Operator Theory
Advances and Applications
Vol. 176

System Theory, the Schur Algorithm and Multidimensional Analysis

Daniel Alpay
Victor Vinnikov
Editors

Birkhäuser

Operator Theory: Advances and Applications
Vol. 176

System Theory, the Schur Algorithm and Multidimensional Analysis

Daniel Alpay
Victor Vinnikov
Editors

Birkhäuser
Basel · Boston · Berlin

Editors:

Daniel Alpay
Victor Vinnikov
Deaprtment of Mathematics
Ben-Gurion University of the Negev
P.O. Box 653
Beer Sheva 84105
Israel

e-mail: dany@math.bgu.ac.il
vinnikov@math.bgu.ac.il

2000 Mathematics Subject Classification 13B99, 15A33, 30E05, 44A60, 46A13, 46G10, 46E15, 47A48, 47A57, 47Bxx, 47L60, 65Fxx

Library of Congress Control Number: 2007922257

Bibliographic information published by Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at <http://dnb.ddb.de>.

ISBN 978-3-7643-8136-3 Birkhäuser Verlag, Basel – Boston – Berlin

Part of Springer Science+Business Media
Printed on acid-free paper produced from chlorine-free pulp. TCF ∞
Cover design: Heinz Hiltbrunner, Basel
Printed in Germany
ISBN-10: 3-7643-8136-1
ISBN-13: 978-3-7643-8136-3
e-ISBN-10: 3-7643-8137-X
e-ISBN-13: 978-3-7643-8137-0

9 8 7 6 5 4 3 2 1

www.birkhauser.ch

Contents

Editorial Introduction

Daniel Alpay and Victor Vinnikov

> *Tout problème profane un mystère; à son tour,*
> *le problème est profané par sa solution.*
> (Cioran, Syllogismes de l'amertume, [5, p. 40])

This present volume contains six papers written on the occasion of the workshop *Operator theory, system theory and scattering theory: multidimensional generalizations and related topics, 2005*, held at the Department of Mathematics of the Ben-Gurion University of the Negev during the period June 26–July 1, 2005. As for the previous conferences held in 2001 and 2003, we wish to thank the participants for a great scientific atmosphere and acknowledge the support of the Center of Advanced Studies in Mathematics of Ben-Gurion University of the Negev, which made the workshop possible. A volume of papers written on the occasion of the 2003 conference appeared in the series *Operator Theory: Advances and Applications*; see [2].

The papers can be divided into the following topics, which had an important place in the conference, namely:

1. Schur analysis
2. Quaternionic analysis
3. Multidimensional operator theory.
4. Moment problems

Schur analysis: Schur analysis originates with the papers [7], [8] where I. Schur associates to a function analytic and bounded by one in modulus in the open unit disk an infinite sequence of numbers in the open unit disk, or if the function is a finite Blaschke product, a finite sequence of numbers in the open unit disk and a number on the unit circle. It is an active field of research, see, e.g., [1] for recent developments. The first paper, **The transformation of Issai Schur and related topics in an indefinite setting**, by A. Dijksma, H. Langer and the first named editor, reviews recent developments in the Schur algorithm in the indefinite case. The authors focus on the scalar case. This hypothesis allows to obtain uniqueness results for the factorization of certain 2×2 matrix functions which are unitary with respect to an indefinite metric on the unit circle or on the real line. The theory of reproducing kernel Pontryagin spaces of the kind introduced by L. de Branges and J. Rovnyak in the Hilbert space case, see [3], [4], play an important role in the theory. The next paper, **A truncated matricial moment problem on a finite interval. The case of an odd number of prescribed moments**, by A. Choque Rivero, Y. Dyukarev, B. Fritzsche and B. Kirstein, deals with the matrix-valued case. The authors use V. Potapov's Fundamental Matrix Inequality method combined with L. Sakhnovich's method of operator identities. The Schur algorithm has been conducive to important applications in signal processing, in particular in relationships with structured matrices. In the paper **Algorithms for hierarchically semi-separable representation**, Z. Sheng, P. Dewilde and S. Chandrasekaran consider hierarchical semi separable matrices, which are structured matrices for which various efficient algorithms can be developed, for instance to compute the Moore-Penrose inverse. The study is in the setting of time-varying systems.

Quaternionic analysis: Linear algebra when the complex field is replaced by the skew field of quaternions has recently been the object of various papers; see for instance [6], [9]. Such results are important *per se* and also in the wider setting of hypercomplex analysis, when the Cauchy operator is replaced by the Cauchy–Fueter (resp. the Dirac) operator and analytic functions are replaced by hypercomplex functions (resp. by regular functions). In **Canonical forms for symmetric and skewsymmetric quaternionic**, L. Rodman studies canonical forms for pairs of quaternionic matrices, and gives applications to various problems where symmetric matrices are involved and to joint numerical cones of pairs of skewsymmetric quaternionic matrices.

Multidimensional operator theory: We discussed at length the various extensions of multidimensional operator theory in the editorial introduction of [2]. In the present volume two directions are represented: the first is the theory of homogeneous operators while the second is related to abstract moment problems. More precisely, in the paper **On the irreducibility of a class of homogeneous operators**, G. Misra and S. Shyam Roy use Hilbert module techniques to study homogeneous d-uples of operators. In **Unbounded normal algebras and spaces of fractions**, F. Vasilescu considers algebras of fractions of continuous functions to study representations of normal algebras and operator moment problems.

Moment problems: Moment problems permeate much of mathematical analysis. They have both motivated many of the more abstract and technical developments during the last hundred years, and benefited in their turn from such developments. In the present volume they are discussed in the first, second and last paper.

References

[1] D. Alpay and I. Gohberg, editors. *Interpolation, Schur functions and moment problems*, volume 165 of *Operator Theory: Advances and Applications*. Birkhäuser Verlag, Basel, 2006.

[2] D. Alpay and V. Vinnikov, editors. *Operator theory, systems theory and scattering theory: multidimensional generalizations*, volume 157 of *Operator Theory: Advances and Applications*. Birkhäuser Verlag, Basel, 2005.

[3] L. de Branges and J. Rovnyak. Canonical models in quantum scattering theory. In C. Wilcox, editor, *Perturbation theory and its applications in quantum mechanics*, pages 295–392. Wiley, New York, 1966.

[4] L. de Branges and J. Rovnyak. *Square summable power series*. Holt, Rinehart and Winston, New York, 1966.

[5] E.M. Cioran. *Syllogismes de l'amertume*. Collection idées. Gallimard, Paris, 1976. First published in 1952.

[6] M. Karow. Self-adjoint operators and pairs of Hermitian forms over the quaternions. *Linear Algebra Appl.*, 299(1-3):101–117, 1999.

[7] I. Schur. Über die Potenzreihen, die im Innern des Einheitkreises beschränkt sind, I. *Journal für die Reine und Angewandte Mathematik*, 147:205–232, 1917. English translation in: I. Schur methods in operator theory and signal processing. (Operator theory: Advances and Applications OT 18 (1986), Birkhäuser Verlag), Basel.

[8] I. Schur. Über die Potenzreihen, die im Innern des Einheitkreises beschränkt sind, II. *Journal für die Reine und Angewandte Mathematik*, 148:122–145, 1918. English translation in: I. Schur methods in operator theory and signal processing. (Operator theory: Advances and Applications OT 18 (1986), Birkhäuser Verlag), Basel.

[9] F. Zhang. Quaternions and matrices of quaternions. *Linear Algebra Appl.*, 251:21–57, 1997.

Daniel Alpay and Victor Vinnikov
Department of Mathematics
Ben-Gurion University of the Negev
Beer-Sheva, Israel
e-mail: `dany@math.bgu.ac.il`
e-mail: `vinnikov@math.bgu.ac.il`

Operator Theory:
Advances and Applications, Vol. 176, 1–98

The Transformation of Issai Schur and Related Topics in an Indefinite Setting

D. Alpay, A. Dijksma and H. Langer

Abstract. We review our recent work on the Schur transformation for scalar generalized Schur and Nevanlinna functions. The Schur transformation is defined for these classes of functions in several situations, and it is used to solve corresponding basic interpolation problems and problems of factorization of rational J-unitary matrix functions into elementary factors. A key role is played by the theory of reproducing kernel Pontryagin spaces and linear relations in these spaces.

Mathematics Subject Classification (2000). Primary 47A48, 47A57, 47B32, 47B50.

Keywords. Schur transform, Schur algorithm, generalized Schur function, generalized Nevanlinna function, Pontryagin space, reproducing kernel, Pick matrix, coisometric realization, self-adjoint realization, J-unitary matrix function, minimal factorization, elementary factor.

Contents

D. Alpay thanks the Earl Katz family for endowing the chair which supports his research and NWO, the Netherlands Organization for Scientific Research (grant B 61-524). The research of A. Dijksma and H. Langer was in part supported by the Center for Advanced Studies in Mathematics (CASM) of the Department of Mathematics, Ben–Gurion University.

1. Introduction

The aim of this survey paper is to review some recent development in Schur analysis for scalar functions in an indefinite setting and, in particular, to give an overview of the papers [7], [8], [9], [10], [11], [15], [16], [17], [18], [125], and [126].

1.1. Classical Schur analysis

In this first subsection we discuss the positive definite case. The starting point is a function $s(z)$ which is analytic and contractive (that is, $|s(z)| \leq 1$) in the open unit disk $\mathbb{D}$; we call such functions *Schur functions*. If $|s(0)| < 1$, by Schwarz' lemma, also the function

$$\widehat{s}(z) = \frac{1}{z} \frac{s(z) - s(0)}{1 - s(z)s(0)^*} \tag{1.1}$$

is a Schur function; here and throughout the sequel * denotes the adjoint of a matrix or an operator and also the complex conjugate of a complex number. The transformation $s(z) \mapsto \widehat{s}(z)$ was defined and studied by I. Schur in 1917–1918 in his papers [116] and [117] and is called the *Schur transformation*. It maps the set of Schur functions which are not identically equal to a unimodular constant into the set of Schur functions. If $\widehat{s}(z)$ is not a unimodular constant, the transformation (1.1) can be repeated with $\widehat{s}(z)$ instead of $s(z)$ etc. In this way, I. Schur associated with a Schur function $s(z)$ a finite or infinite sequence of numbers ρ_j in $\overline{\mathbb{D}}$, called *Schur coefficients*, via the formulas

$$s_0(z) = s(z), \quad \rho_0 = s_0(0),$$

and for $j = 0, 1, \ldots,$

$$s_{j+1}(z) = \widehat{s}_j(z) = \frac{1}{z} \frac{s_j(z) - s_j(0)}{1 - s_j(z)s_j(0)^*}, \quad \rho_{j+1} = s_{j+1}(0). \tag{1.2}$$

The recursion (1.2) is called the *Schur algorithm*. It stops after a finite number of steps if, for some j_0, $|\rho_{j_0}| = 1$. This happens if and only if $s(z)$ is a finite Blaschke product:

$$s(z) = c \prod_{\ell=1}^{n} \frac{z - a_\ell}{1 - z a_\ell^*}, \quad |c| = 1, \quad \text{and} \quad |a_\ell| < 1, \quad \ell = 1, \ldots, n,$$

with $n = j_0$, see [116] and [117].

If for a 2×2 matrix $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $v \in \mathbb{C}$ we define the *linear fractional transform* $\mathcal{T}_M(v)$ by

$$\mathcal{T}_M(v) = \frac{av + b}{cv + d},$$

the transform $\widehat{s}(z)$ in (1.1) can be written as

$$\widehat{s}(z) = \mathcal{T}_{\Phi(z)}(s(z)),$$

where

$$\Phi(z) = \frac{1}{z\sqrt{1-|s(0)|^2}} \begin{pmatrix} 1 & -s(0) \\ -zs(0)^* & z \end{pmatrix}.$$

Then it follows that

$$\Phi(z)^{-1} = \frac{1}{\sqrt{1-|s(0)|^2}} \begin{pmatrix} 1 & s(0) \\ s(0)^* & 1 \end{pmatrix} \begin{pmatrix} z & 0 \\ 0 & 1 \end{pmatrix}, \tag{1.3}$$

and

$$s(z) = \mathcal{T}_{\Phi(z)^{-1}}(\widehat{s}(z)) = \frac{s(0) + z\widehat{s}(z)}{1 + z\widehat{s}(z)s(0)^*}. \tag{1.4}$$

The matrix polynomial $\Phi(z)^{-1}$ in (1.3) is J_c-*inner* with $J_c = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, that is,

$$J_c - \Phi(z)^{-1} J_c \Phi(z)^{-*} \begin{cases} \le 0, & |z| < 1, \\ = 0, & |z| = 1. \end{cases}$$

Note that $\Theta(z) = \Phi(z)^{-1}\Phi(1)$ is of the form

$$\Theta(z) = I_2 + (z-1)\frac{\mathbf{u}\mathbf{u}^* J_c}{\mathbf{u}^* J_c \mathbf{u}}, \quad \mathbf{u} = \begin{pmatrix} 1 \\ s(0)^* \end{pmatrix}. \tag{1.5}$$

Of course $\Phi(z)^{-1}$ in (1.4) can be replaced by $\Theta(z)$, which changes $\widehat{s}(z)$.

Later, see Theorem 5.10, we will see that the matrix function $\Theta(z)$ given by (1.5) is elementary in the sense that it cannot be written as a product of two nonconstant J_c-inner matrix polynomials.

A repeated application of the Schur transformation leads to a representation of $s(z)$ as a linear fractional transformation

$$s(z) = \frac{a(z)\widetilde{s}(z) + b(z)}{c(z)\widetilde{s}(z) + d(z)}, \tag{1.6}$$

where $\widetilde{s}(z)$ is a Schur function and where the matrix function

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

is a J_c-inner matrix polynomial. In fact, this matrix function $\Theta(z)$ can be chosen a finite product of factors of the form (1.5) times a constant J_c-unitary factor. To see this it is enough to recall that the linear fractional transformations $\mathcal{T}_M$ have the semi-group property:

$$\mathcal{T}_{M_1 M_2}(v) = \mathcal{T}_{M_1}(\mathcal{T}_{M_2}(v)),$$

if only the various expressions make sense.

A key fact behind the scene and which hints at the connection with interpolation is the following: *Given a representation* (1.6) *of a Schur function $s(z)$ with a J_c-inner matrix polynomial $\Theta(z)$ and a Schur function $\widetilde{s}(z)$, then the matrix polynomial $\Theta(z)$ depends only on the first $n =$* deg Θ *derivatives of $s(z)$ at the origin.* (Here deg denotes the McMillan degree, see Subsection 3.1.) To see this we use that det $\Theta(z) = e^{it}z^n$ with some $t \in \mathbb{R}$ and $n = \deg \Theta$, see Theorem 3.12. It follows that

$$\begin{aligned} \mathcal{T}_{\Theta(z)}(\widetilde{s}(z)) - \mathcal{T}_{\Theta(z)}(0) &= \frac{\big(a(z)d(z) - b(z)c(z)\big)\widetilde{s}(z)}{\big(c(z)\widetilde{s}(z) + d(z)\big)d(z)} \\ &= \frac{(\det\ \Theta(z))\widetilde{s}(z)d(z)^{-2}}{\big(c(z)\widetilde{s}(z)d(z)^{-1} + 1\big)} = z^n\xi(z) \end{aligned}$$

with

$$\xi(z) = \frac{e^{it}\widetilde{s}(z)d(z)^{-2}}{\big(c(z)\widetilde{s}(z)d(z)^{-1} + 1\big)}.$$

For any nonconstant J_c-inner matrix polynomial $\Theta(z)$ the function $d(z)^{-1}$ is analytic and contractive, and the function $d(z)^{-1}c(z)$ is analytic and strictly contractive, on $\mathbb{D}$, see [79], hence the function $\xi(z)$ is also analytic in the open unit disk and therefore

$$\mathcal{T}_{\Theta(z)}(\widetilde{s}(z)) - \mathcal{T}_{\Theta(z)}(0) = \mathrm{O}(z^n), \quad z \to 0.$$

These relations imply that the Schur algorithm allows to solve recursively the Carathéodory–Fejér interpolation problem: *Given complex numbers* $\sigma_0, \dots, \sigma_{n-1}$*, find all* (*if any*) *Schur functions* $s(z)$ *such that*

$$s(z) = \sigma_0 + z\sigma_1 + \cdots + z^{n-1}\sigma_{n-1} + \mathrm{O}(z^n), \quad z \to 0.$$

The Schur algorithm expresses the fact that one needs to know how to solve this problem only for $n = 1$. We call this problem the *basic interpolation problem.*

The basic interpolation problem: *Given* $\sigma_0 \in \mathbb{C}$*, find all Schur functions* $s(z)$ *such that* $s(0) = \sigma_0$.

Clearly this problem has no solution if $|\sigma_0| > 1$, and, by the maximum modulus principle, it has a unique solution if $|\sigma_0| = 1$, namely the constant function $s(z) \equiv \sigma_0$. If $|\sigma_0| < 1$, then the solution is given by the linear fractional transformation (compare with (1.4))

$$s(z) = \frac{\sigma_0 + z\widetilde{s}(z)}{1 + z\widetilde{s}(z)\sigma_0^*}, \tag{1.7}$$

where $\widetilde{s}(z)$ varies in the set of Schur functions. Note that the solution $s(z)$ is the inverse Schur transform of the parameter $\widetilde{s}(z)$. If we differentiate both sides of (1.7) and put $z = 0$ then it follows that $\widetilde{s}(z)$ satisfies the interpolation condition

$$\widetilde{s}(0) = \frac{\sigma_1}{1 - |\sigma_0|^2}, \quad \sigma_1 = s'(0).$$

Thus if the Carathéodory–Fejér problem is solvable and has more than one solution (this is also called the *nondegenerate case*), these solutions can be obtained by repeatedly solving a basic interpolation problem (namely, first for $s(z)$, then for $\widetilde{s}(z)$, and so on) and are described by a linear fractional transformation of the form (1.6) for some J_c-inner 2×2 matrix polynomial $\Theta(z)$.

The fact that the Carathéodory–Fejér interpolation problem can be solved iteratively via the Schur algorithm implies that any J_c-inner 2×2 matrix polynomial can be written in a unique way as a product of J_c-inner 2×2 matrix polynomials of McMillan degree 1, namely factors of the form

$$\frac{1}{\sqrt{1-|\rho|^2}}\begin{pmatrix}1 & \rho\\ \rho^* & 1\end{pmatrix}\begin{pmatrix}z & 0\\ 0 & 1\end{pmatrix} \tag{1.8}$$

with some complex number ρ, $|\rho| < 1$, and a J_c-unitary constant. These factors of McMillan degree 1 are elementary, see Theorem 5.10, and can be chosen normalized: If one fixes, for instance, the value at $z = 1$ to be I_2, factors $\Theta(z)$ of the form (1.5) come into play. Note that the factor (1.8) is not normalized in this sense when $\rho \neq 0$. Furthermore, the Schur algorithm is also a method which yields this J_c-minimal factorization of a J_c-inner 2×2 matrix polynomial $\Theta(z)$ into elementary factors. Namely, it suffices to take any number τ on the unit circle and to apply the Schur algorithm to the function $s(z) = \mathcal{T}_{\Theta(z)}(\tau)$; the corresponding sequence of elementary J_c-inner 2×2 matrix polynomial gives the J_c-inner minimal factorization of $\Theta(z)$.

Schur's work was motivated by the works of Carathéodory, see [53] and [54], and Toeplitz, see [122], on *Carathéodory functions* which by definition are the analytic functions in the open unit disk which have a nonnegative real part there, see [116, English transl., p. 55].

A sequence of Schur coefficients can also be associated with a Carathéodory function; sometimes these numbers are called *Verblunsky coefficients*, see [86, Chapter 8]. Carathéodory functions $\phi(z)$ play an important role in the study of the trigonometric moment problem via the Herglotz representation formula

$$\phi(z) = ia + \int_0^{2\pi} \frac{e^{it}+z}{e^{it}-z} d\mu(t) = ia + \int_0^{2\pi} d\mu(t) + 2\sum_{\ell=1}^{\infty} z^\ell \int_0^{2\pi} e^{-i\ell t} d\mu(t),$$

where a is a real number and $d\mu(t)$ is a positive measure on $[0, 2\pi)$. A function $\phi(z)$, defined in the open unit disk, is a Carathéodory function if and only if the kernel

$$K_\phi(z,w) = \frac{\phi(z)+\phi(w)^*}{1-zw^*} \tag{1.9}$$

is nonnegative. Similarly, a function $s(z)$, defined in the open unit disk, is a Schur function if and only if the kernel

$$K_s(z,w) = \frac{1-s(z)s(w)^*}{1-zw^*}$$

is nonnegative in $\mathbb{D}$.

In this paper we do not consider Carathéodory functions with associated kernel (1.9), but functions $n(z)$ which are holomorphic or meromorphic in the upper half-plane $\mathbb{C}^+$ and for which the Nevanlinna kernel

$$L_n(z,w) = \frac{n(z)-n(w)^*}{z-w^*}$$

has certain properties. For example, if the kernel $L_n(z,w)$ is nonnegative in $\mathbb{C}^+$, then the function $n(z)$ is called a *Nevanlinna function*. The Schur transformation (1.1) for Schur functions has an analog for Nevanlinna functions in the theory of the Hamburger moment problem and was studied by N.I. Akhiezer, see [4, Lemma 3.3.6] and Subsection 8.1.

To summarize the previous discussion one can say that the Schur transformation, the basic interpolation problem and J_c-inner factorizations of 2×2 matrix polynomials are three different facets of a common object of study, which can be called *Schur analysis*. For more on the original works we refer to [82] and [83]. Schur analysis is presently a very active field, we mention, for example, [75] for scalar Schur functions and [74] and [84] for matrix Schur functions, and the references cited there.

The Schur transform (1.1) for Schur functions is centered at $z_1 = 0$. The Schur transform centered at an arbitrary point $z_1 \in \mathbb{D}$ is defined by

$$\widehat{s}(z) = \frac{1}{b_c(z)}\frac{s(z)-s(z_1)}{1-s(z)s(z_1)^*} \equiv \frac{1}{b_c(z)}\frac{s(z)-\sigma_0}{1-s(z)\sigma_0^*},$$

where $b_c(z)$ denotes the Blaschke factor related to the circle and z_1:

$$b_c(z) = \frac{z-z_1}{1-zz_1^*}.$$

This definition is obtained from (1.1) by changing the independent variable to $\zeta(z) = b_c(z)$, which leaves the class of Schur functions invariant. In this paper we consider the generalization of this transformation to an indefinite setting, that is, to a transformation centered at z_1 of the class of generalized Schur functions with $z_1 \in \mathbb{D}$ and $z_1 \in \mathbb{T}$, and to a transformation centered at z_1 of the class of generalized Nevanlinna functions with $z_1 \in \mathbb{C}^+$ and $z_1 = \infty$ (here also the case $z_1 \in \mathbb{R}$ might be of interest, but it is not considered in this paper). We call this generalized transformation also the *Schur transformation*.

