatrix}.$$

Obviously, $\|G(\zeta)\| = 3$ for a.e. $\zeta \in \mathbb{T}$ and so $\|H_G\| \le \|G\|_{L^\infty} = 3$. We claim $G$ *is not badly approximable.* Assume, on the contrary, that $G$ is badly approximable. In this case, the continuity of $G$ guarantees that the Hankel operator $H_G$ has a (non-zero) maximizing vector $f \in H^2(\mathbb{C}^2)$. In this case, $\|H_G\| = 3$ (e.g. see (1.1.1) above), $Gf \in H^2_-(\mathbb{C}^2)$ and so

$$\|WGf\|_2 = \|Gf\|_2 = \|H_G f\|_2 = 3\|f\|_2 = 3\|V^* f\|_2 \qquad (1.3.6)$$

because the matrix functions $V$ and $W$ are unitary-valued. We see from (1.3.6) that

$$\|v^* f\|_2^2 + 9\|\Theta^t f\|_2^2 = 9\|V^* f\|_2^2 = 9\|v^* f\|_2^2 + 9\|\Theta^t f\|_2^2$$

and so $v^* f = \mathbb{O}$, where $v$ and $\bar{\Theta}$ denote the first and second column functions of $V$. Thus, the fact that

$$Gf = W^* \begin{pmatrix} \mathbb{O} \\ 3\bar{z}\Theta^t f \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} -3\bar{z}\Theta^t f \\ 3\Theta^t f \end{pmatrix}$$

belongs to $H^2_-(\mathbb{C}^2)$ implies that $\Theta^t f = H^2 \cap H^2_- = \{\mathbb{O}\}$ and so $f = VV^* f = \mathbb{O}$, a contradiction. This shows that $G$ cannot be badly approximable.

Let us explain briefly why factorizations of the form

$$\mathcal{W}^* \begin{pmatrix} U & \mathbb{O} \\ \mathbb{O} & \Psi \end{pmatrix} \mathcal{V}^* \tag{1.3.7}$$

are quite useful when dealing with Hankel and Toeplitz operators, where $\mathcal{V}$ and $\mathcal{W}^t$ are $r$-balanced and $U$ is a unitary-valued $r \times r$ matrix function on $\mathbb{T}$ (c.f. (1.3.2)). There are two simple (but fundamental) observations that can be deduced about matrix functions that admit factorizations of this form.

**Theorem 1.3.5.** *Let $G \in L^\infty(\mathbb{M}_{m,n})$ be a matrix function of the form* (1.3.7), *where*

$$\mathcal{V} = (\, \Upsilon \; \overline{\Theta} \,) \quad \textit{and} \quad \mathcal{W}^t = (\, \Omega \; \overline{\Xi} \,). \tag{1.3.8}$$

1. *If $\ker T_{U^*}$ is trivial, then $\ker T_{G^*} = \Xi \ker T_{\Psi^*}$.*

2. *If the Hankel operator $H_G$ has a maximizing vector $f \in H^2(\mathbb{C}^n)$, then $\Theta^t f = \mathbb{O}$ or $\|\Psi\|_{L^\infty} \geq 1$.*

As usual, the Toeplitz operator $T_\Phi : H^2(\mathbb{C}^m) \to H^2(\mathbb{C}^n)$ with symbol $\Phi \in L^\infty(\mathbb{M}_{m,n})$ is defined by

$$T_\Phi f = \mathbb{P}_+ \Phi f, \quad \text{for } f \in H^2(\mathbb{C}^n),$$

where $\mathbb{P}_+$ denotes the orthogonal projection of $L^2(\mathbb{C}^m)$ onto $H^2(\mathbb{C}^m)$.

*Proof.* It is easy to see that $\Omega^t \ker T_{G^*} \subset \ker T_{U^*}$. Therefore, if $f \in \ker T_{G^*}$ and $\ker T_{U^*}$ is trivial, then $\Omega^t f = \mathbb{O}$ and so $f = \mathcal{W}^*\mathcal{W} f = \Xi\eta$, where $\eta = \Xi^* f$. In view of Fact 1.3.1, $\eta \in H^2$. By the same reasonning, it is easy to see that $G^* f = \overline{\Theta}\Psi^*\eta$ belongs to $H^2_-(\mathbb{C}^m)$ if and only if $\eta = \Xi^* f$ belongs to $\ker T_{\Psi^*}$. Thus, we conclude that $f \in \ker T_{G^*}$ if and only if there is a function $\eta \in \ker T_{\Psi^*}$ such that $f = \Xi\eta$. This proves the first assertion.

Suppose now that $H_G$ has a maximizing vector $f \in H^2(\mathbb{C}^n)$. In this case, $Gf \in H^2_-(\mathbb{C}^m)$ and so $\|WGf\|_2 = \|Gf\|_2 = \|H_G f\|_2 = \|f\|_2 = \|V^* f\|_2$. This implies that the column function $b = \Theta^t f$ satisfies $\|\Psi b\|_2 = \|b\|_2$ from which we conclude that the second assertion holds. □

The arguments used in the proof of the previous theorem are contained throughout the literature and are now standard. Actually, the first argument was used by Peller and Young (in [PY1] when $r = 1$) to prove the following very important fact:

**Fact 1.3.6.** *Let $m \le n$. If $G \in L^\infty(\mathbb{M}_{m,n})$ is an admissible very badly approximable function and $t_{m-1}(G) > 0$, then the Toeplitz operator $T_{zG}$ has dense range in $H^2(\mathbb{C}^m)$.*

To prove Fact 1.3.6, Peller and Young used the first assertion in Theorem 1.3.5 (applied to the function $zG$ instead). That assertion allowed them to reduced the verification of density of Range $T_{zG}$ to a similar verification for a matrix function of smaller size and so the result followed by an induction argument.

The converse to Fact 1.3.6 also holds for unitary-valued matrix functions $U$ on $\mathbb{T}$ such that $\|H_U\|_e < 1$. This was proved in [AP1] (see also Chapter 14 in [Pe1]).

**Fact 1.3.7.** *Let $U$ be an $n \times n$ unitary-valued matrix function such that $\|H_U\|_e < 1$. Then $U$ is very badly approximable if and only if the Toeplitz operator $T_{zU}$ has dense range in $H^2(\mathbb{C}^n)$. In this case, the Toeplitz operator $T_U$ is Fredholm and*

$$\operatorname{ind} T_U = \dim \ker T_U .$$

To conclude this introductory chapter, let us mention other notation and terminology that is used in the following chapters.

## 1.4 Other notation and terminology.

Throughout, we also use the following notation and terminology:

$\boldsymbol{m}$ denotes normalized Lebesgue measure on $\mathbb{T}$ so that $\boldsymbol{m}(\mathbb{T}) = 1$;

for $1 \le p \le \infty$, $H^p$ and $H^p_0$ denote the spaces of $L^p$ functions on $\mathbb{T}$ whose Fourier coefficients of negative and non-positive index vanish, respectively;

$C$ denotes the Banach algebra of continuous functions on $\mathbb{T}$;

$\mathbb{O}$ denotes matrix-valued function which equals the zero matrix (whose size will be clear from the context) on $\mathbb{T}$;

$\boldsymbol{S}_p^{m,n}$ denotes the space of $m \times n$ matrices equipped with the Schatten-von Neumann norm $\|\cdot\|_{\boldsymbol{S}_p^{m,n}}$, i.e. for $A \in \mathbb{M}_{m,n}$

$$\|A\|_{\boldsymbol{S}_\infty^{m,n}} \stackrel{\text{def}}{=} \|A\|_{\mathbb{M}_{m,n}} \text{ and } \|A\|_{\boldsymbol{S}_p^{m,n}} \stackrel{\text{def}}{=} \left(\sum_{j\ge 0} s_j^p(A)\right)^{1/p} \quad \text{for } 1 \le p < \infty;$$

$\boldsymbol{S}_p^n \stackrel{\text{def}}{=} \boldsymbol{S}_p^{n,n}$ for $1 \le p \le \infty$;

for $1 \le p \le \infty$, $L^p(\mathbb{M}_{m,n})$ denotes the space of $m \times n$ matrix-valued functions on $\mathbb{T}$ whose entries belong to $L^p$ and equipped with the norm $\|\cdot\|_{L^p(\mathbb{M}_{m,n})}$, where

$$\|F\|^p_{L^p(\mathbb{M}_{m,n})} = \int_{\mathbb{T}} \|F(\zeta)\|^p_{\mathbb{M}_{m,n}} \boldsymbol{dm}(\zeta) \quad \text{for } 1 \le p < \infty$$

$$\text{and} \quad \|F\|_{L^\infty(\mathbb{M}_{m,n})} = \operatorname*{ess\,sup}_{\zeta \in \mathbb{T}} \|F(\zeta)\|_{\mathbb{M}_{m,n}};$$

more generally, if $X$ is a normed space of functions on $\mathbb{T}$ with norm $\|\cdot\|_X$, then $X(\boldsymbol{S}_p^{m,n})$ denotes the space of $m \times n$ matrix functions whose entries belong to $X$ and define

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

if $X$ and $Y$ are normed spaces and $T : X \to Y$ is a bounded linear operator, we say that a non-zero vector $x \in X$ is a *maximizing vector of* $T$ whenever $\|Tx\|_Y = \|T\| \cdot \|x\|_X$;

$\mathcal{H} \ominus \mathcal{K}$ denotes the orthogonal complement of a subspace $\mathcal{K}$ of a Hilbert space $\mathcal{H}$;

$\operatorname{ind} T$ denotes the index of Fredholm operator $T$, i.e. $\operatorname{ind} T = \dim\ker T - \dim\ker T^*$;

if $\mathcal{H}$ and $\mathcal{K}$ are Hilbert spaces and $T : \mathcal{H} \to \mathcal{K}$ is a bounded linear operator, the *singular values* $s_n(T)$, $n \ge 0$, *of* $T$ are defined by

$$s_n(T) = \inf\{\|T - R\| : R \text{ a bounded operator from } \mathcal{H} \text{ to } \mathcal{K}, \operatorname{rank} R \le n\}.$$

and the *essential norm of* $T$ is defined by

$$\|T\|_{\mathrm{e}} = \inf\{\|T - K\| : K \text{ is a compact operator }\}.$$

# Chapter 2

# Monotone thematic factorizations

## 2.1 Introduction

In [PY1], it was observed that thematic indices (see Definition 1.3.2) depend on the choice of the thematic factorization. However, it was conjectured there that the sum of the thematic indices associated with any thematic factorization of a given very badly approximable matrix function $\Phi$ depends only on $\Phi$ (and is therefore independent of the choice of a thematic factorization) whenever $\Phi$ belongs to $(H^\infty + C)(\mathbb{M}_{m,n})$. This conjecture was settled in the affirmative shortly after in [PY2]. Moreover, it was shown in [PT1] that this conjecture remains valid for matrix functions $\Phi$ which are merely admissible.

The result concerning the sum of thematic indices of $\Phi$ leads to the question: *Can one arbitrarily distribute this sum among thematic indices of $\Phi$ by choosing an appropriate thematic factorization?* A partial answer was given in [AP2] in terms of monotone partial thematic factorizations.

**Definition 2.1.1.** A partial thematic factorization of the form (1.3.2) is called *monotone non-increasing* (or non-decreasing) if for any superoptimal singular value $t$, such that $t \geq t_{r-1}$, the thematic indices $k_j, k_{j+1}, \ldots, k_s$ that correspond to all of the superoptimal singular values that are equal to $t$ form a monotone non-increasing sequence (or non-decreasing sequence).

*Remark* 2.1.2. Note that only monotone *non-increasing* partial thematic factorizations were considered in [AP2].

The following result was established in [AP2].

**Fact 2.1.3.** *If $\Phi \in L^\infty(\mathbb{M}_{m,n})$ is an admissible very badly approximable matrix function, then $\Phi$ possesses a* monotone non-increasing *thematic factorization. Moreover, the indices of any monotone non-increasing thematic factorization are uniquely determined by $\Phi$.*

Hence, one cannot arbitrarily distribute the sum of thematic indices of an admissible very badly approximable matrix function among thematic indices in *non-increasing order.* Indeed, thematic indices are uniquely determined when arranged in this way.

We refer the reader to [Pel] for more information and proofs of all previously mentioned facts concerning thematic factorizations.

Before explaining what is done in this chapter, let us consider the following example. Let $G$ be the $2 \times 2$-matrix function defined by

$$G = \frac{1}{\sqrt{2}} \begin{pmatrix} \bar{z} & -1 \\ 1 & z \end{pmatrix} \begin{pmatrix} \bar{z}^2 & \mathbb{O} \\ \mathbb{O} & \bar{z} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \bar{z}^3 & -\bar{z} \\ \bar{z}^2 & 1 \end{pmatrix}.$$

Clearly, $G$ is a very badly approximable continuous (and so admissible) function in its non-increasing monotone thematic factorization with thematic indices 2 and 1. We now ask the question: Does $G$ admit a monotone *non-decreasing* thematic factorization?

It is easy to verify that $G$ can also be factored as

$$G = \begin{pmatrix} -1 & \mathbb{O} \\ \mathbb{O} & 1 \end{pmatrix} \begin{pmatrix} \bar{z} & \mathbb{O} \\ \mathbb{O} & \bar{z}^2 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} -\bar{z}^2 & 1 \\ 1 & z^2 \end{pmatrix}$$

demonstrating that $G$ does admit a monotone non-decreasing thematic factorization with thematic indices 1 and 2. Thus, the natural question arises: *Does every admissible very badly approximable matrix function admit a monotone* non-decreasing *thematic factorization? If so, are the thematic indices in any such factorization uniquely determined by the matrix function itself?* We succeed in providing answers to these questions.

We begin Section 2.2 introducing sufficient conditions under which the Toeplitz operator induced by a unimodular function is invertible. For the reader's convenience, we also state some well-known theorems on the factorization of certain unimodular functions.

In Section 2.3, we establish new results on badly approximable matrix functions. We prove that given a (partial) thematic factorization of a badly approximable matrix function $G$ whose "second" thematic index equals $k$ and an integer $j$ satisfying $1 \leq j \leq k$, it is possible to find a new (partial) thematic factorization of $G$ in which the "first" new thematic index equals $j$. We then give further analysis of the "lower block" obtained in this new factorization of $G$. It is shown that, under rather natural assumptions, the first thematic index of the new lower block is indeed the first thematic index of $G$ in the originally given thematic factorization.

Once these results are available, we argue in Section 2.4 that there is an abundant number of thematic factorizations of an arbitrary (admissible) very badly approximable matrix function. We begin by proving the existence of a monotone non-decreasing thematic factorization for such matrix functions. In contrast to monotone non-increasing thematic factorizations, it is shown that the thematic indices appearing in a monotone non-decreasing thematic factorization are *not* uniquely determined by the matrix function itself. Moreover, we obtain every possible sequence of thematic indices in the case of $2 \times 2$ unitary-valued matrix functions. We further prove that one can obtain various thematic factorizations from a monotone non-increasing thematic factorization while preserving "some structure" of the thematic indices in the case of $m \times n$ matrix functions with $\min\{m, n\} \geq 2$. We close the section by illustrating this with a simple example.

In Section 2.5, we provide an algorithm and demonstrate with an example that the algorithm yields a thematic factorization for any specified sequence of thematic indices of an arbitrary admissible very badly approximable unitary-valued $2 \times 2$ matrix function.

## 2.2 Invertibility of Toeplitz operators and factorization of certain unimodular functions

In this section, we include some useful and perhaps well-known (to those who work with Toeplitz and Hankel operators on the Hardy space $H^2$) results regarding scalar functions that are needed throughout the paper. We begin by introducing sufficient conditions for which a Toeplitz operator $T_w$, where $w$ is a unimodular function on $\mathbb{T}$ (i.e. $w$ has modulus equal to 1 a.e. on $\mathbb{T}$), is invertible on $H^2$. Although a complete description of unimodular functions $w$ for which $T_w$ is invertible is given by the well-known theorem of Devinatz and Widom, the sufficient condition given in Theorem 2.2.2 below is easier to verify.

**Lemma 2.2.1.** *Let $0 < p \leq \infty$. If $h \in H^p$ and $1/h \in H^2$, then the Toeplitz operator $T_{\bar{h}/h}$ has trivial kernel.*

*Proof.* Suppose that $p \geq 2$. Let $f \in \ker T_{\bar{h}/h}$. Since $H^2_- = L^2 \ominus H^2 = \bar{z}\overline{H^2}$, then

$f/h \in (1/\bar{h})\bar{z}\overline{H^2}$. It follows that $f/h \in H^1 \cap \bar{z}\overline{H^1}$ and therefore $\ker T_{\bar{h}/h}$ must be trivial, because $H^1 \cap \bar{z}\overline{H^1}$ is trivial.

Suppose now that $h \in H^p \setminus H^2$ with $0 < p < 2$. Assume, for the sake of contradiction, that $\ker T_{\bar{h}/h}$ is non-trivial. In this case, a simple argument of Hayashi (see the proof of Lemma 5 in [Ha]) shows that there is an outer function $k \in H^2$ such that $\bar{h}/h = \bar{k}/k$, and so there is a $c \in \mathbb{R}$ such that $h = ck$, a contradiction to the assumption that $h \notin H^2$. Thus $T_{\bar{h}/h}$ must have trivial kernel. □

**Theorem 2.2.2.** *Suppose that $h \in H^2$ and $1/h \in H^2$. Then the Toeplitz operator $T_{\bar{h}/h}$ has trivial kernel and dense range. In particular, if $T_{\bar{h}/h}$ is Fredholm, then $T_{\bar{h}/h}$ is invertible.*

*Proof.* By Lemma 2.2.1, we know that $T_{\bar{h}/h}$ has trivial kernel. Now, $h \in H^2$ and $1/h \in H^2$ imply that $h$ is an outer function, and so the fact that $T_{\bar{h}/h}$ has dense range follows from Theorem 4.4.10 in [Pe1]. The rest is obvious. □

We now state a useful converse to Theorem 2.2.2.

**Fact 2.2.3.** *If $w$ is a unimodular function on $\mathbb{T}$ such that $T_w$ is invertible on $H^2$, then $w$ admits a factorization of the form $w = \bar{h}/h$ for some outer function $h$ such that both $h$ and $1/h$ belong to $H^p$ for some $2 < p \leq \infty$.*

This result can be deduced from the theorem of Devinatz and Widom mentioned earlier. A proof can be found in Chapter 3 of [Pe1].

We now state two useful, albeit immediate, implications of Fact 2.2.3.

**Corollary 2.2.4.** *Suppose that $h$ and $1/h$ belong to $H^2$. If the Toeplitz operator $T_{\bar{h}/h}$ is Fredholm, then $h$ and $1/h$ belong to $H^p$ for some $2 < p \leq \infty$.*

**Corollary 2.2.5.** *Let $u$ be a unimodular function on $\mathbb{T}$. If the Toeplitz operator $T_u$ is Fredholm with index $k$, then there is an outer function $h$ such that*

$$u = \bar{z}^k \frac{\bar{h}}{h} \tag{2.2.1}$$

*and both $h$ and $1/h$ belong to $H^p$ for some $2 < p \leq \infty$.*

*Remark* 2.2.6. Even though representation (2.2.1) is very useful (e.g. in the proof of Theorem 2.3.3), it may be difficult to find the function $h$ explicitly, if needed. This is however a very easy task for unimodular functions in the space $\mathcal{R}$ of rational functions with poles outside of $\mathbb{T}$. After all, if $u \in \mathcal{R}$, then there are finite Blaschke products $B_1$ and $B_2$ such that $u = \bar{B}_1 B_2$, by the Maximum Modulus Principle. Thus, $u$ admits a representation of the form (2.2.1) with $k = \deg B_1 - \deg B_2$ for some function $h$ invertible in $H^\infty$ (which is, up to a multiplicative constant, a product of quotients of reproducing kernels of $H^2$).

We also find the classification of admissible scalar badly approximable functions mentioned in Chapter 1 and Remark 2.2.6 useful in proving the next theorem which is part of the lore of our subject.

**Theorem 2.2.7.** *Suppose that $u \in \mathcal{R}$ is a unimodular function on $\mathbb{T}$. Then $u$ is badly approximable if and only if there are finite Blaschke products $B_1$ and $B_2$ such that $\deg B_1 > \deg B_2$ and $u = \bar{B}_1 B_2$ on $\mathbb{T}$. In particular, $u$ admits the representation*

$$u = \bar{z}^k \frac{\bar{h}}{h}$$

*with $k = \operatorname{ind} T_u = \deg B_1 - \deg B_2$ for some function $h$ invertible in $H^\infty$.*

## 2.3 Badly approximable matrix functions

Recall that for $T : X \to Y$, a bounded linear operator between normed spaces $X$ and $Y$, a vector $x \in X$ is called a *maximizing vector of $T$* if $x$ is non-zero and $\|Tx\|_Y = \|T\| \cdot \|x\|_X$.

