Available online at www.sciencedirect.com
SCIENCE DIRECT
J. Math. Anal. Appl. 310 (2005) 97–115
Journal of MATHEMATICAL ANALYSIS AND APPLICATIONS
www.elsevier.com/locate/jmaa

A Grassmannian band method approach to the Nehari–Takagi problem

Orest Iftime [a,b], Marinus Kaashoek [c], Amol Sasane [d,*]

[a] *Delft Center for Systems and Control, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands*
[b] *Faculty of Economics, University of Groningen, Landleven 5, 9747 AD Groningen, The Netherlands*
[c] *Department of Mathematics, Vrije Universiteit, De Boelelaan 1081A, 1081 HV Amsterdam, The Netherlands*
[d] *Department of Mathematics, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom*

Received 22 December 2004

Available online 5 March 2005

Submitted by J.A. Ball

Abstract

The band method is a flexible method for solving a variety of interpolation and extension problems which has evolved into increasing levels of sophistication over the past two decades. This article enhances the Grassmannian version of the band method to handle the Nehari–Takagi problem rather than merely the Nehari problem.

Keywords: Nehari–Takagi problem; Systems theory; Grassmannian band method

1. Introduction

The Grassmannian approach to the Nehari (and Nehari–Takagi) problem goes back to the paper by Ball and Helton [4], and the first attempt at an axiomatization of the Grass-

* Corresponding author.
E-mail addresses: o.v.iftime@dcsc.tudelft.nl (O. Iftime), ma.kaashoek@few.vu.nl (M. Kaashoek), a.j.sasane@lse.ac.uk (A. Sasane).

0022-247X/$ – see front matter
doi:10.1016/j.jmaa.2005.01.048

mannian approach (in the Nehari problem case) was made in Ball [1]. The band method, initiated by Dym and Gohberg [7], provides a unifying framework for solving strictly contractive extension problems; see [10] for an overview and [13] for one of the latest variations. A synthesis of the Grassmannian approach and the band method was developed by Ball, Gohberg and Kaashoek (see [3]), and it is called the Grassmannian approach for extensions problems. The solution to the so-called strictly contractive extension problem in this abstract framework, when applied to a certain concrete case, yielded a complete characterization of solutions to the Nehari problem. The Nehari–Takagi problem is a more general problem, and the Nehari problem can be considered to be a special case of this problem (an excellent exposition of this can be found, for instance in Young [15]). However, this more general Nehari–Takagi problem does not fit a priori in the abstract framework of [3].

The aim of this paper is to suitably enlarge the abstract framework in [3] so as to include the Nehari–Takagi problem as a special case as well. In the more general setting (as compared to the one in [3]) presented in this paper, one has a family of problems (indexed by the nonnegative integers l) and in the special case when $l = 0$, one gets the strictly contractive extension problem of [3]. Thus, our setting can be thought of as a more refined version of [3]. Our abstract setting applies to transfer functions of multi-input multi-output systems, possibly of infinite dimensions, as well as to time varying finite-dimensional systems. A preliminary version of this paper, dealing with the scalar case, has been presented in [11].

In Section 2 we formulate our abstract setting. In Section 3 we give the statement of the strictly contractive extension problem of Nehari–Takagi type in this abstract setting, and present our main theorem which provides a characterization of all solutions to this problem, under certain natural factorization assumptions. The proof of our main theorem, given in Section 5, is based on some properties of linear fractional maps which we present in Section 4. The final section has a review character and illustrates our abstract result on two concrete extension problems, namely

1. The Nehari–Takagi problem for a class of time invariant infinite-dimensional systems.
2. The Nehari–Takagi problem for time varying finite-dimensional linear systems.

The solutions of these problems are known (see, for instance, Sasane [14] and Ball, Mikkola and Sasane [5] for the first problem, and Kaashoek and Kos [12]). We show how these two problems fit into our abstract framework and that one can deal with both of the above problems as special cases of our main result. Thus in some sense our approach captures the essence of the proofs given in [12,14]. We also mention that our main result applies equally well to discrete time systems, both in a time invariant and a time variant setting. A remaining open problem is to use the abstract setting in order to solve the Nehari–Takagi problem for time varying, infinite-dimensional linear systems, which has so far not been considered.

2. Basic objects and their properties

We begin with some preliminaries about matrices over a C^*-algebra. Throughout $\mathcal{R}$ will be a unital C^*-algebra with unit e. Given $p, m \in \mathbb{N}$ we let $\mathcal{R}^{p\times m}$ denote the space of all $p \times m$ matrices with entries from $\mathcal{R}$ and with the corresponding norm $\|\cdot\|_{\mathcal{R}^{p\times m}}$ defined via the Gelfand–Naimark construction, which we recall below (see, for instance, [10]).

Theorem 2.1 (Gelfand–Naimark). *Let $\mathcal{R}$ be a C^*-algebra. Then there is a Hilbert space $\mathcal{H}$ and an isometric *-isomorphism of $\mathcal{R}$ onto a closed *-subalgebra of $\mathcal{L}(\mathcal{H})$.*

In light of this theorem, we equip $\mathcal{R}^{p\times m}$ with the following norm. If $K \in \mathcal{R}^{p\times m}$, then

$$\|K\|_{\mathcal{R}^{p\times m}} = \sup\bigl\{\|Kx\| \mid x \in \mathcal{H}^m,\ \|x\| \leqslant 1\bigr\}. \tag{1}$$

If $K_1 \in \mathcal{R}^{p\times m}$ and $K_2 \in \mathcal{R}^{m\times l}$, then $\|K_1K_2\|_{\mathcal{R}^{p\times l}} \leqslant \|K_1\|_{\mathcal{R}^{p\times m}}\|K_2\|_{\mathcal{R}^{m\times l}}$, where K_1K_2 is defined by the usual matrix multiplication. If $K \in \mathcal{R}^{p\times m}$, then $K^* \in \mathcal{R}^{m\times p}$ is the matrix with (i,j)th entry $K^*(i,j) = (K(j,i))^*$. For $K_1 \in \mathcal{R}^{p\times m}$ and $K_2 \in \mathcal{R}^{m\times l}$, we have $(K_1K_2)^* = K_2^*K_1^*$. From the Gelfand–Naimark theorem, it is easy to see that $\|K^*\|_{\mathcal{R}^{m\times p}} = \|K\|_{\mathcal{R}^{p\times m}}$.

The case $m = p$ is of particular interest. Indeed, $\mathcal{R}^{m\times m}$ equipped with the norm $\|\cdot\|_{\mathcal{R}^{m\times m}}$ and the involution * is again a unital C^*-algebra with the unit E_m being the $m \times m$ matrix with e on the diagonal and zeros elsewhere. The set of invertible elements in $\mathcal{R}^{m\times m}$ is denoted by $G\mathcal{R}^{m\times m}$. An element $K \in \mathcal{R}^{m\times m}$ is said to be *positive definite*, denoted by $K >_{\mathcal{R}^{m\times m}} 0$ if $K = A^*A$ for some $A \in G\mathcal{R}^{m\times m}$. An element $K \in \mathcal{R}^{p\times m}$ is said to be *strictly contractive* if $E_m - K^*K >_{\mathcal{R}^{m\times m}} 0$, which is, in turn, equivalent with $\|K\|_{\mathcal{R}^{p\times m}} < 1$, or equivalent with $E_p - KK^* >_{\mathcal{R}^{p\times p}} 0$. Note that if $K \in \mathcal{R}^{m\times m}$ and $\|K\|_{\mathcal{R}^{m\times m}} < 1$, then $(E_m - K)^{-1} \in \mathcal{R}^{m\times m}$.

The following items 1–4 describe our basic objects, the general setup and introduce the necessary definitions.

1. The triple $(\mathcal{R}, \mathcal{N}_+, \mathcal{N})$. Just as in Ball, Gohberg and Kaashoek [3], the basic objects are a unital C^*-algebra $\mathcal{R}$ with unit e, a subalgebra $\mathcal{N}$ of $\mathcal{R}$, and a subalgebra $\mathcal{N}_+$ of $\mathcal{N}$. We also assume that the unit e of $\mathcal{R}$ belongs to $\mathcal{N}_+$. Given these basic objects we introduce an additional stratification on $\mathcal{N}^{p\times m}$, the set of $p \times m$ matrices with entries in the algebra $\mathcal{N}$.

2. Stratification function μ. For each $p, m \in \mathbb{N}$, we assume the existence of a *stratification function* $\mu : \mathcal{N}^{p\times m} \to \mathbb{N} \cup \{0\}$ that satisfies the properties M1–M3 given below.

M1. $K \in \mathcal{N}^{p\times m}$ is such that $\mu(K) = 0$ if and only if $K \in \mathcal{N}_+^{p\times m}$.
M2. $\mu(K_1 + K_2) \leqslant \mu(K_1) + \mu(K_2)$ for all $K_1, K_2 \in \mathcal{N}^{p\times m}$.
M3. $\mu(K_1K_2) \leqslant \mu(K_1) + \mu(K_2)$ for all $K_1 \in \mathcal{N}^{p\times n}$ and $K_2 \in \mathcal{N}^{n\times m}$.