1.2. Generalized Schur and Nevanlinna functions

In the present paper we consider essentially two classes of scalar functions. The first class consists of the meromorphic functions $s(z)$ on the open unit disc $\mathbb{D}$ for which the kernel

$$K_s(z,w) = \frac{1-s(z)s(w)^*}{1-zw^*}, \quad z,w \in \mathrm{hol}\,(s),$$

has a finite number κ of negative squares (here $\mathrm{hol}\,(s)$ is the domain of holomorphy of $s(z)$), for the definition of negative squares see Subsection 2.1. This is equivalent to the fact that the function $s(z)$ has κ poles in $\mathbb{D}$ but the metric constraint of being not expansive on the unit circle $\mathbb{T}$ (in the sense of nontangential boundary values from $\mathbb{D}$), which holds for Schur functions, remains. We call these functions $s(z)$ *generalized Schur functions with* κ *negative squares*. The second class is the set of *generalized Nevanlinna functions with* κ *negative squares*: These are the meromorphic functions $n(z)$ on $\mathbb{C}^+$ for which the kernel

$$L_n(z,w) = \frac{n(z) - n(w)^*}{z - w^*}, \quad z, w \in \operatorname{hol}(n),$$

has a finite number κ of negative squares. We always suppose that they are extended to the lower half-plane by symmetry: $n(z^*) = n(z)^*$. Generalized Nevanlinna functions $n(z)$ for which the kernel $L_n(z,w)$ has κ negative squares have at most κ poles in the open upper half-plane $\mathbb{C}^+$; they can also have 'generalized poles of nonpositive type' on the real axis, see [99] and [100]. Note that if the kernels $K_s(z,w)$ and $L_n(z,w)$ are nonnegative the functions $s(z)$ and $n(z)$ are automatically holomorphic, see, for instance, [6, Theorem 2.6.5] for the case of Schur functions. The case of Nevanlinna functions can be deduced from this case by using Möbius transformations on the dependent and independent variables, as in the proof of Theorem 7.13 below.

Generalized Schur and Nevanlinna functions have been introduced independently and with various motivations and characterizations by several mathematicians. Examples of functions of bounded type with poles in $\mathbb{D}$ and the metric constraint that the nontangential limits on $\mathbb{T}$ are bounded by 1 were already considered by T. Takagi in his 1924 paper [121] and by N.I. Akhiezer in the maybe lesser known paper [3] of 1930. These functions are of the form

$$s(z) = \frac{p(z)}{z^n p(1/z^*)^*},$$

where $p(z)$ is a polynomial of degree n, and hence are examples of generalized Schur functions. Independently, functions with finitely many poles in $\mathbb{D}$ and the metric constraint on the circle were introduced by Ch. Chamfy, J. Dufresnoy, and Ch. Pisot, see [55] and [78]. It is fascinating that also in the work of these authors there appear functions of the same form, but with polynomials $p(z)$ with integer coefficients, see, for example, [55, p. 249]. In related works of M.-J. Bertin [41] and Ch. Pisot [105] the Schur algorithm is considered where the complex number field is replaced by a real quadratic field or a p-adic number field, respectively. In none of these works any relation was mentioned with the Schur kernel $K_s(z,w)$. The approach using Schur and Nevanlinna kernels was initiated by M.G. Krein and H. Langer in connection with their study of operators in Pontryagin spaces, see [94], [95], [96], [97], [98], and [99]. Their definition in terms of kernels allows to study the classes of generalized Schur and Nevanlinna functions with tools from functional analysis and operator theory (in particular, the theory of reproducing kernel spaces and the theory of operators on spaces with an indefinite inner product), and it leads to connections with realization theory, interpolation theory and other related topics.

1.3. Reproducing kernel Pontryagin spaces

The approach to the Schur transformation in the indefinite case in the present paper is based on the theory of reproducing kernel Pontryagin spaces for scalar and matrix functions, associated, for example, in the Schur case with a Schur function $s(z)$ and a 2×2 matrix function $\Theta(z)$ through the reproducing kernels

$$K_s(z,w) = \frac{1 - s(z)s(w)^*}{1 - zw^*}, \quad K_\Theta(z,w) = \frac{J_c - \Theta(z)J_c\Theta(w)^*}{1 - zw^*}, \; J_c = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix};$$

these spaces are denoted by $\mathcal{P}(s)$ and $\mathcal{P}(\Theta)$, respectively. In the positive case, they have been first introduced by L. de Branges and J. Rovnyak in [50] and [49]. They play an important role in operator models and interpolation theory, see, for instance, [20], [79], and [81]. In the indefinite case equivalent spaces were introduced in the papers by M.G. Krein and H. Langer mentioned earlier.

We also consider the case of generalized Nevanlinna functions $n(z)$ and corresponding 2×2 matrix functions $\Theta(z)$, where the reproducing kernels are of the form

$$L_n(z,w) = \frac{n(z) - n(w)^*}{z - w^*}, \quad K_\Theta(z,w) = \frac{J_\ell - \Theta(z)J_\ell\Theta(w)^*}{z - w^*}, \; J_\ell = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

We denote by $\mathcal{L}(n)$ the reproducing kernel space associated with the first kernel and by $\mathcal{P}(\Theta)$ the reproducing kernel space associated to the second kernel. This space $\mathcal{P}(\Theta)$ differs from the one above, but it should be clear from the context to which reproducing kernel $K_\Theta(z,w)$ it belongs. The questions we consider in this paper are of analytic and of geometric nature. The starting point is the Schur transformation for generalized Schur functions centered at an inner point $z_1 \in \mathbb{D}$ or at a boundary point $z_1 \in \mathbb{T}$, and for generalized Nevanlinna functions centered at an inner point $z_1 \in \mathbb{C}^+$ or at the boundary point ∞. Generalized Schur and Nevanlinna functions are also characteristic functions of certain colligations with a metric constraint, and we study the effect of the Schur transformation on these underlying colligations. We explain this in more detail for generalized Schur functions and an inner point $z_1 \in \mathbb{D}$.

By analytic problems we mean:

- The basic interpolation problem for generalized Schur functions, that is, the problem to determine the set of all generalized Schur functions analytic at a point $z_1 \in \mathbb{D}$ and satisfying $s(z_1) = \sigma_0$. The solution depends on whether $|\sigma_0| < 1$, > 1, or $= 1$, and thus the basic interpolation problem splits into three different cases. It turns out that in the third case more data are needed to get a complete description of the solutions in terms of a linear fractional transformation.
- The problem of decomposing a rational 2×2 matrix function $\Theta(z)$ with a single pole in $1/z_1^*$ and J_c-unitary on $\mathbb{T}$ as a product of elementary factors with the same property. Here the Schur algorithm, which consists of a repeated application of the Schur transformation, gives an explicit procedure to obtain such a factorization. The factors are not only of the form (1.8) as in Subsection 1.1 but may have a McMillan degree > 1. These new types of factors have first been exhibited by Ch. Chamfy in [55] and by Ph. Delsarte, Y. Genin, and Y. Kamp in [63].

By geometric problems we mean in particular

- to give an explicit description of the (unitary, isometric or coisometric) colligation corresponding to $\widehat{s}(z)$ in terms of the colligation of $s(z)$.
- To find the relation between the reproducing kernel spaces $\mathcal{P}(s)$ for $s(z)$ and $\mathcal{P}(\widehat{s}(z))$ for $\widehat{s}(z)$, where $\widehat{s}(z)$ denotes the Schur transform of $s(z)$. In fact, $\mathcal{P}(\widehat{s}(z))$ can be isometrically embedded into $\mathcal{P}(s)$ with orthogonal complement which is an isometric copy of $\mathcal{P}(\Theta)$, for some rational J_c-unitary 2×2 matrix function $\Theta(z)$.

1.4. The general scheme

The Schur transformation for generalized Schur and Nevanlinna functions which we review in this paper can be explained from a general point of view as in [23], [24], and [25]. In fact, consider two analytic functions $a(z)$ and $b(z)$ on a connected set $\Omega \subset \mathbb{C}$ with the property that the sets

$$\begin{aligned}\Omega_+ &= \left\{z \in \Omega \,\middle|\, |a(z)| > |b(z)|\right\},\\ \Omega_- &= \left\{z \in \Omega \,\middle|\, |a(z)| < |b(z)|\right\},\\ \Omega_0 &= \left\{z \in \Omega \,\middle|\, |a(z)| = |b(z)|\right\}\end{aligned}$$

are nonempty; it is enough to require that Ω_+ and Ω_- are nonempty, then $\Omega_0 \neq \{\emptyset\}$ and it contains at least one point $z_0 \in \Omega_0$ for which $a(z_0) \neq 0$ and hence $b(z_0) \neq 0$, see [24, p. 119]. The kernels $K_s(z,w)$ and $L_n(z,w)$ considered above are special cases of the kernel

$$K_X(z,w) = \frac{X(z)JX(w)^*}{a(z)a(w)^* - b(z)b(w)^*}, \tag{1.10}$$

where J is a $p \times p$ signature matrix and $X(z)$ is a meromorphic $1 \times p$ vector function in Ω_+. Indeed, we obtain these kernels by setting $\Omega = \mathbb{C}$, $p = 2$, and

$$X(z) = \begin{pmatrix}1 & -s(z)\end{pmatrix}, \quad a(z) = 1, \quad b(z) = z, \quad J = J_c, \tag{1.11}$$

and

$$X(z) = \begin{pmatrix}1 & -n(z)\end{pmatrix}, \quad a(z) = \frac{1-iz}{\sqrt{2}}, \quad b(z) = \frac{1+iz}{\sqrt{2}}, \quad J = -iJ_\ell, \tag{1.12}$$

respectively, where

$$J_c = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}, \quad J_\ell = \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}; \tag{1.13}$$

here the letters c and ℓ stand for circle and line. We assume that $K_X(z,w)$ has a finite number of negative squares and denote by $\mathcal{B}(X)$ the associated reproducing kernel Pontryagin space. In case of (1.11) we have $\mathcal{B}(X) = \mathcal{P}(s)$ and in case of (1.12) we have $\mathcal{B}(X) = \mathcal{L}(n)$. The Schur transformation centered at a point $z_1 \in \Omega_+ \cup \Omega_0$ in this general setting is defined by means of certain invariant subspaces. To explain this we first restrict the discussion to the case $z_1 \in \Omega_+$ and then briefly discuss the case $z_1 \in \Omega_0$. To construct these invariant subspaces we take the following steps.

Step 1: *Build the linear space $\mathcal{M}(X)$ spanned by the sequence of $p \times 1$ vector functions*

$$f_j(z) = \frac{1}{j!}\frac{\partial^j}{\partial w^{*j}} \frac{JX(w)^*}{a(z)a(w)^* - b(z)b(w)^*}\Big|_{w=z_1}, \quad j = 0, 1, 2, \ldots.$$

Note that the space $\mathcal{M}(X)$ is invariant under the *backward-shift operators* R_ζ:

$$\big(R_\zeta f\big)(z) = \frac{a(z)f(z) - a(\zeta)f(\zeta)}{a(\zeta)b(z) - b(\zeta)a(z)}, \quad f(z) \in \mathcal{M}(X),$$

where $\zeta \in \Omega_+$ is such that $b(\zeta) \neq 0$ and the function $f(z)$ is holomorphic at $z = \zeta$. For $X(z)$ etc. as in (1.11) and (1.12) this reduces to the classical backward-shift invariance. Furthermore a finite-dimensional space is backward-shift invariant if and only if it is spanned by the columns of a matrix function of the form

$$F(z) = C\big(a(z)M - b(z)N\big)^{-1}$$

for suitable matrices M, N, and C.

Step 2: *Define an appropriate inner product on $\mathcal{M}(X)$ such that the map*

$$f(z) \mapsto X(z)f(z), \quad f(z) \in \mathcal{M}(X),$$

is an isometry from $\mathcal{M}(X)$ into the reproducing kernel Pontryagin space $\mathcal{B}(X)$.

We define the inner product on $\mathcal{M}(X)$ by defining it on the subspaces

$$\mathcal{M}_k = \operatorname{span}\big\{f_0(z), \ldots, f_{k-1}(z)\big\}, \quad k = 1, 2, \ldots.$$

The matrix function

$$F(z) = \begin{pmatrix} f_0(z) & f_1(z) & \cdots & f_{k-1}(z) \end{pmatrix}$$

can be written in the form

$$F(z) = C_{z_1}\big(a(z)M_{z_1} - b(z)N_{z_1}\big)^{-1},$$

where with

$$\alpha_j = \frac{a^{(j)}(z_1)}{j!}, \quad \beta_j = \frac{b^{(j)}(z_1)}{j!}, \qquad j = 0, 1, \ldots, k-1,$$

$$M_{z_1} = \begin{pmatrix} \alpha_0 & 0 & \cdots & 0 & 0 \\ \alpha_1 & \alpha_0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \alpha_{k-2} & \alpha_{k-3} & \cdots & \alpha_0 & 0 \\ \alpha_{k-1} & \alpha_{k-2} & \cdots & \alpha_1 & \alpha_0 \end{pmatrix}^*, \quad N_{z_1} = \begin{pmatrix} \beta_0 & 0 & \cdots & 0 & 0 \\ \beta_1 & \beta_0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \beta_{k-2} & \beta_{k-3} & \cdots & \beta_0 & 0 \\ \beta_{k-1} & \beta_{k-2} & \cdots & \beta_1 & \beta_0 \end{pmatrix}^*$$

and

$$C_{z_1} = J\left(X(z_1)^* \quad \frac{X'(z_1)^*}{1!} \quad \cdots \quad \frac{X^{(k-1)}(z_1)^*}{(k-1)!} \right),$$

see [23, (3.11)]. Hence $\mathcal{M}_k$ is backward-shift invariant. To define the restriction of the inner product $\langle\,\cdot\,,\,\cdot\,\rangle$ to $\mathcal{M}_k$ we choose the Gram matrix G associated with these k functions:

$$G = \big(g_{ij}\big)_{i,j=0}^{k-1}, \quad g_{ij} = \langle f_j, f_i\rangle,$$

as the solution of the matrix equation

$$M_{z_1}^* G M_{z_1} - N_{z_1}^* G N_{z_1} = C_{z_1}^* J C_{z_1}, \tag{1.14}$$

see [23, (2.15)]. The solution of (1.14) is unique since $|a(z_1)| > |b(z_1)|$.

Step 3: *Choose the smallest integer $k \geq 1$ such that the inner product space $\mathcal{M}_k$ from* Step 2 *is nondegenerate. It has a reproducing kernel of the form*

$$K_\Theta(z,w) = \frac{J - \Theta(z)J\Theta(w)^*}{a(z)a(w)^* - b(z)b(w)^*}$$

with the $p \times p$ matrix function $\Theta(z)$ given by the formula

$$\Theta(z) = I_p - \big(a(z)a(z_0)^* - b(z)b(z_0)^*\big)F(z)G^{-1}F(z_0)^*J.$$

The statement is a consequence of the following theorem, which describes the structure of certain backward-shift invariant subspaces. Now the matrices M, N, and C are not necessarily of the special form above.

Theorem 1.1. *Let M, N, C be matrices of sizes $m \times m$, $m \times m$, and $p \times m$, respectively, such that*

$$\det\big(a(z_0)M - b(z_0)N\big) \neq 0$$

for some point $z_0 \in \Omega_0$ and that the columns of the $p \times m$ matrix function

$$F(z) = C\big(a(z)M - b(z)N\big)^{-1}$$

are linearly independent in a neighborhood of z_0. Further, let G be an invertible Hermitian $m \times m$ matrix and endow the space $\mathcal{M}$ spanned by the columns of $F(z)$ with the inner product defined by G:

$$\big\langle F\mathbf{c}, F\mathbf{d}\big\rangle = \mathbf{d}^* G\mathbf{c}, \quad \mathbf{c}, \mathbf{d} \in \mathbb{C}^m. \tag{1.15}$$

Then the reproducing kernel for $\mathcal{M}$ is of the form

$$K_\Theta(z,w) = \frac{J - \Theta(z)J\Theta(w)^*}{a(z)a(w)^* - b(z)b(w)^*}$$

if and only if G is a solution of the matrix equation

$$M^*GM - N^*GN = C^*JC.$$

In this case the function $\Theta(z)$ can be chosen as

$$\Theta(z) = I_p - \big(a(z)a(z_0)^* - b(z)b(z_0)^*\big)F(z)G^{-1}F(z_0)^*J. \tag{1.16}$$

For the formula for $\Theta(z)$ and a proof of this theorem, see [23, (2.14)] and [24, Theorem 4.1].

The three steps lead to the following theorem, see [23, Theorem 4.1].

Theorem 1.2. *The following orthogonal decomposition holds*:

$$\mathcal{B}(X) = X\mathcal{P}(\Theta) \oplus \mathcal{B}(X\Theta).$$

The proof of the theorem follows from the decomposition

$$K_X(z,w) = X(z)K_\Theta(z,w)X(w)^* + K_{X\Theta}(z,w)$$

and from the theory of complementation in reproducing kernel Pontryagin spaces, see, for example, [19, Section 5]. We omit further details.

The existence of this minimal integer k and the backward-shift invariance of $\mathcal{M}_k$ in Step 3 are essential ingredients for the definition of the Schur transformation. The matrix function $\Theta(z)$ in Step 3 is elementary in the sense that $\mathcal{P}(\Theta)$ does not contain any proper subspace of the same form, that is, any nontrivial nondegenerate backward-shift invariant subspace. In the sequel we only consider the special cases (1.11) and (1.12). In these cases the space $\mathcal{P}(\Theta)$ is the span of one chain of the backward-shift operator and, by definition, the Schur transformation corresponds to the inverse of the linear fractional transformation $\mathcal{T}_{\Theta U}$ for some J-unitary constant U. The function $X(z)\Theta(z)$ is essentially the Schur transform of $X(z)$; the relation $X(z_1)\Theta(z_1) = 0$ corresponds to the fact that the numerator and the denominator in the Schur transform (1.1) are 0 at $z = z_1 = 0$.

In the boundary case, that is, if $z_1 \in \Omega_0$, $z_1 \neq z_0$, one has to take nontangential boundary values to define the matrices M_{z_1} and N_{z_1}. Then the equation (1.14) has more than one solution; nevertheless a solution G exists such that the required isometry holds.

Special cases of the formula (1.16) for $\Theta(z)$ appear in Section 3 below, see the formulas (3.15), (3.16), (3.23), and (3.24). Specializing to the cases considered in Sections 5 to 8 leads in a systematic way to the elementary J_c- or J_ℓ-unitary factors. The case $z_1 = \infty$ treated in Section 8 corresponds to the Hamburger moment problem for Nevanlinna functions $n(z)$ with finitely many moments given. Taking the nontangential limits alluded to above leads to the fact that the space $\mathcal{B}(X) = \mathcal{L}(n)$ contains functions of the form

$$n(z), \quad z^j n(z) + p_j(z), \quad j = 1, \dots, k-1,$$

where $p_j(z)$ is a polynomial of degree $j-1$, and $\mathcal{M}_k$ in Step 3 is replaced by the span of the functions

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} -p_j(z) \\ z^j \end{pmatrix}, \quad j = 1, \dots, k-1.$$

1.5. Outline of the paper

The following two sections have a preliminary character. In Section 2 we collect some facts about reproducing kernel Pontryagin spaces, in particular about those spaces which are generated by locally analytic kernels. The coefficients in the Taylor expansion of such a kernel lead to the notion of the Pick matrix. We also introduce the classes of generalized Schur and Nevanlinna functions, which are the main objects of study in this paper, the reproducing kernel Pontryagin spaces generated by these functions, and the realizations of these functions as characteristic functions of, for example, unitary or coisometric colligations or as compressed resolvents of self-adjoint operators. In Section 3 we first consider, for a general $p \times p$ signature matrix J, classes of rational J-unitary $p \times p$ matrix functions $\Theta(z)$ on the circle $\mathbb{T}$ and related to the kernel

$$\frac{J - \Theta(z)J\Theta(w)^*}{1 - zw^*},$$

and classes of rational J-unitary $p \times p$ matrix functions $\Theta(z)$ on the line $\mathbb{R}$ and related to the kernel

$$\frac{J - \Theta(z)J\Theta(w)^*}{-\mathrm{i}(z - w^*)}.$$

The special cases in which $p = 2$ and in the circle case $J = J_c$ and in the line case $J = -\mathrm{i}J_\ell$, where J_c and J_ℓ are given by (1.13), play a very important role in this paper. Since these matrix functions are rational, the Hermitian kernels have a finite number of negative and positive squares, and we introduce the finite-dimensional reproducing kernel Pontryagin spaces $\mathcal{P}(\Theta)$ for these kernels. Important notions are those of a factorization and of an elementary factor within the considered classes, and we prove some general factorization theorems. In particular, having in mind the well-known fact that the existence of an invariant subspace of a certain operator corresponds to a factorization of, for example, the characteristic function of this operator, we formulate general factorization theorems for the classes of rational matrix functions considered mainly in this paper.

As was mentioned already, we consider the Schur transformation at z_1 for generalized Schur functions for the cases that the transformation point z_1 is in $\mathbb{D}$ or on the boundary $\mathbb{T}$ of $\mathbb{D}$, and for generalized Nevanlinna functions for the cases that $z_1 \in \mathbb{C}^+$ or $z_1 = \infty$. In accordance with this, the basic interpolation problem, the factorization problem, and the realization problem we have always to consider for each of the four cases. Although the general scheme is in all cases the same, each of these cases has its own features. In particular, there is an essential difference if z_1 is an inner or a boundary point of the considered domain: In the first case we always suppose analyticity in this point, whereas in the second case only a certain asymptotic of the function in z_1 is assumed. (In this paper we only consider these four cases, but it might be of interest to study also the case of functions mapping the open unit disk into the upper half-plane or the upper half-plane into the open unit disk.)

In Section 4 we study the Pick matrices at the point z_1 for all the four mentioned cases. In the following Sections 5–8 we consider the Schur transformation, the basic interpolation problem, the factorization of the rational matrix functions, and the realization of the given scalar functions separately for each of these four cases in one section, which is immediately clear from the headings.

The Schur algorithm in the indefinite case has been studied by numerous authors, see, for example, [1], [21], [56], [57], and [62]. Our purpose here is to take full advantage of the scalar case and to obtain explicit analytical, and not just general formulas. For instance, in [23] and [24] the emphasis is on a general theory; in such a framework the special features of the scalar case and the subtle differences between generalized Schur and generalized Nevanlinna functions remain hidden. In the papers [9], [10], [15], and [18] we considered special cases with proofs specified to the case at hand. The general scheme given in Subsection 1.3 allows one to view these cases in a unified way.

With this survey paper we do not claim to give a historical account of the topics we cover. Besides the papers and books mentioned in the forgoing subsections we suggest the historical review in the book of Shohat and Tamarkin [119] which explains the relationships with the earlier works of Tchebycheff, Stieltjes and Markov, and the recent paper of Z. Sasvari [115]. For more information on the Schur algorithm in the positive scalar case we suggest Khrushchev's paper [92], the papers [61], [66], and [67] for the matrix case and the books [34] and [88].

We also mention that the Schur algorithm was extended to the time-varying case, see [64] and [68], to the case of multiscale processes, see [38] and [39], and to the case of tensor algebras, see [58], [59], and [60].

2. Kernels, classes of functions, and reproducing kernel Pontryagin spaces

In this section we review various facts from reproducing kernel Pontryagin spaces and we introduce the spaces of meromorphic functions needed in this paper.