**Definition 2.3.1.** For a matrix function $\Phi \in L^\infty(\mathbb{M}_{m,n})$ such that $\|H_\Phi\|_e < \|H_\Phi\|$, we define *the space* $\mathcal{M}_\Phi$ *of maximizing vectors of* $H_\Phi$ by

$$\mathcal{M}_\Phi \stackrel{\text{def}}{=} \{f \in H^2(\mathbb{C}^n) : \|H_\Phi f\|_2 = \|H_\Phi\| \cdot \|f\|_2\}.$$

It is easy to show that $M_\Phi$ is a closed subspace which consists of the zero vector and all maximizing vectors of the Hankel operator $H_\Phi$. Moreover, $\mathcal{M}_\Phi$ always contains a maximizing vector of $H_\Phi$ because $\|H_\Phi\|_e < \|H_\Phi\|$; a consequence of the spectral theorem for bounded self-adjoint operators.

We now review results concerning badly approximable matrix functions that are used in this section. Let $G \in L^\infty(\mathbb{M}_{m,n})$ be a badly approximable function such that $\|H_G\|_e < 1$ and $\|H_G\| = 1$. In this case, it is not difficult to show that if $f$ is a non-zero function in $\mathcal{M}_G$, then $Gf \in H^2_-(\mathbb{C}^m)$, $\|G(\zeta)\|_{\mathbb{M}_{m,n}} = 1$ for a.e. $\zeta \in \mathbb{T}$, and $f(\zeta)$ is a maximizing vector of $G(\zeta)$ for a.e. $\zeta \in \mathbb{T}$ (see Theorem 3.2.3 in [Pel] for a proof).

These results can be used to deduce that $G$ admits a factorization of the form

$$W^* \begin{pmatrix} u & \mathbb{O} \\ \mathbb{O} & \Psi \end{pmatrix} V^*, \tag{2.3.1}$$

where $u = \bar{z}\bar{\theta}\bar{h}/h$, $h$ is an outer function in $H^2$, $\theta$ is an inner function, $V = (\, v \;\bar{\Theta}\,)$ and $W^t = (\, w \;\bar{\Xi}\,)$ are thematic, and $\Psi \in L^\infty(\mathbb{M}_{m-1,n-1})$ satisfies $\|\Psi\|_{L^\infty(\mathbb{M}_{m-1,n-1})} \leq 1$. Conversely, it is easy to verify that any matrix function which admits a factorization of this form is badly approximable.

For the same matrix function $G$, it can also be shown that the Toeplitz operator $T_u$ is Fredholm with positive index, $\|H_\Psi\|_e \leq \|H_G\|_e$, and the matrix functions $\Theta$ and $\Xi$ are left-invertible in $H^\infty$, i.e. there are matrix functions $A$ and $B$ in $H^\infty$ such that $A\Theta = I_{n-1}$ and $B\Xi = I_{m-1}$ hold.

We refer the reader to Chapter 2 and Chapter 14 of [Pel] for proofs of the previously mentioned results.

**Lemma 2.3.2.** *Suppose that $G \in L^\infty(\mathbb{M}_{m,n})$ is a matrix function of the form*

$$G = W^* \begin{pmatrix} u & \mathbb{O} \\ \mathbb{O} & \Psi \end{pmatrix} V^*,$$

*where $u$ is a unimodular function such that the Toeplitz operator $T_u$ is Fredholm with* $\operatorname{ind} T_u \geq 0$, $\Psi \in L^\infty(\mathbb{M}_{m-1,n-1})$ *satisfies* $\|\Psi\|_{L^\infty(\mathbb{M}_{m-1,n-1})} \leq 1$, *the matrix functions $V = (\, v \; \bar{\Theta} \,)$ and $W^t = (\, w \; \bar{\Xi} \,)$ are thematic, and the bounded analytic matrix functions $\Theta$ and $\Xi$ are left-invertible in $H^\infty$. Let $A$ and $B$ be left-inverses for $\Theta$ and $\Xi$ in $H^\infty$, respectively, and $\xi \in \ker T_\Psi$.*

1. *If $\xi$ is co-outer, then $A^t\xi + av$ is co-outer for any $a \in H^2$.*

2. *For $a \in H^2$, $A^t\xi + av$ belongs to* $\ker T_G$ *if and only if $a$ satisfies*

$$T_u a = \mathbb{P}_+(w^t B^* \Psi \xi - uv^* A^t \xi). \tag{2.3.2}$$

*Moreover, if $\xi_\# \stackrel{\text{def}}{=} A^t\xi + av$ with $a \in H^2$ satisfying* (2.3.2), *then*

3. *$\eta_\# \stackrel{\text{def}}{=} \bar{z}\bar{G}\bar{\xi}_\#$ is co-outer whenever $\bar{z}\overline{\Psi\xi}$ is co-outer, and*

4. *$\xi_\# \in \mathcal{M}_G$ whenever $\xi \in \mathcal{M}_\Psi$ and $\|H_\Psi\| = 1$.*

*Proof.* Notice that for any $a \in H^2$,

$$\Theta^t \xi_\# = \Theta^t(A^t\xi + av) = \xi, \tag{2.3.3}$$

because $A$ is a left-inverse for $\Theta$ and $V$ is unitary-valued. In particular, if the entries of $\xi$ do not have a common inner divisor, then the entries of $\xi^\#$ do not have a common inner divisor either. This establishes assertion *1.*

Although assertion *2* is contained in [PY1] and [PY2], we provide a proof for future reference. Let $\xi_{\#} = A^t\xi + av$. It follows from (2.3.3) that

$$V^*\xi_{\#} = \begin{pmatrix} v^*\xi_{\#} \\ \Theta^t\xi_{\#} \end{pmatrix} = \begin{pmatrix} v^*\xi_{\#} \\ \xi \end{pmatrix}, \tag{2.3.4}$$

and so $G\xi_{\#} = \bar{w}uv^*\xi_{\#} + \Xi\Psi\xi$. Since $W$ is unitary-valued, then $I_m = \bar{w}w^t + \Xi\Xi^*$ holds and so

$$B^* = I_m B^* = \bar{w}w^t B^* + \Xi\Xi^* B^* = \bar{w}w^t B^* + \Xi.$$

In particular, $\Xi = (I_n - \bar{w}w^t)B^*$ and so

$$G\xi_{\#} = \bar{w}u(v^*A^t\xi + a) + \Xi\Psi\xi = B^*\Psi\xi + \bar{w}(u(v^*A^t\xi + a) - w^t B^*\Psi\xi). \tag{2.3.5}$$

It follows now, from Fact 1.3.1 and (2.3.5), that $G\xi_{\#}$ belongs to $H^2_-(\mathbb{C}^m)$ if and only if $\mathbb{P}_+(u(v^*A^t\xi + a) - w^t B^*\Psi\xi) = \mathbb{O}$ because $\Psi\xi \in H^2_-(\mathbb{C}^{m-1})$ and $w$ is co-outer. Thus, $G\xi_{\#} \in H^2_-(\mathbb{C}^m)$ if and only if $T_u a = \mathbb{P}_+(w^t B^*\Psi\xi - uv^*A^t\xi)$. This completes the proof of *2*.

Henceforth, we fix a function $a_0 \in H^2$ that satisfies (2.3.2). The existence of $a_0$ follows from the fact that $T_u$ is surjective.

To prove *3*, observe that (2.3.5) can be rewritten as

$$G\xi_{\#} = B^*\Psi\xi + \bar{w}\bar{z}\bar{b}_0$$

for some $b_0 \in H^2$ because $\mathbb{P}_+(u(v^*A^t\xi + a_0) - w^t B^*\Psi\xi) = \mathbb{O}$. Let $\eta \stackrel{\text{def}}{=} \bar{z}\overline{\Psi\xi}$. Then $\eta_{\#} = \bar{z}\bar{G}\bar{\xi}^{\#} = B^t\eta + b_0 w$ and so

$$\Xi^t\eta_{\#} = \Xi^t B^t\eta + b_0\Xi^t w = \eta,$$

because $B$ is a left-inverse of $\Xi$ and $W$ is unitary-valued. Hence, $\eta_{\#}$ is co-outer whenever $\eta$ is co-outer.

Finally, we prove *4*. Since $\xi$ is a maximizing vector of $H_\Psi$ and belongs to $\ker T_\Psi$, then $\|\Psi\xi\|_2 = \|H_\Psi \xi\|_2 = \|\xi\|_2$, as $\|H_\Psi\| = 1$. Moreover, since $H_G\xi_{\#} = G\xi_{\#}$, $W$ is unitary-valued, and

$$WG\xi_{\#} = \begin{pmatrix} uv^*\xi_{\#} \\ \Psi\xi \end{pmatrix},$$

we may conclude that

$$\|H_G\xi_{\#}\|_2^2 = \|WG\xi_{\#}\|_2^2 = \|uv^*\xi_{\#}\|_2^2 + \|\Psi\xi\|_2^2 = \|v^*\xi_{\#}\|_2^2 + \|\xi\|_2^2 = \|\xi_{\#}\|_2^2$$

because (2.3.4) holds and $V$ is unitary-valued. Thus $\xi_{\#} \in \mathcal{M}_G$. □

We are now ready to state and prove the main result of this section.

**Theorem 2.3.3.** *Let $m, n \geq 2$ and $G \in L^\infty(\mathbb{M}_{m,n})$ be a matrix function of the form*

$$G = W_0^* \begin{pmatrix} u_0 & \mathbb{O} \\ \mathbb{O} & \Psi_0 \end{pmatrix} V_0^*,$$

*where $u_0$ is a unimodular function such that the Toeplitz operator $T_{u_0}$ is Fredholm with* $\operatorname{ind} T_{u_0} > 0$, $\Psi_0 \in L^\infty(\mathbb{M}_{m-1,n-1})$, *the matrix functions $V_0 = (\, v_0 \;\; \bar{\Theta}\,)$ and $W_0^t = (\, w_0 \;\; \bar{\Xi}\,)$ are thematic, and the bounded analytic matrix functions $\Theta$ and $\Xi$ are left-invertible in $H^\infty$. Suppose that*

$$\Psi_0 = W_1^* \begin{pmatrix} u_1 & \mathbb{O} \\ \mathbb{O} & \Psi_1 \end{pmatrix} V_1^* \tag{2.3.6}$$

*for some unimodular function $u_1$ such that the Toeplitz operator $T_{u_1}$ is Fredholm with* $\operatorname{ind} T_{u_1} > 0$, $\Psi_1 \in L^\infty(\mathbb{M}_{m-2,n-2})$ *such that* $\|\Psi_1\|_{L^\infty(\mathbb{M}_{m-2,n-2})} \leq 1$, *and thematic* *matrix functions* $V_1$ *and* $W_1'^t$. *Then* $G$ *admits a factorization of the form*

$$G = \mathcal{W}^* \begin{pmatrix} u & \mathbb{O} \\ \mathbb{O} & \Delta \end{pmatrix} \mathcal{V}^*$$

*for some unimodular function* $u$ *such that* $T_u$ *is Fredholm with index equal to* 1, *a badly approximable matrix function* $\Delta$ *such that* $\|\Delta\|_{L^\infty(\mathbb{M}_{m-1,n-1})} = 1$, *and thematic matrix functions* $\mathcal{V}$ *and* $\mathcal{W}^t$.

*Proof.* Let $A$ and $B$ be left-inverses of $\Theta$ and $\Xi$ in $H^\infty$, respectively, and $k_j \stackrel{\text{def}}{=} \operatorname{ind} T_{u_j}$ for $j = 0, 1$. By Corollary 2.2.5, there is an outer function $h_j$ such that

$$u_j = \bar{z}^{k_j} \frac{\bar{h}_j}{h_j}$$

and both $h_j$ and $1/h_j$ belong to $H^p$ for some $2 < p \leq \infty$, for $j = 0, 1$. Let $v_1$ denote the first column of $V_1$ and $\xi \stackrel{\text{def}}{=} z^{k_1 - 1} h_1 v_1$. It follows at once from (2.3.6) that $\Psi_0 \xi = \bar{z} \bar{h}_1 \bar{w}_1$. Thus, $\xi$ is a maximizing vector of $H_{\Psi_0}$ and belongs to $\ker T_{\Psi_0}$. In particular, the column function $\eta \stackrel{\text{def}}{=} \bar{z} \bar{\Psi}_0 \bar{\xi} = h_1 w_1$ is co-outer.

Consider the equation

$$T_{u_0} a = \mathbb{P}_+ (w_0^t B^* \Psi_0 \xi - u_0 v_0^* A^t \xi), \; a \in H^2. \tag{2.3.7}$$

It follows from the surjectivity of the Toeplitz operator $T_{u_0}$ that there is an $a_0 \in H^2$ that satisfies (2.3.7). Furthermore, we may assume without loss of generality that $z$ is not an inner divisor $a_0$; otherwise, we consider $a_0 + h_0$ instead of $a_0$.

By Lemma 2.3.2, the column function

$$\xi_\# \stackrel{\text{def}}{=} A^t \xi + a_0 v_0$$

is a maximizing vector of the Hankel operator $H_G$ and belongs to $\ker T_G$, as $\xi$ is a maximizing vector of the Hankel operator $H_{\Psi_0}$ and $\|H_{\Psi_0}\| = 1$. Since $\Theta^t \xi_\# = \xi$ and $h_1 v_1$ is co-outer, then the greatest common inner divisor of the entries of $\xi_\#$ must be an inner divisor of $z^{k_1 - 1}$ by Fact 1.3.1. Therefore, $\xi_\#$ is co-outer whenever $z$ is not an inner divisor of $\xi_\#$. On the other hand, $z$ is an inner divisor of the entries of $\xi_\#$ if and only if $z$ is an inner divisor of $a_0$. Since $z$ is not an inner divisor of $a_0$, it follows that $\xi_\#$ is co-outer.

From (2.3.5) and (2.3.7),

$$G\xi_\# = B^* \Psi_0 \xi + \bar{w}_0 \bar{z} \bar{b}_0,$$

for some $b_0 \in H^2$. Thus the function

$$\eta_\# \stackrel{\text{def}}{=} \bar{z} \bar{G} \bar{\xi}_\# = B^t \eta + b_0 w_0$$

is co-outer as well, by Lemma 2.3.2.

From the remarks following Definition 2.3.1, we deduce that

$$\|\eta_\#(\zeta)\|_{\mathbb{C}^m} = \|G\xi_\#(\zeta)\|_{\mathbb{C}^m} = \|G(\zeta)\|_{\mathbb{M}_{m,n}} \|\xi_\#(\zeta)\|_{\mathbb{C}^n} = \|\xi_\#(\zeta)\|_{\mathbb{C}^n}$$

for a.e. $\zeta \in \mathbb{T}$ because $\xi_\#$ is a maximizing vector of the Hankel operator $H_G$ and belongs to $\ker T_G$. Let $h \in H^2$ be an outer function such that $|h(\zeta)| = \|\xi_\#(\zeta)\|_{\mathbb{C}^n}$ for a.e. $\zeta \in \mathbb{T}$. We obtain that

$$\|\eta_\#(\zeta)\|_{\mathbb{C}^m} = \|\xi_\#(\zeta)\|_{\mathbb{C}^n} = |h(\zeta)| \tag{2.3.8}$$

for a.e. $\zeta \in \mathbb{T}$ and so the column functions

$$\nu \stackrel{\text{def}}{=} \frac{1}{h} \xi_\# \text{ and } \omega \stackrel{\text{def}}{=} \frac{1}{h} \eta_\#$$

are both inner and co-outer.

Consider the unimodular function $u \stackrel{\text{def}}{=} \omega^t G\nu$. It is easy to verify that

$$u = \frac{1}{h^2}(h\omega)^t G(h\nu) = \frac{1}{h^2}\bar{z}|h|^2 = \bar{z}\frac{\bar{h}}{h}, \tag{2.3.9}$$

by (2.3.8), and

$$\|H_u\|_{\mathrm{e}} = \operatorname{dist}_{L^\infty}(u, H^\infty + C) \le \operatorname{dist}_{L^\infty(\mathbb{M}_{m,n})}(G, (H^\infty + C)(\mathbb{M}_{m,n})) < 1,$$

because $\nu$ and $\omega$ are inner and $\|H_G\|_{\mathrm{e}} < 1$. Since $u$ satisfies (2.3.9) and $\|H_u\|_{\mathrm{e}} < 1$, it follows that $u$ is an admissible badly approximable scalar function, and so the Toeplitz operator $T_u$ is Fredholm with positive index (see Chapter 1) and therefore $T_{\bar{h}/h}$ is Fredholm. Since $V_0$ is unitary-valued and

$$V_0^* \xi_\# = \begin{pmatrix} v_0^* \xi_\# \\ \xi \end{pmatrix},$$

then

$$|h(\zeta)|^2 = \|\xi_\#(\zeta)\|^2_{\mathbb{C}^n} = \|V_0^*\xi_\#(\zeta)\|^2_{\mathbb{C}^n} = |(v_0^*\xi_\#)(\zeta)|^2 + \|\xi(\zeta)\|^2_{\mathbb{C}^{n-1}} \ge |h_1(\zeta)|^2$$

holds for a.e. $\zeta \in \mathbb{T}$ and so $1/h \in H^p$. By Theorem 2.2.2, $T_{\bar{h}/h}$ is invertible and so $\operatorname{ind} T_u = 1$.

Let $\mathcal{V}$ and $\mathcal{W}^t$ be thematic matrix functions whose first columns are $\nu$ and $\omega$, respectively. (The existence of such matrix functions was mentioned in Section 1.3.) Since $\omega^t G\nu = u$ is unimodular, it follows that

$$\mathcal{W}G\mathcal{V} = \begin{pmatrix} u & \mathbb{O} \\ \mathbb{O} & \Delta \end{pmatrix}$$

for some bounded matrix function $\Delta \in L^\infty(\mathbb{M}_{m-1,n-1})$ with $L^\infty$-norm equal to 1, which is necessarily badly approximable. This completes the proof. □

**Corollary 2.3.4.** *Suppose that $G$ satisfies the hypothesis of Theorem 2.3.3. If $k$ is an integer satisfying $1 \le k \le \operatorname{ind} T_{u_1}$, then $G$ admits a factorization of the form*

$$G = \mathcal{W}^* \begin{pmatrix} u & \mathbb{O} \\ \mathbb{O} & \Delta \end{pmatrix} \mathcal{V}^*$$

*for some unimodular function $u$ such that $T_u$ is Fredholm with index equal to $k$, a badly approximable matrix function $\Delta$ such that $\|\Delta\|_{L^\infty(\mathbb{M}_{m-1,n-1})} = 1$, and thematic matrix functions $\mathcal{V}$ and $\mathcal{W}^t$.*

*Proof.* Let $k$ be a fixed positive integer satisfying $k \le \operatorname{ind} T_{u_1}$. By Theorem 2.3.3, the matrix function $z^{k-1}G$ admits a factorization of the form

$$z^{k-1}G = \mathcal{W}^* \begin{pmatrix} u & \mathbb{O} \\ \mathbb{O} & \Delta \end{pmatrix} \mathcal{V}^*,$$

where $\operatorname{ind} T_u = 1$, and so

$$G = \mathcal{W}^* \begin{pmatrix} \bar{z}^{k-1}u & \mathbb{O} \\ \mathbb{O} & \bar{z}^{k-1}\Delta \end{pmatrix} \mathcal{V}^*$$

is the desired factorization. □

At this point, we are unsatisfied with the conclusion of Corollary 2.3.4. After all, it does not give any information concerning the matrix function $\Delta$. Therefore, we ask, under some reasonable assumptions, whether the "largest" possible thematic index appearing as the first thematic index in a thematic factorization of $\Delta$ should equal $\operatorname{ind} T_{u_0}$. An affirmative answer is given in Theorem 2.3.5. Prior to stating and proving Theorem 2.3.5, we introduce notation and recall some needed facts.