This produces a stratification on $\mathcal{N}^{p\times m}$ as follows. Define

$$\mathcal{N}_{[l]}^{p\times m} = \bigl\{K \in \mathcal{N}^{p\times m} \mid \mu(K) = l\bigr\}.$$

Then $\mathcal{N}_{[0]}^{p\times m} = \mathcal{N}_+^{p\times m}$, and defining $\mathcal{N}_l^{p\times m} = \bigcup_{k\leqslant l} \mathcal{N}_{[l]}^{p\times m}$, we get the increasing chain of subsets

$$\mathcal{N}_+^{p\times m} = \mathcal{N}_{[0]}^{p\times m} = \mathcal{N}_0^{p\times m} \subset \mathcal{N}_1^{p\times m} \subset \mathcal{N}_2^{p\times m} \subset \cdots$$

of $\mathcal{N}^{p\times m} = \bigcup_{l\geqslant 0} \mathcal{N}_l^{p\times m}$.

3. Index function ν. In order to make "pole-zero counting arguments" in this abstract setting, it turns out that the above stratification function is not quite enough. So we introduce an auxiliary index function and demand that it is homotopically invariant (item 4 below).

Let $m \in \mathbb{N}$. Recall that $\mathcal{N}_+^{m\times m} \cap G\mathcal{R}^{m\times m}$ consists of all elements in $\mathcal{N}_+^{m\times m}$ that are invertible in $\mathcal{R}^{m\times m}$. We assume that

$$K \in \mathcal{N}_+^{m\times m} \cap G\mathcal{R}^{m\times m} \quad \text{implies} \quad K^{-1} \in \mathcal{N}^{m\times m}. \tag{2}$$

This allows to define an index function $\nu : \mathcal{N}_+^{m\times m} \cap G\mathcal{R}^{m\times m} \to \{0, 1, 2, \ldots\}$ as follows:

$$\nu(K) = \mu\big(K^{-1}\big) = \inf\big\{l \mid K^{-1} \in \mathcal{N}_l^{m\times m}\big\}.$$

We also assume that

$$\nu(K_1K_2) = \nu(K_1) + \nu(K_2), \quad K_1, K_2 \in \mathcal{N}_+^{m\times m} \cap G\mathcal{R}^{m\times m}. \tag{3}$$

4. Homotopic invariance of the index. We assume that the following property is satisfied. If $M, N \in \mathcal{N}_+^{m\times m}$, and $tN + M \in G\mathcal{R}^{m\times m}$ for $0 \leqslant t \leqslant 1$, then $\nu(M) = \nu(N + M)$.

We conclude this section with the definition of a coprime factorization. Let $M \in \mathcal{N}_+^{m\times m}$ and $N \in \mathcal{N}_+^{p\times m}$. The pair (M, N) is said to be *right coprime over* $\mathcal{N}_+$ if there exist matrices X, Y with entries in $\mathcal{N}_+$ such that

$$XM - YN = E_m.$$

The element $K \in \mathcal{R}^{p\times m}$ is said to admit a *right coprime factorization* over $\mathcal{N}_+$ if there exist $M \in \mathcal{N}_+^{m\times m}$, $N \in \mathcal{N}_+^{p\times m}$ such that $M \in G\mathcal{R}^{m\times m}$, the pair (M, N) is right coprime over $\mathcal{N}_+$, and $K = NM^{-1}$.

Lemma 2.2. *If $K \in \mathcal{N}^{p\times m}$ admits a right coprime factorization $K = NM^{-1}$ over $\mathcal{N}_+$, then $\nu(M) = \mu(K)$.*

Proof. Since $M \in G\mathcal{R}^{m\times m} \cap \mathcal{N}_+^{m\times m}$, it follows that $M^{-1} \in \mathcal{N}^{m\times m}$, and so $\nu(M) = \mu(M^{-1})$ is defined. Using $K = NM^{-1}$, it follows that

$$\mu(K) = \mu\big(NM^{-1}\big) \leqslant \mu(N) + \mu\big(M^{-1}\big) = 0 + \mu\big(M^{-1}\big) = \nu(M). \tag{4}$$

Here we used that $\mu(N) = 0$ because $N \in \mathcal{N}_+^{p\times m}$. Now let X, Y be matrices with entries in $\mathcal{N}_+$ such that $XM - YN = E_m$. Then $X - YK = M^{-1}$ and so

$$\begin{aligned}\nu(M) = \mu\big(M^{-1}\big) = \mu(X - YK) &\leqslant \mu(X) + \mu(YK) = 0 + \mu(Y) + \mu(K)\\ &= 0 + \mu(K) = \mu(K).\end{aligned} \tag{5}$$

From (4) and (5), it follows that $\mu(K) = \nu(M)$. □

Remark. The assumption (2) is a limitation on the use of the Grassmannian band method. Indeed, it rules out an application to the triple $\mathcal{N}_+ = H^\infty$ (the Hardy space of bounded and analytic functions in the open right-half plane), $\mathcal{N} = \bigcup_{l\geqslant 0} H_l^\infty$ (the space of functions that can be decomposed into the sum of a function in H^∞ and a rational function with all poles in the open right-half plane), and $\mathcal{R} = L^\infty$ (over the imaginary line). As a consequence the abstract scheme in our paper does not cover the classical function theory setting while Ball, Mikkola and Sasane [5] handles this case. This limitation is known and already appears in earlier papers on the band method. For instance, the solution of the abstract contractive band method extension problems works well in a Wiener algebra setting (see [10, Chapter XXXV]) but it does not yield the solution of the classical Nehari problem in a L^∞-setting. On the other hand, as is shown in [8], the abstract positive real band method extension problem applies to the H^∞/L^∞-setting. It is an interesting open problem to extend the general Grassmannian band method to include the Nehari and the Nehari–Takagi problems for the classical function theory set up.

3. Problem statement and the main theorem

Fix $K \in \mathcal{R}^{p\times m}$. An element $F \in \mathcal{R}^{p\times m}$ is said to be a *strictly contractive extension of order l for K* if the following hold:

1. $F - K \in \mathcal{N}^{p\times m}$ and admits a right coprime factorization over $\mathcal{N}_+$, and $\mu(F-K) = l$.
2. $E_m - F^*F >_{\mathcal{R}^{m\times m}} 0$.

Given $K \in \mathcal{R}^{p\times m}$, we wish to derive a linear fractional representation of all strictly contractive extensions of order l for K.

In order to do this, we will assume that there exists a $(p+m)\times(p+m)$ matrix

$$\Theta = \begin{bmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{21} & \Theta_{22} \end{bmatrix} \in \mathcal{R}^{(p+m)\times(p+m)}, \tag{6}$$

with entries in $\mathcal{R}$ that satisfies the following additional conditions:

$$\text{S1:} \quad \Theta \begin{bmatrix} \mathcal{N}_+^{p\times 1} \\ \mathcal{N}_+^{m\times 1} \end{bmatrix} = \begin{bmatrix} E_p & K \\ 0 & E_m \end{bmatrix} \begin{bmatrix} \mathcal{N}_+^{p\times 1} \\ \mathcal{N}_+^{m\times 1} \end{bmatrix},$$

$$\text{S2:} \quad \Theta_{22} \in G\mathcal{R}^{m\times m} \quad \text{and} \quad \nu(\Theta_{22}) = l,$$

$$\text{S3:} \quad \Theta^* \begin{bmatrix} E_p & 0 \\ 0 & -E_m \end{bmatrix} \Theta = \begin{bmatrix} E_p & 0 \\ 0 & -E_m \end{bmatrix}.$$

We remark that in Lemma 4.3 we prove that S1 implies that $\Theta_{22} \in \mathcal{N}_+^{m\times m}$. Thus, by the first part of S2, we have $\Theta_{22} \in \mathcal{N}_+^{m\times m} \cap G\mathcal{R}^{m\times m}$, and hence $\nu(\Theta_{22})$ is well defined. Property S2 requires $\nu(\Theta_{22}) = \mu(\Theta_{22}^{-1}) = l$.

We shall prove the following result, which characterizes all strictly contractive extensions of order l for a given $K \in \mathcal{R}^{p\times m}$.

Theorem 3.1. *Let $K \in \mathcal{R}^{p\times m}$ and let Θ satisfies conditions* S1–S3. *Then given any $Q \in \mathcal{N}_+^{p\times m}$ such that Q is strictly contractive, F defined by*

$$F = (\Theta_{11}Q + \Theta_{12})(\Theta_{21}Q + \Theta_{22})^{-1} \quad (=: \mathcal{F}_\Theta(Q))$$

is a strictly contractive extension of order l for K.

Conversely, if F is a strictly contractive extension of order l for K, there exists a $Q \in \mathcal{N}_+^{p\times m}$ such that Q is strictly contractive and $F = \mathcal{F}_\Theta(Q)$.

To give some further insight into the conditions S1–S3, let us consider the following $(p+m)\times(p+m)$ matrices with entries in $\mathcal{R}$:

$$V=\begin{bmatrix} V_{11} & V_{12}\\ V_{21} & V_{22}\end{bmatrix}:=\begin{bmatrix} E_p & -K\\ 0 & E_m\end{bmatrix}\Theta,\qquad J=\begin{bmatrix} E_p & 0\\ 0 & -E_m\end{bmatrix}. \tag{7}$$

Assume that conditions S1 and S3 are satisfied. As we will see later (Lemma 4.2) this implies that Θ is invertible in $\mathcal{R}^{(p+m)\times(p+m)}$. Thus V is invertible in $\mathcal{R}^{(p+m)\times(p+m)}$, and we can rewrite S3 as

$$\begin{bmatrix} E_p & K\\ K^* & E_m-K^*K\end{bmatrix}=\left(V^{-1}\right)^*JV^{-1}. \tag{8}$$

Next, notice that condition S1 also implies that the entries of V and V^{-1} are in $\mathcal{N}_+$ (see Corollary 4.5). Hence condition S2 requires the spectral factor V to be of a special type, namely the entry V_{22} $(=\Theta_{22})$ is invertible, and $\nu(\Theta_{22})=l$.