2.1. Reproducing kernel Pontryagin spaces

Let p be an integer ≥ 1; in the sequel we mainly deal with $p = 1$ or $p = 2$. A $p \times p$ matrix function $K(z, w)$, defined for z, w in some set Ω, *has κ negative squares* if it is Hermitian:

$$K(z, w) = K(w, z)^*, \quad z, w \in \Omega,$$

and if for every choice of the integer $m \geq 1$, of $p \times 1$ vectors $\mathbf{c}_1, \ldots, \mathbf{c}_m$, and of points $w_1, \ldots, w_m \in \Omega$, the Hermitian $m \times m$ matrix

$$\big(\mathbf{c}_i^* K(w_i, w_j)\mathbf{c}_j\big)_{i,j=1}^m$$

has at most κ negative eigenvalues, and exactly κ negative eigenvalues for some choice of m, $\mathbf{c}_1, \ldots, \mathbf{c}_m$, and $w_1, \ldots, w_m$. In this situation, for $K(z, w)$ we will use the term *kernel* rather than *function* and speak of the *number of negative squares of the kernel* $K(z, w)$. The *number of positive squares of a kernel* $K(z, w)$ is defined accordingly. Associated to a kernel $K(z, w)$ with κ negative squares is a Pontryagin space $\mathcal{P}(K)$ of $p \times 1$ vector functions defined on Ω, which is uniquely determined by the following two properties: For every $w \in \Omega$ and $p \times 1$ vector $\mathbf{c}$, the function $\big(K_w\mathbf{c}\big)(z)$ with

$$\big(K_w\mathbf{c}\big)(z) = K(z, w)\mathbf{c}, \quad z \in \Omega,$$

belongs to $\mathcal{P}(K)$ and for every $f(z) \in \mathcal{P}(K)$,

$$\langle f, K_w \mathbf{c} \rangle_{\mathcal{P}(K)} = \mathbf{c}^* f(w).$$

It follows that the functions $\big(K_w \mathbf{c}\big)(z)$, $w \in \Omega$, $\mathbf{c} \in \mathbb{C}^p$, are dense in $\mathcal{P}(K)$ and

$$\big\langle K_w \mathbf{c}, K_z \mathbf{d} \big\rangle_{\mathcal{P}(K)} = \mathbf{d}^* K(z, w)\mathbf{c}, \quad z, w \in \Omega,\ \mathbf{c}, \mathbf{d} \in \mathbb{C}^p.$$

These two facts can be used to construct via completion the unique reproducing kernel Pontryagin space $\mathcal{P}(K)$ from a given kernel $K(z,w)$. If the kernel $K(z,w)$ has κ negative squares then $\mathrm{ind}_-(\mathcal{P}(K)) = \kappa$, where $\mathrm{ind}_-(\mathcal{P})$ is the negative index of the Pontryagin space $\mathcal{P}$. When $\kappa = 0$ the kernel is called *nonnegative* and the space $\mathcal{P}(K)$ is a Hilbert space.

We recall that any finite-dimensional Pontryagin space $\mathcal{M}$ of functions, which are defined on a set Ω, is a reproducing kernel space with kernel given by

$$K_{\mathcal{M}}(z, w) = \begin{pmatrix} f_1(z) & \cdots & f_m(z) \end{pmatrix} G^{-1} \begin{pmatrix} f_1(w) & \cdots & f_m(w) \end{pmatrix}^*, \tag{2.1}$$

where $f_1(z), \ldots, f_m(z)$ is a basis of $\mathcal{M}$ and G is the corresponding Gram matrix:

$$G = \big(g_{ij}\big)_{i,j=1}^m, \quad g_{ij} = \big\langle f_j, f_i \big\rangle_{\mathcal{M}}.$$

For the nonnegative case, this formula already appears in the work of N. Aronszajn, see, for example, [21, p. 143].

A kernel $K(z,w)$ has κ negative squares if and only if it can be written as

$$K(z, w) = K_+(z, w) + K_-(z, w), \tag{2.2}$$

where $K_+(z,w)$ and $-K_-(z,w)$ are nonnegative kernels on Ω and are such that

$$\mathcal{P}(K_+) \cap \mathcal{P}(-K_-) = \{0\}\,.$$

When $\kappa > 0$ the decomposition is not unique, but for every such decomposition,

$$\dim \mathcal{P}(-K_-) = \kappa.$$

In particular,

$$\mathcal{P}(K) = \mathcal{P}(K_+) \oplus \mathcal{P}(K_-) = \big\{ f(z) = f_+(z) + f_-(z) :\ f_\pm(z) \in \mathcal{P}(K_\pm) \big\} \tag{2.3}$$

with indefinite inner product

$$\langle f, f \rangle_{\mathcal{P}(K)} = \langle f_+, f_+ \rangle_{\mathcal{P}(K_+)} + \langle f_-, f_- \rangle_{\mathcal{P}(K_-)}, \tag{2.4}$$

see [118] and also (2.14) below for an example.

2.2. Analytic kernels and Pick matrices

In this paper we consider $p \times p$ matrix kernels $K(z,w)$ which are defined on some open subset $\Omega = \mathcal{D}$ of $\mathbb{C}$ and are analytic in z and w^* (*Bergman kernels* in W.F. Donoghue's terminology when $K(z,w)$ is nonnegative); we shall call these kernels *analytic* on $\mathcal{D}$.

Lemma 2.1. *If the $p\times p$ matrix kernel $K(z,w)$ is analytic on the open set $\mathcal{D}$ and has a finite number of negative squares, then the elements of $\mathcal{P}(K)$ are analytic $p\times 1$ vector functions on $\mathcal{D}$, and for any nonnegative integer ℓ, any point $w\in\mathcal{D}$ and any $p\times 1$ vector $\mathbf{c}$ the function $\left(\partial^\ell K_w\mathbf{c}/\partial w^{*\ell}\right)(z)$ with*

$$\left(\frac{\partial^\ell K_w\mathbf{c}}{\partial w^{*\ell}}\right)(z)=\frac{\partial^\ell K(z,w)\mathbf{c}}{\partial w^{*\ell}} \tag{2.5}$$

belongs to $\mathcal{P}(K)$ and for every $f(z)\in\mathcal{P}(K)$

$$\left\langle f,\frac{\partial^\ell K_w\mathbf{c}}{\partial w^{*\ell}}\right\rangle_{\mathcal{P}(K)}=\mathbf{c}^*f^{(\ell)}(w),\quad f(z)\in\mathcal{P}(K). \tag{2.6}$$

This fact is well known when $\kappa=0$, but a proof seems difficult to pinpoint in the literature; we refer to [13, Proposition 1.1]. W.F. Donoghue showed that the elements of the space associated to an analytic kernel are themselves analytic, see [73, p. 92] and [19, Theorem 1.1.3]. The decomposition (2.2) or [89, Theorem 2.4], which characterizes norm convergence in Pontryagin spaces by means of the indefinite inner product, allow to extend these results to the case $\kappa>0$, as we now explain. To simplify the notation we give a proof for the case $p=1$. The case $p>1$ is treated in the same way, but taking into account the "directions" $\mathbf{c}$.

Proof of Lemma 2.1. In the proof we make use of [19, pp. 4–10]. The crux of the proof is to show that in the decomposition (2.2) the functions $K_\pm(z,w)$ can be chosen analytic in z and w^*. This reduces the case $\kappa>0$ to the case of zero negative squares. Let $w_1,\ldots,w_m\in\mathcal{D}$ be such that the Hermitian $m\times m$ matrix with ij entry equal to $K(w_i,w_j)$ has κ negative eigenvalues. Since

$$K(w_i,w_j)=\langle K_{w_j},K_{w_i}\rangle_{\mathcal{P}(K)}$$

we obtain from [19, Lemma 1.1.1$'$] that there is a subspace $\mathcal{H}_-$ of the span of the functions $z\mapsto K(z,w_i)$, $i=1,\ldots,m$, which has dimension κ and is negative. Let $f_1(z),\ldots,f_\kappa(z)$ be a basis of $\mathcal{H}_-$ and denote by G the Gram matrix of this basis:

$$G=\left(g_{ij}\right)_{i,j=1}^\kappa,\quad g_{ij}=\left\langle f_j,f_i\right\rangle_{\mathcal{P}(K)}.$$

The matrix G is strictly negative and, by formula (2.1), the reproducing kernel of $\mathcal{H}_-$ is equal to

$$K_-(z,w)=\begin{pmatrix}f_1(z) & \cdots & f_\kappa(z)\end{pmatrix}G^{-1}\begin{pmatrix}f_1(w) & \cdots & f_\kappa(w)\end{pmatrix}^*.$$

By [19, p. 8], the kernel

$$K_+(z,w)=K(z,w)-K_-(z,w)$$

is nonnegative on Ω, and the span of the functions $z\mapsto K_+(z,w)$, $w\in\Omega$, is orthogonal to $\mathcal{H}_-$. Thus (2.3) and (2.4) are in force. The function $K_-(z,w)$ is analytic in z and w^* by construction. Since $K_+(z,w)$ and $-K_-(z,w)$ are nonnegative, it follows from, for example, [13, Proposition 1.1] that for $w\in\Omega$ the functions

$$z\mapsto\frac{\partial^\ell K_\pm(z,w)}{\partial w^{*\ell}}$$

belong to $\mathcal{P}(K_\pm)$. Thus the functions

$$z \mapsto \frac{\partial^\ell K(z,w)}{\partial w^{*\ell}} = \frac{\partial^\ell K_+(z,w)}{\partial w^{*\ell}} + \frac{\partial^\ell K_-(z,w)}{\partial w^{*\ell}}$$

belong to $\mathcal{P}(K)$, and for $f(z) = f_+(z) + f_-(z) \in \mathcal{P}(K)$ we have

$$\begin{aligned}\left\langle f, \frac{\partial^\ell K(\cdot,w)}{\partial w^{*\ell}} \right\rangle_{\mathcal{P}(K)} &= \left\langle f_+, \frac{\partial^\ell K_+(\cdot,w)}{\partial w^{*\ell}} \right\rangle_{\mathcal{P}(K_+)} + \left\langle f_-, \frac{\partial^\ell K_-(\cdot,w)}{\partial w^{*\ell}} \right\rangle_{\mathcal{P}(K_-)} \\ &= f_+^{(\ell)}(w) + f_-^{(\ell)}(w) = f^{(\ell)}(w). \qquad \square\end{aligned}$$

Now let $K(z,w)$ be an analytic scalar kernel on $\mathcal{D} \subset \mathbb{C}$; here $\mathcal{D}$ is always supposed to be simply connected. For $z_1 \in \mathcal{D}$ we consider the Taylor expansion

$$K(z,w) = \sum_{i,j=0}^{\infty} \gamma_{ij}(z-z_1)^i (w-z_1)^{*j}. \tag{2.7}$$

The infinite matrix $\Gamma := (\gamma_{ij})_{i,j=0}^{\infty}$ of the coefficients in (2.7) is called the *Pick matrix of the kernel* $K(z,w)$ *at* z_1; sometimes also its principal submatrices are called Pick matrices at z_1.
For a finite or infinite square matrix $A = (a_{ij})_{i,j\geq 0}$ and any integer $k \geq 1$ not exceeding the number of rows of A, by A_k we denote the principal $k \times k$ submatrix of A. Further, for a finite Hermitian matrix A, $\kappa_-(A)$ is the number of negative eigenvalues of A; if A is an infinite Hermitian matrix we set

$$\kappa_-(A) = \sup\left\{\kappa_-(A_k) \,\middle|\, k = 1, 2, \ldots\right\}.$$

We are mainly interested in situations where this number is finite. Evidently, for any integer $k \geq 1$ we have $\kappa_-(A_k) \leq \kappa_-(A_{k+1})$, if only these submatrices are defined. For a finite or infinite Hermitian matrix A by $k_0(A)$ we denote the smallest integer $k \geq 1$ for which $\det A_k \neq 0$, that is, for which A_k is invertible. In other words, if $k_0(A) = 1$ then $a_{00} \neq 0$ and if $k_0(A) > 1$ then $\det A_1 = \det A_2 = \cdots = \det A_{k_0(A)-1} = 0$, $\det A_{k_0(A)} \neq 0$.

Theorem 2.2. *Let* $K(z,w)$ *be an analytic kernel on the simply connected domain* $\mathcal{D}$ *and* $z_1 \in \mathcal{D}$. *Then the kernel* $K(z,w)$ *has* κ *negative squares if and only if for the corresponding Pick matrix* Γ *of the kernel* $K(z,w)$ *at* $z_1 \in \mathcal{D}$ *we have*

$$\kappa_-(\Gamma) = \kappa. \tag{2.8}$$

We prove this theorem only under the additional assumption that the kernel $K(z,w)$ has a finite number of negative squares, since we shall apply it only in this case, see Corollaries 4.1 and 4.7.

Proof of Theorem 2.2. The relations (2.5) and (2.6) imply for $i, j = 0, 1, \ldots$ and $z, w \in \mathcal{D}$,

$$\frac{\partial^{i+j} K(z,w)}{\partial z^i \partial w^{*j}} = \frac{\partial^{i+j}}{\partial z^i \partial w^{*j}} \langle K_w, K_z \rangle_{\mathcal{P}(K)} = \left\langle \frac{\partial^j K_w}{\partial w^{*j}}, \frac{\partial K_z}{\partial z^{*i}} \right\rangle_{\mathcal{P}(K)},$$

and for the coefficients in (2.7) we get

$$\gamma_{ij} = \frac{1}{i!j!}\left\langle \frac{\partial^j K_w}{\partial w^{*j}}, \frac{\partial^i K_z}{\partial z^{*i}} \right\rangle_{\mathcal{P}(K)}\Big|_{z=w=z_1}. \tag{2.9}$$

It follows that $\kappa_-(\Gamma_m)$ coincides with the negative index of the inner product $\langle\,\cdot\,,\,\cdot\,\rangle_{\mathcal{P}(K)}$ on the span of the elements

$$\frac{\partial^i K_w}{\partial w^{*i}}\Big|_{w=z_1}, \; i = 0, 1, \dots, m-1,$$

in $\mathcal{P}(K)$ and hence $\kappa_-(\Gamma_m) \le \kappa$. The equality follows from the fact that, in view of (2.6), $\mathcal{P}(K)$ is the closed linear span of these elements

$$\frac{\partial^i K_w}{\partial w^{*i}}\Big|_{w=z_1}, \; i = 0, 1, \dots. \qquad \square$$

If z_1 is a boundary point of $\mathcal{D}$ and there exists an m such that the limits

$$\gamma_{ij} = \lim_{z_n' \hat{\to} z_1} \frac{\partial^{i+j} K(z,w)}{\partial z^i \, \partial w^{*j}}\Big|_{z=w=z_n'}, \quad z_n' \in \mathcal{D},$$

exist for $0 \le i, j \le m-1$, then for the corresponding *Pick matrix* Γ_m *of the kernel* $K(z,w)$ *at* z_1 we have

$$\kappa_-(\Gamma_m) \le \kappa. \tag{2.10}$$

This inequality follows immediately from the fact that it holds for the corresponding Pick matrices of $K(z,w)$ at the points z_n'.

2.3. Generalized Schur functions and the spaces $\mathcal{P}(s)$

In this and the following subsection we introduce the concrete reproducing kernel Pontryagin spaces which will be used in this paper. For any integer $\kappa \ge 0$ we denote by $\mathbf{S}_\kappa$ the set of *generalized Schur functions with* κ *negative squares*. These are the functions $s(z)$ which are defined and meromorphic on $\mathbb{D}$ and for which the kernel

$$K_s(z,w) = \frac{1 - s(z)s(w)^*}{1 - zw^*}, \quad z, w \in \operatorname{hol}(s), \tag{2.11}$$

has κ negative squares on $\operatorname{hol}(s)$. In this case we also say that $s(z)$ has κ *negative squares* and write $\operatorname{sq}_-(s) = \kappa$. Furthermore, we set

$$\mathbf{S} = \bigcup_{\kappa \ge 0} \mathbf{S}_\kappa.$$

The elements of $\mathbf{S}$ are called *generalized Schur functions*.

Clearly, the kernel $K_s(z,w)$ determines the function $|s(z)|$ and hence also the function $s(z)$ up to a constant factor of modulus one. We sometimes write the kernel as

$$K_s(z,w) = \frac{\begin{pmatrix} 1 & -s(z) \end{pmatrix} J_c \begin{pmatrix} 1 & -s(w) \end{pmatrix}^*}{1 - zw^*}, \quad J_c = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

By [94, Satz 3.2], $s(z) \in \mathbf{S}_\kappa$ if and only if

$$s(z) = b(z)^{-1} s_0(z), \tag{2.12}$$

where $b(z)$ is a Blaschke product of order κ:

$$b(z) = \prod_{i=1}^{\kappa} \frac{z - \alpha_i}{1 - \alpha_i^* z} \tag{2.13}$$

with zeros $\alpha_i \in \mathbb{D}$, $i = 1, 2, \ldots, \kappa$, and $s_0(z) \in \mathbf{S}_0$ such that $s_0(\alpha_i) \neq 0$, $i = 1, 2, \ldots, \kappa$. Clearly, the points α_i are the poles of $s(z)$ in $\mathbb{D}$, and they appear in (2.13) according to their multiplicities. The decomposition

$$K_s(z,w) = \frac{1}{b(z)b(w)^*} K_{s_0}(z,w) - \frac{1}{b(z)b(w)^*} K_b(z,w) \tag{2.14}$$

is an example of a decomposition (2.2). The functions in the class $\mathbf{S}_0$ are the *Schur functions*: these are the functions which are holomorphic and bounded by 1 on $\mathbb{D}$.

For later reference we observe that

$$s(z) \in \mathbf{S} \text{ and } s(z) \text{ is rational } \Longrightarrow |s(z)| \leq 1 \quad \text{for } z \in \mathbb{T}. \tag{2.15}$$

This follows from (2.12) and the facts that $|b(z)| = 1$ if $z \in \mathbb{T}$ and that a rational Schur function does not have a pole on $\mathbb{T}$. More generally, if $s(z) \in \mathbf{S}$, then for every $\varepsilon > 0$ there is an $r \in (0,1)$ such that $|s(z)| < 1 + \varepsilon$ for all z with $r < |z| < 1$.

As mentioned in Subsection 1.2, there is a difference between the cases $\kappa = 0$ and $\kappa > 0$: When $\kappa = 0$ then the nonnegativity of the kernel $K_s(z,w)$ on an nonempty open set in $\mathbb{D}$ implies that the function $s(z)$ can be extended to an analytic function on $\mathbb{D}$. On the other hand, when $\kappa > 0$, there exist functions $s(z)$ which are not meromorphic in $\mathbb{D}$ and for which the kernel $K_s(z,w)$ has a finite number of negative squares. Such an example is the function which is zero in the whole open unit disk except at the origin, where it takes the value 1, see [19, p. 82]. Such functions were studied in [44], [45], and [46].

We note that the number of negative squares of a function $s(z) \in \mathbf{S}$ is invariant under Möbius transformations

$$\zeta(z) = \frac{z - z_1}{1 - z z_1^*}$$

of the independent variable $z \in \mathbb{D}$, where $z_1 \in \mathbb{D}$. Indeed, since

$$\frac{1 - \zeta(z)\zeta(w)^*}{1 - zw^*} = \frac{1 - |z_1|^2}{(1 - z z_1^*)(1 - w^* z_1)},$$

we have

$$\begin{aligned}
\frac{1 - s\big(\zeta(z)\big)s\big(\zeta(w)\big)^*}{1 - zw^*} &= \frac{1 - s\big(\zeta(z)\big)s\big(\zeta(w)\big)^*}{1 - \zeta(z)\zeta(w)^*} \frac{1 - \zeta(z)\zeta(w)^*}{1 - zw^*} \\
&= \frac{\sqrt{1 - |z_1|^2}}{1 - z z_1^*} \frac{1 - s\big(\zeta(z)\big)s\big(\zeta(w)\big)^*}{1 - \zeta(z)\zeta(w)^*} \frac{\sqrt{1 - |z_1|^2}}{1 - w^* z_1},
\end{aligned}$$

and hence $\mathrm{sq}_-(s\circ\zeta)=\mathrm{sq}_-(s)$. Similarly, if

$$\Theta=\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}$$

is a J_c-unitary constant 2×2 matrix, then the functions $s(z)$ and $\mathcal{T}_\Theta(s(z))$ have the same number of negative squares because

$$\frac{1-\mathcal{T}_\Theta\big(s(z)\big)\mathcal{T}_\Theta\big(s(w)\big)^*}{1-zw^*}=\frac{1}{\gamma s(z)+\delta}\frac{1-s(z)s(w)^*}{1-zw^*}\frac{1}{\big(\gamma s(w)+\delta\big)^*}.$$

If $z_1\in\mathbb{D}$, by $\mathbf{S}_\kappa^{z_1}$ ($\mathbf{S}^{z_1}$, respectively) we denote the functions $s(z)$ from $\mathbf{S}_\kappa$ ($\mathbf{S}$, respectively) which are holomorphic at z_1.

If $z_1\in\mathbb{T}$ we consider also functions which have an asymptotic expansion of the form (recall that $z\hat{\to}z_1$ means that z tends nontangentially from $\mathbb{D}$ to $z_1\in\mathbb{T}$)

$$s(z)=\tau_0+\sum_{i=1}^{2p-1}\tau_j(z-z_1)^i+\mathrm{O}\left((z-z_1)^{2p}\right),\quad z\hat{\to}z_1, \tag{2.16}$$

where the coefficients τ_i, $i=0,1,\dots,2p-1$, satisfy the following assumptions:

(1) $|\tau_0|=1$;
(2) at least one of the numbers $\tau_1,\dots,\tau_p$ is not 0;
(3) the matrix

$$\widehat{P}=\widehat{T}\widehat{B}Q \tag{2.17}$$

with

$$\widehat{T}=\big(t_{ij}\big)_{i,j=0}^{p-1},\quad t_{ij}=\tau_{i+j+1},$$

$$\widehat{B}=\big(b_{ij}\big)_{i,j=0}^{p-1},\quad b_{ij}=z_1^{p+i-j}\binom{p-1-j}{i}(-1)^{p-1-j},$$

and

$$Q=\big(c_{ij}\big)_{i,j=0}^{p-1},\quad c_{ij}=\tau^*_{i+j-(p-1)},$$

is Hermitian.

Here $\widehat{B}$ is a left upper and Q is a right lower triangular matrix. The assumptions (1) and (3) are necessary in order to assure that the asymptotic expansion (2.16) of the function $s(z)$ yields an asymptotic expansion of the kernel $K_s(z,w)$, see (4.15) below. The assumption (2) implies that at least one of the Pick matrices of the kernel $K_s(z,w)$ is invertible, see Theorem 4.6); in the present paper we are interested only in this situation. The set of functions from $\mathbf{S}_\kappa$ ($\mathbf{S}$, respectively) which have an asymptotic expansion (2.16) at $z_1\in\mathbb{T}$ with the properties (1), (2), and (3) we denote by $\mathbf{S}_\kappa^{z_1;2p}$ ($\mathbf{S}^{z_1;2p}$, respectively).

For $s(z)\in\mathbf{S}$ and the corresponding Schur kernel $K_s(z,w)$ from (2.11):

$$K_s(z,w)=\frac{1-s(z)s(w)^*}{1-zw^*},\quad z,w\in\mathrm{hol}\,(s),$$

the reproducing kernel Pontryagin space $\mathcal{P}(K_s)$ is denoted by $\mathcal{P}(s)$.