Suppose that $G$ is a badly approximable matrix function in $L^\infty(\mathbb{M}_{m,n})$ such that $\|H_G\|_e < 1$ and $\|H_G\| = 1$. As mentioned in the remarks following Definition 2.3.1, $G$ admits a representation of the form (2.3.1) for some unimodular function $u$ such that the Toeplitz operator $T_u$ is Fredholm with $\operatorname{ind} T_u > 0$. It turns out that there is an upper bound on the possible values of the index of $T_u$ given by

$$\iota(H_G) \stackrel{\text{def}}{=} \min\{j > 0 : \|H_{z^j G}\| < \|H_G\|\}. \tag{2.3.10}$$

Note that $\iota(H_G)$ is a well-defined non-zero positive integer and depends only on the Hankel operator $H_G$ (and not on the choice of its symbol). Moreover, there exists a (possibly distinct) factorization of $G$ of the form (2.3.1) such that $\operatorname{ind} T_u = \iota(H_G)$ and $\iota(H_\Psi) \le \iota(H_G)$. See [AP2] or Section 10 in Chapter 14 of [Pel] for proofs of these facts.

**Theorem 2.3.5.** *Let $m, n \ge 2$. Suppose that $G \in L^\infty(\mathbb{M}_{m,n})$, $\|H_G\|_e < 1$, and $G$ admits the factorizations*

$$G = W_0^* \begin{pmatrix} u_0 & \mathbb{O} \\ \mathbb{O} & \Psi \end{pmatrix} V_0^* = \mathcal{W}^* \begin{pmatrix} u & \mathbb{O} \\ \mathbb{O} & \Delta \end{pmatrix} \mathcal{V}^*, \tag{2.3.11}$$

*where $u_0$ and $u$ are unimodular functions such that the Toeplitz operators $T_{u_0}$ and $T_u$ are Fredholm with positive indices, $\Psi$ and $\Delta$ are badly approximable functions with $L^\infty$-norm equal to 1, and the matrix function $V_0$, $W_0^t$, $\mathcal{V}$, and $\mathcal{W}^t$ are thematic. If $\operatorname{ind} T_{u_0} = \iota(H_G)$, $\iota(H_\Psi) < \iota(H_G)$ and $\operatorname{ind} T_u \le \iota(H_\Psi)$, then*

$$\iota(H_\Delta) \ge \iota(H_G). \tag{2.3.12}$$

*In addition, if $\operatorname{ind} T_u = \iota(H_\Psi)$, then equality holds in (2.3.12).*

*Proof.* Let $\iota \stackrel{\text{def}}{=} \iota(H_G)$. If $\iota(H_\Delta) < \iota$, then

$$\|H_{z^{\iota-1}\Delta}\| < \|H_\Delta\| = 1 \text{ and } \operatorname{ind} T_{z^{\iota-1}u} \le \iota(H_\Psi) - (\iota - 1) \le 0.$$

It follows that the matrix function

$$z^{\iota-1}G = \mathcal{W}^* \begin{pmatrix} z^{\iota-1}u & \mathbb{O} \\ \mathbb{O} & z^{\iota-1}\Delta \end{pmatrix} \mathcal{V}^*$$

satisfies $\|H_{z^{\iota-1}G}\| < 1 = \|H_G\|$, by Lemma 14.10.7 in [Pe1], and so $\iota(H_G) \le \iota - 1$, a contradiction. This establishes (2.3.12).

Suppose that $\operatorname{ind} T_u = \iota(H_\Psi)$. Let $j \stackrel{\text{def}}{=} \iota(H_\Psi)$ and consider the factorizations

$$z^j G = W_0^* \begin{pmatrix} z^j u_0 & \mathbb{O} \\ \mathbb{O} & z^j \Psi \end{pmatrix} V_0^* = \mathcal{W}^* \begin{pmatrix} z^j u & \mathbb{O} \\ \mathbb{O} & z^j \Delta \end{pmatrix} \mathcal{V}^*.$$

It is easy to see that the sum of the thematic indices of $z^j G$ corresponding to the superoptimal singular value 1 equals $\operatorname{ind} T_{z^j u_0} = \iota(H_G) - \iota(H_\Psi)$, because $\|H_{z^j\Psi}\| < \|H_\Psi\| = 1$.

In order to proceed, we need the following lemma.

**Lemma 2.3.6.** *Let $G \in L^\infty(\mathbb{M}_{m,n})$ be such that $\|H_G\|_{\mathrm{e}} < 1$ and $\|H_G\| = 1$. Suppose that $G$ is a badly approximable matrix function that admits a representation of the form (2.3.1), in which $V$ and $W^t$ are thematic matrix functions, $u$ is a unimodular function, and $\Psi$ is a bounded matrix function. Let $V = (\, v \ \bar{\Theta}\,)$.*

1. *If $f \in \mathcal{M}_G$ satisfies $\Theta^t f = \mathbb{O}$, then $f = \xi v$ for some $\xi \in \ker T_u$.*

2. *If $\Psi$ is a badly approximable matrix function with $L^\infty$-norm equal to 1 and the*

*Toeplitz operator* $T_u$ *is Fredholm with* $\operatorname{ind} T_u \le 0$, *then*

$$\dim \mathcal{M}_G \le \dim \mathcal{M}_\Psi. \tag{2.3.13}$$

*Moreover, if* $\operatorname{ind} T_u = 0$, *then equality holds in* (2.3.13).

We finish the proof of Theorem 2.3.5 before proving Lemma 2.3.6.

As already seen,

$$\iota(H_\Psi) < \iota(H_G) \le \iota(H_\Delta)$$

and so $z^j \Delta$ is a badly approximable matrix function of $L^\infty$-norm equal to 1, since $\|H_{z^j\Delta}\| = \|H_\Delta\| = 1$. Let $\ell \stackrel{\text{def}}{=} \iota(H_{z^j\Delta})$. Then $\|H_{z^\ell z^j \Delta}\| < \|H_{z^j\Delta}\| = 1$ implies that $\ell + j \ge \iota(H_\Delta)$, and therefore

$$\dim \mathcal{M}_{z^j G} = \dim \mathcal{M}_{z^j \Delta} \ge \iota(H_{z^j\Delta}) \ge \iota(H_\Delta) - j = \iota(H_\Delta) - \iota(H_\Psi),$$

by Lemma 2.3.6. Hence

$$\iota(H_G) \ge \iota(H_\Delta),$$

because the sum of the thematic indices of $z^j G$ corresponding to the superoptimal singular value 1, namely $\iota(H_G) - \iota(H_\Psi)$, equals $\dim \mathcal{M}_{z^j G}$ (e.g. see Theorem 14.7.4 of [Pe1]). This completes the proof. □

*Remark* 2.3.7. Note that if the inequality $\operatorname{ind} T_u \le \iota(H_\Psi)$ in Theorem 2.3.5 is strict, then equality in (2.3.12) may not hold. For instance, consider a monotone non-increasing thematic factorization (e.g. see Section 2.4) of any admissible unitary-valued very badly approximable matrix function $G \in L^\infty(\mathbb{M}_2)$ with thematic indices 3 and 2, and any other thematic factorization of $G$ whose first thematic index equals 1.

*Proof of Lemma 2.3.6.* Let $W^t = (\, w \ \bar{\Xi}\,)$. To prove assertion *1*, we may assume that

$f \in \mathcal{M}_G$ is non-zero. Since

$$f = VV^*f = (\, v\, \bar{\Theta}\,) \begin{pmatrix} v^*f \\ \mathbb{O} \end{pmatrix} = v(v^*f),$$

Fact 1.3.1 implies $\xi \stackrel{\text{def}}{=} v^*f \in H^2$, as $v$ is co-outer. It remains to show that $u\xi \in H^2_-$. Since $u\xi\bar{w} = Gf \in H^2_-(\mathbb{C}^m)$, it follows again from Fact 1.3.1 that $u\xi \in H^2_-$ because $w$ is co-outer. Thus, $f = \xi v$ with $\xi \in \ker T_u$.

Suppose now that the functions $\Psi$ and $u$ satisfy the assumptions of assertion *2*. Let $\{f_j\}_{j=1}^N$ be a basis for $\mathcal{M}_G$ and define $g_j = \Theta^t f_j$ for $1 \le j \le N$. Since $\operatorname{ind} T_u \le 0$, then $\ker T_u$ is trivial, and each $g_j$ is a non-zero function in $H^2(\mathbb{C}^{n-1})$ by assertion *1*. Furthermore, $\{g_j\}_{j=1}^N$ is a linearly independent set in $H^2(\mathbb{C}^{n-1})$; after all, if there are scalars $c_1, \ldots, c_n$ such that

$$\mathbb{O} = \sum_{j=1}^{N} c_j g_j = \Theta^t \left( \sum_{j=1}^{N} c_j f_j \right),$$

then $\sum_{j=1}^N c_j f_j = \mathbb{O}$ by assertion *1*, and so $c_j = 0$ for $1 \le j \le N$ because $\{f_j\}_{j=1}^N$ is a linearly independent set. In order to prove (2.3.13), it suffices to show that $g_j$ belongs to $\mathcal{M}_\Psi$ for $1 \le j \le N$. To this end, fix $j_0$ such that $1 \le j_0 \le N$. Since $G$ is badly approximable and admits a factorization of the form (2.3.1), then

$$\|f_{j_0}(\zeta)\|^2_{\mathbb{C}^n} = \|Gf_{j_0}(\zeta)\|^2_{\mathbb{C}^m} = |v^*f_{j_0}(\zeta)|^2 + \|\Psi\Theta^t f_{j_0}(\zeta)\|^2_{\mathbb{C}^{m-1}}$$

for a.e. $\zeta \in \mathbb{T}$. On the other hand,

$$|v^*f_{j_0}(\zeta)|^2 + \|\Theta^t f_{j_0}(\zeta)\|^2_{\mathbb{C}^{n-1}} = \|V^*f_{j_0}(\zeta)\|^2_{\mathbb{C}^n} = \|f_{j_0}(\zeta)\|^2_{\mathbb{C}^n}$$

holds for a.e. $\zeta \in \mathbb{T}$ because $V$ is unitary-valued. Thus, the function $g_{j_0} = \Theta^t f_{j_0}$

satisfies

$$\|\Psi g_{j0}(\zeta)\|_{\mathbb{C}^{m-1}} = \|g_{j0}(\zeta)\|_{\mathbb{C}^{n-1}} \text{ for a.e. } \zeta \in \mathbb{T}.$$

Since $W$ is unitary-valued,

$$WGf_{j0} = \begin{pmatrix} uv^* f_{j0} \\ \Psi\Theta^t f_{j0} \end{pmatrix},$$

and so $\Psi\Theta^t f_{j0} = \Xi^* G f_{j0} \in H^2_-(\mathbb{C}^{m-1})$ by Fact 1.3.1 because $Gf_{j0} \in H^2_-(\mathbb{C}^m)$. Hence, we may conclude that $g_{j0} \in H^2(\mathbb{C}^{m-1})$ satisfies

$$\|H_\Psi g_{j0}\|_2 = \|\Psi g_{j0}\|_2 = \|g_{j0}\|_2,$$

i.e. each $g_{j0}$ is a maximizing vector of the Hankel operator $H_\Psi$. This completes the proof of (2.3.13).

Suppose now that $\operatorname{ind} T_u = 0$. Let $\xi_1, \ldots, \xi_d$ be a basis for $\mathcal{M}_\Psi$. By Lemma 2.3.2, each function $\xi_j$ induces a function $\xi_\#^{(j)} = A^t \xi_j + a_j v \in \mathcal{M}_G$ for some suitable $a_j \in H^2$, $1 \le j \le d$, where $A$ denotes a left-inverse of $\Theta$ in $H^\infty$. It is easy to see from (2.3.4) that $\xi_\#^{(1)}, \ldots, \xi_\#^{(d)}$ form a linearly independent set, as $\{\xi_j\}_{j=1}^d$ is a linearly independent set. Hence $\dim \mathcal{M}_G = \dim \mathcal{M}_\Psi$. $\square$

*Remark* 2.3.8. Notice that the inequality given in 2.3.13 of Lemma 2.3.6 may in fact be strict. For instance, consider the badly approximable matrix function

$$G = \begin{pmatrix} \bar{z} & \mathbb{O} \\ \mathbb{O} & 1 \end{pmatrix}.$$

It is easy to see that $\dim \mathcal{M}_G = 1$ and $G$ admits a factorization of the form

$$G = W^* \begin{pmatrix} z & \mathbb{O} \\ \mathbb{O} & \bar{z}^2 \end{pmatrix} V^*,$$

where

$$W^t = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -\bar{z} \\ z & 1 \end{pmatrix} \quad \text{and} \quad V = \frac{1}{\sqrt{2}} \begin{pmatrix} z^2 & -1 \\ 1 & \bar{z}^2 \end{pmatrix}$$

are thematic, and the Toeplitz operators $T_z$ and $T_{\bar{z}^2}$ are Fredholm with indices $-1$ and 2, respectively. Let $\Psi \stackrel{\text{def}}{=} \bar{z}^2$. It is easy to see that $\dim \mathcal{M}_\Psi = \dim \ker T_\Psi = \operatorname{ind} T_\Psi = 2 > \dim \mathcal{M}_G$ because $\Psi$ is unimodular.

## 2.4 Sequences of thematic indices

We proceed by proving the existence of a monotone non-decreasing thematic factorization and show that other thematic factorizations are induced by a given monotone non-increasing thematic factorization.

**Definition 2.4.1.** Let $G \in L^\infty(\mathbb{M}_{m,n})$ be a badly approximable matrix function whose superoptimal singular values $t_j = t_j(G)$, $j \geq 0$, satisfy

$$\|H_G\|_e < t_{r-1}, \; t_0 = \ldots = t_{r-1}, \;\; \text{and} \;\; t_{r-1} > t_r. \tag{2.4.1}$$

We say that

$$(\, k_0, k_1, k_2, \ldots, k_{r-1} \,)$$

is *a sequence of thematic indices for* $G$ if $G$ admits a partial thematic factorization of the form

$$W_0^* \cdot \ldots \cdot W_{r-1}^* \begin{pmatrix} t_0 u_0 & \mathbb{O} & \ldots & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & t_0 u_1 & \ldots & \mathbb{O} & \mathbb{O} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbb{O} & \mathbb{O} & \ldots & t_0 u_{r-1} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \ldots & \mathbb{O} & \Psi \end{pmatrix} V_{r-1}^* \cdot \ldots \cdot V_0^*, \tag{2.4.2}$$

such that $\operatorname{ind} T_{u_j} = k_j$ and the matrix functions $V_j$ and $W_j$ are of the form (1.3.4) for $0 \le j \le r-1$, and $\Psi$ satisfies (1.3.5).

**Theorem 2.4.2.** *Suppose that $G \in L^\infty(\mathbb{M}_{m,n})$ is a badly approximable matrix function satisfying* (2.4.1). *If $\nu$ equals the sum of the thematic indices corresponding to the superoptimal singular value $t_0(G)$, then*

$$(\underbrace{1, 1, \ldots, 1}_{r-1}, \nu - r + 1) \tag{2.4.3}$$

*is a sequence of thematic indices for $G$. In particular, $G$ admits a monotone non-decreasing thematic factorization.*

*Proof.* Consider any thematic factorization of $\Delta_0 \stackrel{\text{def}}{=} t_0^{-1} G$. It follows from Theorem 2.3.3 that $\Delta_0$ admits a factorization of the form

$$\Delta_0 = \mathcal{W}_0^* \begin{pmatrix} u_0 & \mathbb{O} \\ \mathbb{O} & \Delta_1 \end{pmatrix} \mathcal{V}_0^*$$

where $\operatorname{ind} T_{u_0} = 1$. Similarly, Theorem 2.3.3 implies that $\Delta_1$ also admits a factorization of the form

$$\Delta_1 = \mathcal{W}_1^* \begin{pmatrix} u_1 & \mathbb{O} \\ \mathbb{O} & \Delta_2 \end{pmatrix} \mathcal{V}_1^*$$

where $\operatorname{ind} T_{u_1} = 1$. Continuing in this manner, we obtain matrix functions $\Delta_0$, $\Delta_1$, ..., $\Delta_{r-2}$,$\Delta_{r-1}$ with factorizations of the form

$$\Delta_j = \mathcal{W}_j^* \begin{pmatrix} u_j & \mathbb{O} \\ \mathbb{O} & \Delta_{j+1} \end{pmatrix} \mathcal{V}_j^*,$$

where $\operatorname{ind} T_{u_j} = 1$, for $0 \leq j \leq r-2$. It is easy to see that these matrix functions induce a partial thematic factorization of $G$ in which the first $r-1$ thematic indices equal 1. Since the sum of the thematic indices $\nu$ of $G$ is independent of the partial thematic factorization, it must be that the $r$th thematic index in this induced partial thematic factorization equals $\nu - (r-1)$. □

The following corollaries are immediate.

**Corollary 2.4.3.** *If $G \in L^\infty(\mathbb{M}_{m,n})$ is an admissible very badly approximable matrix function, then $G$ admits a monotone non-decreasing thematic factorization.*

**Corollary 2.4.4.** *If $G \in (H^\infty + C)(\mathbb{M}_{m,n})$ is a very badly approximable matrix function, then $G$ admits a monotone non-decreasing thematic factorization.*

We go on to show that the thematic indices obtained in a monotone non-decreasing thematic factorization are *not* uniquely determined. Moreover, we determine all possible sequences of thematic indices for an admissible very badly approximable unitary-valued $2 \times 2$ matrix function.

**Theorem 2.4.5.** *Let $U \in L^\infty(\mathbb{M}_2)$ be an admissible very badly approximable unitary-valued matrix function. Suppose that $(k_0, k_1)$ is the monotone non-increasing sequence of thematic indices for $U$. Then the collection of sequences of thematic indices for $U$ coincides with the set*

$$\sigma_U \stackrel{\text{def}}{=} \{(k_1 - j, k_0 + j) : 0 \leq j < k_1\} \cup \{(k_0, k_1)\}.$$

*Proof.* Let $0 \leq j < k_1$. By Corollary 2.3.4, $U$ admits a factorization of the form

$$U = \mathcal{W}^* \begin{pmatrix} u_0 & \mathbb{O} \\ \mathbb{O} & u_1 \end{pmatrix} \mathcal{V}^*$$

with $\operatorname{ind} T_{u_0} = k_1 - j$. Since the sum of the thematic indices of $U$ is independent of the thematic factorization, it must be that $\operatorname{ind} T_{u_1} = k_0 + j$. Thus $\sigma_U$ consists of sequences of thematic indices for $U$.

Suppose now that $(a, b)$ is a sequence of thematic indices for $U$ that does not belong to $\sigma_U$. In this case, $U$ admits a factorization of the form

$$U = W^* \begin{pmatrix} u_0 & \mathbb{O} \\ \mathbb{O} & u_1 \end{pmatrix} V^*$$

for some thematic matrix functions $V$ and $W^t$, and unimodular functions $u_0$ and $u_1$ such that $\operatorname{ind} T_{u_0} = a$ and $\operatorname{ind} T_{u_1} = b$. Since $(a, b) \notin \sigma_U$, it follows that $b > a$ and $a > k_1$. Thus,

$$z^{k_1} U = W^* \begin{pmatrix} z^{k_1} u_0 & \mathbb{O} \\ \mathbb{O} & z^{k_1} u_1 \end{pmatrix} V^*$$

is a very badly approximable unitary-valued matrix function. In particular, $z^{k_1} U$ admits a monotone non-increasing thematic sequence, say $(\alpha, \beta)$. Hence, $(\alpha + k_1, \beta + k_1)$ is a monotone non-increasing sequence of thematic indices for $U$ and so, by the uniqueness of a monotone non-increasing sequence, $k_1 = \beta + k_1$ for some $\beta \geq 1$ a contradiction. This completes the proof. □

We now recall how monotone non-increasing thematic factorizations were obtained in [AP2].

Let $G \in L^\infty(\mathbb{M}_{m,n})$ be a badly approximable matrix function such that (2.4.1) holds. In this case, it is known that $G$ admits a monotone *non-increasing* partial thematic factorization and that the thematic indices appearing in any monotone non-increasing partial thematic factorization of $G$ are uniquely determined by $G$. In fact, as discussed in Section 2.3, $G_0 = t_0^{-1}G$ admits a factorization of the form

$$G_0 = W_0^* \begin{pmatrix} u_0 & \mathbb{O} \\ \mathbb{O} & G_1 \end{pmatrix} V_0^*$$

with $\operatorname{ind} T_{u_0} = \iota(H_{G_0})$ and $\iota(H_{G_0}) \geq \iota(H_{G_1})$ (see (2.3.10)). Similarly, for each $1 \leq j \leq r-1$, we obtain a matrix function $G_j$ with a factorization of the form

$$G_j = \breve{W}_j^* \begin{pmatrix} u_j & \mathbb{O} \\ \mathbb{O} & G_{j+1} \end{pmatrix} \breve{V}_j^*$$

with $\operatorname{ind} T_{u_j} = \iota(H_{G_j})$ and $\iota(H_{G_j}) \geq \iota(H_{G_{j+1}})$. Then

$$(\iota(H_{G_0}), \iota(H_{G_1}), \ldots, \iota(H_{G_{r-1}}))$$

is the monotone non-increasing sequence of thematic indices for $G$. (See [AP2] or Section 10 in Chapter 14 of [Pe1].)