Conversely, in order to find a matrix $\Theta\in\mathcal{R}^{(p+m)\times(p+m)}$ satisfying conditions S1–S3, one proceeds as follows. First as in (8), one looks for a J-spectral factorization relative to $\mathcal{N}_+$ of the matrix

$$\begin{bmatrix} E_p & K\\ K^* & E_m-K^*K\end{bmatrix}.$$

Next, one checks whether the factor V can be chosen in such a way that it satisfies the additional property referred to above. If so, then one defines Θ via the first identity in (7). This matrix Θ satisfies conditions S1–S3, and one can apply Theorem 3.1 to get all strictly contractive extensions of order l for K.

4. Preliminaries about linear fractional maps

In this section we collect together a number of results on linear fractional maps of the type appearing in Theorem 3.1. First we show that under appropriate assumptions, the linear fractional map $\mathcal{F}_\Theta$ maps the open unit ball in a one-to-one way onto itself.

Proposition 4.1. *Let*

$$\Theta=\begin{bmatrix} \Theta_{11} & \Theta_{12}\\ \Theta_{21} & \Theta_{22}\end{bmatrix}\in\mathcal{R}^{(p+m)\times(p+m)}.$$

Assume condition S3 *is satisfied, and that Θ and Θ_{22} are invertible in $\mathcal{R}^{(p+m)\times(p+m)}$ and $\mathcal{R}^{m\times m}$, respectively. Then the map*

$$\mathcal{F}_\Theta(Q)=(\Theta_{11}Q+\Theta_{12})(\Theta_{21}Q+\Theta_{22})^{-1}$$

is well defined on $B=\{Q\in\mathcal{R}^{p\times m}\mid E_m-Q^*Q>_{\mathcal{R}^{m\times m}}0\}$, and $\mathcal{F}_\Theta|_{\mathbb{B}}$ is a bijection on $\mathbb{B}$. Furthermore,

$$(\mathcal{F}_\Theta|_{\mathbb{B}})^{-1}=\mathcal{F}_{\Theta^{-1}}|_{\mathbb{B}}. \tag{9}$$

Proof. We divide the proof into four steps.

Step 1. In this step, we show that $\mathcal{F}_\Theta$ is well defined on $\mathbb{B}$. Let

$$J = \begin{bmatrix} E_p & 0 \\ 0 & -E_m \end{bmatrix}.$$

Then assumption S3 gives $\Theta^* J \Theta = J$. Consequently,

$$\Theta J \Theta^* = (\Theta J \Theta^*)(JJ) = (\Theta J \Theta^*)\big(J\big(\Theta\Theta^{-1}\big)J\big) = \Theta J(\Theta^* J\Theta)\Theta^{-1}J$$
$$= \Theta J J \Theta^{-1} J = J.$$

In particular, $\Theta_{21}\Theta_{21}^* - \Theta_{22}\Theta_{22}^* = -E_m$. Since Θ_{22} is invertible, it follows that

$$\Theta_{22}^{-1}\Theta_{21}\Theta_{21}^*\big(\Theta_{22}^*\big)^{-1} - E_m = -\Theta_{22}^{-1}\big(\Theta_{22}^*\big)^{-1}.$$

Hence $\|\Theta_{21}^*(\Theta_{22}^*)^{-1}\|_{\mathcal{R}^{p\times m}} < 1$ and so

$$\big\|\Theta_{22}^{-1}\Theta_{21}\big\|_{\mathcal{R}^{m\times p}} = \big\|\big(\Theta_{22}^{-1}\Theta_{21}\big)^*\big\|_{\mathcal{R}^{p\times m}} = \big\|\Theta_{21}^*\big(\Theta_{22}^*\big)^{-1}\big\|_{\mathcal{R}^{p\times m}} < 1.$$

Using $\|Q\|_{\mathcal{R}^{p\times m}} < 1$, it follows that

$$\big\|\Theta_{22}^{-1}\Theta_{21}Q\big\|_{\mathcal{R}^{m\times m}} \leqslant \big\|\Theta_{22}^{-1}\Theta_{21}\big\|_{\mathcal{R}^{m\times p}}\|Q\|_{\mathcal{R}^{p\times m}} < 1.$$

Hence $\Theta_{21}Q + \Theta_{22} = \Theta_{22}[\Theta_{22}^{-1}\Theta_{21}Q + E_m]$ is invertible, and $F = \mathcal{F}_\Theta(Q)$ is a well-defined element of $\mathcal{R}^{p\times m}$.

Step 2. Next we show that $F = \mathcal{F}_\Theta(Q)$ is strictly contractive if $Q \in \mathbb{B}$. First of all, note that

$$\begin{bmatrix} F \\ E_m \end{bmatrix} = \begin{bmatrix} (\Theta_{11}Q + \Theta_{12})(\Theta_{21}Q + \Theta_{22})^{-1} \\ (\Theta_{21}Q + \Theta_{22})(\Theta_{21}Q + \Theta_{22})^{-1} \end{bmatrix} = \Theta \begin{bmatrix} Q \\ E_m \end{bmatrix} X^{-1}, \tag{10}$$

where $X = \Theta_{21}Q + \Theta_{22}$. Hence, using condition S3, we have that

$$E_m - F^*F = \begin{bmatrix} F^* & E_m \end{bmatrix} \begin{bmatrix} -E_p & 0 \\ 0 & E_m \end{bmatrix} \begin{bmatrix} F \\ E_m \end{bmatrix}$$
$$= \big(X^{-1}\big)^*\begin{bmatrix} Q^* & E_m \end{bmatrix}\Theta^* \begin{bmatrix} -E_p & 0 \\ 0 & E_m \end{bmatrix} \Theta \begin{bmatrix} Q \\ E_m \end{bmatrix} X^{-1}$$
$$= \big(X^{-1}\big)^*\begin{bmatrix} Q^* & E_m \end{bmatrix} \begin{bmatrix} -E_p & 0 \\ 0 & E_m \end{bmatrix} \begin{bmatrix} Q \\ E_m \end{bmatrix} X^{-1}$$
$$= \big(X^{-1}\big)^*\big[E_m - Q^*Q\big]X^{-1} >_{\mathcal{R}^{m\times m}} 0.$$

Thus $F \in \mathbb{B}$.

Step 3. In this step we show that S3 holds with Θ^{-1} instead of Θ, and that the submatrix of Θ^{-1} through rows $p+1$ to $p+m$, and columns $p+1$ to $p+m$ (that is, the $(2,2)$ block entry of Θ^{-1} if partitioned in conformity with Θ) is invertible. To see this, let J denote the matrix as in the first step. Condition S3 and the invertibility of Θ imply that

$$\Theta^{-1} = J\Theta^* J = \begin{bmatrix} \Theta_{11}^* & -\Theta_{21}^* \\ -\Theta_{12}^* & \Theta_{22}^* \end{bmatrix}. \tag{11}$$

Hence the $(2,2)$ block entry of Θ^{-1} is Θ_{22}^*, which is invertible, since Θ_{22} is invertible. Furthermore,

$$\left(\Theta^{-1}\right)^* J\Theta^{-1} = (\Theta^*)^{-1} J(J\Theta^* J) = J.$$

Hence S3 holds with Θ^{-1} instead of Θ.

From what has been proved in the previous paragraph, we conclude that the results of the first two steps also hold with Θ^{-1} instead of Θ. Thus $\mathcal{F}_{\Theta^{-1}}$ is well defined on $\mathbb{B}$, and maps $\mathbb{B}$ into itself. Therefore, in order to complete the proof, it remains to show that

$$\mathcal{F}_{\Theta^{-1}}\left(\mathcal{F}_\Theta(Q)\right) = Q \quad \text{for all } Q \in \mathbb{B}, \quad \text{and} \tag{12}$$

$$\mathcal{F}_\Theta\left(\mathcal{F}_{\Theta^{-1}}(F)\right) = F \quad \text{for all } F \in \mathbb{B}. \tag{13}$$

Actually, since Θ^{-1} has the same properties as Θ, it suffices to prove (12). We will do this in the next step.