For a function $s(z) \in \mathbf{S}_\kappa$ there exists a realization of the form

$$s(z) = \gamma + b_c(z)\big\langle\big(1 - b_c(z)T\big)^{-1}u, v\big\rangle, \quad b_c(z) = \frac{z - z_1}{1 - zz_1^*}, \tag{2.18}$$

with a complex number γ: $\gamma = s(z_1)$, a bounded operator T in some Pontryagin space $(\mathcal{P}, \langle\,\cdot\,,\,\cdot\,\rangle)$, and elements u and $v \in \mathcal{P}$. With the entries of (2.18) we form the operator matrix

$$\mathcal{V} = \begin{pmatrix} T & u \\ \langle\,\cdot\,, v\rangle & \gamma \end{pmatrix} : \begin{pmatrix} \mathcal{P} \\ \mathbb{C} \end{pmatrix} \to \begin{pmatrix} \mathcal{P} \\ \mathbb{C} \end{pmatrix}. \tag{2.19}$$

Then the following statements are equivalent, see [19]:

(a) $s(z) \in \mathbf{S}^{z_1}$.

(b) $s(z)$ admits the realization (2.18) such that the operator matrix $\mathcal{V}$ in (2.19) is isometric in $\begin{pmatrix} \mathcal{P} \\ \mathbb{C} \end{pmatrix}$ and *closely innerconnected*, that is,

$$\mathcal{P} = \overline{\operatorname{span}}\,\big\{T^j v \,\big|\, j = 0, 1, \dots\big\}.$$

(c) $s(z)$ admits the realization (2.18) such that the operator matrix $\mathcal{V}$ in (2.19) is coisometric (that is, its adjoint is isometric) in $\begin{pmatrix} \mathcal{P} \\ \mathbb{C} \end{pmatrix}$ and *closely outerconnected*, which means that

$$\mathcal{P} = \overline{\operatorname{span}}\,\big\{T^{*i} v \,\big|\, i = 0, 1, \dots\big\}.$$

(d) $s(z)$ admits the realization (2.18) such that the operator matrix $\mathcal{V}$ in (2.19) is unitary in $\begin{pmatrix} \mathcal{P} \\ \mathbb{C} \end{pmatrix}$ and *closely connected*, that is,

$$\mathcal{P} = \overline{\operatorname{span}}\,\big\{T^{*i} v, T^j u \,\big|\, i, j = 0, 1, \dots\big\}.$$

The realizations in (b), (c), and (d) are unique up to isomorphisms (unitary equivalence) of the spaces and of the operators and elements. The connectedness condition in (b), (c), and (d) implies that $\operatorname{sq}_-(s) = \operatorname{ind}_-(\mathcal{P})$. For example, the closely outerconnected coisometric realization in (b) with $z_1 = 0$ can be chosen as follows: $\mathcal{P}$ is the reproducing kernel space $\mathcal{P}(s)$, T is the operator

$$\big(Tf\big)(z) = \frac{1}{z}\big(f(z) - f(0)\big), \quad f(z) \in \mathcal{P}(s),$$

and u and v are the elements

$$u(z) = \frac{1}{z}\big(s(z) - s(0)\big), \quad v(z) = K_s(z, 0).$$

This is the *backward-shift realization* but here the emphasis is on the metric structure of the realization (that is, the coisometry property) rather than the minimality, see [37] and [85].

2.4. Generalized Nevanlinna functions and the spaces $\mathcal{L}(n)$

For any integer $\kappa \geq 0$ we denote by $\mathbf{N}_\kappa$ the set of *generalized Nevanlinna functions with κ negative squares.* These are the meromorphic functions $n(z)$ on $\mathbb{C}^+$ for which the kernel

$$L_n(z,w) = \frac{n(z) - n(w)^*}{z - w^*}, \quad z, w \in \operatorname{hol}(n), \tag{2.20}$$

has κ negative squares on $\operatorname{hol}(n)$. In this case we also say that $n(z)$ has κ *negative squares* and we write $\operatorname{sq}_-(n) = \kappa$. We sometimes write the kernel as

$$L_n(z,w) = \frac{\begin{pmatrix} 1 & -n(z) \end{pmatrix} J_\ell \begin{pmatrix} 1 & -n(w) \end{pmatrix}^*}{z - w^*}, \quad J_\ell = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

For $\kappa = 0$ the class $\mathbf{N}_0$ consists of all Nevanlinna functions $n(z)$: these are the functions which are holomorphic on $\mathbb{C}^+$ and satisfy $\operatorname{Im} n(z) \geq 0$ for $z \in \mathbb{C}^+$. By a result of [70], $n(z) \in \mathbf{N}_\kappa$ admits the representation

$$n(z) = \frac{\prod_{i=1}^{\kappa_1}(z - \alpha_i)(z - \alpha_i^*)}{\prod_{j=1}^{\kappa_2}(z - \beta_j)(z - \beta_j^*)}\, n_0(z),$$

where κ_1 and κ_2 are integers ≥ 0 with $\kappa = \max(\kappa_1, \kappa_2)$, α_i and β_j are points from $\mathbb{C}^+ \cup \mathbb{R}$ such that $\alpha_i \neq \beta_j$, and $n_0(z) \in \mathbf{N}_0$. A function $n(z) \in \mathbf{N}_\kappa$ is always considered to be extended to the open lower half-plane by symmetry:

$$n(z^*) = n(z)^*, \quad z \in \operatorname{hol}(n), \tag{2.21}$$

and to those points of the real axis into which it can be continued analytically. The kernel $L_n(z,w)$ extended by (2.20) to all these points if $w \neq z^*$ and set equal to $n'(z)$ when $w = z^*$ still has κ negative squares. Accordingly, $\operatorname{hol}(n)$ now stands for the largest set on which $n(z)$ is holomorphic. We set

$$\mathbf{N} = \bigcup_{\kappa \geq 0} \mathbf{N}_\kappa.$$

The elements of $\mathbf{N}$ are called *generalized Nevanlinna functions.*

If $z_1 \in \mathbb{C}^+$, by $\mathbf{N}_\kappa^{z_1}$ ($\mathbf{N}^{z_1}$, respectively) we denote the functions $n(z)$ from $\mathbf{N}_\kappa$ ($\mathbf{N}$, respectively) which are holomorphic at z_1.

We consider also functions $n(z) \in \mathbf{N}$ which have for some integer $p \geq 1$ an asymptotic expansion at $z_1 = \infty$ of the form

$$n(z) = -\frac{\mu_0}{z} - \frac{\mu_1}{z^2} - \cdots - \frac{\mu_{2p-1}}{z^{2p}} + \mathrm{O}\left(\frac{1}{z^{2p+1}}\right), \quad z = \mathrm{i}y,\ y \uparrow \infty,$$

where

(1) $\mu_j \in \mathbb{R}$, $j = 0, 1, \ldots, 2p-1$, and
(2) at least one of the coefficients $\mu_0, \mu_1, \ldots, \mu_{p-1}$ is not equal to 0.

The fact that $\lim_{y\uparrow\infty} n(iy) = 0$, and items (1) and (2) are the analogs of items (1), (3), and (2), respectively, in Subsection 2.3. Here the reality of the coefficients is needed in order to assure that the asymptotic expansion of the function $n(z)$

implies an asymptotic expansion of the Nevanlinna kernel (2.20). The asymptotic expansion above is equivalent to the asymptotic expansion

$$n(z) = -\frac{\mu_0}{z} - \frac{\mu_1}{z^2} - \cdots - \frac{\mu_{2p-1}}{z^{2p}} - \frac{\mu_{2p}}{z^{2p+1}} + \mathrm{o}\left(\frac{1}{z^{2p+1}}\right), \quad z = \mathrm{i}y, \ y \uparrow \infty,$$

for some additional real number μ_{2p}, see [94, Bemerkung 1.11]). The set of all functions $n(z) \in \mathbf{N}_\kappa$ ($n(z) \in \mathbf{N}$, respectively) which admit expansions of the above forms with properties (1) and (2) is denoted by $\mathbf{N}_\kappa^{\infty;2p}$ ($\mathbf{N}^{\infty;2p}$, respectively). Note that any rational function of the class $\mathbf{N}$ which vanishes at ∞ belongs to $\mathbf{N}^{\infty;2p}$ for all sufficiently large integers p.

If $n(z) \in \mathbf{N}$ and

$$L_n(z,w) = \frac{n(z) - n(w)^*}{z - w^*}$$

is the kernel from (2.20), then the reproducing kernel space $\mathcal{P}(L_n)$ is denoted by $\mathcal{L}(n)$.

A function $n(z)$ is a generalized Nevanlinna function if and only it admits a representation of the form

$$n(z) = n(z_0)^* + (z - z_0^*)\big\langle\big(I + (z - z_0)(A - z)^{-1}\big)u_0, u_0\big\rangle_{\mathcal{P}}, \tag{2.22}$$

where $\mathcal{P}$ is a Pontryagin space, A is a self-adjoint relation in $\mathcal{P}$ with a nonempty resolvent set $\rho(A)$, $z_0 \in \rho(A)$, and $u_0 \in \mathcal{P}$. The representation is called a *self-adjoint realization centered at* z_0. The realization can always be chosen such that

$$\overline{\mathrm{span}}\,\big\{\big(I + (z - z_0)(A - z)^{-1}\big)u_0 \,\big|\, z \in (\mathbb{C} \setminus \mathbb{R}) \cap \rho(A)\big\} = \mathcal{P}.$$

If this holds we say that the realization is *minimal.* Minimality implies that the self-adjoint realization of $n(z)$ is unique up to unitary equivalence, and also that $\mathrm{hol}\,(n) = \rho(A)$ and $\mathrm{sq}_-(n) = \mathrm{ind}_-(\mathcal{P})$, see [69].

An example of a minimal self-adjoint realization of a generalized Nevanlinna function $n(z)$ is given by (2.22), where

(a) $\mathcal{P} = \mathcal{L}(n)$, the reproducing kernel Pontryagin space with kernel $L_n(z,w)$, whose elements are locally holomorphic functions $f(\zeta)$ on $\mathrm{hol}\,(n)$,
(b) A is the self-adjoint relation in $\mathcal{L}(n)$ with resolvent given by $(A - z)^{-1} = R_z$, the difference-quotient operator defined by

$$(R_z f)(\zeta) = \begin{cases} \dfrac{f(\zeta) - f(z)}{\zeta - z}, & \zeta \neq z, \\ f'(z), & \zeta = z, \end{cases} \qquad f(\zeta) \in \mathcal{L}(n),$$

(c) $u_0(\zeta) = cL_n(\zeta, z_0^*)$ or $u_0(\zeta) = cL_n(\zeta, z_0)$ with $c \in \mathbb{C}$ and $|c| = 1$.

For a proof and further details related to this example, we refer to [71, Theorem 2.1].

2.5. Additional remarks and references

The first comprehensive paper on the theory of reproducing kernel spaces is the paper [29] by N. Aronszajn's which appeared in 1943. We refer to [30] and [108] for accounts on the theory of reproducing kernel Hilbert spaces and to [89] and [33] for more information on Pontryagin spaces. Reproducing kernel Pontryagin spaces (and reproducing kernel Krein spaces) appear in the general theory developed by L. Schwartz in [118] but have been first explicitly studied by P. Sorjonen [120].

One of the first examples of kernels with a finite number of negative squares was considered by M.G. Krein in [93]: He studied continuous functions $f(t)$ on $\mathbb{R}$ for which the kernel

$$K_f(s,t) = f(s-t)$$

has this property. Nonnegative kernels ($\kappa = 0$) were first defined by J. Mercer in the setting of integral equations, see [104]. For a historical discussion, see, for example, [40, p. 84].

When $\kappa = 0$ it is well known, see [118] and [108], that there is a one-to-one correspondence between nonnegative $p \times p$ matrix kernels and reproducing kernel Hilbert spaces of $p \times 1$ vector functions defined in Ω. This result was extended by P. Sorjonen [120] and L. Schwartz [118] to a one-to-one correspondence between kernels with κ negative squares and reproducing kernel Pontryagin spaces.

If $s(z) \in \mathbf{S}_0$, the space $\mathcal{P}(s)$ is contractively included in the Hardy space $\mathbf{H}_2$ on $\mathbb{D}$: this means that $\mathcal{P}(s) \subset \mathbf{H}_2$ and that the inclusion map is a contraction, see [50]. If, moreover, $s(z)$ is *inner*, that is, its boundary values on $\mathbb{T}$ have modulus 1 almost everywhere, then

$$\mathcal{P}(s) = \mathbf{H}_2 \ominus s\mathbf{H}_2,$$

see, for instance, [21, Theorem 3.5]. The theory of reproducing kernel Pontryagin spaces of the form $\mathcal{P}(s)$ can be found in [19], see also [12].

3. Some classes of rational matrix functions

In this section we review the main features of the theory of rational functions needed in the sequel. Although there we mostly deal with rational scalar or 2×2 matrix functions, we start with the case of $p \times p$ matrix functions for any integer $p \geq 1$. In the general setting discussed in the Subsection 1.4, the results we present correspond to the choices of $a(z)$ and $b(z)$ for the open unit disk and the open upper half-plane and to $F(z)$ of the form

$$F(z) = C(zI - A)^{-1} \quad \text{or} \quad F(z) = C(I - zA)^{-1}.$$

We often use straightforward arguments and not the general results of [23], [24], and [25].

3.1. Realizations and McMillan degree of rational matrix functions

Recall that any rational matrix function $R(z)$ which is analytic at zero can be written as

$$R(z) = D + zC(I - zA)^{-1}B, \tag{3.1}$$

where A, B, C, and D are matrices of appropriate sizes; evidently, $D = R(0)$. If $R(z)$ a rational matrix function which is analytic at ∞, then it can be written as

$$R(z) = D + C(zI - A)^{-1}B; \tag{3.2}$$

now $D = R(\infty)$. The realization (3.1) or (3.2) is called *minimal* if the size of the square matrix A is as small as possible. Equivalently, see [37], it is minimal if it is both *observable*, which means that

$$\bigcap_{\ell=0}^{\infty} \ker CA^{\ell} = \{0\}\,,$$

and *controllable*, that is, if A is an $m \times m$ matrix, then

$$\bigcup_{\ell=0}^{\infty} \operatorname{ran} A^{\ell}B = \mathbb{C}^m.$$

Minimal realizations are unique up to a similarity matrix: If, for example, (3.1) is a minimal realization of $R(z)$, then any other minimal realization of $R(z) = D + C_1(zI - A_1)^{-1}B_1$ is related to the realization (3.1) by

$$\begin{pmatrix} A_1 & B_1 \\ C_1 & D \end{pmatrix} = \begin{pmatrix} S & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} S^{-1} & 0 \\ 0 & I \end{pmatrix}$$

for some uniquely defined invertible matrix S.

The size of the matrix A in a minimal representation (3.1) or (3.2) is called the *McMillan degree of* $R(z)$ and denoted by $\deg R$. In fact, the original definition of the McMillan degree uses the local degrees of the poles of the function: If $R(z)$ has a pole at w with principal part

$$\sum\nolimits_{j=1}^{n} \frac{R_j}{(z-w)^j},$$

then the *local degree of* $R(z)$ *at* w is defined by

$$\deg_w R = \operatorname{rank} \begin{pmatrix} R_n & 0 & \cdots & 0 & 0 \\ R_{n-1} & R_n & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ R_2 & R_3 & \cdots & R_n & 0 \\ R_1 & R_2 & \cdots & R_{n-1} & R_n \end{pmatrix}. \tag{3.3}$$

The *local degree at* ∞ is by definition the local degree at $z = 0$ of the function $R(1/z)$. The McMillan degree of $R(z)$ is equal to the sum of the local degrees at all poles $w \in \mathbb{C} \cup \{\infty\}$. In particular, if $R(z)$ has a single pole at w (as will often be the case in the present work) the McMillan degree of $R(z)$ is given by (3.3). We refer to [37, Section 4.1] and [90] for more information.

3.2. J-unitary matrix functions and the spaces $\mathcal{P}(\Theta)$: the line case

We begin with a characterization of $\mathcal{P}(\Theta)$ spaces. Let J be any $p \times p$ signature matrix, that is, J is self-adjoint and $J^2 = I_p$. A rational $p \times p$ matrix function $\Theta(z)$ is called *J-unitary on the line*, if

$$\Theta(x)^* J\Theta(x) = J, \quad x \in \mathbb{R} \cap \operatorname{hol}(\Theta),$$

and *J-unitary on the circle* if

$$\Theta(e^{\mathrm{i}t})^* J\Theta(e^{\mathrm{i}t}) = J, \quad t \in [0, 2\pi),\, e^{\mathrm{i}t} \in \operatorname{hol}(\Theta).$$

If $\Theta(z)$ is rational and J-unitary on the circle, the kernel

$$K_\Theta(z, w) = \frac{J - \Theta(z)J\Theta(w)^*}{1 - zw^*} \tag{3.4}$$

has a finite number of positive and of negative squares and we denote by $\mathcal{P}(\Theta)$ the corresponding reproducing kernel Pontryagin space $\mathcal{P}(K_\Theta)$; similarly, if $\Theta(z)$ is rational and J-unitary on the line the kernel

$$K_\Theta(z, w) = \frac{J - \Theta(z)J\Theta(w)^*}{-\mathrm{i}(z - w^*)} \tag{3.5}$$

has a finite number of positive and of negative squares and the corresponding reproducing kernel Pontryagin space is also denoted by $\mathcal{P}(\Theta)$. Evidently, both spaces $\mathcal{P}(\Theta)$ are finite-dimensional, see also [21, Theorem 6.9].

Both kernels could be treated in a unified way using the framework of kernels with denominator of the form $a(z)a(w)^* - b(z)b(w)^*$ as mentioned in Subsection 1.4. We prefer, however, to consider both cases separately, and begin with the line case. The following theorem characterizes the J-unitarity on the line of a rational matrix function $\Theta(z)$ in terms of a minimal realization of $\Theta(z)$.

Recall that R_ζ denotes the *backward-shift* (or the *difference-quotient*) operator based on the point $\zeta \in \mathbb{C}$:

$$\big(R_\zeta f\big)(z) = \frac{f(z) - f(\zeta)}{z - \zeta}.$$

A set $\mathcal{M}$ of analytic vector functions on an open set Ω is called *backward-shift invariant* if for all $\zeta \in \Omega$ we have $R_\zeta \mathcal{M} \subset \mathcal{M}$.

Theorem 3.1. *Let $\mathcal{P}$ be a finite-dimensional reproducing kernel Pontryagin space of analytic $p \times 1$ vector functions on an open set Ω which is symmetric with respect to the real line. Then $\mathcal{P}$ is a $\mathcal{P}(\Theta)$ space with reproducing kernel $K_\Theta(z, w)$ of the form* (3.5) *if and only if the following conditions are satisfied.*

(a) *$\mathcal{P}$ is backward-shift invariant.*
(b) *For every ζ, $\omega \in \Omega$ and $f(z)$, $g(z) \in \mathcal{P}$ the de Branges identity holds:*

$$\big\langle R_\zeta f, g\big\rangle_{\mathcal{P}} - \big\langle f, R_\omega g\big\rangle_{\mathcal{P}} - (\zeta - \omega^*)\big\langle R_\zeta f, R_\omega g\big\rangle_{\mathcal{P}} = \mathrm{i}\, g(\omega)^* J f(\zeta). \tag{3.6}$$

In this case $\dim \mathcal{P} = \deg \Theta$.

The identity (3.6) first appears in [47]. A proof of the if and only if statement in this theorem can be found in [21] and a proof of the last equality in [26]. The finite-dimensionality and the backward-shift invariance of $\mathcal{P}$ force the elements of $\mathcal{P}$ to be rational: A basis of $\mathcal{P}$ is given by the columns of a matrix function of the form

$$F(z) = C(T - zA)^{-1}.$$

If the elements of $\mathcal{P}$ are analytic in a neighborhood of the origin one can choose $T = I$, that is, $F(z) = C(I - zA)^{-1}$. Since $R_0F(z) = C(I - zA)^{-1}A$, the choice $\zeta = \omega = 0$ in (3.6) shows that the Gram matrix G associated with $F(z)$:

$$\big\langle F\mathbf{c}, F\mathbf{d}\big\rangle_{\mathcal{P}} = \mathbf{d}^*G\mathbf{c}, \quad \mathbf{c}, \mathbf{d} \in \mathbb{C}^m,$$

satisfies the Lyapunov equation

$$GA - A^*G = \mathrm{i}C^*JC. \tag{3.7}$$

It follows from Theorem 3.1 that $\Theta(z)$ is rational and J-unitary on the real line. We now study these functions using realization theory.

Theorem 3.2. *Let $\Theta(z)$ be a $p \times p$ matrix function which is analytic at infinity and let $\Theta(z) = D + C(zI - A)^{-1}B$ be a minimal realization of $\Theta(z)$. Then $\Theta(z)$ is J-unitary on $\mathbb{R}$ if and only if the following conditions are satisfied.*

(a) *The matrix D is J-unitary: $DJD^* = J$.*

(b) *There exists a unique Hermitian invertible matrix G such that*

$$GA - A^*G = -\mathrm{i}C^*JC, \qquad B = -\mathrm{i}G^{-1}C^*JD. \tag{3.8}$$

If (a) *and* (b) *hold, then $\Theta(z)$ can be written as*

$$\Theta(z) = \big(I_p - \mathrm{i}C(zI - A)^{-1}G^{-1}C^*J\big)D, \tag{3.9}$$

for $z, w \in \mathrm{hol}\ (\Theta)$ we have

$$K_\Theta(z, w) = \frac{J - \Theta(z)J\Theta(w)^*}{-\mathrm{i}(z - w^*)} = C(zI - A)^{-1}G^{-1}(wI - A)^{-*}C^*, \tag{3.10}$$

and the space $\mathcal{P}(\Theta)$ is spanned by the columns of the matrix function

$$F(z) = C(zI - A)^{-1}.$$

The matrix G is called the *associated Hermitian matrix* for the given realization. It is invertible, and its numbers of negative and of positive eigenvalues are equal to the numbers of negative and positive squares of the kernel (3.5). The latter follows from the formula (3.10). We outline the proof of Theorem 3.2 as an illustration of the state space method; for more information, see [26, Theorem 2.1], where functions are considered, which are J-unitary on the imaginary axis rather than on the real axis.

Proof of Theorem 3.2. We first rewrite the J-unitarity of $\Theta(z)$ on the real line as

$$\Theta(z) = J\Theta(z^*)^{-*}J. \tag{3.11}$$

By analyticity, this equality holds for all complex numbers z, with the possible exception of finitely many. Let $\Theta(z) = D + C(zI - A)^{-1}B$ be a minimal realization of $\Theta(z)$. Since $\Theta(\infty)$ is J-unitary we have $D^*JD = J$ and, in particular, D is invertible. A minimal realization of $\Theta(z)^{-1}$ is given by

$$\Theta(z)^{-1} = D^{-1} - D^{-1}C(zI - zA^{\times})^{-1}BD^{-1}, \quad \text{with} \quad A^{\times} = A - BD^{-1}C,$$

see [37, pp. 6,7]. Thus (3.11) can be rewritten as

$$D + C(zI - A)^{-1}B = JD^{-*}J - JD^{-*}B^*\big(zI - (A^{\times})^*\big)^{-1}C^*D^{-*}J.$$

This is an equality between two minimal realizations of a given rational function and hence there exists a unique matrix S such that

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} S^{-1} & 0 \\ 0 & I_p \end{pmatrix} \begin{pmatrix} A^* - C^*D^{-*}B^* & C^*D^{-*}J \\ -JD^{-*}B^* & JD^{-*}J \end{pmatrix} \begin{pmatrix} S & 0 \\ 0 & I_p \end{pmatrix}.$$

This equation is equivalent to the J-unitarity of D together with the equations

$$\begin{aligned} SA - A^*S &= -C^*D^{-*}B^*S, \\ SB &= C^*D^{-*}J, \\ C &= -JD^{-*}B^*S. \end{aligned}$$

The first two equations lead to

$$SA - A^*S = C^*JC.$$

Both S and $-S^*$ are solution of the above equations, and hence, since S is unique, $S = -S^*$. Setting $S = \mathrm{i}\,G$, we obtain $G = G^*$, the equalities (3.8) and the equality (3.9).