Note that, in the general setting of $m \times n$ matrix functions, at least two sequences of thematic indices for $G$ exist; the monotone non-increasing sequence and the sequence in (2.4.3). The question remains: *Are there any others?*

**Theorem 2.4.6.** *Suppose $G \in L^\infty(\mathbb{M}_{m,n})$ is a badly approximable matrix function satisfying* (2.4.1). *If*

$$(\, k_0, k_1, k_2, \ldots, k_{r-1} \,)$$

*is the monotone non-increasing sequence of thematic indices for $G$, then*

$$(\, k_1, k_0, k_2, \ldots, k_{r-1} \,)$$

*is also sequence of thematic indices for $G$.*

*Proof.* Without loss of generality, we may assume that $t_0 = 1$ and $k_0 > k_1$. By Theorems 2.3.3 and 2.3.5, $G$ admits a thematic factorization of the form

$$G = \mathcal{W}^* \begin{pmatrix} u & \mathbb{O} \\ \mathbb{O} & \Delta \end{pmatrix} \mathcal{V}^*,$$

where $\operatorname{ind} T_u = k_1$ and $\iota(H_\Delta) = k_0$. Let $(\kappa_1, \kappa_2, \ldots, \kappa_{r-1})$ be the monotone non-increasing sequence of thematic indices for $\Delta$. In particular, $\kappa_1 = k_0$ and

$$(k_1, k_0, \kappa_2, \ldots, \kappa_{r-1}) \tag{2.4.4}$$

is a sequence of thematic indices for $G$. We claim that

$$\kappa_j = k_j \text{ for } 2 \le j \le r - 1.$$

By considering the monotone non-increasing sequence for $G$, it is easy to see that the sum of the thematic indices corresponding to the superoptimal singular value 1 of $z^{k_2} G$ equals

$$(k_0 - k_2) + (k_1 - k_2).$$

On the other hand, this sum is also equal to

$$(k_1 - k_2) + (k_0 - k_2) + \sum_{\{j \ge 2:\ \kappa_j \ge k_2\}} (\kappa_j - k_2),$$

because the sequence in (2.4.4) is a sequence of thematic indices for $G$. This implies that $\kappa_2 \le k_2$. Now, by considing the matrix function $z^{\kappa_2} G$, the same argument reveals that $k_2 \le \kappa_2$. Therefore $\kappa_2 = k_2$.

Let $2 \le \ell < r - 1$. Suppose we have already shown that $\kappa_j = k_j$ for $2 \le j \le \ell$. In the same manner, the sum of the thematic indices corresponding to the superoptimal singular value 1 of $z^{k_{\ell+1}}G$ equals

$$\sum_{j=0}^{\ell}(k_j - k_{\ell+1}) \text{ and } \sum_{j=0}^{\ell}(k_j - k_{l+1}) + \sum_{\{j\geq \ell+1:\, \kappa_j \geq k_{l+1}\}} (\kappa_j - k_{l+1}).$$

This implies that $\kappa_{\ell+1} \leq k_{l+1}$, and a similar argument shows $k_{l+1} \leq \kappa_{l+1}$. Hence we must have that $\kappa_j = k_j$ for $2 \leq j \leq r-1$. □

Theorem 2.4.6 provides a stronger conclusion than one might think. Loosely speaking, it says that we can always interchange the highest two adjacent thematic indices in any monotone non-increasing sequence of thematic indices and still obtain another sequence of thematic indices for the same matrix function. Let us illustrate this with the following example.

**Example 2.4.7.** For simplicity, consider the very badly approximable function

$$G = \begin{pmatrix} \bar{z}^3 & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \bar{z}^2 & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \bar{z} \end{pmatrix}.$$

Clearly, $(3,2,1)$ is the monotone non-increasing sequence of thematic indices for $G$. Our results imply that there are many other sequences of thematic indices for $G$. Indeed, by considering the subsequence $(2,1)$, Theorem 2.4.6 implies that $(3,1,2)$ is also a sequence of thematic indices for $G$. Similarly, it is easy to see that $(2,3,1)$ and $(2,1,3)$ are also sequences of thematic indices for $G$. On the other hand, it follows from Theorem 2.4.2 that $(1,1,4)$ is a sequence of thematic indices for $G$.

This leads us to ask: *Are there other sequences of thematic indices in which the first index is equal to 1?*

It can be verified that $G$ admits the following thematic factorizations:

$$
\begin{aligned}
G &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \sqrt{2} \\ 1 & -1 & \mathbb{O} \end{pmatrix} \begin{pmatrix} \bar{z} & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \bar{z}^3 & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \bar{z}^2 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} \bar{z}^2 & \mathbb{O} & 1 \\ 1 & \mathbb{O} & -z^2 \\ \mathbb{O} & \sqrt{2} & \mathbb{O} \end{pmatrix} \\
&= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -z & \mathbb{O} \\ \bar{z} & 1 & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \sqrt{2} \end{pmatrix} \begin{pmatrix} \bar{z} & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & \bar{z}^4 & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \bar{z} \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} \bar{z}^2 & 1 & \mathbb{O} \\ -1 & z^2 & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \sqrt{2} \end{pmatrix}.
\end{aligned}
$$

Thus, $(1,3,2)$ and $(1,4,1)$ are sequences of thematic indices for $G$ as well. These sequences induce two others by considering the subsequences $(3,2)$ and $(4,1)$; namely $(1,2,3)$ and $(1,1,4)$. Thus, the matrix function $G$ admits at least 8 different sequences of thematic indices, namely

$$
(3,2,1),\ (3,1,2),\ (2,3,1),\ (2,1,3),\ (1,3,2),
$$
$$
(1,2,3),\ (1,4,1),\ \text{and}\ (1,1,4).
$$

It is easy to verify that these are all possible sequences of thematic indices for $G$.

## 2.5 Unitary-valued very badly approximable $2 \times 2$ matrix functions

The problem of finding all possible sequences of thematic indices for an arbitrary admissible very badly approximable matrix function seems rather difficult for $m \times n$ matrix functions with $\min\{m, n\} > 2$. However, in the case of unitary-valued $2 \times 2$ matrix functions, the problem has a straightforward solution provided by Theorem 2.4.5. In this section, we introduce a simple algorithm that yields thematic factorizations with desired thematic indices for such matrix functions.

## Algorithm

Let $U$ be an admissible very badly approximable unitary-valued $2 \times 2$ matrix function on $\mathbb{T}$ and $(k_0, k_1)$ denote the monotone non-increasing sequence of thematic indices for $U$. Suppose $U$ admits a monotone non-increasing thematic factorization of the form

$$U = (\, \bar{w}_0 \ \Xi \,) \begin{pmatrix} \bar{z}^{k_0} \frac{\bar{h}_0}{h_0} & \mathbb{O} \\ \mathbb{O} & \bar{z}^{k_1} \frac{\bar{h}_1}{h_1} \end{pmatrix} \begin{pmatrix} v^* \\ \Theta^t \end{pmatrix}, \tag{2.5.1}$$

where $h_0$, $h_1$, and their respective inverses belong to $H^p$ for some $2 < p \leq \infty$. For each integer $j$ satisfying $1 \leq j \leq k_1$, a thematic factorization of $U$ with thematic indices $(j, k_0 + k_1 - j)$ can be obtained as follows.

1. Find left-inverses $A$ and $B$ in $H^\infty$ for $\Theta$ and $\Xi$, respectively.

2. Set $u_0 = \bar{z}^{k_0 - j + 1} \dfrac{\bar{h}_0}{h_0}$ and $\Psi = \bar{z}^{k_1 - j + 1} \dfrac{\bar{h}_1}{h_1}$.

3. Let $\xi = z^{k_1 - j} h_1$. Find a solution $a_0 \in H^2$ to the equation

$$T_u a_0 = \mathbb{P}_+ (w^t B^* \Psi - u v^* A^t) \xi.$$

   If $j < k_1$, we require, in addition, that $z$ is not an inner divisor of $a_0$. (Note that if $z$ is an inner divisor of $a_0$, then it suffices to replace $a_0$ with $a_0 + h_0$.)

4. Let $\xi_\# = A^t \xi + a_0 v$ and $\eta_\# = \bar{z} \bar{G} \bar{\xi}_\#$. Choose an outer function $h \in H^2$ such that $|h(\zeta)| = \|\xi_\#(\zeta)\|_{\mathbb{C}^2}$ for a.e. $\zeta \in \mathbb{T}$.

5. Let $\nu = h^{-1} \xi_\#$ and $\omega = h^{-1} \eta_\#$. Find thematic completions

$$\mathcal{V} = (\, \nu \ \bar{\Upsilon} \,) \text{ and } \mathcal{W}^t = (\, \omega \ \bar{\Omega} \,)$$

   to $\nu$ and $\omega$, respectively.

6. The desired thematic factorization for $G$ is given by

$$G = \mathcal{W}^* \begin{pmatrix} u & \mathbb{O} \\ \mathbb{O} & \Delta \end{pmatrix} \mathcal{V}^* \tag{2.5.2}$$

where

$$u = \bar{z}^j \frac{\bar{h}}{h} \quad \text{and} \quad \Delta = \Omega^* G \bar{\Upsilon}.$$

**End of algorithm**

The validity of this algorithm is justified by the proof of Theorem 2.3.3 and Corollary 2.3.4.

For matrix functions $G \in \mathcal{R}(\mathbb{M}_2)$, the badly approximable scalar functions appearing in the diagonal factor of (2.5.1) also belong to $\mathcal{R}$. This is a consequence of the results in [PY1] (see also Sections 5 and 12 of Chapter 14 in [Pe1]). As mentioned in Remark 2.2.6, the outer functions $h_0$ and $h_1$ are (up to a multiplicative constant) products of quotients of reproducing kernels of $H^2$. Therefore, steps 1 through 6 of the algorithm are more easily implemented if $G \in \mathcal{R}(\mathbb{M}_2)$.

**Example 2.5.1.** Consider the matrix function

$$G = \frac{1}{\sqrt{2}} \begin{pmatrix} \bar{z}^4 & -\bar{z}^2 \\ \bar{z}^3 & \bar{z} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \bar{z} & -1 \\ 1 & z \end{pmatrix} \begin{pmatrix} \bar{z}^3 & \mathbb{O} \\ \mathbb{O} & \bar{z}^2 \end{pmatrix}. \tag{2.5.3}$$

Let

$$w = \frac{1}{\sqrt{2}} \begin{pmatrix} z \\ 1 \end{pmatrix}, \ \Xi = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ z \end{pmatrix}, \ v = \begin{pmatrix} 1 \\ \mathbb{O} \end{pmatrix}, \ \Theta = \begin{pmatrix} \mathbb{O} \\ 1 \end{pmatrix},$$

$$h_0 = h_1 = 1, \ B = \sqrt{2}\begin{pmatrix} -1 & \mathbb{O} \end{pmatrix}, \ \text{and} \ A = \begin{pmatrix} \mathbb{O} & 1 \end{pmatrix}. \tag{2.5.4}$$

We find thematic factorizations with sequences of indices $(2,3)$ and $(1,4)$.

1. *A thematic factorization for $G$ with sequence of indices* $(2,3)$*:*

Let

$$u_0 = \bar{z}^2,\ \Psi = \bar{z},\ \xi = 1,\ \text{and}\ a_0 = -z^2.$$

In this case, it is easy to verify that

$$\xi_{\#} = \begin{pmatrix} -z^2 \\ 1 \end{pmatrix} \quad \text{and} \quad \eta_{\#} = \frac{1}{\sqrt{2}} \begin{pmatrix} -2 \\ \mathbb{O} \end{pmatrix}.$$

Since $\|\xi_{\#}(\zeta)\|_{\mathbb{C}^2} = 2$ on $\mathbb{T}$, we may take $h(\zeta) = \sqrt{2}$ for $\zeta \in \mathbb{T}$. Then

$$\nu = \frac{1}{\sqrt{2}} \begin{pmatrix} -z^2 \\ 1 \end{pmatrix} \quad \text{and} \quad \omega = \begin{pmatrix} -1 \\ \mathbb{O} \end{pmatrix}$$

have thematic completions

$$\mathcal{V} = (\,\nu\ \bar{\Upsilon}\,) \ \text{and}\ \mathcal{W}^t = (\,\omega\ \bar{\Omega}\,),$$

where

$$\Upsilon = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ z^2 \end{pmatrix} \quad \text{and} \quad \Omega = \begin{pmatrix} \mathbb{O} \\ 1 \end{pmatrix}.$$

Thus, $G$ admits the factorization

$$G = \begin{pmatrix} -1 & \mathbb{O} \\ \mathbb{O} & 1 \end{pmatrix} \begin{pmatrix} \bar{z}^2 & \mathbb{O} \\ \mathbb{O} & \bar{z}^3 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} -\bar{z}^2 & 1 \\ 1 & z^2 \end{pmatrix},$$

with sequence of thematic indices $(2,3)$ as desired.

2. *A thematic factorization for $G$ with sequence of indices* $(1,4)$*:*

Let

$$u_0 = \bar{z}^3,\ \Psi = \bar{z}^2,\ \xi = z, \text{ and } a_0 = 1 - z^3.$$

It is easy to verify that

$$\xi_\# = \begin{pmatrix} 1 - z^3 \\ z \end{pmatrix} \text{ and } \eta_\# = \frac{1}{\sqrt{2}} \begin{pmatrix} z^3 - 2 \\ z^2 \end{pmatrix}.$$

Since $\|\xi_\#(\zeta)\|^2_{\mathbb{C}^2} = 3 - \bar{z}^3 - z^3$ on $\mathbb{T}$, we may choose

$$h = a^2 - \frac{1}{a^2} z^3, \text{ where } a = \frac{-1 + \sqrt{5}}{2}.$$

Let

$$\nu = \frac{1}{h} \begin{pmatrix} 1 - z^3 \\ z \end{pmatrix} \text{ and } \omega = \frac{1}{h\sqrt{2}} \begin{pmatrix} z^3 - 2 \\ z^2 \end{pmatrix}.$$

We may now take

$$\Upsilon = \frac{1}{h} \begin{pmatrix} -z \\ 1 - z^3 \end{pmatrix} \text{ and } \Omega = \frac{1}{h\sqrt{2}} \begin{pmatrix} -z^2 \\ z^3 - 2 \end{pmatrix},$$

so that the matrix functions

$$\mathcal{V} = (\,\nu\ \bar{\Upsilon}\,) \text{ and } \mathcal{W}^t = (\,\omega\ \bar{\Omega}\,)$$

are thematic. Since

$$\Omega^* G \bar{\Upsilon} = \frac{\bar{z}^4}{\bar{h}^2}(3 - \bar{z}^3 - z^3) = \frac{\bar{z}^4}{\bar{h}^2}|h|^2 = \bar{z}^4 \frac{h}{\bar{h}},$$

$G$ admits the factorization

$$G = \mathcal{W}^* \begin{pmatrix} z\frac{\bar{h}}{h} & \mathbb{O} \\ \mathbb{O} & z^4\frac{h}{\bar{h}} \end{pmatrix} \mathcal{V}^*$$

with sequence of thematic indices $(1, 4)$ as desired.

# Chapter 3

# On the sum of superoptimal singular values

## 3.1 An extremal problem

In this chapter, we study the following extremal problem and its relevance to the sum of the superoptimal singular values of a matrix function:

*Extremal Problem* 3.1.1. Let $m, n > 1$ and $1 \leq k \leq \min\{m, n\}$. Given a matrix function $\Phi \in L^\infty(\mathbb{M}_{m,n})$, when is there a matrix function $\Psi_*$ in the set $\mathcal{A}_k^{n,m}$ such that

$$\int_{\mathbb{T}} \operatorname{trace}(\Phi(\zeta)\Psi_*(\zeta))d\boldsymbol{m}(\zeta) = \sigma_k(\Phi)?$$

The set $\mathcal{A}_k^{n,m}$ is defined by

$$\mathcal{A}_k^{n,m} \stackrel{\text{def}}{=} \left\{\Psi \in H_0^1(\mathbb{M}_{n,m}) : \|\Psi\|_{L^1(\mathbb{M}_{n,m})} \leq 1, \operatorname{rank}\Psi(\zeta) \leq k \text{ a.e. } \zeta \in \mathbb{T}\right\}$$

and $\sigma_k(\Phi)$ is defined by

$$\sigma_k(\Phi) \stackrel{\text{def}}{=} \sup_{\Psi \in \mathcal{A}_k^{n,m}} \left| \int_{\mathbb{T}} \operatorname{trace}(\Phi(\zeta)\Psi(\zeta))d\boldsymbol{m}(\zeta) \right|. \tag{3.1.1}$$

Whenever $n = m$, we use the notation $\mathcal{A}_k^n \stackrel{\text{def}}{=} \mathcal{A}_k^{n,m}$ instead.

The importance of this problem arose from the following observation due to Peller [Pe3].

**Theorem 3.1.1.** *Let* $1 \le k \le \min\{m, n\}$. *If* $\Phi \in L^\infty(\mathbb{M}_{m,n})$ *is admissible, then*

$$\sigma_k(\Phi) \le t_0(\Phi) + \ldots + t_{k-1}(\Phi). \tag{3.1.2}$$

*Proof.* Let $\Psi \in \mathcal{A}_k^{n,m}$. We may assume, without loss of generality, that $\Phi$ is very badly approximable. Indeed,

$$\int_{\mathbb{T}} \operatorname{trace}(\Phi(\zeta)\Psi(\zeta))\boldsymbol{dm}(\zeta) = \int_{\mathbb{T}} \operatorname{trace}((\Phi - Q)(\zeta)\Psi(\zeta))\boldsymbol{dm}(\zeta)$$

holds for any $Q \in H^\infty(\mathbb{M}_{m,n})$, and so we may replace $\Phi$ with $\Phi - Q$ if necessary, where $Q$ is the superoptimal approximation in $H^\infty(\mathbb{M}_{m,n})$ to $\Phi$.