Step 4. Take $Q \in \mathbb{B}$, and define $F = \mathcal{F}_\Theta(Q)$, and $G = \mathcal{F}_{\Theta^{-1}}(F)$. By using (10) for Θ as well as Θ^{-1}, we have

$$\begin{bmatrix} F \\ E_m \end{bmatrix} = \Theta \begin{bmatrix} Q \\ E_m \end{bmatrix} (\Theta_{21} Q + \Theta_{22})^{-1},$$

$$\begin{bmatrix} G \\ E_m \end{bmatrix} = \Theta^{-1} \begin{bmatrix} F \\ E_m \end{bmatrix} \left(-\Theta_{12}^* F + \Theta_{22}^*\right)^{-1}.$$

Now observe that

$$\begin{aligned} -\Theta_{12}^* F + \Theta_{22}^* &= [\, 0_{m\times p} \ \ E_m \,] \begin{bmatrix} \Theta_{11}^* F - \Theta_{21}^* \\ -\Theta_{12}^* F + \Theta_{22}^* \end{bmatrix} = [\, 0_{m\times p} \ \ E_m \,]\Theta^{-1} \begin{bmatrix} F \\ E_m \end{bmatrix} \\ &= [\, 0_{m\times p} \ \ E_m \,]\Theta^{-1}\Theta \begin{bmatrix} Q \\ E_m \end{bmatrix} (\Theta_{21} Q + \Theta_{22})^{-1} \\ &= [\, 0_{m\times p} \ \ E_m \,] \begin{bmatrix} Q \\ E_m \end{bmatrix} (\Theta_{21} Q + \Theta_{22})^{-1} = (\Theta_{21} Q + \Theta_{22})^{-1}. \end{aligned}$$

In particular, $(\Theta_{21} Q + \Theta_{22})^{-1}(-\Theta_{12}^* F + \Theta_{22}^*)^{-1} = E_m$. But then

$$\begin{aligned} G &= [\, E_p \ \ 0_{p\times m} \,] \begin{bmatrix} G \\ E_m \end{bmatrix} = [\, E_p \ \ 0_{p\times m} \,]\Theta^{-1} \begin{bmatrix} F \\ E_m \end{bmatrix} \left(-\Theta_{12}^* F + \Theta_{22}^*\right)^{-1} \\ &= [\, E_p \ \ 0_{p\times m} \,]\Theta^{-1}\Theta \begin{bmatrix} Q \\ E_m \end{bmatrix} (\Theta_{21} Q + \Theta_{22})^{-1}\left(-\Theta_{12}^* F + \Theta_{22}^*\right)^{-1} \\ &= [\, E_p \ \ 0_{p\times m} \,] \begin{bmatrix} Q \\ E_m \end{bmatrix} = Q. \end{aligned}$$

Thus (12) is proved, and the proof is complete. □

As consequences of the assumptions S1 and S3 we state some simple lemmas, which will be used in the next section to prove the main theorem.

Lemma 4.2. *Assume that conditions* S1 *and* S3 *are satisfied. Then the matrix Θ is invertible in $\mathcal{R}^{(p+m)\times(p+m)}$, and Θ^{-1} is given by*

$$\Theta^{-1} = J\Theta^* J = \begin{bmatrix} \Theta_{11}^* & -\Theta_{21}^* \\ -\Theta_{12}^* & \Theta_{22}^* \end{bmatrix}.$$

Proof. From assumption S1, it follows that

$$\Theta \begin{bmatrix} \mathcal{N}_+^{p\times k} \\ \mathcal{N}_+^{m\times k} \end{bmatrix} = \begin{bmatrix} E_p & K \\ 0 & E_m \end{bmatrix} \begin{bmatrix} \mathcal{N}_+^{p\times k} \\ \mathcal{N}_+^{m\times k} \end{bmatrix}, \quad k = 1, 2, \ldots.$$

Since e belongs to $\mathcal{N}_+$, we have E_{p+m} belongs to $\mathcal{N}_+^{(p+m)\times(p+m)}$, and hence the above identity (with $k = p + m$) shows that there exists a matrix $X \in \mathcal{N}_+^{(p+m)\times(p+m)}$ such that

$$\Theta X = \begin{bmatrix} E_p & K \\ 0 & E_m \end{bmatrix}.$$

The term in the right-hand side of the previous identity is invertible. Hence in $\mathcal{R}^{(p+m)\times(p+m)}$ the matrix Θ has a right inverse, Θ' say, that is, $\Theta\Theta' = E_{p+m}$. Next define

$$\Theta'' = \begin{bmatrix} -E_p & 0 \\ 0 & E_m \end{bmatrix} \Theta^* \begin{bmatrix} -E_p & 0 \\ 0 & E_m \end{bmatrix} = \begin{bmatrix} \Theta_{11}^* & -\Theta_{21}^* \\ -\Theta_{12}^* & \Theta_{22}^* \end{bmatrix}.$$

Then assumption S3 yields that $\Theta''\Theta = E_{p+m}$. Thus

$$\Theta'' = \Theta'' E_{p+m} = \Theta''(\Theta\Theta') = (\Theta''\Theta)\Theta' = E_{p+m}\Theta' = \Theta'.$$

So Θ is invertible in $\mathcal{R}^{(p+m)\times(p+m)}$ with $\Theta^{-1} = \Theta''$. □

Lemma 4.3. *Assume that condition* S1 *is satisfied. Then $\Theta_{21} \in \mathcal{N}_+^{m\times p}$ and $\Theta_{22} \in \mathcal{N}_+^{m\times m}$.*

Proof. From assumption S1, it follows that

$$\begin{aligned} \begin{bmatrix} 0_{m\times p} & E_m \end{bmatrix} \begin{bmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{21} & \Theta_{22} \end{bmatrix} \begin{bmatrix} \mathcal{N}_+^{p\times 1} \\ \mathcal{N}_+^{m\times 1} \end{bmatrix} &= \begin{bmatrix} 0_{m\times p} & E_m \end{bmatrix} \begin{bmatrix} E_p & K \\ 0 & E_m \end{bmatrix} \begin{bmatrix} \mathcal{N}_+^{p\times 1} \\ \mathcal{N}_+^{m\times 1} \end{bmatrix} \\ &= \mathcal{N}_+^{m\times 1}. \end{aligned}$$

In particular,

$$\begin{aligned} \Theta_{21}\mathcal{N}_+^{p\times 1} &= \begin{bmatrix} \Theta_{21} & \Theta_{22} \end{bmatrix} \begin{bmatrix} \mathcal{N}_+^{p\times 1} \\ 0_{m\times 1} \end{bmatrix} \subset \begin{bmatrix} 0_{m\times p} & E_m \end{bmatrix} \begin{bmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{21} & \Theta_{22} \end{bmatrix} \begin{bmatrix} \mathcal{N}_+^{p\times 1} \\ \mathcal{N}_+^{m\times 1} \end{bmatrix} \\ &= \mathcal{N}_+^{m\times 1}. \end{aligned}$$

Since $e, 0 \in \mathcal{N}_+$, we have that $\Theta_{21} \in \mathcal{N}_+^{m\times p}$. Similarly

$$\Theta_{22}\mathcal{N}_+^{m\times 1} = [\Theta_{21} \ \ \Theta_{22}] \begin{bmatrix} 0_{p\times 1} \\ \mathcal{N}_+^{m\times 1} \end{bmatrix} \subset [0_{m\times p} \ \ E_m] \begin{bmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{21} & \Theta_{22} \end{bmatrix} \begin{bmatrix} \mathcal{N}_+^{p\times 1} \\ \mathcal{N}_+^{m\times 1} \end{bmatrix} = \mathcal{N}_+^{m\times 1},$$

and so $\Theta_{22}\mathcal{N}_+^{m\times 1} \subset \mathcal{N}_+^{m\times 1}$. Since $e, 0 \in \mathcal{N}_+$, we obtain also that $\Theta_{22} \in \mathcal{N}_+^{m\times m}$. □

Lemma 4.4. *Assume that conditions* S1 *and* S3 *are satisfied. Then* $\Theta_{11}^* \in \mathcal{N}_+^{p\times p}$ *and* $\Theta_{12}^* \in \mathcal{N}_+^{m\times p}$.

Proof. Condition S3 may be rewritten as

$$\begin{bmatrix} -\Theta_{11}^* & \Theta_{21}^* \\ -\Theta_{12}^* & \Theta_{22}^* \end{bmatrix} \begin{bmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{21} & \Theta_{22} \end{bmatrix} = \begin{bmatrix} -E_p & 0 \\ 0 & E_m \end{bmatrix},$$

and using condition S1 we have that

$$\begin{bmatrix} -E_p & 0 \\ 0 & E_m \end{bmatrix} \begin{bmatrix} \mathcal{N}_+^{p\times 1} \\ \mathcal{N}_+^{m\times 1} \end{bmatrix} = \begin{bmatrix} -\Theta_{11}^* & \Theta_{21}^* \\ -\Theta_{12}^* & \Theta_{22}^* \end{bmatrix} \begin{bmatrix} E_p & K \\ 0 & E_m \end{bmatrix} \begin{bmatrix} \mathcal{N}_+^{p\times 1} \\ \mathcal{N}_+^{m\times 1} \end{bmatrix}.$$

In particular,

$$-\Theta_{11}^* \mathcal{N}_+^{p\times 1} = \begin{bmatrix} -\Theta_{11}^* & \Theta_{21}^* \end{bmatrix} \begin{bmatrix} E_p & K \\ 0 & E_m \end{bmatrix} \begin{bmatrix} \mathcal{N}_+^{p\times 1} \\ 0_{m\times 1} \end{bmatrix} \subset \mathcal{N}_+^{p\times 1}$$

and

$$-\Theta_{12}^* \mathcal{N}_+^{p\times 1} = \begin{bmatrix} -\Theta_{12}^* & \Theta_{22}^* \end{bmatrix} \begin{bmatrix} E_p & K \\ 0 & E_m \end{bmatrix} \begin{bmatrix} \mathcal{N}_+^{p\times 1} \\ 0_{m\times 1} \end{bmatrix} \subset \mathcal{N}_+^{p\times 1}.$$

Since $e, 0 \in \mathcal{N}_+$, we obtain $\Theta_{11}^* \in \mathcal{N}_+^{p\times p}$ and $\Theta_{12}^* \in \mathcal{N}_+^{m\times p}$. □

Define $V \in \mathcal{R}^{(p+m)\times(p+m)}$ by

$$V = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} := \begin{bmatrix} E_p & -K \\ 0 & E_m \end{bmatrix} \Theta. \tag{14}$$

Corollary 4.5. *Let* $V \in \mathcal{R}^{(p+m)\times(p+m)}$ *be defined as in* (14). *If* Θ *satisfies* S1 *and* S3, *then* V *is invertible in* $\mathcal{R}^{(p+m)\times(p+m)}$, *and the entries of* V *and* V^{-1} *belong to* $\mathcal{N}_+$.