To prove the converse statement we prove (3.10) using (3.9). We have

$$\begin{aligned} &\Theta(z)J\Theta(w)^* - J \\ &\quad = \big(I_p - \mathrm{i}\,C(zI - A)^{-1}G^{-1}C^*J\big)J\big(I_p + \mathrm{i}\,JCG^{-1}(wI - A)^{-*}C^*\big) - J \\ &\quad = -\mathrm{i}\,C(zI - A)^{-1}G^{-1}C^* + \mathrm{i}\,CG^{-1}(wI - A)^{-*}C^* \\ &\qquad - \mathrm{i}\,C(zI - A)^{-1}G^{-1}\mathrm{i}\,C^*JCG^{-1}(wI - A)^{-*}C^* \\ &\quad = C(zI - A)^{-1}\left\{-\mathrm{i}\,G^{-1}(w^*I - A^*) + \mathrm{i}\,(zI - A)G^{-1}\right. \\ &\qquad \left. + G^{-1}C^*JCG^{-1}\right\}(wI - A)^{-*}C^*. \end{aligned}$$

By (3.8), the sum insides the curly brackets is equal to

$$-\mathrm{i}\,w^*G^{-1} + \mathrm{i}\,G^{-1}A^* + \mathrm{i}\,zG^{-1} + \mathrm{i}\,AG^{-1} - \mathrm{i}G^{-1}(GA - A^*G)G^{-1} = \mathrm{i}\,(z - w^*)G^{-1},$$

and equation (3.10) follows. That $\mathcal{P}(\Theta)$ is spanned by the columns of $F(z)$ follows from (3.10) and the minimality of the realization of $\Theta(z)$. $\square$

In Section 8 we will need the analog of Theorem 3.2 for spaces of polynomials (which in particular are analytic at the origin but not at infinity). Note that the equations in (3.12) and (3.13) below differ by a minus sign from their counterparts (3.8) and (3.9) above.

Theorem 3.3. *Let $\Theta(z)$ be a $p \times p$ matrix function which is analytic at the origin and let $\Theta(z) = D + zC(I - zA)^{-1}B$ be a minimal realization of $\Theta(z)$. Then $\Theta(z)$ is J-unitary on $\mathbb{R}$ if and only if the following conditions are satisfied.*

(a) *The matrix D is J-unitary*: $DJD^* = J$.

(b) *There exists a unique Hermitian invertible matrix G such that*

$$GA - A^*G = \mathrm{i}C^*JC, \quad B = \mathrm{i}G^{-1}C^*JD. \tag{3.12}$$

In this case, $\Theta(z)$ can be written as

$$\Theta(z) = \big(I_p + z\mathrm{i}C(I - zA)^{-1}G^{-1}C^*J\big)D, \tag{3.13}$$

for $z, w \in \mathrm{hol}\,(\Theta)$ we have

$$K_\Theta(z,w) = \frac{J - \Theta(z)J\Theta(w)^*}{-\mathrm{i}(z - w^*)} = C(I - zA)^{-1}G^{-1}(I - wA)^{-*}C^*, \tag{3.14}$$

and the space $\mathcal{P}(\Theta)$ is spanned by the columns of $F(z) = C(I - zA)^{-1}$.

The proof is a direct consequence of Theorem 3.2. Indeed, consider $\Psi(z) = \Theta(1/z)$. It is analytic at infinity and admits the minimal realization

$$\Psi(z) = D + C(zI - A)^{-1}B.$$

Furthermore, $\Psi(z)$ and $\Theta(z)$ are simultaneously J-unitary on the real line. If we apply Theorem 3.2 to $\Psi(z)$, we obtain an invertible Hermitian matrix G' such that both equalities in (3.8) hold. Replacing z, w, and Θ in (3.10) by $1/z$, $1/w$, and Ψ we obtain:

$$K_\Theta(z,w) = -C(I - zA)^{-1}(G')^{-1}(I - wA)^{-*}C^*.$$

It remains to set $G = -G'$. Then, evidently, the number of negative and positive squares of the kernel $K_\Theta(z,w)$ is equal to the number of negative and positive eigenvalues of the matrix G.

The next two theorems are special cases of Theorem 1.1. The first theorem is also a consequence of formula (3.10). It concerns spaces spanned by functions which are holomorphic at ∞.

Theorem 3.4. *Let (C, A) be an observable pair of matrices of sizes $p \times m$ and $m \times m$ respectively, denote by $\mathcal{M}$ the space spanned by the columns of the $p \times m$ matrix function $F(z) = C(zI - A)^{-1}$, and let G be a nonsingular Hermitian $m \times m$ matrix which defines the inner product* (1.15):

$$\big\langle F\mathbf{c}, F\mathbf{d}\big\rangle_{\mathcal{M}} = \mathbf{d}^*G\mathbf{c}, \quad \mathbf{c}, \mathbf{d} \in \mathbb{C}^m.$$

Then $\mathcal{M}$ is a $\mathcal{P}(\Theta)$ space with kernel $K_\Theta(z,w)$ of the form (3.5) *if and only if G is a solution of the Lyapunov equation in* (3.8)*. In this case, a possible choice of $\Theta(z)$ is given by the formula*

$$\Theta_{z_0}(z) = I_p + \mathrm{i}(z - z_0)C(zI - A)^{-1}G^{-1}(z_0I - A)^{-*}C^*J, \tag{3.15}$$

where $z_0 \in \mathrm{hol}\,(\Theta) \cap \mathbb{R}$. Any other choice of $\Theta(z)$ differs from $\Theta_{z_0}(z)$ by a J-unitary constant factor on the right.

Letting $z_0 \to \infty$ we obtain from (3.15) formula (3.9) with $D = I_p$.

Similarly, the next theorem is also a consequence of (3.14) and Theorem 3.3. Its formulation is almost the same as the one of the previous theorem except that, since now we consider spaces of functions which are holomorphic at $z = 0$, the Lyapunov equation in (3.8) is replaced by the Lyapunov equation (3.7), which differs from it by a minus sign.

Theorem 3.5. *Let (C, A) be an observable pair of matrices of sizes $p \times m$ and $m \times m$ respectively, denote by $\mathcal{M}$ the space spanned by the columns of the $p \times m$ matrix function $F(z) = C(I - zA)^{-1}$, and let G be a nonsingular Hermitian $m \times m$ matrix which defines the inner product* (1.15)*:*

$$\langle F\mathbf{c}, F\mathbf{d}\rangle_{\mathcal{M}} = \mathbf{d}^* G\mathbf{c}, \quad \mathbf{c}, \mathbf{d} \in \mathbb{C}^m.$$

Then $\mathcal{M}$ is a $\mathcal{P}(\Theta)$ space with kernel $K_\Theta(z, w)$ of the form (3.5) *if and only if G is a solution of the Lyapunov equation* (3.7)*. In this case, a possible choice of $\Theta(z)$ is given by the formula*

$$\Theta_{z_0}(z) = I_p + \mathrm{i}(z - z_0)C(I - zA)^{-1}G^{-1}(I - z_0A)^{-*}C^*J, \tag{3.16}$$

where $z_0 \in \mathrm{hol}\,(\Theta) \cap \mathbb{R}$. Any other choice of $\Theta(z)$ differs from $\Theta_{z_0}(z)$ by a J-unitary constant factor on the right.

If we set $z_0 = 0$ we obtain from (3.16) formula (3.13) with $D = I_p$.

The following theorem will be used to prove factorization results in Subsection 3.4.

Theorem 3.6. *Let $\Theta(z)$ be a rational $p \times p$ matrix function which is analytic at infinity and J-unitary on $\mathbb{R}$ and let $\Theta(z) = D + C(zI - A)^{-1}B$ be a minimal realization of $\Theta(z)$. Then*

$$\det \Theta(z) = \frac{\det(zI - A^*)}{\det(zI - A)} \det D. \tag{3.17}$$

In particular, if $\Theta(z)$ has only one pole $w \in \mathbb{C}$, then

$$\det \Theta(z) = c\left(\frac{z - w^*}{z - w}\right)^{\deg\ \Theta}$$

for some unimodular constant c.

Proof. By Theorem 3.2, $\det D \neq 0$, and thus we have

$$\begin{aligned}
\det \Theta(z) &= \det\,\left(I + C(zI - A)^{-1}BD^{-1}\right)\,\det D \\
&= \det\,\left(I + (zI - A)^{-1}BD^{-1}C\right)\,\det D \\
&= \det\,(zI - A)^{-1}\,\det\,\left(zI - A + BD^{-1}C\right)\,\det D \\
&= \frac{\det\,(zI - A^\times)}{\det\,(zI - A)}\,\det D.
\end{aligned}$$

In view of (3.8),

$$\begin{aligned} A^{\times} &= A - BD^{-1}C \\ &= A + iG^{-1}C^*JDD^{-1}C \\ &= A + G^{-1}(GA - A^*G) \\ &= G^{-1}A^*G, \end{aligned}$$

and hence

$$\det(zI - A^{\times}) = \det(zI - A^*),$$

which proves (3.17). To prove the second statement, it suffices to note that the minimality implies that if $\Theta(z)$ has only one pole in w then A is similar to a direct sum of Jordan blocks all based on the same point and the size of A is the degree of $\Theta(z)$. □

3.3. J-unitary matrix functions and the spaces $\mathcal{P}(\Theta)$: the circle case

We now turn to the characterization of a reproducing kernel Pontryagin space as a $\mathcal{P}(\Theta)$ space in the circle case.

Theorem 3.7. *Let $\mathcal{P}$ be a finite-dimensional reproducing kernel space of analytic vector functions on an open set Ω, which is symmetric with respect to the unit circle. Then it is a $\mathcal{P}(\Theta)$ space with reproducing kernel $K_\Theta(z,w)$ of the form* (3.4) *if and only if the following conditions are satisfied.*

(a) *$\mathcal{P}$ is backward-shift invariant.*
(b) *For every ζ, $\omega \in \Omega$ and $f(z)$, $g(z) \in \mathcal{P}$ the de Branges–Rovnyak identity holds:*

$$\langle f,g\rangle_{\mathcal{P}}+\zeta\langle R_\zeta f,g\rangle_{\mathcal{P}}+\omega^*\langle f,R_\omega g\rangle_{\mathcal{P}}-(1-\zeta\omega^*)\langle R_\zeta f,R_\omega g\rangle_{\mathcal{P}} = g(\omega)^*Jf(\zeta). \quad (3.18)$$

In this case $\dim\mathcal{P} = \deg\Theta$.

The identity (3.18) first appears in [107]. A proof of this theorem can be found in [21] and [26]. If the elements of $\mathcal{P}$ are analytic in a neighborhood of the origin, a basis of the space is given by the columns of a matrix function of the form $F(z) = C(I - zA)^{-1}$ and the choice $\zeta = \omega = 0$ in (3.18) leads to the Stein equation

$$G - A^*GA = C^*JC \quad (3.19)$$

for the Gram matrix G associated with $F(z)$.

The function $\Theta(z)$ in Theorem 3.7 is rational and J-unitary on the circle. To get a simple characterization in terms of minimal realizations of such functions $\Theta(z)$ we assume analyticity both at the origin and at infinity. This implies in particular that the matrix A in the next theorem is invertible. The theorem is the circle analog of Theorems 3.2 and 3.3; for a proof see [26, Theorem 3.1].

Theorem 3.8. *Let $\Theta(z)$ be a rational $p \times p$ matrix function analytic both at the origin and at infinity and let $\Theta(z) = D + C(zI - A)^{-1}B$ be a minimal realization* *of* $\Theta(z)$*. Then* $\Theta(z)$ *is* J*-unitary on the unit circle if and only if there exists an invertible Hermitian matrix* G *such that*

$$\begin{pmatrix} G & 0 \\ 0 & -J \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}^* \begin{pmatrix} G & 0 \\ 0 & -J \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix}. \tag{3.20}$$

We note that the matrix G in (3.20) satisfies the Stein equation

$$G - A^*GA = -C^*JC, \tag{3.21}$$

and that the formula

$$K_\Theta(z,w) = C(zI - A)^{-1}G^{-1}(wI - A)^{-*}C^* \tag{3.22}$$

holds, see [26, (3.17)]. This formula and the minimality of the realization of Θ imply that the space $\mathcal{P}(\Theta)$ is spanned by the columns of the matrix function $F(z) = C(zI - A)^{-1}$.

The next two theorems are particular cases of Theorem 1.1. They are the analogs of Theorems 3.4 and 3.5, respectively. The first one concerns spaces of functions which are holomorphic at ∞, the second one concerns spaces of functions which are holomorphic at 0. Their formulations are the same, except for the Stein equations: they differ by a minus sign.

Theorem 3.9. *Let* (C, A) *be an observable pair of matrices of sizes* $p \times m$ *and* $m \times m$ *respectively, and let* G *be an invertible Hermitian* $m \times m$ *matrix. Then the linear span* $\mathcal{M}$ *of the columns of the* $p \times m$ *matrix function* $F(z) = C(zI - A)^{-1}$ *endowed with the inner product*

$$\langle F\mathbf{c}, F\mathbf{d}\rangle = \mathbf{d}^*G\mathbf{c}, \quad \mathbf{c}, \mathbf{d} \in \mathbb{C}^m,$$

is a $\mathcal{P}(\Theta)$ *space with reproducing kernel* $K_\Theta(z,w)$ *of the form in* (3.4) *if and only if* G *is a solution of the Stein equation* (3.21)*:*

$$G - A^*GA = -C^*JC.$$

In this case, one can choose

$$\Theta(z) = I_p - (1 - zz_0^*)C(zI - A)^{-1}G^{-1}(z_0 I - A)^{-*}C^*J, \tag{3.23}$$

where $z_0 \in \mathbb{T} \cap \rho(A)$*.*

If A is invertible, Theorem 3.9 can also be proved using Theorem 3.8 and formula (3.22). Theorem 3.9 cannot be applied to backward-shift invariant spaces of polynomials; these are the spaces spanned by the columns of the matrix function $F(z) = C(I - zA)^{-1}$ where A is a nilpotent matrix. The next theorem holds in particular for such spaces.

Theorem 3.10. *Let* (C, A) *be an observable pair of matrices of sixes* $p \times m$ *and* $m \times m$ *respectively, and let* G *be an invertible Hermitian* $m \times m$ *matrix. Then the linear span* $\mathcal{M}$ *of the columns of the* $p \times m$ *matrix function* $F(z) = C(I - zA)^{-1}$ *endowed with the inner product*

$$\langle F\mathbf{c}, F\mathbf{d}\rangle = \mathbf{d}^*G\mathbf{c}, \quad \mathbf{c}, \mathbf{d} \in \mathbb{C}^m,$$

is a $\mathcal{P}(\Theta)$ space with reproducing kernel $K_\Theta(z,w)$ of the form in (3.4) *if and only if G is a solution of the Stein equation* (3.19)*:*

$$G - A^*GA = C^*JC.$$

If this is the case, one can choose

$$\Theta(z) = I_p - (1 - zz_0^*)C(I - zA)^{-1}G^{-1}(I - z_0A)^{-*}C^*J, \tag{3.24}$$

where $z_0 \in \mathbb{T}$ is such that $z_0^* \in \rho(A)$.

Assume that the spectral radius of A is strictly less than 1. Then the Stein equation (3.19) has a unique solution which can be written as

$$G = \sum_{i=0}^{\infty} A^{*i}C^*JCA^i.$$

This means that the space $\mathcal{M}$ is isometrically included in the Krein space $\mathbf{H}_{2,J}$ of $p \times 1$ vector functions with entries in the Hardy space $\mathbf{H}_2$ of the open unit disk equipped with the indefinite inner product

$$\langle f, g\rangle_{\mathbf{H}_{2,J}} = \langle f, Jg\rangle_{\mathbf{H}_2}.$$

The above discussion provides the key to the following theorem.

Theorem 3.11. *Let J be a $p \times p$ signature matrix and let $\Theta(z)$ be a rational $p \times p$ matrix function which is J-unitary on the unit circle and has no poles on the closed unit disk. Then*

$$\mathcal{P}(\Theta) = \mathbf{H}_{2,J} \ominus \Theta\mathbf{H}_{2,J}.$$

The McMillan degree is invariant under Möbius transformations, see [37]. This allows to state the counterpart of Theorem 3.6.

Theorem 3.12. *Let $\Theta(z)$ be a rational $p \times p$ matrix function which is J-unitary on the unit circle and has a unique pole at the point $1/w^*$ including, possibly, $w = 0$. Then*

$$\det\Theta(z) = c\left(\frac{z-w}{1-zw^*}\right)^{\deg\,\Theta},$$

where c is a unimodular constant.

In the sequel we shall need only the case $p = 2$. Then the signature matrices are $J = J_c$ for the circle and $J = -\mathrm{i}J_\ell$ for the line, where J_c and J_ℓ are given by (1.13). The above formulas for $\Theta(z)$ are the starting point of our approach in this paper. In each of the cases we consider, the matrix A is a Jordan block and the space $\mathcal{P}(\Theta)$ has no G-nondegenerate A invariant subspaces (besides the trivial ones). Under these assumptions we obtain analytic formulas for the functions $\Theta(z)$. Thus in the Sections 5 to 8 using the reproducing kernel space methods we obtain explicit formulas for $\Theta(z)$ in special cases.

3.4. Factorizations of J-unitary matrix functions

The product or the factorization (depending on the point of view)

$$R(z) = R_1(z)R_2(z),$$

where $R(z)$, $R_1(z)$, and $R_2(z)$ are rational $p \times p$ matrix functions is called *minimal* if

$$\deg\ R_1R_2 = \deg\ R_1 + \deg\ R_2.$$

The factorization is called *trivial*, if at least one of the factors is a constant matrix. The rational function $R(z)$ is called *elementary*, if it does not admit nontrivial minimal factorizations. One of the problems studied in this paper is the *factorization of certain classes of rational matrix functions into elementary factors*. Note that:

- A given rational matrix function may lack nontrivial factorizations, even if its McMillan degree is greater than 1.
- The factorization, if it exists, need not be unique.

As an example for the first assertion, consider the function

$$R(z) = \begin{pmatrix} 1 & z^2 \\ 0 & 1 \end{pmatrix}.$$

Its McMillan degree equals 2. One can check that it does not admit nontrivial minimal factorization by using the characterization of such factorizations proved in [37]. We give here a direct argument. Assume it admits a nontrivial factorization into factors of degree 1: $R(z) = R_1(z)R_2(z)$. In view of (3.1) these are of the form

$$R_i(z) = D_i + \frac{z\mathbf{c}_i\mathbf{b}_i^*}{1 - za_i}, \quad i = 1, 2,$$

where $a_i \in \mathbb{C}$, $\mathbf{b}_i$ and $\mathbf{c}_i$ are 2×1 vectors, $i = 1, 2$. We can assume without loss of generality that $D_i = I_2$. Then $a_1 = a_2 = 0$ since the factorization is minimal (or by direct inspection of the product) and so we have

$$\begin{aligned} \begin{pmatrix} 1 & z^2 \\ 0 & 1 \end{pmatrix} &= \left(I_2 + z\mathbf{c}_1\mathbf{b}_1^*\right)\left(I_2 + z\mathbf{c}_2\mathbf{b}_2^*\right) \\ &= I_2 + z\left(\mathbf{c}_1\mathbf{b}_1^* + \mathbf{c}_2\mathbf{b}_2^*\right) + z^2\left(\mathbf{b}_1^*\mathbf{c}_2\right)\mathbf{c}_1\mathbf{b}_2^*. \end{aligned}$$

Thus $\mathbf{b}_1^*\mathbf{c}_2 \neq 0$ and

$$\mathbf{c}_1\mathbf{b}_1^* + \mathbf{c}_2\mathbf{b}_2^* = 0_{2\times 2}. \tag{3.25}$$

On the other hand, taking determinants of both sides of the above factorization leads to

$$1 = \left(1 + z\mathbf{b}_1^*\mathbf{c}_1\right)\left(1 + z\mathbf{b}_2^*\mathbf{c}_2\right)$$

and so

$$\mathbf{b}_1^*\mathbf{c}_1 = \mathbf{b}_2^*\mathbf{c}_2 = 0.$$

Multiplying (3.25) by $\mathbf{c}_2$ on the right we obtain $\mathbf{c}_1 = 0$, and thus $R(z) = R_2(z)$, which is a contradiction to the fact that there is a nontrivial factorization.

If, additionally, the function $R(z)$ is J-unitary (on the real axis or on the unit circle) the factorization $R(z) = R_1(z)R_2(z)$ is called *J-unitary* if both factors are J-unitary. Then the following problems arise:

- Describe all rational elementary J-unitary matrix functions, these are the rational J-unitary matrix functions which do not admit nontrivial minimal J-unitary factorizations.
- Factor a given rational J-unitary matrix function into a minimal product of rational elementary J-unitary matrix functions.

We note that a rational J-unitary matrix function may admit nontrivial minimal factorizations, but lack nontrivial minimal J-unitary factorizations. Examples can be found in [5], [21, pp. 148–149], and [26, p. 191]. One such example is presented after Theorem 3.14. In the positive case, the first instance of uniqueness is the famous result of L. de Branges on the representation of J-inner entire functions when J is a 2×2 matrix with signature $(1, 1)$, see [48]. Related uniqueness results in the matrix case have been proved by D. Arov and H. Dym, see [31] and [32].

As a consequence of Theorems 3.6 and 3.12, in special cases products of rational J-unitary matrix functions are automatically minimal.

Theorem 3.13. *Let $z_1 \in \mathbb{C}^+$ ($\in \mathbb{D}$, respectively) and let $\Theta_1(z)$, $\Theta_2(z)$ be 2×2 matrix functions which both are J-unitary on the real line (the unit circle, respectively) and have a single pole at z_1. Then the product $\Theta_1(z)\Theta_2(z)$ is minimal and*

$$\mathcal{P}(\Theta_1\Theta_2) = \mathcal{P}(\Theta_1) \oplus \Theta_1\mathcal{P}(\Theta_2), \tag{3.26}$$

where the sum is direct and orthogonal.