It follows from the well-known inequality $|\operatorname{trace}(A)| \le \|A\|_{\boldsymbol{S}_1^m}$ that the inequalities

$$|\operatorname{trace}(\Phi(\zeta)\Psi(\zeta))| \le \|\Phi(\zeta)\Psi(\zeta)\|_{\boldsymbol{S}_1^m} \le \left(\sum_{j=0}^{k-1} s_j(\Phi(\zeta))\right) \|\Psi(\zeta)\|_{\mathbb{M}_{n,m}}$$

hold for a.e. $\zeta \in \mathbb{T}$. Thus,

$$\begin{aligned}
\left|\int_{\mathbb{T}} \operatorname{trace}(\Phi(\zeta)\Psi(\zeta))\boldsymbol{dm}(\zeta)\right| &\le \int_{\mathbb{T}} \left(\sum_{j=0}^{k-1} s_j(\Phi(\zeta))\right) \|\Psi(\zeta)\|_{\mathbb{M}_{n,m}}\boldsymbol{dm}(\zeta) \\
&\le \int_{\mathbb{T}} \left(\sum_{j=0}^{k-1} t_j(\Phi)\right) \|\Psi(\zeta)\|_{\mathbb{M}_{n,m}}\boldsymbol{dm}(\zeta) \\
&\le \left(\sum_{j=0}^{k-1} t_j(\Phi)\right) \|\Psi\|_{L^1(\mathbb{M}_{n,m})} \\
&\le \sum_{j=0}^{k-1} t_j(\Phi).
\end{aligned} \tag{3.1.3}$$

because the singular values of $\Phi$ satisfy $s_j(\Phi(\zeta)) = t_j(\Phi)$ for a.e. $\zeta \in \mathbb{T}$ since $\Phi$ is very badly approximable. $\square$

Before proceeding, let us observe that equality holds in (3.1.2) for some simple cases. Let $r$ be a positive integer and $t_0, t_1, \ldots, t_{r-1}$ be positive numbers satisfying

$$t_0 \geq t_1 \geq \ldots \geq t_{r-1}.$$

Suppose $\Phi$ is an $n \times n$ matrix function of the form

$$\Phi \stackrel{\text{def}}{=} \begin{pmatrix} t_0 u_0 & \mathbb{O} & \ldots & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & t_1 u_1 & \ldots & \mathbb{O} & \mathbb{O} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbb{O} & \mathbb{O} & \ldots & t_{r-1} u_{r-1} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \ldots & \mathbb{O} & \Phi_{\#} \end{pmatrix}, \tag{3.1.4}$$

where $\|\Phi_{\#}\|_{L^\infty} \leq t_{r-1}$ and $u_j$ is a unimodular function of the form $u_j = \bar{z}\bar{\theta}_j \bar{h}/h$ with $\theta_j$ an inner function for $0 \leq j \leq r-1$ and $h$ an outer function in $H^2$. Without loss of generality, we may assume that $\|h\|_{L^2} = 1$. It can be seen that if

$$\Psi \stackrel{\text{def}}{=} \begin{pmatrix} z\theta_0 h^2 & \mathbb{O} & \ldots & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & z\theta_1 h^2 & \ldots & \mathbb{O} & \mathbb{O} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbb{O} & \mathbb{O} & \ldots & z\theta_{r-1} h^2 & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \ldots & \mathbb{O} & \mathbb{O} \end{pmatrix}, \tag{3.1.5}$$

then $\Psi \in H_0^1(\mathbb{M}_n)$, rank $\Psi(\zeta) = r$ a.e. on $\mathbb{T}$, $\|\Psi\|_{L^1(\mathbb{M}_n)} = 1$, and

$$\int_{\mathbb{T}} \text{trace}(\Phi(\zeta)\Psi(\zeta))\, d\boldsymbol{m}(\zeta) = t_0 + \ldots + t_{r-1}.$$

Thus we obtain that

$$\sigma_r(\Phi) = t_0(\Phi) + \ldots + t_{r-1}(\Phi).$$

On the other hand, one cannot expect the inequality (3.1.2) to become an equality in general. After all, by the Hahn-Banach Theorem,

$$\operatorname{dist}_{L^\infty(\boldsymbol{S}_1^n)}(\Phi, H^\infty(\mathbb{M}_n)) = \sigma_n(\Phi), \tag{3.1.6}$$

and there are admissible very badly approximable $2 \times 2$ matrix functions $\Phi$ for which the strict inequality

$$\operatorname{dist}_{L^\infty(\boldsymbol{S}_1^2)}(\Phi, H^\infty(\mathbb{M}_2)) < t_0(\Phi) + t_1(\Phi)$$

holds. For instance, consider the matrix function

$$\Phi = \begin{pmatrix} \bar{z} & \mathbb{O} \\ \mathbb{O} & \bar{z} \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & \bar{z} \\ -z & 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \bar{z} & \bar{z}^2 \\ -1 & \bar{z} \end{pmatrix}.$$

Clearly, $\Phi$ has superoptimal singular values $t_0(\Phi) = t_1(\Phi) = 1$. Let

$$F = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbb{O} & \mathbb{O} \\ -1 & \mathbb{O} \end{pmatrix}.$$

It is not difficult to verify that

$$s_0((\Phi - F)(\zeta)) = \frac{1}{2}\sqrt{3 + \sqrt{5}} \text{ and } s_1((\Phi - F)(\zeta)) = \frac{1}{2}\sqrt{3 - \sqrt{5}}$$

for all $\zeta \in \mathbb{T}$. Therefore

$$\operatorname{dist}_{L^\infty(\boldsymbol{S}_1^2)}(\Phi, H^\infty(\mathbb{M}_2)) \le \|\Phi - F\|_{L^\infty(\boldsymbol{S}_1^2)} < 2 = t_0(\Phi) + t_1(\Phi). \tag{3.1.7}$$

By virtue of Theorem 3.1.1 and the remarks following it, one may ask whether it is possible to characterize the matrix functions $\Phi$ for which (3.1.2) becomes an equality. So, let $\Phi$ be an admissible $n \times n$ matrix function with a superoptimal approximant $Q$ in $H^\infty(\mathbb{M}_n)$ for which equality in Theorem 3.1.1 holds with $k = n$. In this case, it must be that

$$\operatorname{dist}_{L^\infty(\boldsymbol{S}_1^n)}(\Phi, H^\infty(\mathbb{M}_n)) = \sum_{j=0}^{n-1} t_j(\Phi) = \sum_{j=0}^{n-1} s_j((\Phi - Q)(\zeta)) = \|\Phi - Q\|_{L^\infty(\boldsymbol{S}_1^n)}$$

by (3.1.6) and thus the superoptimal approximant $Q$ must be a best approximant to $\Phi$ under the $L^\infty(\boldsymbol{S}_1^n)$ norm as well. Hence, we are led to investigate the following problems:

1. For which matrix functions $\Phi$ does Extremal problem 3.1.1 have a solution?
2. If $Q_\$$ is a best approximant to $\Phi$ under the $L^\infty(\boldsymbol{S}_1^n)$-norm, when does it follow that $Q_\$$ is the superoptimal approximant to $\Phi$ in $L^\infty(\mathbb{M}_n)$?
3. Can we find necessary and sufficient conditions on $\Phi$ to obtain equality in (3.1.2) of Theorem 3.1.1?

Before addressing these problems, we recall certain standard principles of functional analysis in Section 3.2 that are used throughout this chapter. In particular, we give their explicit formulation for the spaces $L^p(\boldsymbol{S}_q^{m,n})$.

In Section 3.3, we introduce the Hankel-type operators $H_\Phi^{\{k\}}$ on spaces of matrix functions and $k$-extremal functions, and prove that the number $\sigma_k(\Phi)$ equals the operator norm of $H_\Phi^{\{k\}}$. We also show that Extremal problem 3.1.1 has a solution if and only if the Hankel-type operator $H_\Phi^{\{k\}}$ has a maximizing vector, and thus answer question 1 in terms Hankel-type operators.

In Section 3.4, we establish the main results of this chapter concerning best approximation under the $L^\infty(\boldsymbol{S}_1^{m,n})$ norm (Theorem 3.4.7) and the sum of superoptimal singular values (Theorem 3.4.13). The latter result characterizes the smallest number $k$ for which

$$\int_{\mathbb{T}} \operatorname{trace}(\Phi(\zeta)\Psi(\zeta))d\boldsymbol{m}(\zeta)$$

equals the sum of all non-zero superoptimal singular values for some function $\Psi \in \mathcal{A}_k^{n,m}$. These results serve as partial solutions to problems 2 and 3.

Lastly, in Section 3.5, we restrict our attention to unitary-valued very badly approximable matrix functions. For any such matrix function $U$, we provide a representation of any function $\Psi$ for which the formula

$$\int_{\mathbb{T}} \operatorname{trace}(U(\zeta)\Psi(\zeta))d\boldsymbol{m}(\zeta) = n$$

holds.

## 3.2 Best approximation in $L^q(\boldsymbol{S}_p^{m,n})$ and dual extremal problems

We now provide explicit formulation of some basic results concerning best approximation in $H^q(\boldsymbol{S}_p^{m,n})$ for functions in $L^q(\boldsymbol{S}_p^{m,n})$ and the corresponding dual extremal problem. We first consider the general setting.

**Definition 3.2.1.** Let $X$ be a normed space, $M$ be a closed subspace of $X$, and $x_0 \in X$. We say that *$m_0$ is a best approximant to $x_0$ in $M$* if $m_0 \in M$ and

$$\|x_0 - m_0\|_X = \operatorname{dist}(x_0, M) \stackrel{\text{def}}{=} \inf\{\|x_0 - m\|_X : m \in M\}.$$

It is known that *if $X$ is a reflexive Banach space and $M$ is a closed subspace of $X$, then each $x_0 \in X \setminus M$ has a best approximant $m_0$ in $M$.*

Two standard principles from functional analysis are used throughout this chapter.

Namely, if $X$ is a normed space with a linear subspace $M$, then for any $\Lambda_0 \in X^*$ and $x_0 \in X$

$$\sup_{m \in M, \|m\| \le 1} |\Lambda_0(m)| = \min\left\{\|\Lambda_0 - \Lambda\| : \Lambda \in M^\perp\right\} \quad \text{and}$$

$$\max_{\Lambda \in M^\perp, \|\Lambda\| \le 1} |\Lambda(x_0)| = \operatorname{dist}(x_0, M) \quad \text{whenever } M \text{ is closed.}$$

We now discuss these results in the case of the spaces $L^q(\boldsymbol{S}_p^{m,n})$.

Let $1 \le q < \infty$ and $1 \le p \le \infty$. *Here $p'$ denotes the conjugate exponent to $p$*, i.e. $p' = p/(p-1)$.

It is known that the dual space of $L^q(\boldsymbol{S}_p^{m,n})$ is isometrically isomorphic to $L^{q'}(\boldsymbol{S}_{p'}^{n,m})$ via the mapping $\Phi \mapsto \Lambda_\Phi$, where $\Phi \in L^{q'}(\boldsymbol{S}_{p'}^{n,m})$ and

$$\Lambda_\Phi(\Psi) = \int_{\mathbb{T}} \operatorname{trace}(\Phi(\zeta)\Psi(\zeta))d\boldsymbol{m}(\zeta) \quad \text{for } \Psi \in L^q(\boldsymbol{S}_p^{m,n}).$$

In particular, it follows that the annihilator of $H^q(\boldsymbol{S}_p^{m,n})$ in $L^q(\boldsymbol{S}_p^{m,n})$ is given by $H_0^{q'}(\boldsymbol{S}_{p'}^{n,m})$, and so

$$\operatorname{dist}_{L^q(\boldsymbol{S}_p^{m,n})}(\Phi, H^q(\boldsymbol{S}_p^{m,n})) = \max_{\|\Psi\|_{H_0^{q'}(\boldsymbol{S}_{p'}^{n,m})} \le 1} \left| \int_{\mathbb{T}} \operatorname{trace}(\Phi(\zeta)\Psi(\zeta))d\boldsymbol{m}(\zeta) \right|.$$

by our remarks at the beginning of this section. Moreover, if $1 < q < \infty$, then $\Phi \in L^q(\boldsymbol{S}_p^{m,n})$ has a best approximant $Q$ in $H^q(\boldsymbol{S}_p^{m,n})$ (as $L^q(\boldsymbol{S}_p^{m,n})$ is reflexive); that is,

$$\|\Phi - Q\|_{L^q(\boldsymbol{S}_p^{m,n})} = \operatorname{dist}_{L^q(\boldsymbol{S}_p^{m,n})}(\Phi, H^q(\boldsymbol{S}_p^{m,n})).$$

The situation is similar in the case of $L^\infty(\boldsymbol{S}_p^{m,n})$. Indeed, $L^\infty(\boldsymbol{S}_p^{m,n})$ is a dual space, and so there is a $Q \in H^\infty(\boldsymbol{S}_p^{m,n})$ such that

$$\|\Phi - Q\|_{L^\infty(\boldsymbol{S}_p^{m,n})} = \operatorname{dist}_{L^\infty(\boldsymbol{S}_p^{m,n})}(\Phi, H^\infty(\boldsymbol{S}_p^{m,n})).$$

Again, it also follows from our remarks at the beginning of this section that

$$\operatorname{dist}_{L^\infty(\boldsymbol{S}_p^{m,n})}(\Phi, H^\infty(\boldsymbol{S}_p^{m,n})) = \sup_{\|\Psi\|_{H_0^1(\boldsymbol{S}_{p'}^{n,m})}\le 1} \left| \int_{\mathbb{T}} \operatorname{trace}(\Phi(\zeta)\Psi(\zeta))\,d\boldsymbol{m}(\zeta) \right|.$$

However, an extremal function may fail to exist in this case even if $\Phi$ is a scalar-valued function. An example can be deduced from Section 1 of Chapter 1 in [Pe1].

## 3.3 $\sigma_k(\Phi)$ as the norm of a Hankel-type operator and $k$-extremal functions

We now introduce the Hankel-type operators $H_\Phi^{\{k\}}$ which act on spaces of matrix functions. We prove that the number $\sigma_k(\Phi)$ equals the operator norm of $H_\Phi^{\{k\}}$ and characterize when $H_\Phi^{\{k\}}$ has a maximizing vector.

We begin by establishing the following lemma.

**Lemma 3.3.1.** *Let* $1 \le k \le \min\{m, n\}$. *If* $\Psi \in H^1(\mathbb{M}_{n,m})$ *is such that* $\operatorname{rank}\Psi(\zeta) = k$ *for a.e.* $\zeta \in \mathbb{T}$, *then there are functions* $R \in H^2(\mathbb{M}_{n,k})$ *and* $Q \in H^2(\mathbb{M}_{k,m})$ *such that* $R(\zeta)$ *has rank equal to* $k$ *for almost every* $\zeta \in \mathbb{T}$,

$$\Psi = RQ \;\; and \;\; \|R(\zeta)\|^2_{\mathbb{M}_{n,k}} = \|Q(\zeta)\|^2_{\mathbb{M}_{k,m}} = \|\Psi(\zeta)\|_{\mathbb{M}_{n,m}} \; for\ a.e.\ \zeta \in \mathbb{T}.$$

*Proof.* Consider the set

$$\mathscr{A} = \operatorname{clos}_{L^1(\mathbb{C}^n)}\{f \in H^1(\mathbb{C}^n) : \; f(\zeta) \in \operatorname{Range}\Psi(\zeta) \;\text{ a.e. on } \mathbb{T}\}.$$

Since $\mathscr{A}$ is a non-trivial invariant subspace of $H^1(\mathbb{C}^n)$ under multiplication by $z$, there is an $n \times r$ inner function $\Theta$ such that $\mathscr{A} = \Theta H^1(\mathbb{C}^r)$. We first show that $r = k$. Let $\{e_j\}_{j=1}^r$ be an orthonormal basis for $\mathbb{C}^r$. Then for almost every $\zeta \in \mathbb{T}$, we have that $\{\Theta(\zeta)e_j\}_{j=1}^r$ is a linearly independent set, since $\Theta$ is inner. Moreover, $\{\Theta(\zeta)e_j\}_{j=1}^r$ is a basis for $\operatorname{Range}\Theta(\zeta) = \operatorname{Range}\Psi(\zeta)$ for a.e. $\zeta \in \mathbb{T}$. Since $\dim\operatorname{Range}\Psi(\zeta) = k$ a.e. on $\mathbb{T}$, it follows that $r = \dim\operatorname{Range}\Theta(\zeta) = \dim\operatorname{Range}\Psi(\zeta) = k$. In particular, we obtain that

$$\mathscr{A} = \Theta H^1(\mathbb{C}^k).$$

Therefore, $\Psi = \Theta F$ for some $k \times m$ matrix function $F \in H^1(\mathbb{M}_{k,m})$, because the columns of $\Psi$ belong to $\mathscr{A}$.

Let $h$ be an outer function in $H^2$ such that $|h(\zeta)| = \|\Psi(\zeta)\|_{\mathbb{M}_{n,m}}^{1/2}$ for a.e. $\zeta \in \mathbb{T}$. (The existence of $h$ is a consequence of the fact that $\log\|\Psi(\zeta)\|_{\mathbb{M}_{n,m}} \in L^1$ as $\Psi \in H^1(\mathbb{M}_{n,m})$.) Thus, the matrix functions

$$R = h\Theta \quad \text{and} \quad Q = h^{-1}F$$

have the desired properties. □

**Definition 3.3.2.** Let $\Phi \in L^\infty(\mathbb{M}_{m,n})$, $1 \le k \le \min\{m,n\}$, and $\rho : L^2(\boldsymbol{S}_1^{m,k}) \to L^2(\boldsymbol{S}_1^{m,k})/H^2(\boldsymbol{S}_1^{m,k})$ denote the natural quotient map. We define the *Hankel-type operator* $H_\Phi^{\{k\}} : H^2(\mathbb{M}_{n,k}) \to L^2(\boldsymbol{S}_1^{m,k})/H^2(\boldsymbol{S}_1^{m,k})$ by setting

$$H_\Phi^{\{k\}}F \stackrel{\text{def}}{=} \rho(\Phi F) \quad \text{for } F \in H^2(\mathbb{M}_{n,k}).$$

The norm in the quotient space $L^2(\boldsymbol{S}_1^{m,k})/H^2(\boldsymbol{S}_1^{m,k})$ is the natural one; that is, the norm of a coset equals the infimum of the $L^2(\boldsymbol{S}_1^{m,k})$-norms of its elements.

**Theorem 3.3.3.** *Let $1 \le k \le \min\{m, n\}$. If $\Phi \in L^\infty(\mathbb{M}_{m,n})$, then*

$$\sigma_k(\Phi) = \left\| H_\Phi^{\{k\}} \right\|_{H^2(\mathbb{M}_{n,k}) \to L^2(\boldsymbol{S}_1^{m,k})/H^2(\boldsymbol{S}_1^{m,k})}.$$

*Proof.* Consider the collection

$$\mathcal{B}_k^{n,m} = \{RQ : \|R\|_{H^2(\mathbb{M}_{n,k})} \le 1, \|Q\|_{H_0^2(\mathbb{M}_{k,m})} \le 1\}.$$

We claim that $\mathcal{B}_k^{n,m} = \mathcal{A}_k^{n,m}$. Indeed if $\Psi \in \mathcal{A}_k$ satisfies $\operatorname{rank} \Psi(\zeta) = j$ for $\zeta \in \mathbb{T}$, where $1 \le j \le k$, then by Lemma 3.3.1 there are functions $R \in H^2(\mathbb{M}_{n,j})$ and $Q \in H_0^2(\mathbb{M}_{j,m})$ such that $R(\zeta)$ has rank equal to $j$ for almost every $\zeta \in \mathbb{T}$,

$$\Psi = RQ \text{ and } \|R(\zeta)\|^2_{\mathbb{M}_{n,j}} = \|Q(\zeta)\|^2_{\mathbb{M}_{j,m}} = \|\Psi(\zeta)\|_{\mathbb{M}_{n,m}} \text{ for a.e. } \zeta \in \mathbb{T}.$$

We may now add zeros, if necessary, to obtain $n \times k$ and $k \times m$ matrix functions

$$R_\# = \begin{pmatrix} R & \mathbb{O} \end{pmatrix} \text{ and } Q_\# = \begin{pmatrix} Q \\ \mathbb{O} \end{pmatrix},$$

respectively, from which it follows that $\Psi = R_\# Q_\# \in \mathcal{B}_k^{n,m}$. Therefore $\mathcal{A}_k^{n,m} \subset \mathcal{B}_k^{n,m}$. The reverse inclusion is trivial and so these sets are equal.

Hence

$$\begin{aligned}
\sigma_k(\Phi) &= \sup_{\|R\|_{H^2(\mathbb{M}_{n,k})} \le 1} \; \sup_{\|Q\|_{H_0^2(\mathbb{M}_{k,m})} \le 1} \left| \int_{\mathbb{T}} \operatorname{trace}(\Phi(\zeta) R(\zeta) Q(\zeta)) d\boldsymbol{m}(\zeta) \right| \\
&= \sup_{\|R\|_{H^2(\mathbb{M}_{n,k})} \le 1} \operatorname{dist}_{L^2(\boldsymbol{S}_1^{m,k})}(\Phi R, H^2(\mathbb{M}_{m,k})) \\
&= \|H_\Phi^{\{k\}}\|_{H^2(\mathbb{M}_{n,k}) \to L^2(\boldsymbol{S}_1^{m,k})/H^2(\boldsymbol{S}_1^{m,k})}. \quad \square
\end{aligned}$$

**Definition 3.3.4.** Let $\Phi \in L^\infty(\mathbb{M}_{m,n})$ and $1 \le k \le \min\{m, n\}$. We say that $\Psi$ is *a*

*$k$-extremal function for* $\Phi$ if $\Psi \in \mathcal{A}_k^{n,m}$ and

$$\sigma_k(\Phi) = \int_{\mathbb{T}} \text{trace}(\Phi(\zeta)\Psi(\zeta))d\boldsymbol{m}(\zeta).$$

Thus a matrix function $\Phi$ has a $k$-extremal function if and only if Extremal problem 3.1.1 has a positive solution.

We can now describe matrix functions that have a $k$-extremal function in terms of Hankel-type operators.