Proof. Since S1 and S3 are satisfied, we know that Θ is invertible in $\mathcal{R}^{(p+m)\times(p+m)}$ (Lemma 4.2), and hence V is invertible in $\mathcal{R}^{(p+m)\times(p+m)}$. But then S1 can be rewritten in the following equivalent form:

$$V \begin{bmatrix} \mathcal{N}_+^{p\times 1} \\ \mathcal{N}_+^{m\times 1} \end{bmatrix} = \begin{bmatrix} \mathcal{N}_+^{p\times 1} \\ \mathcal{N}_+^{m\times 1} \end{bmatrix},$$

which is equivalent to the statement that the entries of V and V^{-1} belong to $\mathcal{N}_+$. □

5. Proof of the main theorem

In this section we prove Theorem 3.1. Throughout $K \in \mathcal{R}^{p\times m}$, and Θ is the $(p+m)\times(p+m)$ matrix with entries in $\mathcal{R}$, partitioned as in (6). We assume that conditions S1–S3 are satisfied. Recall that $\mathbb{B} = \{Q \in \mathcal{R}^{p\times m} \mid E_m - Q^*Q >_{\mathcal{R}^{m\times m}} 0\}$.

Proof of Theorem 3.1. From Lemma 4.2, we know that Θ is invertible in $\mathcal{R}^{(p+m)\times(p+m)}$. Furthermore, according to condition S2, the element Θ_{22} is invertible in $\mathcal{R}^{m\times m}$. But then we can apply Proposition 4.1 to show that the map $\mathcal{F}_\Theta$ in Theorem 3.1 is well defined on $\mathbb{B}$, and maps $\mathbb{B}$ in a one-to-one way onto itself.

Step 1. Let $Q \in \mathbb{B} \cap \mathcal{N}_+^{p\times m}$. We already know that $F = \mathcal{F}_\Theta(Q)$ is strictly contractive. According to Lemma 4.3, the entries of Θ_{21} and Θ_{22} belong to $\mathcal{N}_+$. We claim that

$$\nu(\Theta_{21}Q + \Theta_{22}) = l. \tag{15}$$

To see this, we apply the assumption of homotopic invariance of the index function (item 4 in Section 2) with $N = \Theta_{21}Q$ and $M = \Theta_{22}$. Indeed, for $t \in [0,1]$, we have $tQ \in \mathbb{B}$, and hence we know from Proposition 4.1 that $t\Theta_{21}Q + \Theta_{22} = \Theta_{21}(tQ) + \Theta_{22} \in G\mathcal{R}^{m\times m}$ for $t \in [0,1]$. According to condition S2 we have $\nu(\Theta_{22}) = l$, and thus we obtain (15) by applying homotopic invariance.

Let V be given by (14). Thus

$$\begin{bmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{21} & \Theta_{22} \end{bmatrix} = \begin{bmatrix} E_p & K \\ 0 & E_m \end{bmatrix} \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} = \begin{bmatrix} V_{11} + KV_{21} & V_{12} + KV_{22} \\ V_{21} & V_{22} \end{bmatrix}.$$

It follows that

$$\begin{aligned} F - K &= (\Theta_{11}Q + \Theta_{12})(\Theta_{21}Q + \Theta_{22})^{-1} - K \\ &= \big((V_{11} + KV_{21})Q + V_{12} + KV_{22}\big)(V_{21}Q + V_{22})^{-1} - K \\ &= \big(V_{11}Q + V_{12} + K(V_{21}Q + V_{22})\big)(V_{21}Q + V_{22})^{-1} - K \\ &= (V_{11}Q + V_{12})(V_{21}Q + V_{22})^{-1}. \end{aligned}$$

Now define $N = V_{11}Q + V_1$ and $M = V_{21}Q + V_{22} (= \Theta_{21}Q + \Theta_{22})$. Since $Q \in \mathcal{N}_+^{p\times m}$ and the entries of V are also in $\mathcal{N}_+$, both N and M are matrices with entries in $\mathcal{N}_+$. Since $M \in \mathcal{N}_+^{m\times m} \cap G\mathcal{R}^{m\times m}$, we have $M^{-1} \in \mathcal{N}^{m\times m}$, and thus it follows that $F - K \in \mathcal{N}^{p\times m}$.

Next we show that the pair (M, N) is coprime over $\mathcal{N}_+$. To do this, let $\Lambda = V^{-1}$. Then from Corollary 4.5 the entries $\Lambda(i,j)$, $i, j \in \{1, \dots, p+m\}$ of Λ are in $\mathcal{N}_+$. Partition Λ is in conformity with the partition of V:

$$\Lambda = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \Lambda_{22} \end{bmatrix}.$$

Then we have

$$\begin{aligned} \Lambda_{21}N + \Lambda_{22}M &= \Lambda_{21}(V_{11}Q + V_{12}) + \Lambda_{22}(V_{21}Q + V_{22}) \\ &= (\Lambda_{21}V_{11} + \Lambda_{22}V_{21})Q + (\Lambda_{21}V_{12} + \Lambda_{22}V_{22}) = E_m. \end{aligned}$$

Hence the pair (M, N) is coprime over $\mathcal{N}_+$, and thus $F - K = NM^{-1}$ admits a right coprime factorization over $\mathcal{N}_+$. So using Lemma 2.2, $\mu(F - K) = \mathrm{ind}(M) = l$. This completes the proof of the first part.

Step 2. We prove the converse in this part. Left F be strictly contractive extension of order l for K. Since F is strictly contractive, from Proposition 4.1 we know that there exists a unique $Q \in \mathcal{R}^{p\times m}$ such that $E_m - Q^*Q$ is positive definite in $\mathcal{R}^{m\times m}$ and $F = \mathcal{F}_\Theta(Q)$. In fact, using (9) and (11), this Q is given by

$$Q = \mathcal{F}_{\Theta^{-1}}(F) = \big(\Theta_{11}^* F - \Theta_{21}^*\big)\big(-\Theta_{12}^* F + \Theta_{22}^*\big)^{-1}.$$

Thus $Q = RS^{-1}$, where

$$\begin{bmatrix} R \\ S \end{bmatrix} = \Theta^{-1} \begin{bmatrix} F \\ E_m \end{bmatrix}. \tag{16}$$

It follows that

$$\begin{aligned}
\begin{bmatrix} R \\ S \end{bmatrix} &= \Theta^{-1} \begin{bmatrix} K \\ E_m \end{bmatrix} + \Theta^{-1} \begin{bmatrix} F-K \\ E_m \end{bmatrix} \\
&= \Theta^{-1} \begin{bmatrix} E_p & K \\ 0 & E_m \end{bmatrix} \begin{bmatrix} 0_{p\times m} \\ E_m \end{bmatrix} + \Theta^{-1} \begin{bmatrix} E_p \\ 0_{m\times p} \end{bmatrix} (F-K) \\
&= V^{-1} \begin{bmatrix} 0_{p\times m} \\ E_m \end{bmatrix} + \begin{bmatrix} \Theta_{11}^* \\ -\Theta_{12}^* \end{bmatrix} (F-K) \\
&= \begin{bmatrix} \Lambda_{12} \\ \Lambda_{22} \end{bmatrix} + \begin{bmatrix} \Theta_{11}^* \\ -\Theta_{12}^* \end{bmatrix} (F-K).
\end{aligned}$$

Here we used that V is defined by (14) and that Λ denotes the inverse of V which has been partitioned in conformity with the partitioning of V. From the above, we see that

$$R = \Lambda_{12} + \Theta_{11}^*(F-K) \quad \text{and} \quad S = \Lambda_{22} - \Theta_{12}^*(F-K). \tag{17}$$

Now let $F-K$ admit the right coprime factorization $F-K = NM^{-1}$ over $\mathcal{N}_+$. Then from Lemma 2.2, we have $\nu(M) = l$. Using $F-K = NM^{-1}$ in (17), we get $RM = \Lambda_{12}M + \Theta_{11}^*N$ and $SM = \Lambda_{22}M - \Theta_{12}^*N$. From Corollary 4.5 we know that Λ_{12} and Λ_{22} have entries in $\mathcal{N}_+$. Furthermore, according to Lemma 4.4, the entries of Θ_{11}^* and Θ_{12}^* belong to $\mathcal{N}_+$. Finally, since M and N also have entries in $\mathcal{N}_+$, we conclude that RM and SM have entries in $\mathcal{N}_+$ as well. Since S and M are invertible in $\mathcal{R}^{m\times m}$, it follows that SM is also invertible in $\mathcal{R}^{m\times m}$. So we know that $SM \in \mathcal{N}_+^{m\times m} \cap G\mathcal{R}^{m\times m}$. Now we show that $\nu(SM) = 0$.