Proof. We prove the theorem only for the line case; the proof for the circle case is similar. According to Theorem 3.6, with $c = c_\Theta$,

$$\begin{aligned} c_{\Theta_1\Theta_2}\left(\frac{z - z_1}{z - z_1^*}\right)^{\deg\ \Theta_1\Theta_2} &= \det\ (\Theta_1\Theta_2)(z) \\ &= \big(\det\ \Theta_1(z)\big)\big(\det\ \Theta_2(z)\big) \\ &= c_{\Theta_1}\left(\frac{z - z_1}{z - z_1^*}\right)^{\deg\ \Theta_1} c_{\Theta_2}\left(\frac{z - z_1}{z - z_1^*}\right)^{\deg\ \Theta_2} \\ &= c_{\Theta_1}c_{\Theta_2}\left(\frac{z - z_1}{z - z_1^*}\right)^{\deg\ \Theta_1 + \deg\ \Theta_2}. \end{aligned}$$

Therefore

$$\deg\ \Theta_1\Theta_2 = \deg\ \Theta_1 + \deg\ \Theta_2,$$

and the product $\Theta_1(z)\Theta_2(z)$ is minimal. The formula (3.26) follows from the kernel decomposition

$$K_{\Theta_1\Theta_2}(z, w) = K_{\Theta_1}(z, w) + \Theta_1(z)K_{\Theta_2}(z, w)\Theta_1(w)^*$$

and the minimality of the product, which implies that the dimensions of the spaces on both sides of the equality (3.26) coincide (recall that $\dim\ \mathcal{P}(\Theta) = \deg\ \Theta$, see Theorem 3.1). □

The following theorem is crucial for the proofs of the factorization theorems we give in the sequel, see Theorems 5.2, 6.4, 7.9, and 8.4.

Theorem 3.14. *Let $\Theta(z)$ be a rational $p \times p$ matrix function which is J-unitary on the unit circle or on the real axis. Then there is a one-to-one correspondence* (*up to constant J-unitary factors*) *between J-unitary minimal factorizations of $\Theta(z)$ and nondegenerate subspaces of $\mathcal{P}(\Theta)$ which are backward-shift invariant.*

For proofs see [21, Theorem 8.2] and [26]. Since we are in the finite-dimensional case, if $\dim \mathcal{P}(\Theta) > 1$, the backward-shift operator R_ζ always has proper invariant subspaces. However, they can be all degenerated with respect to the inner product of $\mathcal{P}(\Theta)$. Furthermore, R_ζ may have different increasing sequences of nondegenerate subspaces, leading to different J-unitary decompositions.

We give an example of a function lacking J-unitary factorizations. Let J be a $p \times p$ signature matrix (with nontrivial signature) and take two $p \times 1$ vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ such that

$$\mathbf{u}_1^* J \mathbf{u}_1 = \mathbf{u}_2^* J \mathbf{u}_2 = 0, \quad \mathbf{u}_1^* J \mathbf{u}_2 \neq 0.$$

Furthermore, choose $\alpha, \beta \in \mathbb{D}$ such that $\alpha \neq \beta$ and define the $p \times p$ matrices

$$W_{ij} = \frac{\mathbf{u}_i \mathbf{u}_j^*}{\mathbf{u}_j^* J \mathbf{u}_i}, \quad i, j = 1, 2, \ i \neq j.$$

Then the $p \times p$ matrix function

$$\Theta(z) = \left(I_p - \frac{1 - z(1 - \alpha^* \beta)}{(1 - z\alpha^*)(1 - \beta)} W_{12} \right) \left(I_p - \frac{1 - z(1 - \beta^* \alpha)}{(1 - z\beta^*)(1 - \alpha)} W_{21} \right)$$

is J-unitary on the unit circle and admits a nontrivial factorization but has no nontrivial J-unitary factorizations. The fact behind this is that the space $\mathcal{P}(\Theta)$ is spanned by the functions

$$\frac{\mathbf{u}_1}{1 - z\alpha^*}, \quad \frac{\mathbf{u}_2}{1 - z\beta^*}$$

and does not admit nondegenerate R_ζ invariant subspaces.

The four types of J-unitary rational matrix functions $\Theta(z)$, which are studied in the present paper, see Subsections 5.3, 6.3, 7.3, and 8.3, have a single singularity and are 2×2 matrix-valued. This implies that the underlying spaces $\mathcal{P}(\Theta)$ have a unique sequence of backward-shift invariant subspaces. Therefore the factorization into elementary factors is in all these cases either trivial or unique. We shall make this more explicit in the rest of this subsection. To this end we introduce some more notation. All the matrix functions $\Theta(z)$ in the following sections are rational 2×2 matrix functions; we denote the set of these functions by $\mathcal{U}$. Recall from (1.13) that

$$J_c = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad J_\ell = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Further, $\mathcal{U}_c$ denotes the set of all $\Theta(z) \in \mathcal{U}$ which are J_c-unitary on the circle, that is, they satisfy

$$\Theta(z) J_c \Theta(z)^* = J_c, \quad z \in \mathbb{T} \cap \mathrm{hol}\,(\Theta),$$

and $\mathcal{U}_\ell$ denotes the set of all $\Theta(z) \in \mathcal{U}$ which are J_ℓ-unitary on the line, that is, they satisfy

$$\Theta(z)J_\ell\Theta(z)^* = J_\ell, \quad z \in \mathbb{R} \cap \mathrm{hol}\,(\Theta).$$

Finally, if $z_1 \in \mathbb{C} \cup \{\infty\}$, then $\mathcal{U}_c^{z_1}$ stands for the set of those matrix functions in $\mathcal{U}_c$ which have a unique pole at $1/z_1^*$, and $\mathcal{U}_\ell^{z_1}$ stands for the set of those matrix functions from $\mathcal{U}_\ell$ which have a unique pole at z_1^*. Here we adhere to the convention that $1/0 = \infty$ and $1/\infty = 0$. Thus the elements of these classes admit the following representations:

(i) If $\Theta(z) \in \mathcal{U}_c^0 \cup \mathcal{U}_\ell^\infty$, then $\Theta(z)$ is a polynomial in z:

$$\Theta(z) = \sum_{j=0}^{n} T_j z^j.$$

(ii) If $\Theta(z) \in \mathcal{U}_c^{z_1}$ with $z_1 \neq 0, \infty$, then it is of the form

$$\Theta(z) = \sum_{j=0}^{n} \frac{T_j}{(1 - zz_1^*)^j}.$$

(iii) If $\Theta(z) \in \mathcal{U}_\ell^{z_1}$ with $z_1 \neq \infty$, then it is of the form

$$\Theta(z) = \sum_{j=0}^{n} \frac{T_j}{(z - z_1^*)^j}.$$

In all these cases n is an integer ≥ 0 and T_j are 2×2 matrices, $j = 0, 1, \dots, n$. The sets $\mathcal{U}_c^{z_1}$ etc. are all closed under multiplication. Moreover, the McMillan degree of $\Theta(z)$ in (i)–(iii) is given by (3.3) (with R_j replaced by T_j).

It is well known that the J_c-*unitary constants*, that is, the constant J_c-unitary matrices, are of the form

$$\frac{1}{1 - |\rho|^2} \begin{pmatrix} 1 & \rho \\ \rho^* & 1 \end{pmatrix} \begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix}$$

with ρ, c_1, $c_2 \in \mathbb{C}$ such that $|\rho| < 1$ and $|c_1| = |c_2| = 1$, and the J_ℓ-unitary constants are of the form

$$\mathrm{e}^{i\theta} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

with $\theta, \alpha, \beta, \gamma, \delta \in \mathbb{R}$ such that $\alpha\delta - \beta\gamma = 1$.

By Theorem 3.13, products in the sets $\mathcal{U}_\ell^{z_1}$ with $z_1 \in \mathbb{C}^+$ and $\mathcal{U}_c^{z_1}$ with $z_1 \in \mathbb{D}$ are automatically minimal. This is not the case with products in the sets $\mathcal{U}_\ell^{z_1}$ with $z_1 \in \mathbb{R} \cup \{\infty\}$ and $\mathcal{U}_c^{z_1}$ with $z_1 \in \mathbb{T}$, since these sets are closed under taking inverses. One can say more, but first a definition: We say that the matrix function $\Theta(z) \in \mathcal{U}$ is *normalized at the point* z_0 if $\Theta(z_0) = I_2$, the 2×2 identity matrix. In the sequel we normalize

$$\begin{array}{rl} \Theta(z) \in \mathcal{U}_c^{z_1},\ z_1 \in \mathbb{D}, & \text{in } z_0 \in \mathbb{T}, \\ \Theta(z) \in \mathcal{U}_c^{z_1},\ z_1 \in \mathbb{T}, & \text{in } z_0 \in \mathbb{T} \setminus \{z_1\}, \\ \Theta(z) \in \mathcal{U}_\ell^{\infty} & \text{in } z_0 = 0, \\ \Theta(z) \in \mathcal{U}_\ell^{z_1},\ z_1 \in \mathbb{C}^+, & \text{in } z_0 = \infty. \end{array}$$

For each of the four classes of matrix functions in the forthcoming sections we describe the normalized elementary factors and the essentially unique factorizations in terms of these factors. The factorizations are unique in that the constant matrix U is the last factor in the product. It could be positioned at any other place of the product, then, however, the elementary factors may change. In this sense we use the term *essential uniqueness*.

Theorem 3.15. *Let $z_1 \in \mathbb{D}$ ($\in \mathbb{C}^+$, respectively) and let $\Theta(z) \in \mathcal{U}_c^{z_1}$ ($\in \mathcal{U}_\ell^{z_1}$, respectively) be normalized and such that $\Theta(z_1) \neq 0_{2\times 2}$. Then $\Theta(z)$ admits a unique minimal factorization into normalized elementary factors.*

Proof. We consider the line case, the circle case is treated in the same way. It is enough to check that the space $\mathcal{P}(\Theta)$ is made of one chain and then to use Theorem 3.14. The space $\mathcal{P}(\Theta)$ consists of rational 2×1 vector functions which have only a pole in z_1^*, see Lemma 2.1. It is backward-shift invariant and therefore has a basis of elements of Jordan chains based on the point z_1^*. The beginning of each such chain is of the form

$$\frac{\mathbf{u}}{z - z_1^*}$$

for some 2×1 vector $\mathbf{u}$. Assume that there is more than one chain, that is, that there are two chains with first elements

$$f(z) = \frac{\mathbf{u}}{z - z_1^*}, \quad g(z) = \frac{\mathbf{v}}{z - z_1^*},$$

such that the 2×1 vectors $\mathbf{u}$ and $\mathbf{v}$ are linearly independent. Equation (3.6) implies that we have for $c, d, c', d' \in \mathbb{C}$,

$$\langle cf + dg, c'f + d'g \rangle_{\mathcal{P}(\Theta)} = \mathrm{i}\frac{(c'\mathbf{u} + d'\mathbf{v})^* J(c\mathbf{u} + d\mathbf{v})}{z_1 - z_1^*}.$$

Hence the space spanned by $f(z)$ and $g(z)$ is a nondegenerate backward-shift invariant subspace of $\mathcal{P}(\Theta)$, and therefore a $\mathcal{P}(\Theta_1)$ space, where $\Theta_1(z)$ is a factor of $\Theta(z)$. The special forms of $f(z)$ and $g(z)$ imply that

$$\Theta_1(z) = \frac{z - z_1}{z - z_1^*} I_2.$$

Hence $\Theta(z_1) = 0$, which contradicts the hypothesis. □

The last argument in the proof can be shortened. One can show that

$$\mathbf{u}^* J\Theta(z_1) = \mathbf{v}^* J\Theta(z_1) = 0,$$

and hence that $\Theta(z_1) = 0$ by using the following theorem, which is the analog of Theorem 3.11. Now $\mathbf{H}_{2,J}$ denotes the Krein space of $p \times 1$ vector functions with entries in the Hardy space $\mathbf{H}_2$ of the open upper half-plane equipped with the indefinite inner product

$$\langle f, g \rangle_{\mathbf{H}_{2,J}} = \langle f, Jg \rangle_{\mathbf{H}_2}. \tag{3.27}$$

Theorem 3.16. *Let $\Theta(z)$ be a rational $p \times p$ matrix function which is J-unitary on $\mathbb{R}$, and assume that $\Theta(z)$ does not have any poles in the closed upper half-plane. Then*

$$\mathcal{P}(\Theta) = \mathbf{H}_{2,J} \ominus \Theta\mathbf{H}_{2,J} \tag{3.28}$$

in the sense that the spaces are equal as vector spaces and that, moreover,

$$\langle f, g\rangle_{\mathcal{P}(\Theta)} = \frac{1}{2\pi}\langle f, g\rangle_{\mathbf{H}_{2,J}}$$

Even though, as mentioned above, products need not be minimal in $\mathcal{U}_\ell^{z_1}$ with $z_1 \in \mathbb{R} \cup \{\infty\}$ and in $\mathcal{U}_c^{z_1}$ with $z_1 \in \mathbb{T}$, the analog of Theorem 3.15 holds true:

Theorem 3.17. *Let $z_1 \in \mathbb{T}$ ($\in \mathbb{R} \cup \{\infty\}$, respectively) and let $\Theta(z)$ be a normalized element in $\mathcal{U}_c^{z_1}$ (in $\mathcal{U}_\ell^{z_1}$, respectively). Then $\Theta(z)$ admits a unique minimal factorization into normalized elementary factors.*

Proof. As in the case of Theorem 3.15 we consider the line case and $z_1 \in \mathbb{R}$. We show that the space $\mathcal{P}(\Theta)$ is spanned by the elements of only one chain. Suppose, on the contrary, that it is spanned by the elements of more than one chain. Then it contains elements of the form

$$\frac{\mathbf{u}}{z - z_1}, \quad \frac{\mathbf{v}}{z - z_1},$$

where $\mathbf{u}$ and $\mathbf{v}$ are 2×1 vectors. Then, by equation (3.6),

$$\mathbf{u}^* J\mathbf{v} = \mathbf{u}^* J\mathbf{u} = \mathbf{v}^* J\mathbf{v} = 0.$$

Thus $\mathbf{u}$ and $\mathbf{v}$ span a neutral space of the space $\mathbb{C}^2$ endowed with the inner product $\mathbf{y}^* J\mathbf{x}$, $\mathbf{x}, \mathbf{y} \in \mathbb{C}^2$. Since $J = -\mathrm{i}J_\ell$, it follows that every nontrivial neutral subspace has dimension 1 and thus there is only one chain in $\mathcal{P}(\Theta)$. The rest is plain from Theorem 3.14. $\square$

3.5. Additional remarks and references

We refer to [37] for more information on realization and minimal factorization of rational matrix functions. Rational matrix functions which are J-unitary on the unit circle or on the real line were studied in [21] and [26].

For connections between the structural identities (3.6) and (3.18) and the Lyapunov and Stein equations, see [13] and [80].

As remarked in [22, §8] most of the computations related to finite-dimensional spaces $\mathcal{P}(\Theta)$ would still make sense if one replaces the complex numbers by an arbitrary field and conjugation by a field isomorphism. For possible applications to coding theory, see the discussion in Subsection 8.5 and see also the already mentioned papers of M.-J. Bertin [41] and Ch. Pisot [105].

4. Pick matrices

In this section we introduce the Pick matrices at the point z_1 for the four classes of generalized Schur and Nevanlinna functions: $\mathbf{S}^{z_1}$ with $z_1 \in \mathbb{D}$, $\mathbf{S}^{z_1;2p}$ with $z_1 \in \mathbb{T}$ and an integer $p \geq 1$, $\mathbf{N}^{z_1}$ with $z_1 \in \mathbb{C}^+$, and $\mathbf{N}^{\infty;2p}$ with an integer $p \geq 1$. In fact, in each case only the smallest nondegenerate Pick matrix at z_1 is of actual interest.

4.1. Generalized Schur functions: $z_1 \in \mathbb{D}$

We consider a function $s(z) \in \mathbf{S}^{z_1}$. Recall that this means that $s(z)$ belongs to some generalized Schur class $\mathbf{S}_\kappa$ and is holomorphic at $z_1 \in \mathbb{D}$. The Taylor expansion of $s(z)$ at z_1 we write as

$$s(z) = \sum_{i=0}^{\infty} \sigma_i (z - z_1)^i. \tag{4.1}$$

The kernel $K_s(z,w) = \dfrac{1 - s(z)s(w)^*}{1 - zw^*}$ is holomorphic in z and w^* at $z = w = z_1$ with Taylor expansion

$$K_s(z,w) = \frac{1 - s(z)s(w)^*}{1 - zw^*} = \sum_{i,j=0}^{\infty} \gamma_{ij} (z - z_1)^i (w - z_1)^j. \tag{4.2}$$

Here and elsewhere in the paper where we consider a Taylor expansion we are only interested in the Taylor coefficients, so we do not (need to) specify the domain of convergence of the expansion. Recall that the Pick matrix of the kernel $K_s(z,w)$ at z_1 is $\Gamma = (\gamma_{ij})_{i,j=0}^{\infty}$, which we call also the *Pick matrix of the function* $s(z)$ *at* z_1. As a consequence of Theorem 2.7 we have the following result.

Corollary 4.1. *For* $s(z) \in \mathbf{S}^{z_1}$ *it holds that*

$$s(z) \in \mathbf{S}_\kappa^{z_1} \iff \kappa_-(\Gamma) = \kappa.$$

If we write (4.2) as

$$\begin{aligned} 1 - s(z)s(w)^* = \big(1 - |z_1|^2 - (z - z_1)z_1^* - z_1(w - z_1)^* - (z - z_1)(w - z_1)^*\big) \\ \times \sum_{i,j=0}^{\infty} \gamma_{ij} (z - z_1)^i (w - z_1)^{*j}, \end{aligned}$$

insert for $s(z)$ the expansion (4.1), and compare coefficients we see that (4.2) is equivalent to the following equations for the numbers γ_{ij}:

$$\begin{aligned} \big(1 - |z_1|^2\big)\gamma_{ij} - z_1^*\gamma_{i-1,j} - z_1\gamma_{i,j-1} - \gamma_{i-1,j-1} = -\sigma_i\sigma_j^*, \\ i,j = 0,1,\dots, \quad i + j > 0, \end{aligned} \tag{4.3}$$

and the “initial conditions”

$$\gamma_{00} = \frac{1 - |\sigma_0|^2}{1 - |z_1|^2}; \quad \gamma_{i,-1} = \gamma_{-1,j} = 0, \quad i,j = 0,1,\dots. \tag{4.4}$$

With the shift matrix

$$S = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ 0 & 0 & 0 & 0 & \ddots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

the Toeplitz matrix

$$\Sigma = (\sigma_{j-k})_{j,k=0}^{\infty} = \begin{pmatrix} \sigma_0 & 0 & 0 & 0 & \cdots \\ \sigma_1 & \sigma_0 & 0 & 0 & \cdots \\ \sigma_2 & \sigma_1 & \sigma_0 & 0 & \cdots \\ \sigma_3 & \sigma_2 & \sigma_1 & \sigma_0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

of the Taylor coefficients σ_j of the generalized Schur function $s(z)$ in (4.1) (setting $\sigma_j = 0$ if $j < 0$), and the vectors

$$\mathbf{s} = \begin{pmatrix} \sigma_0 & \sigma_1 & \sigma_2 & \cdots \end{pmatrix}^{\top}, \qquad \mathbf{e}_0 = \begin{pmatrix} 1 & 0 & 0 & \cdots \end{pmatrix}^{\top},$$

the relations (4.3), (4.4) can be written as the *Stein equation*

$$(1-|z_1|^2)\Gamma - z_1^* S^*\Gamma - z_1\Gamma S - S^*\Gamma S = \begin{pmatrix} \mathbf{e}_0 & \mathbf{s} \end{pmatrix} J_c \begin{pmatrix} \mathbf{e}_0 & \mathbf{s} \end{pmatrix}^* \ \left(= \mathbf{e}_0\mathbf{e}_0^* - \mathbf{s}\,\mathbf{s}^* \right). \quad (4.5)$$

To obtain an explicit formula for the Pick matrix Γ we first consider the corresponding expansion for the particular case $s(z) = 0$, that is, for the kernel $\dfrac{1}{1-zw^*}$:

$$\frac{1}{1-zw^*} = \sum\nolimits_{i,j=0}^{\infty} \gamma_{ij}^0 (z-z_1)^i (w-z_1)^{*j}, \quad z, w \in \mathbb{D}.$$

Specifying (4.3), (4.4) for $s(z) = 0$, we obtain that the coefficients γ_{ij}^0 are the unique solutions of the difference equations

$$\left(1-|z_1|^2\right)\gamma_{ij}^0 - z_1^*\gamma_{i-1,j}^0 - z_1\gamma_{i,j-1}^0 - \gamma_{i-1,j-1}^0 = 0, \quad i,j = 0,1,\ldots,\ i+j > 0, \quad (4.6)$$

with the initial conditions

$$\gamma_{00}^0 = \frac{1}{1-|z_1|^2}; \quad \gamma_{i,-1}^0 = \gamma_{-1,j}^0 = 0, \quad i,j = 0,1,\ldots.$$

or, in matrix form,

$$\left(1-|z_1|^2\right)\Gamma^0 - z_1^* S^*\Gamma^0 - z_1\Gamma^0 S - S^*\Gamma^0 S = \mathbf{e_0}\mathbf{e_0}^*. \quad (4.7)$$

Lemma 4.2. *The entries of the matrix* $\Gamma^0 = \left(\gamma_{ij}^0\right)_{i,j=0}^{\infty}$ *are given by*

$$\gamma_{ij}^0 = \left(1-|z_1|^2\right)^{-(i+j+1)} \left(D^*D\right)_{ij},$$

where D *is the matrix*

$$D = \left(d_{ij}\right)_{i,j=0}^{\infty} \quad \text{with} \quad d_{ij} = \binom{j}{i} z_1^{j-i}, \quad i,j = 0,1,\ldots.$$

The matrix Γ^0 *is positive in the sense that all its principal submatrices are positive.*

Since the binomial coefficients $\binom{j}{i}$ vanish for $j < i$ the matrix D is right upper triangular.

Proof of Lemma 4.2. We introduce the numbers

$$\beta_{ij} = \bigl(1 - |z_1|^2\bigr)^{i+j+1} \gamma_{ij}^0, \quad i, j = 0, 1, \ldots .$$

Then $\beta_{00} = 1$, $\beta_{i,-1} = \beta_{-1,j} = 0$, $i, j = 0, 1, \ldots$, and the difference equations (4.6) imply

$$\beta_{ij} = \beta_{i-1,j}\, z_1^* + \beta_{i,j-1}\, z_1 \bigl(1 - |z_1|^2\bigr)\, \beta_{i-1,j-1}, \quad i, j = 0, 1, \ldots, \; i + j > 0. \tag{4.8}$$

We show that the numbers $\beta_{ij} = \bigl(D^* D\bigr)_{ij}$ satisfy the relations (4.8). We have

$$\beta_{ij} = \sum_{k=0}^{j} d_{ki}^*\, d_{kj} = \sum_{k=0}^{j} \binom{i}{k} z_1^{*i-k} \binom{j}{k} z_1^{j-k},$$

and it is to be shown that this expression equals

$$\sum_{k=0}^{j} \binom{i-1}{k} (z_1^*)^{i-k} \binom{j}{k} z_1^{j-k} + \sum_{k=0}^{j-1} \binom{i}{k} (z_1^*)^{i-k} \binom{j-1}{k} z_1^{j-k}$$
$$+ \bigl(1 - |z_1|^2\bigr) \sum_{k=0}^{j-1} \binom{i-1}{k} (z_1^*)^{i-1-k} \binom{j-1}{k} z_1^{j-1-k}.$$

Comparing coefficients of $(z_1^*)^{i-k} z_1^{j-k}$ it turns out that we have to prove the relation

$$\binom{i-1}{k}\binom{j}{k} + \binom{i}{k}\binom{j-1}{k} + \binom{i-1}{k-1}\binom{j-1}{k-1} - \binom{i-1}{k}\binom{j-1}{k} = \binom{i}{k}\binom{j}{k}.$$

If the identity

$$\binom{\mu}{\nu} - \binom{\mu-1}{\nu} = \binom{\mu-1}{\nu-1} \tag{4.9}$$

is applied to the first and the last term of the left-hand side, and to the second term of the left-hand side and the term on the right-hand side we get

$$\binom{i-1}{k}\binom{j-1}{k-1} - \binom{i}{k}\binom{j-1}{k-1} + \binom{i-1}{k-1}\binom{j}{k-1} = 0,$$

and another application of (4.9) gives the desired result.