**Theorem 3.3.5.** *Let $\Phi \in L^\infty(\mathbb{M}_{m,n})$. The matrix function $\Phi$ has a $k$-extremal function if and only if the Hankel-type operator $H_\Phi^{\{k\}}$ has a maximizing vector.*

*Proof.* To simplify notation, let

$$\left\|H_\Phi^{\{k\}}\right\| \stackrel{\text{def}}{=} \left\|H_\Phi^{\{k\}}\right\|_{H^2(\mathbb{M}_{n,k}) \to L^2(\boldsymbol{S}_1^{m,k})/H^2(\boldsymbol{S}_1^{m,k})}.$$

Suppose $\Psi$ is a $k$-extremal function for $\Phi$. Let $j \in \mathbb{N}$ be such that $j \le k$ and

$$\text{rank}\,\Psi(\zeta) = j \text{ for a.e. } \zeta \in \mathbb{T}.$$

By Lemma 3.3.1, there is an $R \in H^2(\mathbb{M}_{n,j})$ and a $Q \in H_0^2(\mathbb{M}_{j,m})$ such that

$$\Psi = RQ \text{ and } \|R(\zeta)\|_{\mathbb{M}_{n,j}}^2 = \|Q(\zeta)\|_{\mathbb{M}_{j,m}}^2 = \|\Psi(\zeta)\|_{\mathbb{M}_{n,m}} \text{ for a.e. } \zeta \in \mathbb{T}.$$

As before, adding zeros if necessary, we obtain $n \times k$ and $k \times m$ matrix functions

$$R_\# = \begin{pmatrix} R & \mathbb{O} \end{pmatrix} \text{ and } Q_\# = \begin{pmatrix} Q \\ \mathbb{O} \end{pmatrix},$$

respectively, so that $\Psi = R_{\#}Q_{\#}$ and

$$\|Q_{\#}(\zeta)\|^2_{\mathbb{M}_{k,m}} = \|Q(\zeta)\|^2_{\mathbb{M}_{j,m}} = \|\Psi(\zeta)\|_{\mathbb{M}_{n,m}} \text{ for a.e. } \zeta \in \mathbb{T}.$$

Let us show that $R_{\#}$ is a maximizing vector for $H_{\Phi}^{\{k\}}$. Since $Q_{\#}$ belongs to $H_0^2(\mathbb{M}_{k,m})$, we have that for any $F \in H^2(\boldsymbol{S}_1^{m,k})$

$$\begin{aligned}\sigma_k(\Phi) &= \int_{\mathbb{T}} \text{trace}(\Phi(\zeta)\Psi(\zeta))d\boldsymbol{m}(\zeta) = \int_{\mathbb{T}} \text{trace}(\Phi(\zeta)R_{\#}(\zeta)Q_{\#}(\zeta))d\boldsymbol{m}(\zeta)\\ &= \int_{\mathbb{T}} \text{trace}((\Phi R_{\#} - F)(\zeta)Q_{\#}(\zeta))d\boldsymbol{m}(\zeta).\end{aligned}$$

and so

$$\begin{aligned}\sigma_k(\Phi) &= \left|\int_{\mathbb{T}} \text{trace}((\Phi R_{\#} - F)(\zeta)Q_{\#}(\zeta))d\boldsymbol{m}(\zeta)\right|\\ &\le \int_{\mathbb{T}} \left|\text{trace}((\Phi R_{\#} - F)(\zeta)Q_{\#}(\zeta))\right| d\boldsymbol{m}(\zeta)\\ &\le \int_{\mathbb{T}} \|(\Phi R_{\#} - F)(\zeta)Q_{\#}(\zeta)\|_{\boldsymbol{S}_1^m} d\boldsymbol{m}(\zeta)\\ &\le \int_{\mathbb{T}} \|(\Phi R_{\#} - F)(\zeta)\|_{\boldsymbol{S}_1^{m,k}} \|Q_{\#}(\zeta)\|_{\mathbb{M}_{k,m}} d\boldsymbol{m}(\zeta)\\ &\le \|\Phi R_{\#} - F\|_{L^2(\boldsymbol{S}_1^{m,k})} \|Q_{\#}\|_{L^2(\mathbb{M}_{k,m})}\\ &= \|\Phi R_{\#} - F\|_{L^2(\boldsymbol{S}_1^{m,k})} \|\Psi\|_{L^1(\mathbb{M}_{n,m})}\\ &\le \|\Phi R_{\#} - F\|_{L^2(\boldsymbol{S}_1^{m,k})}.\end{aligned}$$

By Theorem 3.3.3, we obtain that

$$\sigma_k(\Phi) \le \left\|H_{\Phi}^{\{k\}} R_{\#}\right\|_{L^2(\boldsymbol{S}_1^{m,k})/H^2(\boldsymbol{S}_1^{m,k})} \le \left\|H_{\Phi}^{\{k\}}\right\| = \sigma_k(\Phi),$$

and therefore

$$\left\|H_{\Phi}^{\{k\}}\right\| = \left\|H_{\Phi}^{\{k\}} R_{\#}\right\|_{L^2(\boldsymbol{S}_1^{m,k})/H^2(\boldsymbol{S}_1^{m,k})}.$$

Thus, $R_{\#}$ is a maximizing vector of $H_\Phi$.

Conversely, suppose the Hankel-type operator $H_\Phi^{\{k\}}$ has a maximizing vector $R \in H^2(\mathbb{M}_{n,k})$. Without loss of generality, we may assume that $\|R\|_{L^2(\mathbb{M}_{n,k})} = 1$. Then

$$\operatorname{dist}_{L^2(\boldsymbol{S}_1^{m,k})}(\Phi R, H^2(\boldsymbol{S}_1^{m,k})) = \left\|H_\Phi^{\{k\}}\right\|.$$

By the remarks in Section 3.2, there is a function $G \in H_0^2(\mathbb{M}_{k,m})$ such that $\|G\|_{L^2(\mathbb{M}_{k,m})} \le 1$ and

$$\int_{\mathbb{T}} \operatorname{trace}((\Phi R)(\zeta)G(\zeta))d\boldsymbol{m}(\zeta) = \operatorname{dist}_{L^2(\boldsymbol{S}_1^{m,k})}(\Phi R, H^2(\boldsymbol{S}_1^{m,k})).$$

On the other hand, since $R$ is a maximizing vector of $H_\Phi^{\{k\}}$, it follows from Theorem 3.3.3 that

$$\int_{\mathbb{T}} \operatorname{trace}(\Phi(\zeta)(RG)(\zeta))d\boldsymbol{m}(\zeta) = \left\|H_\Phi^{\{k\}}\right\| = \sigma_k(\Phi).$$

Hence $\Psi \overset{\text{def}}{=} RG$ is a $k$-extremal function for $\Phi$. □

Before stating the next result, let us recall that the Hankel operator $H_\Phi : H^2(\mathbb{C}^n) \to H^2_-(\mathbb{C}^m)$ is defined by $H_\Phi f = \mathbb{P}_-\Phi f$ for $f \in H^2(\mathbb{C}^n)$. The following is an immediate consequence of the previous theorem when $k = 1$.

**Corollary 3.3.6.** *Let $\Phi \in L^\infty(\mathbb{M}_{m,n})$. The Hankel operator $H_\Phi$ has a maximizing vector if and only if $\Phi$ has a 1-extremal function.*

*Proof.* By Theorem 3.3.5, $\Phi$ has a 1-extremal function if and only if the Hankel-type operator $H_\Phi^{\{1\}} : H^2(\mathbb{C}^n) \to L^2(\mathbb{C}^m)/H^2(\mathbb{C}^m)$ has a maximizing vector. The conclusion now follows by considering the "natural" isometric isomorphism between the spaces $H^2_-(\mathbb{C}^m) = L^2(\mathbb{C}^m) \ominus H^2(\mathbb{C}^m)$ and $L^2(\mathbb{C}^m)/H^2(\mathbb{C}^m)$. □

*Remark* 3.3.7. It is worth mentioning that if a matrix function $\Phi$ is such that the Hankel operator $H_\Phi$ has a maximizing vector (e.g. $\Phi \in (H^\infty + C)(\mathbb{M}_n)$), then any

1-extremal function $\Psi$ of $\Phi$ satisfies

$$\int_{\mathbb{T}} \text{trace}(\Phi(\zeta)\Psi(\zeta))dm(\zeta) = \|H_\Phi\| = t_0(\Phi).$$

This is a consequence of Corollary 3.3.6 and Theorem 3.3.3.

*Remark* 3.3.8. There are other characterizations of the class of bounded matrix functions $\Phi$ such that the Hankel operator $H_\Phi$ has a maximizing vector. These involve "dual" extremal functions and "thematic" factorizations. We refer the interested reader to [Pe2] for details.

**Corollary 3.3.9.** *Let $1 \le k \le \ell \le n$ and $\Phi \in L^\infty(\mathbb{M}_n)$. Suppose that $\sigma_k(\Phi) = \sigma_\ell(\Phi)$. If $H_\Phi^{\{k\}}$ has a maximizing vector, then $H_\Phi^{\{\ell\}}$ also has a maximizing vector.*

*Proof.* This is an immediate consequence of Theorem 3.3.5. □

## 3.4 How about the sum of superoptimal singular values?

In this section, we prove in Theorem 3.4.7 that equality is obtained in (3.1.2) under some natural conditions.

For the rest of this chapter, we assume that $m = n$.

Consider the non-decreasing sequence $\sigma_1(\Phi), \dots, \sigma_n(\Phi)$. Recall that

$$\sigma_n(\Phi) = \text{dist}_{L^\infty(\boldsymbol{S}_1^n)}(\Phi, H^\infty(\mathbb{M}_n))$$

and the distance on the right-hand side is in fact always attained, i.e. a best approximant $Q$ to $\Phi$ under the $L^\infty(\boldsymbol{S}_1^n)$ norm always exists as explained in Section 3.2.

**Theorem 3.4.1.** *Let $\Phi \in L^\infty(\mathbb{M}_n)$ and $1 \le k \le n$. Suppose $Q$ is a best approximant to $\Phi$ in $H^\infty(\mathbb{M}_n)$ under the $L^\infty(\boldsymbol{S}_1^n)$-norm. If the Hankel-type operator $H_\Phi^{\{k\}}$ has a maximizing vector $\mathcal{F}$ in $H^2(\mathbb{M}_{n,k})$ and $\sigma_k(\Phi) = \sigma_n(\Phi)$, then*

1. *$Q\mathcal{F}$ is a best approximant to $\Phi\mathcal{F}$ in $H^2$ under the $L^2(\boldsymbol{S}_1^{n,k})$-norm,*
2. *for each $j \ge 0$,*

$$s_j((\Phi - Q)(\zeta)\mathcal{F}(\zeta)) = s_j((\Phi - Q)(\zeta))\|\mathcal{F}(\zeta)\|_{\mathbb{M}_{n,k}} \text{ for a.e. } \zeta \in \mathbb{T},$$

3. *$\sum_{j=0}^{k-1} s_j((\Phi - Q)(\zeta) = \sigma_k(\Phi)$ holds for a.e. $\zeta \in \mathbb{T}$, and*
4. *$s_j((\Phi - Q)(\zeta)) = 0$ holds for a.e. $\zeta \in \mathbb{T}$ whenever $j \ge k$.*

*Proof.* By our assumptions,

$$\begin{aligned}
\|H_\Phi^{\{k\}}\|^2\|\mathcal{F}\|^2_{L^2(\mathbb{M}_{n,k})} &= \|H_\Phi^{\{k\}}\mathcal{F}\|^2_{L^2(\boldsymbol{S}_1^{n,k})/H^2(\boldsymbol{S}_1^{n,k})} = \|\rho(\Phi\mathcal{F})\|^2 \\
&= \|\rho((\Phi - Q)\mathcal{F})\|^2 \\
&\le \|(\Phi - Q)\mathcal{F}\|^2_{L^2(\boldsymbol{S}_1^{n,k})} = \int_{\mathbb{T}} \|(\Phi - Q)(\zeta)\mathcal{F}(\zeta)\|^2_{\boldsymbol{S}_1^{n,k}} \boldsymbol{dm}(\zeta) \\
&\le \int_{\mathbb{T}} \|(\Phi - Q)(\zeta)\|^2_{\boldsymbol{S}_1^n}\|\mathcal{F}(\zeta)\|^2_{\mathbb{M}_{n,k}} \boldsymbol{dm}(\zeta) \\
&\le \|\Phi - Q\|^2_{L^\infty(\boldsymbol{S}_1^n)}\|\mathcal{F}\|^2_{L^2(\mathbb{M}_{n,k})} = \sigma_k(\Phi)^2\|\mathcal{F}\|^2_{L^2(\mathbb{M}_{n,k})}.
\end{aligned}$$

It follows from Theorem 3.3.3 that all inequalities are equalities. In particular, we obtain that $Q\mathcal{F}$ is a best approximant to $\Phi Q$ under the $L^2(\boldsymbol{S}_1^{n,k})$-norm since the first inequality is actually an equality. For almost every $\zeta \in \mathbb{T}$,

$$\begin{aligned}
\|(\Phi - Q)(\zeta)\mathcal{F}(\zeta)\|_{\boldsymbol{S}_1^n} &= \|(\Phi - Q)(\zeta)\|_{\boldsymbol{S}_1^n}\|\mathcal{F}(\zeta)\|_{\mathbb{M}_{n,k}} \text{ and} \qquad (3.4.1) \\
\|(\Phi - Q)(\zeta)\|_{\boldsymbol{S}_1^n} &= \|\Phi - Q\|_{L^\infty(\boldsymbol{S}_1^n)} = \sigma_k(\Phi),
\end{aligned}$$

because the second and third inequalities are equalities as well. It follows from (3.4.1) that for each $j \geq 0$,

$$s_j((\Phi - Q)(\zeta)\mathcal{F}(\zeta)) = s_j((\Phi - Q)(\zeta))\|\mathcal{F}(\zeta)\|_{\mathbb{M}_{n,k}} \text{ for a.e. } \zeta \in \mathbb{T}.$$

We claim that if $j \geq k$, then $s_j((\Phi - Q)(\zeta)) = 0$ for a.e. $\zeta \in \mathbb{T}$. By Theorem 3.3.5, we can choose a $k$-extremal function, say $\Psi$, for $\Phi$. Since $\Psi$ belongs to $H_0^1(\mathbb{M}_n)$,

$$\begin{aligned}\sigma_k(\Phi) &= \int_{\mathbb{T}} \operatorname{trace}(\Phi(\zeta)\Psi(\zeta))\boldsymbol{dm}(\zeta) = \int_{\mathbb{T}} \operatorname{trace}((\Phi - Q)(\zeta)\Psi(\zeta))\boldsymbol{dm}(\zeta) \\ &\leq \int_{\mathbb{T}} \|(\Phi - Q)(\zeta)\Psi(\zeta)\|_{\boldsymbol{S}_1^n}\boldsymbol{dm}(\zeta) \leq \int_{\mathbb{T}} \|(\Phi - Q)(\zeta)\|_{\boldsymbol{S}_1^n}\|\Psi(\zeta)\|_{\mathbb{M}_n}\boldsymbol{dm}(\zeta) \\ &\leq \|\Phi - Q\|_{L^\infty(\boldsymbol{S}_1^n)}\|\Psi\|_{L^1(\mathbb{M}_n)} \leq \|\Phi - Q\|_{L^\infty(\boldsymbol{S}_1^n)} = \sigma_k(\Phi),\end{aligned}$$

and so all inequalities are equalities. It follows that

$$|\operatorname{trace}((\Phi - Q)(\zeta)\Psi(\zeta))| = \|(\Phi - Q)(\zeta)\|_{\boldsymbol{S}_1^n}\|\Psi(\zeta)\|_{\mathbb{M}_n} \text{ for a.e. } \zeta \in \mathbb{T}. \quad (3.4.2)$$

In order to complete the proof, we need the following lemma.

**Lemma 3.4.2.** *Let $A \in \mathbb{M}_n$ and $B \in \mathbb{M}_n$. Suppose that $A$ and $B$ satisfy*

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

*If* $\operatorname{rank} A \leq k$, *then* $\operatorname{rank} B \leq k$ *as well.*

We first finish the proof of Theorem 3.4.1 before proving Lemma 3.4.2.

It follows from (3.4.2) and Lemma 3.4.2 that

$$\operatorname{rank}((\Phi - Q)(\zeta)) \leq k \text{ for a.e. } \zeta \in \mathbb{T}.$$

In particular, if $j \geq k$, then

$$s_j((\Phi - Q)(\zeta)) = 0 \text{ for a.e. } \zeta \in \mathbb{T},$$

and so

$$\sum_{j=0}^{k-1} s_j((\Phi - Q)(\zeta)) = \|(\Phi - Q)(\zeta)\|_{\boldsymbol{S}_1^n} = \sigma_k(\Phi) \text{ for a.e. } \zeta \in \mathbb{T}.$$

This completes the proof. □

*Remark* 3.4.3. Lemma 3.4.2 is a slight modification of Lemma 4.6 in [BNP]. Although the proof of Lemma 3.4.2 given below is almost the same as that given in [BNP] for Lemma 4.6, we include it for the convenience of the reader.

*Proof of Lemma 3.4.2.* Let $B$ have polar decomposition $B = UP$ and set $C = AU$, where $P = (B^*B)^{1/2}$. Let $e_1, \ldots, e_n$ be an orthonormal basis of eigenvectors for $P$ and $Pe_j = \lambda_j e_j$. It is easy to see that the following inequalities hold:

$$\begin{aligned}
|\operatorname{trace}(AB)| = |\operatorname{trace}(CP)| &= \left|\sum_{j=1}^{n}(Pe_j, C^*e_j)\right| = \left|\sum_{j=1}^{n}\lambda_j(e_j, C^*e_j)\right| \\
&= \left|\sum_{j=1}^{n}\lambda_j(Ce_j, e_j)\right| \leq \sum_{j=1}^{n}\lambda_j\,|(Ce_j, e_j)| \leq \sum_{j=1}^{n}\lambda_j\|Ce_j\| \\
&\leq \|C\|_{\mathbb{M}_n}\sum_{j=1}^{n}\lambda_j.
\end{aligned}$$

On the other hand,

$$\|A\|_{\mathbb{M}_n}\|B\|_{\boldsymbol{S}_1^n} = \|C\|_{\mathbb{M}_m}\|P\|_{\boldsymbol{S}_1^n} = \|C\|_{\mathbb{M}_n}\sum_{j=1}^{n}\lambda_j$$

and so, by the assumption $|\operatorname{trace}(AB)| = \|A\|_{\mathbb{M}_n}\|B\|_{S_1^n}$, it follows that

$$\sum_{j=1}^{n} \lambda_j \|Ce_j\| = \|C\|_{\mathbb{M}_n} \sum_{j=1}^{n} \lambda_j.$$

Therefore $\lambda_j\|Ce_j\| = \|C\|_{\mathbb{M}_n}\lambda_j$ for each $j$. However, if $\operatorname{rank} A \le k$, then $\operatorname{rank} C \le k$. Thus there are at most $k$ vectors $e_j$ such that $\|Ce_j\| = \|C\|_{\mathbb{M}_n}$. In particular, there are at least $n-k$ vectors $e_j$ such that $\|Ce_j\| < \|C\|_{\mathbb{M}_n}$. Thus, $\lambda_j = 0$ for those $n-k$ vectors $e_j$, $\operatorname{rank} P \le k$, and so $\operatorname{rank} B \le k$. □

*Remark* 3.4.4. Note that the distance function $d_\Phi$ defined on $\mathbb{T}$ by

$$d_\Phi(\zeta) \stackrel{\text{def}}{=} \|(\Phi - Q)(\zeta)\|_{S_1^n}$$

equals $\sigma_k(\Phi)$ for almost every $\zeta \in \mathbb{T}$ and is therefore independent of the choice of the best approximant $Q$. This is an immediate consequence of Theorem 3.4.1. A similar phenomenon occurs in the case of matrix functions $\Phi \in L^p(\mathbb{M}_n)$ for $2 < p < \infty$. We refer the reader to [BNP] for details.