Using (16), we have

$$\Theta_{21}R + \Theta_{22}S = [\,\Theta_{21}\ \ \Theta_{22}\,]\Theta^{-1} \begin{bmatrix} F \\ E_m \end{bmatrix} = [\,0_{m\times p}\ \ E_m\,] \begin{bmatrix} F \\ E_m \end{bmatrix} = E_m.$$

It follows that $M = \Theta_{21}RM + \Theta_{22}SM$. Define $\tilde{N} = -\Theta_{21}RM$. Then

$$t\tilde{N} + M = (1-t)\Theta_{21}RM + \Theta_{22}SM.$$

By Lemma 4.3, Θ_{21} has entries in $\mathcal{N}_+$, and from the previous paragraph, we know that the same holds for RM. Thus $\tilde{N}$ has entries in $\mathcal{N}_+$. If $t \in [0,1]$, then $(1-t)Q \in \mathbb{B}$, and thus $\Theta_{21}(1-t)Q + \Theta_{22}$ is invertible in $\mathcal{R}^{m\times m}$. As we have seen in the previous paragraph, SM is also invertible in $\mathcal{R}^{m\times m}$. Consequently

$$t\tilde{N} + M = (1-t)\Theta_{21}RM + \Theta_{22}SM = \big(\Theta_{21}(1-t)Q + \Theta_{22}\big)SM \in G\mathcal{R}^{m\times m}$$

where we used the fact that $Q = RM(SM)^{-1}$. So $t\tilde{N} + M \in \mathcal{N}_+^{m\times m} \cap G\mathcal{R}^{m\times m}$ for $t \in [0, 1]$, and applying the assumption of the homotopic invariance of the index (item 4 in Section 2) with $N = \tilde{N}$ and $M = M$, we obtain

$$l = \nu(M) = \nu(\Theta_{22}SM) = \nu(\Theta_{22}) + \nu(SM) = l + \nu(SM).$$

Hence $\nu(SM) = 0$. In other words, $\mu((SM)^{-1}) = 0$ which implies that $(SM)^{-1} \in \mathcal{N}_+^{m\times m}$, and so $Q = RM(SM)^{-1} \in \mathcal{N}_+^{p\times m}$. This completes the proof. □

6. Examples

This section has a review character. We illustrate our main theorem on two concrete examples, namely the Nehari–Takagi problem for a class of time invariant infinite-dimensional systems, and the Nehari–Takagi problem for a class of time varying, finite-dimensional linear systems. The first example is taken from the paper [14], and the second from [12].

6.1. Nehari–Takagi problem for time invariant infinite-dimensional linear systems

1. The triple $(\mathcal{R}, \mathcal{N}_+, \mathcal{N})$. We take $\mathcal{R}$ to be the (commutative) C^*-algebra of continuous functions k on the imaginary axis $i\mathbb{R}$ for which the limits $\lim_{\omega\to\pm\infty} k(i\omega)$ exist and are equal.

$\mathcal{N}_+$ is the subalgebra of functions $k : \overline{\mathbb{C}_+} \to \mathbb{C}$ continuous on the closed right-half plane $\overline{\mathbb{C}_+} := \mathbb{C}_+ \cup i\mathbb{R} \cup \{\infty\}$, analytic in $\mathbb{C}_+$, and bounded in $\overline{\mathbb{C}_+}$. (These are exactly the analytic functions $k(s)$ obtained via a Möbius transformation $z \mapsto \frac{s+1}{s-1}$ on the domain of functions in the disc algebra.)

$\mathcal{N}$ is the subalgebra of functions $k : \overline{\mathbb{C}_+} \to \mathbb{C}$ defined in the closed right-half plane $\overline{\mathbb{C}_+}$ such that $k = g + h$, where $h \in \mathcal{N}_+$ and g is a strictly proper rational function with all its poles contained in $\mathbb{C}_+$.

In the Gelfand–Naimark construction one can take $\mathcal{H}$ to be the Hilbert space $L^2(i\mathbb{R}, \mathbb{C})$, and the norm on $\mathcal{R}^{p\times m}$ is then the usual norm

$$\|K\|_{\mathcal{R}^{p\times m}} = \sup_{\omega\in\mathbb{R}} \big\|K(i\omega)\big\|_{\mathcal{L}(\mathbb{C}^m,\mathbb{C}^p)}.$$

2. Stratification function μ. From the proofs of Lemmas 2.5.1, 2.5.3 in [14], it follows that every $K \in \mathcal{N}^{p\times m}$ admits a coprime factorization $K = NM^{-1}$ over $\mathcal{N}_+$ where M is rational, with $\det(M)$ being a proper rational function with all its zeros contained in $\mathbb{C}_+$. We define $\mu(K)$ to be the number of zeros of $\det(M)$. Then M1–M3 can be verified easily.

The function μ then produces a stratification on $\mathcal{N}^{p\times m}$: the set $\mathcal{N}_{[l]}^{p\times m}$ consists of all functions $K : \overline{\mathbb{C}_+} \to \mathbb{C}^{p\times m}$ such that $K = G + H$, where $H \in \mathcal{N}_+^{p\times m}$ and G is the rational transfer function of a finite-dimensional system with MacMillan degree l and having all its poles in $\mathbb{C}_+$. The set $\mathcal{N}_l^{p\times m}$ comprises the set of all functions $K : \overline{\mathbb{C}_+} \to \mathbb{C}^{p\times m}$ such that $K = G + H$, where $H \in \mathcal{N}_+^{p\times m}$ and G is the rational transfer function of a finite-dimensional system with MacMillan degree *at most* equal to l and having all its poles in $\mathbb{C}_+$.

3. Index function ν. We first show that $\mathcal{N}^{m\times m}$ is inverse closed. Since every $K\in\mathcal{N}^{m\times m}$ admits a coprime factorization over $\mathcal{N}_+$ (see the proof of [14, Lemma 2.5.3]), it suffices to show that if $K\in G\mathcal{R}^{m\times m}\cap\mathcal{N}_+^{m\times m}$, then $K^{-1}\in\mathcal{N}^{m\times m}$. The latter follows from the proof of [14, Lemma 2.5.10].

The index of $K=NM^{-1}$ (where (N,M) is a coprime over $\mathcal{N}_+$) is then the difference between the number of zeros of N in $\mathbb{C}_+$ (the number of zeros of K in $\mathbb{C}_+$) and the number of zeros of M in $\mathbb{C}_+$ (the number of poles of K in $\mathbb{C}_+$).

In order to show additivity of the index, one can proceed as follows. Let $K_1,K_2\in\mathcal{N}^{m\times m}$, and let $K_1=N_1M_1^{-1}$, $K_2=N_2M_2^{-1}$ be coprime factorizations over $\mathcal{N}_+$. We note that

$$K_1K_2=\big(N_1\,\mathrm{adj}(M_1)N_2\,\mathrm{adj}(M_1)\big)\big(\det(M_1)\det(M_2)E_m\big)^{-1}$$

is a right coprime factorization over $\mathcal{N}_+$ for K_1K_2, and

$$K_2^{-1}K_1^{-1}=\big(M_2\,\mathrm{adj}(N_2)M_1\,\mathrm{adj}(N_1)\big)\big(\det(N_2)\det(N_1)E_m\big)^{-1}$$

is a right coprime factorization over $\mathcal{N}_+$ for $K_2^{-1}K_1^{-1}$. Consequently, from the proof of Lemma 2.2, it follows that

$$\mu(K_1K_2)=\mu\big(\det(M_1)\det(M_2)E_m\big)\quad\text{and}\quad\mu\big(K_2^{-1}K_1^{-1}\big)=\mu\big(\det(N_1)\det(N_2)E_m\big).$$

Thus

$$\begin{aligned}\nu(K_1K_2)&=\mu\big(\det(N_1)\big)-\mu\big(\det(M_1)\big)+\mu\big(\det(N_2)\big)-\mu\big(\det(M_2)\big)\\&=\nu(K_1)+\nu(K_2).\end{aligned}$$

4. Homotopic invariance of the index. We note that $\nu(tN+M)=\mu(\det(tN+M))$, and so the homotopic invariance of the index is a consequence of the homotopic invariance of the Nyquist index (see, for instance, Curtain and Zwart [6, Lemma A.1.18, p. 570]): indeed, let $\delta>0$ be such that $\mathbb{C}_{\delta,+}:=\{s\in\mathbb{C}\mid\mathrm{Re}(s)>\delta\}$ contains all the zeros of $\det(M)$ and $\det(N+M)$. Let $\epsilon\in(0,\delta)$. Applying [6, Lemma A.1.18] with $h(s,t):=\det(tN(s+\epsilon)+M(s+\epsilon))$, it follows that the number of zeros of $\det(M)$ and $\det(N+M)$ are the same in $\mathbb{C}_{\epsilon,+}$. But the choice of $\epsilon\in(0,\delta)$ was arbitrary, and so it follows that $\det(M)$ and $\det(N+M)$ have the same number of zeros in $\mathbb{C}_+$. But they do not have any zeros on the imaginary axis, and so the result follows.