To prove the last statement we use that a Hermitian matrix is positive if and only if the determinant of each of its principal submatrices is positive. Applying the elementary rules to calculate determinants, we find that for all integers $j \geq 1$,

$$\det \Gamma_j^0 = \bigl(1 - |z_1|^2\bigr)^{k^2} \det\bigl(\, D_j^*\, D_j\bigr) = \bigl(1 - |z_1|^2\bigr)^{k^2} > 0.$$

It follows that all principal submatrices of Γ^0 are positive. □

Clearly, for the special case $z_1 = 0$ we obtain $\Gamma^0 = I$, the infinite identity matrix.

Now it is easy to give an explicit expression for the Pick matrix Γ.

Theorem 4.3. *The Pick matrix Γ of the function $s(z) \in \mathbf{S}^{z_1}$ at the point z_1, that is, the solution of the Stein equation (4.5), is given by the relation*

$$\Gamma = \Gamma^0 - \Sigma\,\Gamma^0\Sigma^*. \tag{4.10}$$

Proof. Inserting $\Gamma = \Gamma^0 - \Sigma\,\Gamma^0\Sigma^*$ in (4.5) the left-hand side becomes

$$\begin{aligned}&\bigl(1-|z_1|^2\bigr)\Gamma^0 - z_1^*S^*\Gamma^0 - z_1\Gamma^0 S - S^*\Gamma^0 S\\ &\quad - \bigl(1-|z_1|^2\bigr)\Sigma\Gamma^0\Sigma^* + z_1^*S^*\Sigma\Gamma^0\Sigma^* + z_1\Sigma\Gamma^0\Sigma^*S + S^*\Sigma\Gamma^0\Sigma^*S.\end{aligned}$$

By (4.7), the terms on the first line add up to $\mathbf{e_0}\mathbf{e_0}^*$. If we observe the relations $S^*\Sigma = \Sigma S^*$, $\Sigma\mathbf{e_0}\mathbf{e_0}^*\Sigma^* = \mathbf{s}\mathbf{s}^*$, and again (4.7), the second line becomes

$$-\Sigma\Bigl(\bigl(1-|z_1|^2\bigr)\Gamma^0 - z_1^*S^*\Gamma^0 - z_1\Gamma^0 S - S^*\Gamma^0 S\Bigr)\Sigma^* = -\mathbf{s}\,\mathbf{s}^*. \qquad \square$$

In the particular case $z_1 = 0$ the Pick matrix Γ for $s(z) \in \mathbf{S}^0$ at 0 becomes

$$\Gamma = I - \Sigma\,\Sigma^*, \tag{4.11}$$

and the Stein equation (4.5) reads as

$$\Gamma - S^*\Gamma S = \begin{pmatrix}\mathbf{e}_0 & \mathbf{s}\end{pmatrix} J_c \begin{pmatrix}\mathbf{e}_0 & \mathbf{s}\end{pmatrix}^* \ \bigl(= \mathbf{e}_0\mathbf{e}_0^* - \mathbf{s}\,\mathbf{s}^*\bigr). \tag{4.12}$$

The relations (4.11) and (4.12) imply for $m = 1, 2, \ldots$

$$\Gamma_m = I_m - \Sigma_m\,\Sigma_m^*, \qquad \text{and} \qquad \Gamma_m - S_m^*\Gamma_m S_m = C^*J_cC;$$

here S_m is the principal $m \times m$ submatrix of the shift matrix S and C is the $2 \times m$ matrix

$$C = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \sigma_0^* & \sigma_1^* & \cdots & \sigma_{m-1}^* \end{pmatrix}.$$

Recall that for the Pick matrix Γ the smallest positive integer j such that the principal submatrix Γ_j is invertible is denoted by $k_0(\Gamma)$:

$$k_0(\Gamma) := \min\bigl\{j \bigm| \det\Gamma_j \neq 0\bigr\}.$$

Theorem 4.4. *For the function $s(z) \in \mathbf{S}^{z_1}$ which is not identically equal to a unimodular constant and its Pick matrix Γ at z_1 we have*

$$|\sigma_0| \neq 1 \quad\Longleftrightarrow\quad k_0(\Gamma) = 1;$$

if $|\sigma_0| = 1$ then $k_0(\Gamma) = 2k$ where k is the smallest integer $k \geq 1$ such that $\sigma_k \neq 0$.

Proof. The first claim follows from the relation

$$\gamma_{00} = \frac{1-|\sigma_0|^2}{1-|z_1|^2}. \tag{4.13}$$

If $|\sigma_0| = 1$ and k is the smallest positive integer such that $\sigma_k \neq 0$ then we write the principal $2k \times 2k$ submatrices of Γ^0 and of Σ as 2×2 block matrices:

$$\Gamma^0_{2k} = \begin{pmatrix} A & B \\ B^* & D \end{pmatrix}, \quad \Sigma_{2k} = \begin{pmatrix} \sigma_0 I_k & 0 \\ \Delta & \sigma_0 I_k \end{pmatrix},$$

where all blocks are $k \times k$ matrices, $A = A^* = \Gamma^0_k$, $D = D^*$, I_k is the $k \times k$ unit matrix, and

$$\Delta = \begin{pmatrix} \sigma_k & 0 & \cdots & 0 & 0 \\ \sigma_{k+1} & \sigma_k & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \sigma_{2k-2} & \sigma_{2k-3} & \cdots & \sigma_k & 0 \\ \sigma_{2k-1} & \sigma_{2k-2} & \cdots & \sigma_{k+1} & \sigma_k \end{pmatrix}.$$

Then

$$\Gamma_{2k} = \begin{pmatrix} 0 & \sigma_0 A \Delta^* \\ \sigma_0^* \Delta A & \Delta A \Delta^* + \sigma_0 B^* \Delta^* + \sigma_0^* \Delta B \end{pmatrix}$$

and, since A, by Lemma 4.2, and Δ are invertible, Γ_{2k} is invertible and no principal submatrix of Γ_{2k} is invertible. Hence $k_0(\Gamma) = 2k$. $\square$

4.2. Generalized Schur functions: $z_1 \in \mathbb{T}$

We consider a function $s(z) \in \mathbf{S}$ which for some integer $p \geq 1$ has at $z_1 \in \mathbb{T}$ an asymptotic expansion of the form

$$s(z) = \tau_0 + \sum_{i=1}^{2p-1} \tau_i (z - z_1)^i + \mathrm{O}\left((z - z_1)^{2p}\right), \quad z \hat{\to} z_1. \tag{4.14}$$

Theorem 4.5. *Suppose that the function $s(z) \in \mathbf{S}$ has the asymptotic expansion* (4.14) *with $|\tau_0| = 1$. Then the following statements are equivalent.*

(1) *The matrix $\widehat{P}$ in* (2.17) *is Hermitian.*
(2) *The kernel $K_s(z, w)$ has an asymptotic expansion of the form*

$$\begin{aligned} K_s(z,w) \quad &= \quad \textstyle\sum_{0 \le i+j \le 2p-2} \gamma_{ij} (z - z_1)^i (w - z_1)^{*j} \\ &\quad\quad + \mathrm{O}\left((\max\{|z - z_1|, |w - z_1|\})^{2p-1}\right), \quad z, w \hat{\to} z_1. \end{aligned} \tag{4.15}$$

If (1) *and* (2) *hold, then for the Pick matrix $\Gamma_p = \left(\gamma_{ij}\right)_{i,j=0}^{p-1}$ at z_1 we have $\Gamma_p = \widehat{P}$.*

For a proof, see [18, Lemma 2.1]. We mention that the coefficients γ_{ij} satisfy the relations (compare with (4.3))

$$z_1^* \gamma_{i-1,j} + z_1 \gamma_{i,j-1} + \gamma_{i-1,j-1} = \tau_i \tau_j^*, \quad i, j = 0, 1, \ldots, 2p-2, \ 1 \le i+j \le 2p-2,$$

where, if i or $j = -1$, γ_{ij} is set equal to zero, and that the Pick matrix Γ_p satisfies the Stein equation

$$\Gamma_p - A_p^* \Gamma_p A_p = \; C^* J_c C,$$

where $A_p = z_1^* I_p + S_p$ and C is the $2 \times p$ matrix

$$C = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \tau_0^* & \tau_1^* & \cdots & \tau_{p-1}^* \end{pmatrix}.$$

If statements (1) and (2) of Theorem 4.5 hold, we are interested in the smallest integer $k_0 := k_0(\Gamma_p) \geq 1$, for which the principal $k_0 \times k_0$ submatrix $\Gamma_{k_0} := (\Gamma_p)_{k_0}$ of Γ_p is invertible. Recall that $\mathbf{S}^{z_1;2p}$ is the class of functions $s(z) \in \mathbf{S}$ which have an asymptotic expansion of the form (4.14) with the properties that $|\tau_0| = 1$, not all coefficients $\tau_1, \ldots, \tau_p$ vanish, and the statements (1) and (2) of Theorem 4.5 hold. The second property implies that k_0 exists and $1 \leq k_0 \leq p$.

Theorem 4.6. *Suppose that the function $s(z) \in \mathbf{S}^{z_1;2p}$ has the asymptotic expansion (4.14). Then $k_0 = k_0(\Gamma_p)$ coincides with the smallest integer $k \geq 1$ such that $\tau_k \neq 0$, and we have*

$$\Gamma_{k_0} = \Gamma_k = \tau_0^* \Delta B, \tag{4.16}$$

where

$$\Delta = \begin{pmatrix} \tau_k & 0 & \cdots & 0 & 0 \\ \tau_{k+1} & \tau_k & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \tau_{2k-2} & \tau_{2k-3} & \cdots & \tau_k & 0 \\ \tau_{2k-1} & \tau_{2k-2} & \cdots & \tau_{k+1} & \tau_k \end{pmatrix}$$

and B is the right lower matrix

$$B = \begin{pmatrix} 0 & 0 & \cdots & 0 & (-1)^{k-1}\binom{k-1}{0} z_1^{2k-1} \\ 0 & 0 & \cdots & (-1)^{k-2}\binom{k-2}{0} z_1^{2k-3} & (-1)^{k-1}\binom{k-1}{1} z_1^{2k-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & -\binom{1}{0} z_1^3 & \cdots & (-1)^{k-2}\binom{k-2}{k-3} z_1^{k} & (-1)^{k-1}\binom{k-1}{k-2} z_1^{k+1} \\ z_1 & -\binom{1}{1} z_1^2 & \cdots & (-1)^{k-2}\binom{k-2}{k-2} z_1^{k-1} & (-1)^{k-1}\binom{k-1}{k-1} z_1^{k} \end{pmatrix}.$$

This theorem is proved in [18, Lemma 2.1]. The matrix Γ_k in (4.16) is right lower triangular. The entries on the second main diagonal are given by

$$\gamma_{i,k-1-i} = (-1)^{k-1-i} z_1^{2k-1-2i} \tau_0^* \tau_k, \quad i = 0, 1, \ldots, k-1, \tag{4.17}$$

hence, because τ_0, τ_k, $z_1 \neq 0$, Γ_k is invertible. Since Γ_k is Hermitian, by (4.17), $z_1^k \tau_0^* \tau_k$ is purely imaginary if k is even and real if k is odd, and the number of negative eigenvalues of Γ_k is equal to

$$\kappa_-(\Gamma_k) = \begin{cases} k/2, & k \text{ is even}, \\ (k-1)/2, & k \text{ is odd and } (-1)^{(k-1)/2} z_1^k \tau_0^* \tau_k > 0, \\ (k+1)/2, & k \text{ is odd and } (-1)^{(k-1)/2} z_1^k \tau_0^* \tau_k < 0. \end{cases} \tag{4.18}$$

Under the assumptions of the theorem we have

$$\mathrm{sq}_-(s) \geq \kappa_-(\Gamma_k). \tag{4.19}$$

This follows from the asymptotic expansion (4.15) of the kernel $K_s(z,w)$ and the inequality (2.10).

4.3. Generalized Nevanlinna functions: $z_1 \in \mathbb{C}^+$

We consider $n(z) \in \mathbf{N}^{z_1}$ and write its Taylor expansion at z_1 as

$$n(z) = \sum_{i=0}^{\infty} \nu_i (z - z_1)^i; \tag{4.20}$$

the series converges in a neighborhood of z_1, which we need not specify, because we are only interested in the Taylor coefficients of $n(z)$. The kernel $L_n(z,w)$ is holomorphic in z and in w^* at $z = w = z_1$ with Taylor expansion

$$L_n(z,w) = \frac{n(z) - n(w)^*}{z - w^*} = \sum_{i,j=0}^{\infty} \gamma_{ij} (z - z_1)^i (w - z_1)^{*j}.$$

We call the Pick matrix $\Gamma = \left(\gamma_{ij}\right)_{i,j=0}^{\infty}$ of this kernel also the *Pick matrix for the function* $n(z)$ *at* z_1. We readily obtain the following corollary of Theorem 2.7.

Corollary 4.7. *If* $n(z) \in \mathbf{N}^{z_1}$, *then*

$$n(z) \in \mathbf{N}_\kappa^{z_1} \iff \kappa_-(\Gamma) = \kappa.$$

The entries γ_{ij} of the Pick matrix of $n(z)$ at z_1 satisfy the equations

$$\gamma_{00} = \frac{\nu_0 - \nu_0^*}{z_1 - z_1^*} = \operatorname{Im} \nu_0 / \operatorname{Im} z_1,$$

$$(z_1 - z_1^*)\gamma_{i0} + \gamma_{i-1,0} = \nu_i,\ i \geq 1, \quad (z_1 - z_1^*)\gamma_{0j} - \gamma_{0,j-1} = -\nu_j^*,\ j \geq 1,$$

and

$$(z_1 - z_1^*)\gamma_{ij} + \gamma_{i-1,j} - \gamma_{i,j-1} = 0, \quad i,j \geq 1.$$

In matrix form these equations can be written as the *Lyapunov equation*

$$(z_1 - z_1^*)\Gamma + S^*\Gamma - \Gamma S = \begin{pmatrix} \mathbf{n} & \mathbf{e}_0 \end{pmatrix} J_\ell \begin{pmatrix} \mathbf{n} & \mathbf{e}_0 \end{pmatrix}^* \ \left(= \mathbf{n}\mathbf{e}_0^* - \mathbf{e}_0\mathbf{n}^*\right), \tag{4.21}$$

where

$$\mathbf{n} = \begin{pmatrix} \nu_0 & \nu_1 & \nu_2 & \cdots \end{pmatrix}^\top, \qquad \mathbf{e}_0 = \begin{pmatrix} 1 & 0 & 0 & \cdots \end{pmatrix}^\top.$$

To find a formula for the Pick matrix Γ, in analogy to Subsection 4.1, we first consider the Taylor expansion of the simpler kernel for the function $n(z) \in \mathbf{N}_0$ with $n(z) \equiv \mathrm{i}$, $\operatorname{Im} z > 0$:

$$\frac{2\,\mathrm{i}}{z - w^*} = \sum_{i,j=0}^{\infty} \gamma_{ij}^0 (z - z_1)^i (w - z_1)^{*j}.$$

We obtain

$$\gamma_{00}^0 = \frac{2\,\mathrm{i}}{z_1 - z_1^*} = 1/\operatorname{Im} z_1 \tag{4.22}$$

and

$$(z_1 - z_1^*)\gamma_{ij}^0 + \gamma_{i-1,j}^0 - \gamma_{i,j-1}^0 = 0, \quad i,j = 0,1,\ldots,\ i+j \geq 1, \tag{4.23}$$

where $\gamma_{ij}^0 = 0$ if $i = -1$ or $j = -1$, or, in explicit form,

$$\gamma_{ij}^0 = \frac{\partial^{i+j}}{\partial z^i \partial w^{*j}} \left. \frac{2\,\mathrm{i}}{z - w^*} \right|_{z=w=z_1} = \binom{i+j}{i} \frac{2\,\mathrm{i}(-1)^i}{(z_1 - z_1^*)^{i+j+1}}, \quad i,j = 0,1,\ldots.$$

From this formula one can derive the following result.

Lemma 4.8. *All principal submatrices of* $\Gamma^0 = \left(\gamma_{ij}^0\right)_{i,j=0}^{\infty}$ *are positive.*

Proof. By induction one can prove that the determinant of the $\ell \times \ell$ matrix

$$A_\ell = \left(a_{ij}\right)_{i,j=0}^{\ell-1}, \quad a_{ij} = \binom{i+j}{i},$$

is equal to 1. Using elementary rules for calculating determinants we find that for all integers $\ell \geq 1$,

$$\Gamma_\ell^0 = (-1)^{[\ell/2]} \frac{(2\,\mathrm{i})^\ell}{(z_1 - z_1^*)^{\ell^2}} \det A_\ell = \frac{1}{2^{\ell(\ell-1)}(\operatorname{Im} z_1)^{\ell^2}},$$

and hence, see the proof of Lemma 4.2, Γ^0 is positive. □

The relations (4.22) and (4.23) can be written in matrix form as (compare with (4.21))

$$(z_1 - z_1^*)\Gamma^0 + S^*\Gamma^0 - \Gamma^0 S = 2\mathrm{i}\mathbf{e}_0\mathbf{e}_0^*.$$

Now the Pick matrix Γ in (4.21) becomes

$$\Gamma = \frac{1}{2\,\mathrm{i}}\left(\Sigma\,\Gamma^0 - \Gamma^0\Sigma^*\right),$$

where

$$\Sigma = \begin{pmatrix} \nu_0 & 0 & 0 & 0 & \cdots \\ \nu_1 & \nu_0 & 0 & 0 & \cdots \\ \nu_2 & \nu_1 & \nu_0 & 0 & \cdots \\ \nu_3 & \nu_2 & \nu_1 & \nu_0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

We also need the analog of Theorem 4.4, that is, the smallest positive integer $k_0 = k_0(\Gamma)$ such that for $n(z) \in \mathbf{N}$ the principal submatrix Γ_{k_0} is invertible.

Theorem 4.9. *For the function* $n(z) \in \mathbf{N}^{z_1}$ *which is not identically equal to a real constant and its Pick matrix* Γ *at* z_1 *we have*

$$\operatorname{Im}\nu_0 \neq 0 \Longleftrightarrow k_0(\Gamma) = 1;$$

if $\nu_0 \in \mathbb{R}$ *and* k *is the smallest integer* ≥ 1 *such that* $\nu_k \neq 0$, *then* $k_0(\Gamma) = 2k$.

Proof. The first statement follows from the formula

$$\Gamma_1 = \gamma_{00} = \operatorname{Im}\nu_0/\operatorname{Im} z_1.$$

We prove the second statement. Assume $\nu_0 \in \mathbb{R}$ and let k be the smallest integer ≥ 1 such that $\nu_k \neq 0$. Then the principal $2k \times 2k$ submatrices of Γ^0 and Σ can be written as the block matrices

$$\Gamma^0_{2k} = \begin{pmatrix} A & B \\ B^* & D \end{pmatrix}, \quad \Sigma_{2k} = \begin{pmatrix} \nu_0 I_k & 0 \\ \Delta & \nu_0 I_k \end{pmatrix},$$

where all blocks are $k \times k$ matrices, $A = A^* = \Gamma^0_k$, $D = D^*$, and Δ is the triangular matrix

$$\Delta = \begin{pmatrix} \nu_k & 0 & \cdots & 0 & 0 \\ \nu_{k+1} & \nu_k & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \nu_{2k-2} & \nu_{2k-3} & \cdots & \nu_k & 0 \\ \nu_{2k-1} & \nu_{2k-2} & \cdots & \nu_{k+1} & \nu_k \end{pmatrix}.$$

Then

$$\Gamma_{2k} = \frac{1}{2\mathrm{i}} \begin{pmatrix} 0 & -A\Delta^* \\ \Delta A & \Delta B - B^*\Delta^* \end{pmatrix}, \tag{4.24}$$

and, since A, by Lemma 4.8, and Δ are invertible, Γ_{2k} is invertible and no principal submatrix of Γ_{2k} is invertible. $\square$

4.4. Generalized Nevanlinna functions: $z_1 = \infty$

In this subsection we consider a function $n(z)$ from the class $\mathbf{N}^{\infty;2p}$ for some integer $p \geq 1$. This means, see Subsection 2.4, that $n(z)$ belongs to the class $\mathbf{N}$ and has an asymptotic expansion of the form

$$n(z) = -\frac{\mu_0}{z} - \frac{\mu_1}{z^2} - \cdots - \frac{\mu_{2p-1}}{z^{2p}} - \frac{\mu_{2p}}{z^{2p+1}} + \mathrm{o}\left(\frac{1}{z^{2p+1}}\right), \quad z = \mathrm{i}y,\ y \uparrow \infty,$$

with

(i) $\mu_j \in \mathbb{R}$, $j = 0, 1, \ldots, 2p$, and
(ii) not all coefficients $\mu_0, \mu_1, \ldots, \mu_{p-1}$ vanish.

The asymptotic expansion of $n(z)$ yields an asymptotic expansion for the kernel $L_n(z, w)$:

$$\begin{aligned} L_n(z,w) &= \frac{n(z) - n(w)^*}{z - w^*} \\ &= \sum_{0 \leq i+j \leq 2p} \frac{\gamma_{ij}}{z^{i+1} w^{*(j+1)}} + \mathrm{o}\left(\max\left(|z|^{-(2p+2)}, |w|^{-(2p+2)}\right)\right), \end{aligned} \tag{4.25}$$

$$z = \mathrm{i}y,\ w = \mathrm{i}\eta,\ y, \eta \uparrow \infty,$$

with $\gamma_{ij} = \mu_{i+j}$, $0 \leq i + j \leq 2p$. The Pick matrix is therefore the $(p+1) \times (p+1)$ Hankel matrix

$$\Gamma_{p+1} = \left(\gamma_{ij}\right)_0^p \quad \text{with} \quad \gamma_{ij} = \mu_{i+j},\ 0 \leq i, j \leq p.$$

This implies the following theorem. Recall that $k_0(\Gamma_{p+1})$ is the smallest integer $k \geq 1$ for which the principal $k \times k$ submatrix $\Gamma_k = (\Gamma_{p+1})_k$ of Γ_{p+1} is invertible. Condition (ii) implies that $1 \leq k_0(\Gamma_{p+1}) \leq p$.