**Corollary 3.4.5.** *Let $\Phi \in L^\infty(\mathbb{M}_n)$ be an admissible matrix function and $1 \le k \le n$. If the Hankel-type operator $H_\Phi^{\{k\}}$ has a maximizing vector and $\sigma_k(\Phi) = \sigma_n(\Phi)$, then*

$$\sum_{j=0}^{k-1} s_j((\Phi - Q)(\zeta)) \le \sum_{j=0}^{k-1} t_j(\Phi)$$

*for any best approximation $Q$ of $\Phi$ in $H^\infty(\mathbb{M}_n)$ under the $L^\infty(S_1^n)$-norm.*

*Proof.* This is an immediate consequence of Theorems 3.1.1 and 3.4.1. □

**Definition 3.4.6.** A matrix function $\Phi \in L^\infty(\mathbb{M}_n)$ is said to have *order $\ell$* if $\ell$ is the smallest number such that $H_\Phi^{\{\ell\}}$ has a maximizing vector and

$$\sigma_\ell(\Phi) = \operatorname{dist}_{L^\infty(S_1^n)}(\Phi, H^\infty(\mathbb{M}_n)).$$

If no such number $\ell$ exists, we say that $\Phi$ is *inaccessible.*

The interested reader should compare this definition of "order" with the one made in [BNP] for matrix functions in $L^p(\mathbb{M}_n)$ for $2 < p < \infty$. Also, due to Corollary 3.3.9, it is clear that if $\Phi \in L^\infty(\mathbb{M}_n)$ has order $\ell$, then the Hankel-type operator $H_\Phi^{\{k\}}$ has a maximizing vector and

$$\sigma_k(\Phi) = \text{dist}_{L^\infty(\boldsymbol{S}_1^n)}(\Phi, H^\infty(\mathbb{M}_n))$$

holds for each $k \geq \ell$.

**Theorem 3.4.7.** *Let $\Phi \in L^\infty(\mathbb{M}_n)$ be an admissible matrix function of order $k$. The following statements are equivalent.*

1. *$Q \in H^\infty$ is a best approximant to $\Phi$ under the $L^\infty(\boldsymbol{S}_1^n)$-norm and the functions*

$$\zeta \mapsto s_j((\Phi - Q)(\zeta)),\ 0 \leq j \leq k-1,$$

*are constant almost everywhere on $\mathbb{T}$.*

2. *$Q$ is the superoptimal approximant to $\Phi$, $t_j(\Phi) = 0$ for $j \geq k$, and*

$$\sigma_k(\Phi) = t_0(\Phi) + \ldots + t_{k-1}(\Phi).$$

*Proof.* We first prove that *1* implies *2.* By Corollary 3.4.5, we have that, for almost every $\zeta \in \mathbb{T}$,

$$\sum_{j=0}^{k-1} s_j((\Phi - Q)(\zeta)) \leq \sum_{j=0}^{k-1} t_j(\Phi) \leq \sum_{j=0}^{k-1} \operatorname*{ess\,sup}_{\zeta\in\mathbb{T}} s_j((\Phi - Q)(\zeta)) = \sum_{j=0}^{k-1} s_j((\Phi - Q)(\zeta)).$$

This implies that

$$t_j(\Phi) = \operatorname*{ess\,sup}_{\zeta\in\mathbb{T}} s_j((\Phi - Q)(\zeta)) = s_j((\Phi - Q)(\zeta)) \quad \text{for } 0 \le j \le k-1,$$

$Q \in \Omega_{k-1}(\Phi)$, and

$$\sum_{j=0}^{k-1} t_j(\Phi) = \sum_{j=0}^{k-1} s_j((\Phi - Q)(\zeta)) = \sigma_k(\Phi).$$

Moreover, Theorem 3.4.1 gives that $s_j((\Phi - Q)(\zeta)) = 0$ a.e. on $\mathbb{T}$ for $j \ge k$, and so $t_j(\Phi) = 0$ for $j \ge k$, as $Q \in \Omega_{k-1}(\Phi)$. Hence, $Q$ is the superoptimal approximant to $\Phi$.

Let us show that *2* implies *1*. Clearly, it suffices to show that if *2* holds, then $Q$ is a best approximant to $\Phi$ under the $L^\infty(\boldsymbol{S}_1^n)$-norm. Suppose *2* holds. In this case, we must have that

$$\sigma_k(\Phi) = \sum_{j=0}^{k-1} t_j(\Phi) = \sum_{j=0}^{k-1} s_j((\Phi - Q)(\zeta)) = \|\Phi - Q\|_{L^\infty(\boldsymbol{S}_1^n)}.$$

Since $\Phi$ has order $k$, it follows that

$$\sigma_n(\Phi) = \|\Phi - Q\|_{L^\infty(\boldsymbol{S}_1^n)}$$

and so the proof is complete. □

For the rest of this section, we restrict ourselves to admissible matrix functions $\Phi$ which are also very badly approximable. Recall that, in this case, the function $\zeta \mapsto s_j(\Phi(\zeta))$ equals $t_j(\Phi)$ a.e. on $\mathbb{T}$ for $0 \le j \le n-1$, as mentioned in Section 1.1. The next result follows at once from Theorem 3.4.7.

**Corollary 3.4.8.** *Let $\Phi$ be an admissible very badly approximable $n \times n$ matrix function of order $k$. The zero matrix function is a best approximant to $\Phi$ under the*

$L^\infty(\boldsymbol{S}_1^n)$*-norm if and only if* $t_j(\Phi) = 0$ *for* $j \geq k$ *and*

$$\sigma_k(\Phi) = t_0(\Phi) + \ldots + t_{k-1}(\Phi).$$

It is natural to question at this point whether or not the collection of admissible very badly approximable matrix functions of order $k$ is non-empty. It turns out that one can easily construct examples of admissible very badly approximable matrix functions of order $k$ (see Examples 3.4.14 and 3.4.15). Theorem 3.4.10 below gives a simple sufficient condition for determining when a very badly approximable matrix function has order $k$. We first need the following lemma.

**Lemma 3.4.9.** *Let* $\Phi \in L^\infty(\mathbb{M}_n)$. *Suppose there is* $\Psi \in \mathcal{A}_k^n$ *such that*

$$\int_{\mathbb{T}} \operatorname{trace}(\Phi(\zeta)\Psi(\zeta))d\boldsymbol{m}(\zeta) = \|\Phi\|_{L^\infty(\boldsymbol{S}_1^n)}.$$

*Then* $\Psi$ *is a* $k$*-extremal function for* $\Phi$, $\sigma_k(\Phi) = \sigma_n(\Phi)$, *and the zero matrix function is a best approximant to* $\Phi$ *under the* $L^\infty(\boldsymbol{S}_1^n)$*-norm.*

*Proof.* By the assumptions on $\Psi$, we have

$$\|\Phi\|_{L^\infty(\boldsymbol{S}_1^n)} = \int_{\mathbb{T}} \operatorname{trace}(\Phi(\zeta)\Psi(\zeta))d\boldsymbol{m}(\zeta) \leq \sigma_k(\Phi).$$

On the other hand,

$$\sigma_k(\Phi) \leq \operatorname{dist}_{L^\infty(\boldsymbol{S}_1^n)}(\Phi, H^\infty) \leq \|\Phi\|_{L^\infty(\boldsymbol{S}_1^n)}$$

always holds. Since all the previously mentioned inequalities are equalities, the conclusion follows. □

**Theorem 3.4.10.** *Let* $\Phi \in L^\infty(\mathbb{M}_n)$ *be an admissible very badly approximable matrix* *function. Suppose there is* $\Psi \in \mathcal{A}_k^n$ *such that*

$$\int_{\mathbb{T}} \text{trace}(\Phi(\zeta)\Psi(\zeta))\boldsymbol{dm}(\zeta) = t_0(\Phi) + \ldots + t_{n-1}(\Phi).$$

*If* $t_{k-1}(\Phi) > 0$, *then* $\Phi$ *has order* $k$ *and the zero matrix function is a best approximant to* $\Phi$ *under the* $L^\infty(\boldsymbol{S}_1^n)$*-norm.*

*Proof.* By the remarks preceding Corollary 3.4.8, it is easy to see that

$$\|\Phi\|_{L^\infty(\boldsymbol{S}_1^n)} = t_0(\Phi) + \ldots + t_n(\Phi).$$

It follows from Lemma 3.4.9 that $\Psi$ is a $k$-extremal function for $\Phi$, $\sigma_k(\Phi) = \sigma_n(\Phi)$, and the zero matrix function is a best approximant to $\Phi$ under the $L^\infty(\boldsymbol{S}_1^n)$-norm. Thus $\|\Phi\|_{L^\infty(\boldsymbol{S}_1^n)} = \sigma_k(\Phi)$. Moreover, by Theorem 3.1.1,

$$\sigma_{k-1}(\Phi) \le t_0(\Phi) + \ldots + t_{k-2}(\Phi) < t_0(\Phi) + \ldots + t_{k-1}(\Phi) \le \|\Phi\|_{L^\infty(\boldsymbol{S}_1^n)}.$$

Therefore $\sigma_{k-1}(\Phi) < \sigma_k(\Phi)$. ☐

*Remark* 3.4.11. Notice that under the hypotheses of Theorem 3.4.10, one also obtains that $t_{k-1}(\Phi)$ is the smallest non-zero superoptimal singular value of $\Phi$. This is an immediate consequence of Corollary 3.4.8.

We now formulate the corresponding result for admissible very badly approximable unitary-valued matrix functions. These functions are considered in greater detail in Section 3.5.

**Corollary 3.4.12.** *Let* $U \in L^\infty(\mathbb{M}_n)$ *be an admissible very badly approximable unitary-valued matrix function. If there is* $\Psi \in \mathcal{A}_n^n$ *such that*

$$\int_{\mathbb{T}} \text{trace}(U(\zeta)\Psi(\zeta))\boldsymbol{dm}(\zeta) = n,$$

*then $U$ has order $n$ and the zero matrix function is a best approximant to $U$ under the $L^\infty(\boldsymbol{S}_1^n)$-norm.*

*Proof.* This is a trivial consequence of Theorem 3.4.10 and the fact that

$$t_j(U) = 1 \text{ for } 0 \le j \le n-1. \qquad \square$$

We are now ready to state the main result of this section.

**Theorem 3.4.13.** *Let $\Phi$ be an admissible very badly approximable $n \times n$ matrix function. The following statements are equivalent:*

1. *$k$ is the smallest number for which there exists $\Psi \in \mathcal{A}_k^n$ such that*

$$\int_{\mathbb{T}} \text{trace}(\Phi(\zeta)\Psi(\zeta))d\boldsymbol{m}(\zeta) = t_0(\Phi) + \ldots + t_{n-1}(\Phi);$$

2. *$\Phi$ has order $k$, $t_j(\Phi) = 0$ for $j \ge k$ and*

$$\sigma_k(\Phi) = t_0(\Phi) + \ldots + t_{k-1}(\Phi).$$

*Proof.* Let

$$\kappa(\Phi) \stackrel{\text{def}}{=} \inf\{\, j \ge 0 : \text{there exists a } \Psi \in \mathcal{A}_j^n \text{ such that} \\ \int_{\mathbb{T}} \text{trace}(\Phi(\zeta)\Psi(\zeta))d\boldsymbol{m}(\zeta) = t_0(\Phi) + \ldots + t_{n-1}(\Phi)\,\}$$

Clearly, $\kappa(\Phi)$ may be infinite for arbitrary $\Phi$.

Suppose $\kappa = \kappa(\Phi)$ is finite. Then Lemma 3.4.9 implies that $\Phi$ has a $\kappa$-extremal function, $\sigma_\kappa(\Phi) = \sigma_n(\Phi)$, and the zero matrix function is a best approximant to $\Phi$ under the $L^\infty(\boldsymbol{S}_1^n)$-norm. In particular, $\Phi$ has order $k \le \kappa(\Phi)$, $t_j(\Phi) = 0$ for $j \ge k$, and

$$\sigma_k(\Phi) = t_0(\Phi) + \ldots + t_{k-1}(\Phi),$$

by Corollary 3.4.8.

On the other hand, if $\Phi$ has order $k$, $t_j(\Phi) = 0$ for $j \geq k$, and

$$\sigma_k(\Phi) = t_0(\Phi) + \ldots + t_{k-1}(\Phi),$$

then $\Phi$ has a $k$-extremal function $\Psi \in \mathcal{A}_k^n$ such that

$$\int_{\mathbb{T}} \text{trace}(\Phi(\zeta)\Psi(\zeta))dm(\zeta) = \sigma_k(\Phi) = t_0(\Phi) + \ldots + t_{k-1}(\Phi).$$

Since $t_j(\Phi) = 0$ for $j \geq k$, it follows that

$$\int_{\mathbb{T}} \text{trace}(\Phi(\zeta)\Psi(\zeta))dm(\zeta) = t_0(\Phi) + \ldots + t_{n-1}(\Phi).$$

Thus $\kappa(\Phi) \leq k$.

Hence, if either $\kappa(\Phi)$ is finite or $\Phi$ satisfies *2*, then $k = \kappa(\Phi)$. □

We end this section by illustrating existence of very badly approximable matrix functions of order $k$ by giving two simple examples; a $2 \times 2$ matrix function of order 2 and a $3 \times 3$ matrix function of order 2.

**Example 3.4.14.** Let

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \bar{z}^2 & \mathbb{O} \\ \mathbb{O} & \bar{z} \end{pmatrix}.$$

It is easy to see that $\Phi$ is a continuous (and hence admissible) unitary-valued very badly approximable matrix function with superoptimal singular values $t_0(\Phi) = t_1(\Phi) = 1$. We claim that $\Phi$ has order 2. Indeed, the matrix function

$$\Psi = \begin{pmatrix} z^2 & \mathbb{O} \\ \mathbb{O} & z \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

satisfies

$$\int_{\mathbb{T}} \text{trace}(\Phi(\zeta)\Psi(\zeta))dm(\zeta) = 2,$$

and so $\Phi$ has order 2 by Corollary 3.4.12.

**Example 3.4.15.** Let $t_0$ and $t_1$ be two positive numbers satisfying $t_0 \geq t_1$. Let

$$\Phi = \begin{pmatrix} t_0\bar{z}^a & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & t_1\bar{z}^b & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{O} \end{pmatrix}$$

where $a$ and $b$ are positive integers. It is easy to see that $\Phi$ is a continuous (and hence admissible) very badly approximable matrix function with superoptimal singular values $t_0(\Phi) = t_0$, $t_1(\Phi) = t_1$, and $t_2(\Phi) = 0$. Again, we have that $\Phi$ has order 2. After all, the matrix function

$$\Psi = \begin{pmatrix} z^a & \mathbb{O} & \mathbb{O} \\ \mathbb{O} & z^b & \mathbb{O} \\ \mathbb{O} & \mathbb{O} & \mathbb{O} \end{pmatrix}$$

satisfies

$$\int_{\mathbb{T}} \text{trace}(\Phi(\zeta)\Psi(\zeta))dm(\zeta) = t_0 + t_1 = t_0(\Phi) + t_1(\Phi) + t_2(\Phi),$$

and so $\Phi$ has order 2 by Theorem 3.4.10, since $t_1(\Phi) = t_1 > 0$.

## 3.5 Unitary-valued very badly approximable matrix functions

We lastly consider the class $\mathcal{U}_n$ of admissible very badly approximable unitary-valued matrix functions of size $n \times n$ and provide a representation of any $n$-extremal function

$\Psi$ for a function $U \in \mathcal{U}_n$ such that

$$\int_{\mathbb{T}} \operatorname{trace}(U(\zeta)\Psi(\zeta))d\boldsymbol{m}(\zeta) = t_0(U) + \ldots + t_{n-1}(U) \tag{3.5.1}$$

holds. Note that for any such $U$ we have that $t_j(U) = 1$ for $0 \le j \le n-1$.

Recall that, for a matrix function $\Phi \in L^\infty(\mathbb{M}_{m,n})$, the *Toeplitz operator* $T_\Phi$ is defined by

$$T_\Phi f = \mathbb{P}_+ \Phi f, \ \text{ for } f \in H^2(\mathbb{C}^n),$$

where $\mathbb{P}_+$ denotes the orthogonal projection from $L^2(\mathbb{C}^n)$ onto $H^2(\mathbb{C}^m)$.

It is well-known that, for any function $U \in \mathcal{U}_n$, the Toeplitz operator $T_U$ is Fredholm and $\operatorname{ind} T_U > 0$. In particular, the Toeplitz operator $T_{\det U}$ is Fredholm and

$$\operatorname{ind} T_{\det U} = \operatorname{ind} T_U.$$

We refer the reader to Chapter 14 in [Pe1] for more information concerning functions in $\mathcal{U}_n$.

**Theorem 3.5.1.** *Suppose $U \in \mathcal{U}_n$ has an $n$-extremal function $\Psi$ such that* (3.5.1) *holds. Then $\Psi$ admits a representation of the form*

$$\Psi = zh^2\Theta,$$

*where $h \in H^2$ is an outer function such that $\|h\|_{L^2} = 1$ and $\Theta$ is a finite Blaschke-Potapov product. Moreover, the scalar functions* $\det(U\Theta)$ *and* $\operatorname{trace}(U\Theta)$ *are admissible badly approximable functions that admit the factorizations*

$$\det(U\Theta) = \bar{z}^n \frac{\bar{h}^n}{h^n} \ \textit{ and } \ \operatorname{trace}(U\Theta) = n\bar{z}\frac{\bar{h}}{h}.$$

*Proof.* It follows from (3.5.1) that all inequalities in (3.1.3) are equalities and so

$$\operatorname{trace}(U(\zeta)\Psi(\zeta)) = \|U(\zeta)\Psi(\zeta)\|_{S_1^n} = n\|\Psi(\zeta)\|_{\mathbb{M}_n} \tag{3.5.2}$$

holds for a.e. $\zeta \in \mathbb{T}$. Since $U$ is unitary-valued, then

$$\|U(\zeta)\Psi(\zeta)\|_{S_1^n} = \|\Psi(\zeta)\|_{S_1^n},$$

and so

$$\|\Psi(\zeta)\|_{S_1^n} = n\|\Psi(\zeta)\|_{\mathbb{M}_n}$$

must hold for a.e. $\zeta \in \mathbb{T}$. Therefore

$$s_j(\Psi(\zeta)) = \|\Psi(\zeta)\|_{\mathbb{M}_n} \text{ for a.e. } \zeta \in \mathbb{T},\ 0 \leq j \leq n-1.$$

By the Singular Value Decomposition Theorem for matrices (or, more generally, the Schmidt Decomposition Theorem), it follows that

$$\Psi(\zeta) = \|\Psi(\zeta)\|_{\mathbb{M}_n} V(\zeta) \text{ for a.e. } \zeta \in \mathbb{T}, \tag{3.5.3}$$

for some unitary-valued matrix function $V$. Let $h \in H^2$ be an outer function such that

$$|h(\zeta)| = \|\Psi(\zeta)\|_{\mathbb{M}_n}^{1/2} \text{ on } \mathbb{T}.$$

Consider also the matrix function $\Xi \stackrel{\text{def}}{=} h^{-2}\Psi$. It follows from (3.5.3) that

$$(\Xi^*\Xi)(\zeta) = \frac{1}{|h(\zeta)|^4}(\Psi^*\Psi)(\zeta) = I_n \text{ for a.e. } \zeta \in \mathbb{T},$$

and so $\Xi$ is an inner function. Thus $\Psi$ admits the factorization

$$\Psi = zh^2\Theta$$

for some $n \times n$ unitary-valued inner function $\Theta$ and an outer function $h \in H^2$ such that $\|h\|_{L^2} = 1$.

Note that the first equality in (3.5.2) indicates that the scalar function $\varphi \stackrel{\text{def}}{=} \operatorname{trace}(U\Theta)$ satisfies

$$zh^2\varphi = n|h|^2 \text{ on } \mathbb{T},$$

or equivalently

$$\varphi = n\bar{z}\frac{\bar{h}}{h}.$$

Moreover, $\|H_{U\Theta}\|_{\mathrm{e}} \leq \|H_U\|_{\mathrm{e}} < 1$, hence $\|H_\varphi\|_{\mathrm{e}} < n = \|H_\varphi\|$ implying that $\varphi$ is an admissible badly approximable scalar function on $\mathbb{T}$. We conclude that the Toeplitz operator $T_\varphi$ is Fredholm and $\operatorname{ind} T_\varphi > 0$ (c.f. Theorem 7.5.5 in [Pel]).