Statement of the problem. Let K be the transfer function of an infinite-dimensional system with generating operators (A,B,C) such that A is exponentially stable, infinitesimal generator of a strongly continuous semigroup $(e^{tA})_{t\geqslant 0}$ on the Hilbert space X, $B\in\mathcal{L}(\mathbb{C}^m,X)$, and $C\in\mathcal{L}(X,\mathbb{C}^p)$. Then $K\in\mathcal{N}_+^{p\times m}$. Let l be a nonnegative integer, and suppose that $\sigma_{l+1}(H_K)<1<\sigma_l(H_K)$, where $\sigma_r(\cdot)$ denotes the rth singular value of a bounded linear operator, and $H_K:L^2((0,\infty),\mathbb{C}^m)\to L^2((0,\infty),\mathbb{C}^p)$ denotes the Hankel operator corresponding to the impulse response $h(t)=Ce^{tA}B$, $t\geqslant 0$, and given by

$$(H_Ku)(t)=\int_0^\infty Ce^{(t+\tau)A}Bu(\tau)\,d\tau,\quad t\geqslant 0.$$

Then find a *strictly contractive extension F of order l for K.*

Let

$$\Lambda(s) := \begin{bmatrix} I_p & 0 \\ 0 & I_m \end{bmatrix} + \begin{bmatrix} -CL_B \\ B^* \end{bmatrix} (I - L_C L_B)^{-1} (sI + A^*)^{-1} [\, C^* \;\; L_C B \,],$$

where L_B and L_C denote the controllability and observability Gramians, respectively, of the infinite-dimensional system given by the triple (A, B, C) (see, for instance, Curtain and Zwart [6, Theorem 4.1.23, p. 160]). Then, from [14, Chapter 4], it follows that Θ defined by

$$\Theta = \begin{bmatrix} I_p & K \\ 0 & I_m \end{bmatrix} \Lambda(-\,\cdot)^{-1} \tag{18}$$

satisfies S1–S3. Moreover, Θ is invertible. Hence applying Theorem 3.1, we obtain the following result.

Theorem 6.1. *Let A be the exponentially stable, infinitesimal generator of a strongly continuous semigroup $(e^{tA})_{t\geqslant 0}$ on the Hilbert space X, $B \in \mathcal{L}(\mathbb{C}^m, X)$, and $C \in \mathcal{L}(X, \mathbb{C}^p)$. Let $K(s) = C(sI - A)^{-1}B \in \mathcal{N}_+^{p\times m}$ and suppose that there exists an l such that $\sigma_{l+1}(H_K) < 1 < \sigma_l(H_K)$, where $\sigma_r(H_K)$ denote the Hankel singular values. Let Θ be given by* (18). *Then $F \in \mathcal{R}^{p\times m}$ is a strictly contractive extension of order l for K if and only if $F = (\Theta_{11}Q + \Theta_{12})(\Theta_{21}Q + \Theta_{22})^{-1}$, for some $Q \in \mathcal{N}_+^{p\times m}$ such that $\|Q\|_{\mathcal{R}^{p\times m}} < 1$.*

6.2. Nehari–Takagi problem for time varying, finite-dimensional linear systems

The problem considered in this section deals with integral operators on $L^2(\mathbb{R}, \mathbb{C}^m)$ which are input–output operators of time varying finite-dimensional systems of the form

$$\Sigma\colon\quad \begin{cases} \left(\frac{d}{dt}x\right)(t) = A(t)x(t) + B(t)u(t), \\ y(t) = C(t)x(t) + D(t)u(t), \end{cases} \qquad t \in \mathbb{R}, \tag{19}$$

here $A \in L^\infty(\mathbb{R}, \mathbb{C}^{n\times n})$, $B \in L^\infty(\mathbb{R}, \mathbb{C}^{n\times m})$, $C \in L^\infty(\mathbb{R}, \mathbb{C}^{m\times n})$, and $D \in L^\infty(\mathbb{R}, \mathbb{C}^{m\times m})$. The differential equation

$$\left(\frac{d}{dt}x\right)(t) = A(t)x(t), \qquad t \in \mathbb{R}, \tag{20}$$

is assumed to have a dichotomy P_Σ. The input–output operator T_Σ of the system Σ, acting on $L^2(\mathbb{R}, \mathbb{C}^m)$, is given by

$$(T_\Sigma u)(t) = D(t)\varphi(t) + \int_{-\infty}^{\infty} C(t)\gamma_A(t, s)B(s)u(s)\,ds, \qquad t \in \mathbb{R},\; u \in L^2\big(\mathbb{R}, \mathbb{C}^m\big),$$

where $C(t)\gamma_A(t, s)B(s)$ is the uniquely defined bounded weighting pattern which is obtained from the evolution matrix and the dichotomy P_Σ, see [12]. In what follows we shall always assume that $D(t) = dI_m$ for each $t \in \mathbb{R}$, where d is a complex number.

If $T = T_\Sigma$, where Σ is as in the preceding paragraph with $D(t) \equiv dI_m$ and with the homogeneous equation (20) possessing a dichotomy P_Σ, then we refer to Σ as an *admissible realization* of T, and we call T the *input–output operator* of the *admissible* system Σ.

Statement of the problem. The starting point for the Nehari–Takagi problem treated in this section is a lower triangular integral operator K on $L^2(\mathbb{R}, \mathbb{C}^m)$ which is the input–output operator of the admissible system $\Sigma(K)$ given by

$$\Sigma(K):\quad \begin{cases} \left(\frac{d}{dt}x\right)(t) = A_K(t)x(t) + B_K(t)u(t), \\ y(t) = C_K(t)x(t), \end{cases} \quad t \in \mathbb{R}. \tag{21}$$

Given such an operator K and a nonnegative integer l, the problem is to find all operators $T = T_\Sigma$ with Σ an admissible system (19) satisfying rank $P_\Sigma \leqslant l$ such that $\|K + T\| < 1$. Here the norm $\|\cdot\|$ is the L^2-induced operator norm (see item 1 below).

Let us now put this problem into the abstract set up of Section 2.

1. The triple $(\mathcal{R}, \mathcal{N}_+, \mathcal{N})$. In this subsection $\mathcal{R}$ is the C^*-algebra of all bounded operators on $L^2(\mathbb{R})$. Thus $\mathcal{R}^{m\times m}$ is the set of all bounded operators T on $L^2(\mathbb{R}, \mathbb{C}^m)$. The norm on $\mathcal{R}^{m\times m}$ is the L^2-induced norm, that is,

$$\|T\|_{\mathcal{R}^{m\times m}} = \sup\left\{\|Tu\|_{L^2(\mathbb{R},\mathbb{C}^p)} \mid u \in L^2\left(\mathbb{R}, \mathbb{C}^m\right),\ \|u\|_{L^2(\mathbb{R},\mathbb{C}^m)} = 1\right\}.$$

The class of all integral operators with an admissible realization is denoted by $\mathcal{N}^{m\times m}$. The subset $\mathcal{N}_+^{m\times m}$ comprises all $T \in \mathcal{N}^{m\times m}$ that are upper triangular integral operators. Thus $T \in \mathcal{N}_+^{m\times m}$ if and only if $T = T_\Sigma$ with dichotomy $P_\Sigma = 0$, that is, the admissible realization Σ is backward-stable. From [12, Sections 1.2 and 1.3] it follows that $\mathcal{N} = \mathcal{N}^{1\times 1}$ is a subalgebra of $\mathcal{R}$, and $\mathcal{N}_+ = \mathcal{N}_+^{1\times 1}$ is a subalgebra of $\mathcal{N}$. Furthermore, the unit of $\mathcal{R}$, which is equal to the identity operator on $L^2(\mathbb{R})$, belongs to $\mathcal{N}_+$.

2. Stratification function μ. For $K \in \mathcal{N}^{m\times m}$ we define

$$\mu(T) = \min\{\text{rank}\, P_\Sigma \mid \Sigma \text{ is an admissible realization of } T\}.$$

Thus $\mu(T) \leqslant l$ if and only if T is the input–output operator of an admissible system Σ with the rank of the dichotomy P_Σ being less than or equal to l. With this definition the properties M1–M3 follow from [12, formula (1.13)].

The function μ then produces a stratification on $\mathcal{N}^{m\times m}$. Indeed, the set $\mathcal{N}_l^{m\times m}$ comprises the set of all input–output operators T in $\mathcal{N}^{m\times m}$ of systems Σ with dichotomy P_Σ with rank$(P_\Sigma) \leqslant l$. Furthermore, $\mathcal{N}_{[l]}^{m\times m}$ consists of all operators T in $\mathcal{N}^{m\times m}$ that have an admissible realization Σ with dichotomy P_Σ of rank l and no admissible realization with a dichotomy of strictly smaller rank.