Returning to (3.5.2), it also follows that each eigenvalue of $U(\zeta)\Psi(\zeta)$ equals $\|\Psi(\zeta)\|_{\mathbf{M}_n} = |h(\zeta)|^2$ for a.e. $\zeta \in \mathbb{T}$ . In particular,

$$|h(\zeta)|^{2n} = \det U(\zeta)\Psi(\zeta) = (z^n h^{2n})(\zeta) \cdot \det U(\zeta) \cdot \det \Theta(\zeta)$$

holds a.e. $\zeta \in \mathbb{T}$. By setting

$$\theta \stackrel{\text{def}}{=} \det \Theta \text{ and } u \stackrel{\text{def}}{=} \det U,$$

we have that $u$ admits the factorization

$$u = \bar{\theta}\bar{z}^n\frac{\bar{h}^n}{h^n} = \bar{\theta}\omega^n,$$

where $\omega \stackrel{\text{def}}{=} \bar{z}\bar{h}/h = \varphi/n$. Since the Toeplitz operator $T_\omega$ is Fredholm with positive index, $T_{u\bar{\omega}^n}$ is Fredholm as well. Since $\ker T_\theta = \{\mathbb{O}\}$ and $u\bar{\omega}^n = \bar{\theta}$, then

$$\dim(H^2 \ominus \theta H^2) = \dim\ker T_\theta^* = \dim\ker T_{\bar{\theta}} = \operatorname{ind} T_{\bar{\theta}} < \infty$$

and so $\theta$ is a finite Blaschke product. The conclusion follows from the well-known lemma stated below. □

**Lemma 3.5.2.** *If $\Theta$ is a unitary-valued inner function such that* $\det\Theta$ *is a finite Blaschke product, then $\Theta$ is a Blaschke-Potapov product.*

*Proof.* Let $\theta = \det\Theta$. It is easy to see that $\Theta^*\theta$ is an inner function. Since $B \stackrel{\text{def}}{=} \theta I_n$ is a finite Blaschke-Potapov product and $BH^2(\mathbb{C}^n) \subset \Theta H^2(\mathbb{C}^n)$, then $\Theta H^2(\mathbb{C}^n)$ has finite codimension, and so $\Theta$ must be a finite Blaschke-Potapov product. □

**Corollary 3.5.3.** *Suppose $U \in \mathcal{U}_2$ has a 2-extremal function $\Psi$ such that* (3.5.1) *holds. If $U$ is a rational matrix function such that* $\operatorname{ind} T_U = 2$, *then $\Theta$ is a unitary constant on $\mathbb{T}$.*

*Proof.* Due to the results of [PY1], $U$ admits a (thematic) factorization of the form

$$U = \begin{pmatrix} \bar{w}_1 & -w_2 \\ \bar{w}_2 & w_1 \end{pmatrix} \begin{pmatrix} u_0 & \mathbb{O} \\ \mathbb{O} & u_1 \end{pmatrix} \begin{pmatrix} \bar{v}_1 & \bar{v}_2 \\ -v_2 & v_1 \end{pmatrix},$$

where $v_1, v_2, w_1$ and $w_2$ are scalar rational functions such that

$$|v_1|^2 + |v_2|^2 = |w_1|^2 + |w_2|^2 = 1 \text{ a.e. on } \mathbb{T},$$

$v_1$ and $v_2$ have no common zeros in the unit disk $\mathbb{D}$, $w_1$ and $w_2$ have no common zeros in $\mathbb{D}$, and $u_0$ and $u_1$ are scalar badly approximable rational unimodular functions on $\mathbb{T}$. These results may also be found in Sections 5 and 12 from Chapter 14 of [Pe1].

Suppose $\Psi = zh^2\Theta$ is an $n$-extremal function for $U$ such that (3.5.1) holds as in the conclusion of Theorem 3.5.1. Assume, for the sake of contradiction, that $\Theta$ is not a unitary constant.

Since $u_j$ is a scalar badly approximable rational unimodular function on $\mathbb{T}$, it admits a factorization of the form

$$u_j = c_j \bar{z}^{k_j} \frac{\bar{h}_j}{h_j},$$

where $c_j$ is a unimodular constant, the function $h_j$ is $H^\infty$-invertible, and $k_j = \operatorname{ind} T_{u_j}$, for $j = 0, 1$. In particular, we have

$$u\theta = c_0 c_1 \bar{z}^2 \theta \frac{\bar{h}_0}{h_0} \frac{\bar{h}_1}{h_1},$$

as $k_0 + k_1 = \operatorname{ind} T_U = 2$, where $\theta \stackrel{\text{def}}{=} \det \Theta$ and $u \stackrel{\text{def}}{=} \det U$.

On the other hand, by Theorem 3.5.1,

$$u\theta = \bar{z}^2 \frac{\bar{h}^2}{h^2}$$

and so the function $h^2 h_0^{-1} h_1^{-1}$ and its conjugate

$$\frac{\bar{h}^2}{\bar{h}_0 \bar{h}_1} = c_0 c_1 \theta \frac{h^2}{h_0 h_1}$$

belong to $H^1$. Therefore $h^2 h_0^{-1} h_1^{-1}$ equals a constant and so $\theta$ equals a constant as well. Thus, the conclusion follows from the fact that $\theta\Theta^*$ is an inner function. $\square$

We end this section with an example to illustrate some of our main results.

**Example 3.5.4.** Consider the matrix function

$$U = \begin{pmatrix} \bar{z} & \mathbb{O} \\ \mathbb{O} & \bar{z} \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & \bar{z} \\ -z & 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \bar{z} & \bar{z}^2 \\ -1 & \bar{z} \end{pmatrix}.$$

Clearly, $U$ belongs to $\mathcal{U}_2$ and it has superoptimal singular values $t_0(U) = t_1(U) = 1$.

We ask the question, *is there a 2-extremal function $\Psi$ for $U$ such that* (3.5.1) *holds with $n = 2$?* Let us assume for the moment that such a function $\Psi$ exists. In this case, Corollary 3.5.3 implies that $\Psi$ must be of the form $\Psi = zh^2\Theta$, where

$$\Theta = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is a unitary constant and $h$ is an outer function in $H^2$ such that $\|h\|_{L^2} = 1$. Since

$$\bar{z}^2 \frac{\bar{h}^2}{h^2} = \det(U\Theta) = \bar{z}^2(ad - bc),$$

it is easy to see that $h^2$ and its conjugate belong to $H^1$, and so $h^2$ is a constant of modulus 1. Relabeling the scalars $a, b, c$, and $d$, we may assume that $h^2$ equals 1 a.e. on $\mathbb{T}$. Thus,

$$2\bar{\zeta} = \text{trace}(U(\zeta)\Theta(\zeta)) = \frac{1}{\sqrt{2}}\left(a\bar{\zeta} + c\bar{\zeta}^2 - b + d\bar{\zeta}\right)$$

holds for a.e. $\zeta \in \mathbb{T}$, and so $b = c = 0$ and $a + d = 2\sqrt{2}$. However, $\Theta$ is unitary valued so it must be the case that $|a| = |d| = 1$, and so

$$2\sqrt{2} = a + d = |a + d| \leq |a| + |d| = 2,$$

which is a contradiction. Thus no such $\Psi$ exists. In particular, we must have that $\Phi$ does not have order 2 or $\sigma_2(\Phi) < t_0(\Phi) + t_1(\Phi) = 2$ by Theorem 3.4.13.

Actually, we have already shown that *the zero matrix function is not a best approximant to $U$ under the $L^\infty(\boldsymbol{S}_1^2)$ norm*, i.e. $\sigma_2(\Phi) < 2$. Indeed, we have

$$\text{dist}_{L^\infty(\boldsymbol{S}_1^2)}(U, H^\infty(\mathbb{M}_2)) < t_0(U) + t_1(U) = \|U\|_{L^\infty(\boldsymbol{S}_1^2)}.$$

by (3.1.7).

We now ask, *does $U$ have order 1, order 2, or is $U$ inaccessible?* It is clear that $U$ has a 1-extremal function by Remark 3.3.7. In fact, it is easy to check that the matrix function

$$\Psi_1 = \frac{z}{\sqrt{2}} \begin{pmatrix} 1 & \mathbb{O} \\ z & \mathbb{O} \end{pmatrix}$$

defines a 1-extremal function for $U$ and

$$\sigma_1(U) = \int_{\mathbb{T}} \text{trace}(U(\zeta)\Psi_1(\zeta))d\boldsymbol{m}(\zeta) = \|H_U\| = t_0(U) = 1.$$

However, $U$ *does not have order 1.* Indeed, one can see that the matrix function

$$\Psi_* = \frac{z}{\sqrt{3}} \begin{pmatrix} 1 & \mathbb{O} \\ z & 1 \end{pmatrix}$$

belongs to $H_0^1(\mathbb{M}_2)$, $\|\Psi_*\|_{L^1(\mathbb{M}_2)} \leq 1$, and

$$1 < \sqrt{\frac{3}{2}} = \int_{\mathbb{T}} \text{trace}(U(\zeta)\Psi(\zeta))d\boldsymbol{m}(\zeta) \leq \sigma_2(U).$$

Therefore, either $U$ has order 2 or $U$ is inaccessible. This matter requires further investigation.

# Chapter 4

# An index formula in connection with meromorphic approximation

Let $\varphi$ be an (essentially) bounded measurable function defined on the unit circle $\mathbb{T}$. Recall that, for $k \geq 0$, $H^{\infty}_{(k)}$ denotes the collection of meromorphic functions in $\mathbb{D}$ which are bounded near $\mathbb{T}$ and have at most $k$ poles in $\mathbb{D}$ (counting multiplicities). The *Nehari-Takagi problem* is to find a $q \in H^{\infty}_{(k)}$ which is closest to $\varphi$ with respect to the $L^{\infty}$-norm, i.e. to find $q \in H^{\infty}_{(k)}$ such that

$$\|\varphi - q\|_{\infty} = \text{dist}_{L^{\infty}}(\varphi, H^{\infty}_{(k)}) = \inf_{f \in H^{\infty}_{(k)}} \|\varphi - f\|_{\infty}.$$

Any such function $q$ is called a *best approximation in* $H^{\infty}_{(k)}$ *to* $\varphi$. Although uniqueness of a best approximation in $H^{\infty}_{(k)}$ need not hold in general, if $\varphi$ satisfies $\|H_{\varphi}\|_{\text{e}} < s_k(H_{\varphi})$, then uniqueness does hold. Here $H_{\varphi}$ denotes the Hankel operator with symbol $\varphi$, $s_k(H_{\varphi})$ is the $k$th singular value of $H_{\varphi}$, and $\|H_{\varphi}\|_{\text{e}}$ denotes the essential norm of $H_{\varphi}$ (precise definitions will be given below). Moreover, under these assumptions, it can be shown that the function defined by $u = s_k^{-1}(\varphi - q)$ has modulus 1 a.e. on

$\mathbb{T}$, the Toeplitz operator $T_u$ is Fredholm and

$$\operatorname{ind} T_u = \dim\ker T_u = 2k + \mu, \tag{4.0.1}$$

where $\mu$ denotes the multiplicity of the singular value $s_k \stackrel{\text{def}}{=} s_k(H_\varphi)$ of the Hankel operator $H_\varphi$ (e.g. see Chapter 4 in [Pe1]).

In view of these results, it seems natural to ask whether analogous results hold in the case of matrix-valued functions.

Suppose $\Phi$ is a $k$-admissible $n \times n$ matrix function with superoptimal approximation $Q$ in $H^\infty_{(k)}(\mathbb{M}_n)$ and $t^{(k)}_{n-1}(\Phi) > 0$. In this case, the matrix function $G = \Phi - Q$ is very badly approximable, the Toeplitz operator $T_G$ is Fredholm and

$$\operatorname{ind} T_G = \dim\ker T_G.$$

Therefore, we are led to ask: *Is it true that*

$$\dim\ker T_{\Phi-Q} = 2k + \mu? \tag{4.0.2}$$

Notice that the validity of (4.0.2) is well-known when $k = 0$. Indeed, the left-hand side equals the sum of all thematic indices that correspond to the superoptimal singular value $t_0(G) = \|H_G\|$ of the matrix function $G$ and the right-hand side equals the multiplicity of the singular value $s_0(H_G) = \|H_G\|$.

Actually, for arbitrary $k \geq 0$, it is easy to see that the dimension of $\ker T_{\Phi-Q}$ must be at most $2k + \mu$. After all, if this conclusion fails, then the singular value $s_0(H_{\Phi-Q}) = s_k(H_\Phi)$ of the Hankel operator $H_{\Phi-Q}$ must have multiplicity strictly greater $2k+\mu$ and so

$$s_k(H_\Phi) = s_0(H_{\Phi-Q}) = s_{2k+\mu}(H_{\Phi-Q}) \leq s_{k+\mu}(H_\Phi) + s_k(H_Q) = s_{k+\mu}(H_\Phi)$$
$$< s_{k+\mu-1}(H_\Phi) = s_k(H_\Phi)$$

holds, because $Q \in H^\infty_{(k)}(\mathbb{M}_n)$, a contradiction. Therefore,

$$\dim\ker T_{\Phi-Q} \leq 2k+\mu.$$

However, the following example shows that equality in (4.0.2) may fail in general.

**Example 4.0.5.** Consider the matrix function

$$\Phi = \frac{1}{\sqrt{2}}\begin{pmatrix} \bar{z}^5 + \frac{1}{3}\bar{z} & -\frac{1}{3}\bar{z}^2 \\ \bar{z}^4 & \frac{1}{3}\bar{z} \end{pmatrix}.$$

It is not difficult to verify that the nonzero singular values of the Hankel operator $H_\Phi$ are

$$s_0(H_\Phi) = \frac{\sqrt{10}}{3},\ s_1(H_\Phi) = s_2(H_\Phi) = s_3(H_\Phi) = 1, s_4(H_\Phi) = \frac{1}{\sqrt{2}},\ \text{and}\ \ s_5(H_\Phi) = \frac{1}{3}.$$

In particular, if $\mu$ denotes the multiplicity of the singular value $s_1(H_\Phi) = 1$ of the Hankel operator $H_\Phi$, we have $2k+\mu = 5$.

We now proceed to find the superoptimal approximation in $H^\infty_{(1)}(\mathbb{M}_2)$ to $\Phi$ by using an algorithm due to Peller and Young (see [PY3]) and following the notation used in Section 17 of Chapter 14 in [Pel].

Consider the vector functions $f, g \in H^2(\mathbb{C}^2)$ defined by

$$f = \begin{pmatrix} z \\ \mathbb{O} \end{pmatrix} \ \text{and}\ \ g = \frac{z^2}{\sqrt{2}}\begin{pmatrix} z \\ 1 \end{pmatrix}.$$

It is easy to check that $f$ is a Schmidt vector corresponding to the singular value $s_1(H_\Phi) = 1$ of $H_\Phi$ and $g = \bar{z}\overline{H_\Phi f}$. Moreover, $f$ and $g$ admit inner-outer factorizations $f = zv$ and $g = z^2 w$, where

$$v = \begin{pmatrix} 1 \\ \mathbb{O} \end{pmatrix} \quad \text{and} \quad w = \frac{1}{\sqrt{2}} \begin{pmatrix} z \\ 1 \end{pmatrix}.$$

Choose inner and co-outer matrix functions

$$\Theta = \begin{pmatrix} \mathbb{O} \\ 1 \end{pmatrix} \quad \text{and} \quad \Xi = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ z \end{pmatrix}$$

so that $V = (\, v\ \overline{\Theta}\,)$ and $W^t = (\, w\ \overline{\Xi}\,)$ are thematic. Let

$$Q_\# = \frac{1}{\sqrt{2}} \begin{pmatrix} \frac{1}{3}\bar{z} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} \end{pmatrix}.$$

It is easy to check that

$$W(\Phi - Q_\#)V = \begin{pmatrix} \bar{z}^4 & \mathbb{O} \\ \mathbb{O} & \frac{1}{3}\bar{z}^2 \end{pmatrix}. \tag{4.0.3}$$

Set $\Psi_\# = \frac{1}{3}\bar{z}^2$. Now, it can be verified that one can choose the following functions in the algorithm:

$$B = \begin{pmatrix} z & \mathbb{O} \\ \mathbb{O} & 1 \end{pmatrix}, \; K = \frac{1}{\sqrt{2}} \begin{pmatrix} \frac{1}{3} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} \end{pmatrix}, \; D_\# = \mathbb{O}, \; \Pi = z, \; \Sigma = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix},$$

$$x = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \; \Lambda = 1, \; Y = \mathbb{O}, \quad \text{and} \quad L_\# = \mathbb{O}.$$

In particular, any matrix function $Q \in \Omega_0^{(1)}(\Phi)$ must satisfy

$$W(\Phi - Q)V = \begin{pmatrix} \bar{z}^4 & \mathbb{O} \\ \mathbb{O} & \Psi \end{pmatrix},$$

where $\Psi = \Psi_{\#} + \Pi^*(L_{\#} - L)\Lambda^* = \bar{z}\left(\frac{1}{3}\bar{z} - L\right)$ has $L^\infty$-norm at most 1 and $L \in H^\infty$. Thus, the superoptimal approximation $Q$ to $\Phi$ in $H^\infty_{(1)}(\mathbb{M}_2)$ is determined by finding the best approximant to $\frac{1}{3}\bar{z}$ in $H^\infty$. Therefore the superoptimal approximation in $H^\infty_{(1)}(\mathbb{M}_2)$ to $\Phi$ is given by

$$Q = \frac{1}{\sqrt{2}}\begin{pmatrix} \bar{z} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} \end{pmatrix}.$$

By (4.0.3), we can see that

$$\dim \ker T_{\Phi - Q} = 4$$

even though $2k + \mu = 5$. Hence, the equality in (4.0.2) may fail in general.

# BIBLIOGRAPHY


[AAK1] V.M. Adamyan, D.Z. Arov, and M.G. Krein. On infinite Hankel matrices and generalized problems of Carathéodory-Fejér and F. Riesz. Funktsional. Anal. i Prilozhen. **2:1** (1968), 1–19; English transl.: Functional Anal. Appl. **2:1** (1968).

[AAK2] V.M. Adamyan, D.Z. Arov, and M.G. Krein. On infinite Hankel matrices and generalized problems of Carathéodory-Fejér and I. Schur. Funktsional. Anal. i Prilozhen. **2:2** (1968), 1–17; English transl.: Functional Anal. Appl. **2:2** (1968).

[AP1] R.B. Alexeev and V.V. Peller. Badly approximable matrix functions and canonical factorizations. Indiana Univ. Math. J. **49** (2000), 1247–1285.

[AP2] R.B. Alexeev and V.V. Peller. Invariance properties of thematic factorizations of matrix functions. J. Funct. Anal. **179** (2001), 309–332.

[BNP] L. Baratchart, F.L. Nazarov, and V.V. Peller. Analytic approximation of matrix functions in $L^p$. To appear in J. Approx. Theory

[F] P.A. Fuhrmann. *Linear Systems and Operators in Hilbert Space.* McGraw-Hill, New York, 1981. To appear in J. Approx. Theory

[Ha] E. Hayashi. The solution sets of extremal problems in $H^1$. Proc. Amer. Math. Soc. **93**, No. 4 (1985), 690–696.

[Kh] S. Khavinson. On some extremal problems of the theory of analytic functions. Uchen. Zap. Mosk. Univ. Matem. **144**, No. 4 (1951), 133–143. English Translation: Amer. Math. Soc. Transl. (2) **32** (1963), 139–154.

[Ne] Z. Nehari. On bounded bilinear forms. Ann. Math., **65** (1957), 153–162.

[Pe1] V.V. Peller. *Hankel Operators and Their Applications.* Springer Monographs in Mathematics. Springer, New York, 2003.

[Pe2] V.V. Peller. Analytic approximation of matrix functions and dual extremal functions. To appear in Proc. Amer. Math. Soc.

[Pe3] V.V. Peller. Personal communication.

[PK] V.V. Peller and S.V. Khruschëv. Hankel operators, best approximation and stationary Gaussian processes. Uspekhi Mat. Nauk **37:1** (1982), 53–124; English transl.: Russian Math. Surveys **37** (1982), 53–124.

[Po] S.J. Poreda. A characterization of badly approximable functions. Trans. Amer. Math. Soc. **169** (1972), 249–256.

[PY1] V.V. Peller and N. J. Young. Superoptimal analytic approximation of matrix functions. J. Funct. Anal. **120** (1994), 300–343.

[PY2] V.V. Peller and N. J. Young. Superoptimal singular values and indices of matrix functions. Int. Eq. Op. Theory. **20** (1994), 350–363.

[PY3] V.V. Peller and N.J. Young. Superoptimal approximation by meromorphic functions. Math. Proc. Camb. Phil. Soc. **119** (1996), 497–511.

[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. Sel. math., New ser. **11** (2005), 127–154.

[Tr1] S.R. Treil. The Adamyan-Arov-Krein theorem: a vector version. Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) **141** (1985), 56–71 (Russian)

[Tr2] S.R. Treil. On superoptimal approximation by analytic and meromorphic matrix-valued functions. J. Funct. Anal. **131** (1995), 386–414.

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