3. Index function ν. It can be shown that $\mathcal{N}^{m\times m}$ is inverse closed. More precisely, if T is the input–output operator of an admissible system Σ (with dichotomy P_Σ), then T^{-1} is the input–output operator of a some system (denoted by, say Σ^{-1}) with dichotomy $P_{\Sigma^{-1}}$. In particular, if $T \in \mathcal{N}_+^{m\times m} \cap G\mathcal{R}^{m\times m}$, then $T^{-1} \in \mathcal{N}^{m\times m}$. This allows us to define the index function ν:

$$\nu(T) = \mu\left(T^{-1}\right), \quad T \in \mathcal{N}_+^{m\times m} \cap G\mathcal{R}^{m\times m}.$$

One can show that $\nu(T)$ is equal to the Fredholm index of the Fredholm operator $\tilde{T}$ acting on $L^2([0,\infty),\mathbb{C}^m)$, which is the compression of the input–output operator T to the half-line, that is,

$$(\tilde{T}u)(t) = (T\tilde{u})(t), \quad t \geqslant 0,\ u \in L^2\big([0,\infty),\mathbb{C}^m\big),$$

with

$$\tilde{u}(t) = \begin{cases} u(t) & \text{if } t \geqslant 0, \\ 0 & \text{if } t < 0. \end{cases}$$

The additivity of the index then follows from the additivity of the Fredholm index.

4. Homotopic invariance of the index. Let $N \in \mathcal{N}_+^{m\times m}$, $M \in \mathcal{N}_+^{m\times m} \cap G\mathcal{R}^{m\times m}$, and assume that $tN + M \in G\mathcal{R}^{m\times m}$ for $0 \leqslant t \leqslant 1$. For $t \in (0,1]$, we have

$$\nu(N+M) = \operatorname{ind}\left(\frac{1}{t}M\big(tNM^{-1} + E_m\big)\right) = \nu\left(\frac{1}{t}M\right) + \nu\big(tNM^{-1} + E_m\big)$$
$$= \nu(M) + \nu\big(tNM^{-1} + E_m\big).$$

But for t small enough, we can ensure that the input–output operator corresponding to tNM^{-1} has norm < 1. Hence from the stability of the index of Fredholm operators under perturbations that are small in operator norm (see [9, Section XI.4]), it follows that $\nu(tNM^{-1} + E_m) = \nu(E_m) = 0$. From this it can be seen that item 4 of Section 2 holds.

Conclusion.

Now let $K = K_\Sigma$ be the lower triangular integral operator. Suppose that

$$\sup_{\tau\in\mathbb{R}} \sigma_{l+1}\big(H_K(\tau)\big) < 1 < \inf_{\tau\in\mathbb{R}} \sigma_l\big(H_K(\tau)\big),$$

where $\sigma_r(\cdot)$ denotes the rth singular value of a bounded linear operator, and for $\tau \in \mathbb{R}$, $H_K(\tau)$ denotes the generalized Hankel operator on $L^2([0,\infty),\mathbb{C}^m)$ corresponding to the integral operator K, and is given by

$$\big(H_K(\tau)u\big)(t) = \int_0^\infty C(t+\tau)\gamma_A(t+\tau,\tau-s)B(\tau-s)u(s)\,ds, \quad t \geqslant 0.$$

Define

$$G_K(t) = U_K(t)^{-*}\int_t^\infty U_K(s)^*C_K(s)^*C_K(s)U_K(s)\,ds\,U_K(t)^{-1}, \quad t \in \mathbb{R}, \tag{22}$$

$$Z_K(t) = G_K(t)^{-1} - U_K(t)\int_{-\infty}^t U_K(s)^{-1}B_K(s)B_K(s)^*U_K(s)^{-*}\,ds\,U_K(t)^*, \tag{23}$$

$t \in \mathbb{R}$.

Assume that $G_K(t)$ is uniformly positive definite, and assume that for each $t \in \mathbb{R}$ the matrix $Z_K(t)$ is non-singular, and $(\det Z_K(\cdot)^{-1}) \neq 0$. So we may consider

$$A_\Theta = \begin{bmatrix} A_K(t) & 0 \\ 0 & -A_K(t)^* \end{bmatrix}, \tag{24}$$

$$B_\Theta = \begin{bmatrix} I - G_K(t)^{-1}Z_K(t)^{-1} & G_K(t)^{-1}Z_K(t)^{-1} \\ Z_K(t)^{-1} & -Z_K(t)^{-1} \end{bmatrix} \begin{bmatrix} -G_K(t)^{-1}C_K(t)^* & 0 \\ 0 & B_K(t) \end{bmatrix}, \tag{25}$$

$$C_\Theta = \begin{bmatrix} C_K(t) & 0 \\ 0 & B_K(t)^* \end{bmatrix}, \tag{26}$$

$$D_\Theta = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{27}$$

for $t \in \mathbb{R}$, and let Θ be the input–output operator of the system $\Sigma(\Theta)$, where

$$\Sigma(\Theta): \quad \begin{cases} \left(\frac{d}{dt}x\right)(t) = A_\Theta(t)x(t) + B_\Theta(t)u(t), \\ y(t) = C_\Theta(t)x(t) + D_\Theta(t)u(t), \end{cases} \quad t \in \mathbb{R}. \tag{28}$$

Using [12, Lemmas 5.2 and 5.3], one can prove that S1–S3 hold. We remark that Θ is also invertible. Indeed, we have $\Theta^* J \Theta = J$ and $\Theta J \Theta^* = J$ (this follows from Ball, Mikkola and Sasane [2, Theorem 2.1]), so that Θ has both a right inverse and a left inverse, and so it is invertible. Hence the Theorem 3.1 applies and we have the following.

Theorem 6.2. *Let K be the input–output operator of a forward stable admissible system of the form* (21). *Suppose that*

$$\sup_{\tau \in \mathbb{R}} \sigma_{l+1}\big(H_K(\tau)\big) < 1 < \inf_{\tau \in \mathbb{R}} \sigma_l\big(H_K(\tau)\big),$$

where $H_K(\tau)$ denotes the generalized Hankel operator corresponding to the integral operator K. Let Θ be the input–output operator of the system given by (28). *Then $F \in \mathcal{N}^{m\times m}$ is a strictly contractive extension of order l for K if and only if $F = (\Theta_{11}Q + \Theta_{12})(\Theta_{21}Q + \Theta_{22})^{-1}$, for some $Q \in \mathcal{N}_+^{m\times m}$ such that $\|Q\| < 1$.*

Acknowledgments

The authors thank Professor Joseph Ball (Virginia Tech., Blacksburg, USA) and Henrik Sandberg (Lunds Tekniska Högskola, Lund, Sweden) for several useful discussions.

References

[1] J.A. Ball, Nevanlinna–Pick interpolation: Generalizations and applications, in: J.B. Conway, B.B. Morrel (Eds.), Recent Results in Operator Theory, vol. I, Longman, 1988, pp. 51–94.

[2] J.A. Ball, I. Gohberg, M.A. Kaashoek, Nevanlinna–Pick interpolation for time varying input–output maps: The discrete time case, in: I. Gohberg (Ed.), Time-Variant Systems and Interpolation Operator Theory, in: Operator Theory, vol. 56, Birkhäuser, 1992, pp. 1–51.

[3] J.A. Ball, I. Gohberg, M.A. Kaashoek, The band method and Grassmannian approach for completion and extension problems, in: Oper. Theory Adv. Appl., vol. 87, Birkhäuser, 1996.

[4] J.A. Ball, J.W. Helton, A Beurling–Lax theorem for the Lie group $U(m,n)$ which contains most classical interpolation, J. Operator Theory 9 (1983) 107–142.

[5] J.A. Ball, K.M. Mikkola, A.J. Sasane, State-space formulas for the Nehari–Takagi problem for nonexponentially stable infinite-dimensional systems, SIAM J. Control Optim. (2004), in press.
[6] R.F. Curtain, H.J. Zwart, An Introduction to Infinite-Dimensional Linear Systems Theory, Springer, 1995.
[7] H. Dym, I. Gohberg, Extensions of kernels of Fredholm operators, J. Anal. Math. 42 (1982/1983) 51–97.
[8] R.L. Ellis, I. Gohberg, M.A. Kaashoek, Extensions with a positive real part. A new version of the abstract band method with applications, Integral Equations Operator Theory 16 (1993) 360–384.
[9] I. Gohberg, S. Goldberg, M.A. Kaashoek, Classes of Linear Operators, vol. I, Operator Theory, vol. 49, Birkhäuser, 1990.
[10] I. Gohberg, S. Goldberg, M.A. Kaashoek, Classes of Linear Operators, vol. II, Operator Theory, vol. 63, Birkhäuser, 1993.
[11] O.V. Iftime, M.A. Kaashoek, H. Sandberg, A.J. Sasane, A Grassmannian approach to the Hankel norm approximation problem, in: Proceedings of the Sixteenth International Symposium on Mathematical Theory of Networks and Systems, Leuven, Belgium, July, 2004, on CD-ROM.
[12] M.A. Kaashoek, J. Kos, The Nehari–Takagi problem for input–output operators of time varying continuous time systems, Integral Equations Operator Theory 18 (1994) 435–467.
[13] L. Rodman, I.M. Spitkovsky, H.J. Woerdeman, Abstract Band Method via Factorization, Positive and Band Extensions of Multivariable almost Periodic Matrix Functions, and Spectral Estimation, Mem. Amer. Math. Soc., vol. 762, Amer. Math. Soc., Providence, RI, 2002.
[14] A.J. Sasane, Hankel-Norm Approximation for Infinite-Dimensional Systems, Lecture Notes in Control and Information Sciences, vol. 277, Springer, 2002.
[15] N.J. Young, The Nehari problem and the optimal Hankel norm approximation, in: Analysis and Optimization of Systems: State and Frequency Domain Approaches for Infinite-Dimensional Systems, in: Lecture Notes in Control and Inform. Sci., vol. 185, Springer, 1993, pp. 199–221.