Theorem 4.10. *The index $k_0(\Gamma_{p+1})$ is determined by the relation*

$$k_0(\Gamma_{p+1}) = k \geq 1 \quad \Longleftrightarrow \quad \mu_0 = \mu_1 = \cdots = \mu_{k-2} = 0,\ \mu_{k-1} \neq 0.$$

If $k = 1$ in the theorem, then the first condition on the right-hand side of the arrow should be discarded. With $k = k_0(\Gamma_{p+1})$ and $\varepsilon = \operatorname{sgn} \mu_{k-1}$ we have

$$\kappa_-(\Gamma_k) = \begin{cases} [k/2], & \varepsilon_{k-1} > 0, \\ [(k+1)/2], & \varepsilon_{k-1} < 0. \end{cases} \tag{4.26}$$

The analog of the inequality (4.19) for Nevanlinna functions reads: If $n(z) \in \mathbf{N}^{\infty:2p}$ and k is as in Theorem 4.10, then

$$\operatorname{ind}_-(n) \geq \kappa_-(\Gamma_k). \tag{4.27}$$

This follows from the asymptotic expansion (4.25) of the kernel $L_n(z,w)$ and the inequality (2.10). A geometric proof can be given via formula (8.15) in Subsection 8.4 below.

4.5. Additional remarks and references

If R_{ij}, $i,j = 0,1,2,\ldots$, is the covariance $\ell \times \ell$ matrix function of a discrete second order $\ell \times 1$ vector-valued stochastic process, the $\ell \times \ell$ matrix function

$$S(z,w) = \sum_{i,j \geq 0} R_{ij} z^i w^{*j}$$

is called the *covariance generating function*. It is a nonnegative kernel in the open unit disk but in general it has no special structure. T. Kailath and H. Lev-Ari, see [101], [102], and [103], considered such functions when they are of the form

$$\sum_{i,j \geq 0} R_{ij} z^i w^{*j} = \frac{X(z)JX(w)^*}{1 - zw^*},$$

where $X(z)$ is a $\ell \times p$ matrix function and J is a $p \times p$ signature matrix. The corresponding stochastic processes contain as special cases the class of second order wide sense stationary stochastic processes, and are in some sense close to stationary stochastic processes. The case where $X(z)JX(w)^* = \varphi(z) + \varphi(w)^*$ (compare with (1.9)) corresponds to the case of wide sense stationary stochastic processes. Without loss of generality we may assume that

$$J = \begin{pmatrix} I_r & -I_s \end{pmatrix}, \quad p = r + s.$$

If $X(z)$ is of bounded type and written as

$$X(z) = \begin{pmatrix} A(z) & B(z) \end{pmatrix},$$

where $A(z)$ and $B(z)$ are $\ell \times r$ and $\ell \times s$ matrix functions, then we have using Leech's theorem

$$X(z) = A(z) \begin{pmatrix} I_r & -S(z) \end{pmatrix},$$

where $S(z)$ is a Schur $r \times s$ matrix function, which allows to write the kernel

$$\frac{X(z)JX(w)^*}{1-zw^*} = \alpha(z)\frac{\varphi(z)+\varphi(w)^*}{1-zw^*}\alpha(w)^*$$

for functions $\alpha(z)$ and $\varphi(z)$ of appropriate sizes, that is, the process is $\alpha(z)$-stationary in the sense of T. Kailath and H. Lev-Ari. For further details and more see [14, Section 4]. The assumption that $X(z)$ is of bounded type is in fact superfluous. Using Nevanlinna–Pick interpolation one also gets to the factorization $B(z) = -A(z)S(z)$.

In [21, Theorem 2.1] the following result is proved: Let J be a $p \times p$ signature matrix and let $X(z)$ be a $\ell \times p$ matrix function which is analytic in a neighborhood of the origin. Then the kernel

$$\frac{X(z)JX(w)^*}{1-zw^*} = \sum_{i,j\geq 0} R_{ij} z^i w^{*j}$$

is nonnegative if and only if all the finite sections $(R_{ij})_{i,j=0,\ldots,n}$, $n = 0, 1, \ldots$, are nonnegative, compare with Theorem 2.2.

We also mention the following result.

Theorem 4.11. *If $z_1 \in \mathbb{D}$, the formal power series*

$$\sum_{i=0}^{\infty} \sigma_i (z-z_1)^i$$

is the Taylor expansion of a generalized Schur function $s(z) \in \mathbf{S}_\kappa^{z_1}$ if and only if the matrix Γ, determined by the coefficients σ_i, $i = 0, 1, \ldots$, according to (4.10)*, has the property $\kappa_-(\Gamma) = \kappa$.*

For $\kappa = 0$, this result goes back to I. Schur ([116] and [117]) who proved it using the Schur transformation. If $\kappa > 0$ and the power series is convergent, this was proved in [96]. The general result for $\kappa \geq 0$ was proved by M. Pathiaux-Delefosse in [42, Theorem 3.4.1]) who used a generalized Schur transformation for generalized Schur functions, which we define in Subsection 5.1. Theorem 4.11 also appears in a slightly different form in [57, Theorem 3.1].

5. Generalized Schur functions: $z_1 \in \mathbb{D}$

5.1. The Schur transformation

Recall that J_c and $b_c(z)$ stand for the 2×2 signature matrix and the Blaschke factor related to the circle and $z_1 \in \mathbb{D}$:

$$J_c = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad b_c(z) = \frac{z-z_1}{1-zz_1^*}.$$

Suppose that $s(z) \in \mathbf{S}$ is not identically equal to a unimodular constant. The *Schur transform* $\widehat{s}(z)$ of $s(z)$ depends on whether $s(z)$ has a pole or not at the point z_1, and, if $s(z)$ is holomorphic at z_1, that is, $s(z) \in \mathbf{S}^{z_1}$, also on the first terms of the Taylor expansion (4.1)

$$s(z) = \sum_{i=0}^{\infty} \sigma_i (z - z_1)^i$$

of $s(z)$. It is defined as follows.

(i) If $s(z) \in \mathbf{S}^{z_1}$ and $|\sigma_0| < 1$, then

$$\widehat{s}(z) = \frac{1}{b_c(z)} \frac{s(z) - \sigma_0}{1 - s(z)\sigma_0^*}. \tag{5.1}$$

(ii) If $s(z) \in \mathbf{S}^{z_1}$ and $|\sigma_0| > 1$, then

$$\widehat{s}(z) = b_c(z) \frac{1 - s(z)\sigma_0^*}{s(z) - \sigma_0}. \tag{5.2}$$

(iii) If $s(z) \in \mathbf{S}^{z_1}$ and $|\sigma_0| = 1$, then

$$\widehat{s}(z) = \frac{\big(q(z) - (z - z_1)^k (1 - z z_1^*)^k\big) s(z) - \sigma_0 q(z)}{\sigma_0^* q(z) s(z) - \big(q(z) + (z - z_1)^k (1 - z z_1^*)^k\big)}, \tag{5.3}$$

where k is the smallest integer ≥ 1 such that $\sigma_k \neq 0$, and the polynomial $q(z)$ of degree $2k$ is defined as follows. Consider the polynomial $p(z)$ of degree $\leq k - 1$ determined by

$$p(z)\big(s(z) - \sigma_0\big) = \sigma_0 (z - z_1)^k (1 - z z_1^*)^k + \mathrm{O}\big((z - z_1)^{2k}\big), \quad z \to z_1, \tag{5.4}$$

and set $q(z) = p(z) - z^{2k} p(1/z^*)^*$.

(iv) If $s(z) \in \mathbf{S} \setminus \mathbf{S}^{z_1}$, that is, if $s(z) \in \mathbf{S}$ has a pole at z_1, then

$$\widehat{s}(z) = b_c(z) s(z). \tag{5.5}$$

This definition for $z_1 = 0$ of the Schur transformation first appears in the works [55], [63], [76], [78], and [42, Definition 3.3.1]. Note that $\widehat{s}(z)$ in (5.1) is holomorphic at z_1 whereas in the other cases $\widehat{s}(z)$ may have a pole at z_1. The function $\widehat{s}(z)$ in (5.2) is holomorphic at z_1 if and only if $\sigma_1 \neq 0$; it has a pole of order $q \geq 1$ if and only if $\sigma_1 = \cdots = \sigma_q = 0$ and $\sigma_{q+1} \neq 0$. As to item (iii): The integer $k \geq 1$ with $\sigma_k \neq 0$ exists because, by hypothesis, $s(z) \not\equiv \sigma_0$. The polynomial $q(z)$ satisfies

$$q(z) + z^{2k} q(1/z^*)^* = 0. \tag{5.6}$$

By substituting

$$p(z) = c_0 + c_1(z - z_1) + \cdots + c_{k-1}(z - z_1)^{k-1}$$

we see that it can equivalently and more directly be defined as follows:

$$\begin{aligned} q(z) = {} & c_0 + c_1(z - z_1) + \cdots + c_{k-1}(z - z_1)^{k-1} \\ & - \big(c_{k-1}^* z^{k+1}(1 - z z_1^*)^{k-1} + c_{k-2}^* z^{k+2}(1 - z z_1^*)^{k-2} + \cdots + c_0^* z^{2k}\big) \end{aligned}$$

with the coefficients $c_0, c_1, \ldots, c_{k-1}$ given by

$$c_0 \sigma_{k+i} + \cdots + c_i \sigma_k = \sigma_0 \binom{k}{i} (-z_1^*)^i (1 - |z_1|^2)^{k-i}, \quad i = 0, 1, \ldots, k-1. \tag{5.7}$$

Note that $c_0 \neq 0$. The denominator of the quotient in (5.3) has the asymptotic expansion:

$$\sigma_0^* q(z)\big(s(z)-\sigma_0\big) - (z-z_1)^k(1-zz_1^*)^k = \mathrm{O}\big((z-z_1)^{2k}\big), \quad z \to z_1. \tag{5.8}$$

We claim that it is not identically equal to 0. Indeed, otherwise we would have

$$s(z) = \sigma_0\left(1 + \frac{(1-z^{-1}z_1)^k(1-zz_1^*)^k}{z^{-k}p(z) - z^k p(1/z^*)^*}\right)$$

and therefore, since the quotient on the right-hand side is purely imaginary on $|z|=1$, the function $s(z)$ would not be bounded by 1 on $\mathbb{T}$, see (2.15). Hence $\widehat{s}(z)$ in (5.3) is well defined. From writing it in the form

$$\widehat{s}(z) = \sigma_0 - \frac{(z-z_1)^k(1-zz_1^*)^k(s(z)-\sigma_0)}{\sigma_0^* q(z)(s(z)-\sigma_0) - (z-z_1)^k(1-zz_1^*)^k}$$

and using that

$$s(z) - \sigma_0 = \sigma_k(z-z_1)^k + \mathrm{O}\big((z-z_1)^{k+1}\big), \quad z \to z_1,$$

where $\sigma_k \neq 0$, we readily see that it has a pole at z_1 of order q if and only if the denominator has the Taylor expansion

$$\sigma_0^* q(z)(s(z)-\sigma_0) - (z-z_1)^k(1-zz_1^*)^k = t_{2k+q}(z-z_1)^{2k+q} + \cdots \tag{5.9}$$

in which the coefficient t_{2k+q} is a nonzero complex number. The Schur transform $\widehat{s}(z)$ is holomorphic at z_1 if and only if the expansion (5.9) holds with $q=0$ or, equivalently, if

$$t_{2k} = \sigma_0^* \sum_{l=0}^{k-1} c_l \sigma_{2k-l} \neq 0. \tag{5.10}$$

It can be shown that necessarily $1 \leq k \leq \kappa$ and $0 \leq q \leq \kappa - k$.

If in the cases (ii) and (iii) the Schur transform $\widehat{s}(z)$ of $s(z)$ has a pole at z_1 of order q then by q times applying the Schur transformation to it according to case (iv), that is, by multiplying it by $b_c(z)^q$, we obtain a function

$$b_c(z)^q\,\widehat{s}(z)$$

which is holomorphic at z_1. We shall call this function the $q+1$-*fold composite Schur transform of* $s(z)$. In this definition we allow setting $q=0$: the 1-fold composite Schur transform of $s(z)$ exists if $\widehat{s}(z)$ is holomorphic at z_1 and then it equals $\widehat{s}(z)$.

The following theorem implies that the Schur transformation maps the set of functions of $\mathbf{S}$, which are not unimodular constants, into $\mathbf{S}$. In the cases (i) the negative index is retained, in the cases (ii)–(iv) it is reduced.

Theorem 5.1. *Let $s(z) \in \mathbf{S}$ and assume that it is not a unimodular constant. For its Schur transform $\widehat{s}(z)$ the following holds in the cases* (i)–(iv) *as above.*

(i) $s(z) \in \mathbf{S}_\kappa^{z_1} \quad \Longrightarrow \quad \widehat{s}(z) \in \mathbf{S}_\kappa^{z_1}$.

(ii) $s(z) \in \mathbf{S}_\kappa^{z_1} \quad \Longrightarrow \quad \kappa \geq 1$ and $\widehat{s}(z) \in \mathbf{S}_{\kappa-1}$.

(iii) $s(z) \in \mathbf{S}_\kappa^{z_1} \quad \Longrightarrow \quad 1 \le k \le \kappa$ and $\widehat{s}(z) \in \mathbf{S}_{\kappa-k}$.
(iv) $s(z) \in \mathbf{S}_\kappa \setminus \mathbf{S}^{z_1} \quad \Longrightarrow \quad \kappa \ge 1$ and $\widehat{s}(z) \in \mathbf{S}_{\kappa-1}$.

This theorem appears without proof in [76]. In [7] the theorem is proved by using realization theorems as in Subsection 5.5 and in [10] it is proved by applying Theorem 1.2 with $X(z)$ etc. given by (1.11).

The formulas (5.1)–(5.5) are all of the form

$$\widehat{s}(z) = \mathcal{T}_{\Phi(z)}(s(z))$$

for some rational 2×2 matrix function $\Phi(z)$. Indeed, in case (i) the matrix function $\Phi(z)$ can be chosen as

$$\Phi(z) = \begin{pmatrix} \dfrac{1}{b_c(z)} & -\dfrac{\sigma_0}{b_c(z)} \\ -\sigma_0^* & 1 \end{pmatrix},$$

in case (ii) as

$$\Phi(z) = \begin{pmatrix} -\sigma_0^* & 1 \\ \dfrac{1}{b_c(z)} & -\dfrac{\sigma_0}{b_c(z)} \end{pmatrix},$$

in case (iii) as

$$\begin{aligned}\Phi(z) &= \frac{1}{(z-z_1)^{2k}} \begin{pmatrix} -q(z) + (z-z_1)^k(1-zz_1^*)^k & \sigma_0 q(z) \\ -\sigma_0^* q(z) & q(z) - (z-z_1)^k(1-zz_1^*)^k \end{pmatrix} \\ &= \frac{1}{b_c(z)^k} I_2 - \frac{q(z)}{(z-z_1)^{2k}} \begin{pmatrix} 1 & -\sigma_0 \\ \sigma_0^* & -1 \end{pmatrix},\end{aligned}$$

and, finally, in case (iv) as

$$\Phi(z) = \begin{pmatrix} 1 & 0 \\ 0 & \dfrac{1}{b_c(z)} \end{pmatrix}.$$

In the following it is mostly not the Schur transformation itself but the matrix function $\Phi(z)$ and its inverse, normalized at some point $z_0 \in \mathbb{T}$ and chosen such that it has a pole at z_1^*, which plays a decisive role. Recall that for a 2×2 matrix function $\Psi(z)$ the range of the linear fractional transformation $\mathcal{T}_{\Psi(z)}$, when applied to all elements of a class $\mathbf{S}_\kappa$, is invariant if $\Psi(z)$ is replaced by $\alpha(z)\Psi(z)U$ where $\alpha(z)$ is a nonzero scalar function and U is a J_c-unitary constant. We normalize the four matrix functions $\Phi(z)$ considered above with some $z_0 \in \mathbb{T}$ and set

$$\Theta(z) = \Phi(z)^{-1}\Phi(z_0).$$

Theorem 5.2. *In cases* (i) *and* (ii)

$$\Theta(z) = I_2 + \left(\frac{b_c(z)}{b_c(z_0)} - 1\right)\frac{\mathbf{u}\mathbf{u}^* J_c}{\mathbf{u}^* J_c \mathbf{u}}, \quad \mathbf{u} = \begin{pmatrix} 1 \\ \sigma_0^* \end{pmatrix}, \tag{5.11}$$

in case (iii)

$$\Theta(z) = \left(\frac{b_c(z)}{b_c(z_0)}\right)^k I_2 + \frac{q_1(z)}{(1-zz_1^*)^{2k}}\mathbf{u}\mathbf{u}^* J_c, \quad \mathbf{u} = \begin{pmatrix} 1 \\ \sigma_0^* \end{pmatrix}, \tag{5.12}$$

where

$$q_1(z) = \frac{1}{b_c(z_0)^k} q(z) - \frac{q(z_0)}{(z_0 - z_1)^{2k}}(z-z_1)^k(1-zz_1^*)^k$$

is a polynomial of degree $\leq 2k$ *having the properties* $q_1(z_0) = 0$ *and*

$$b_c(z_0)^k q_1(z) + b_c(z_0)^{*k} z^{2k} q_1\left(1/z^*\right)^* = 0, \tag{5.13}$$

and, finally, in case (iv)

$$\Theta(z) = \begin{pmatrix} 1 & 0 \\ 0 & \dfrac{b_c(z)}{b_c(z_0)} \end{pmatrix} = I_2 + \left(\frac{b_c(z)}{b_c(z_0)} - 1\right)\frac{\mathbf{u}\mathbf{u}^* J_c}{\mathbf{u}^* J_c \mathbf{u}}, \quad \mathbf{u} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{5.14}$$

The proof of the theorem is straightforward. Property (5.13) of the polynomial $q_1(z)$ follows from (5.6). Note that in the cases (i), (ii), and (iii) in Theorem 5.2, we have

$$\mathbf{u}^* J_c \mathbf{u} = 1 - |\sigma_0|^2,$$

which is positive, negative, and $= 0$, respectively. In the latter case we have that if a and b are complex numbers with $a \neq 0$, then

$$\left(aI_2 + b\mathbf{u}\mathbf{u}^* J_c\right)^{-1} = \frac{1}{a} I_2 - \frac{b}{a^2}\mathbf{u}\mathbf{u}^* J_c$$

and this equality can be useful in proving formula (5.12). In case (iv) we have $\mathbf{u}^* J_c \mathbf{u} = -1$ and note that the formula for $\Theta(z)$ is the same as in cases (i) and (ii), but with a different 2×1 vector $\mathbf{u}$. The connection between $\Theta(z)$ and $s(z)$ in the next theorem follows from Theorem 1.1 with $X(z)$ etc. defined by (1.11):

$$X(z) = \begin{pmatrix} 1 & -s(z) \end{pmatrix}, \quad a(z) = 1, \quad b(z) = z, \quad J = J_c,$$

and hence from Theorem 3.10.

Theorem 5.3. *The four matrix functions* $\Theta(z)$ *in Theorem* 5.2 *can be chosen according to* (3.24) *as*

$$\Theta(z) = I_2 - (1 - zz_0^*)C(I - zA)^{-1}G^{-1}(I - z_0A)^{-*}C^* J_c,$$

with in cases (i) *and* (ii)

$$C = \begin{pmatrix} 1 \\ \sigma_0^* \end{pmatrix}, \quad A = z_1^*, \quad G = \frac{1 - |\sigma_0|^2}{1 - |z_1|^2},$$

in case (iii), *where* $\sigma_1 = \sigma_2 = \cdots = \sigma_{k-1} = 0$ *and* $\sigma_k \neq 0$,

$$C = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \sigma_0^* & 0 & \cdots & 0 & \sigma_k^* & \cdots & \sigma_{2k-1}^* \end{pmatrix}, \quad A = z_1^* I_{2k} + S_{2k},$$

S_{2k} being the $2k \times 2k$ shift matrix

$$S_{2k} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix},$$

and $G = \Gamma_{2k}$ the $2k \times 2k$ principal minor of the Pick matrix Γ in (4.10), *and finally, in case* (iv)

$$C = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad A = z_1^*, \quad G = \frac{-1}{1 - |z_1|^2}.$$

The proof of this theorem can be found in [9]. In cases (i), (ii), and (iii) G is the smallest invertible submatrix of the Pick matrix Γ, see Theorem 4.4 and formula (4.13). Clearly, the matrix functions $\Theta(z)$ in Theorems 5.2 and 5.3 are normalized elements in the class $\mathcal{U}_c^{z_1}$.

5.2. The basic interpolation problem

The *basic interpolation problem* for the class $\mathbf{S}_\kappa^{z_1}$ in its simplest form can be formulated as follows:

Problem 5.4. *Given $\sigma_0 \in \mathbb{C}$ and an integer $\kappa \geq 0$. Determine all functions $s(z) \in \mathbf{S}_\kappa^{z_1}$ with $s(z_1) = \sigma_0$.*

However, in this paper we seek the solution of this problem by means of the generalized Schur transformation of Subsection 5.1. Therefore it is natural to formulate it in a more complicated form in the cases (ii) and (iii) of Subsection 5.1.

In case (i), that is, if $|\sigma_0| < 1$, a formula for the set of all solutions $s(z) \in \mathbf{S}^{z_1}$ of Problem 5.4 can be given, see Theorem 5.5 below. If $|\sigma_0| \geq 1$ more information can be (or has to be) prescribed in order to get a compact solution formula. For example, in the case $|\sigma_0| > 1$ also an integer $k \geq 1$, and in the case $|\sigma_0| = 1$ additionally the following first $2k-1$ Taylor coefficients of the solution $s(z)$ with $k-1$ among them being zero and another nonnegative integer q can be prescribed. Therefore sometimes we shall speak of an *interpolation problem with augmented data*. In the following we always assume that the solution $s(z) \in \mathbf{S}^{z_1}$ has the Taylor expansion (4.1). We start with the simplest case.

Case (i): $|\sigma_0| < 1$.
For every integer $\kappa \geq 0$ the Problem 5.4 has infinitely many solutions $s(z) \in \mathbf{S}_\kappa^{z_1}$ as the following theorem shows.

Theorem 5.5. *If $|\sigma_0| < 1$ and κ is a nonnegative integer, then the formula*

$$s(z) = \frac{b_c(z)\widetilde{s}(z) + \sigma_0}{b_c(z)\sigma_0^*\widetilde{s}(z) + 1} \tag{5.15}$$

gives a one-to-one correspondence between all solutions $s(z) \in \mathbf{S}_\kappa^{z_1}$ *of Problem* 5.4 *and all parameters* $\widetilde{s}(z) \in \mathbf{S}_\kappa^{z_1}$.

If $s(z)$ and $\widetilde{s}(z)$ are related by (5.15), from the relation

$$s(z) - \sigma_0 = \frac{\left(1 - |\sigma_0|^2\right) b_c(z) \widetilde{s}(z)}{b_c(z)\sigma_0^* \widetilde{s}(z) + 1}$$

it follows that the function $s(z) - \sigma_0$ has a zero of order k at $z = z_1$ if and only if $\widetilde{s}(z)$ has a zero of order $k-1$ at $z = z_1$. From this and the fact, that for any integer $k \geq 1$ we have

$$\widetilde{s}(z) \in \mathbf{S}_\kappa^{z_1} \iff b_c(z)^k \widetilde{s}(z) \in \mathbf{S}_\kappa^{z_1},$$

it follows that the formula

$$s(z) = \frac{b_c(z)^k \widetilde{s}(z) + \sigma_0}{b_c(z)^k \sigma_0^* \widetilde{s}(z) + 1}$$

gives a one-to-one correspondence between all solutions $s(z) \in \mathbf{S}_\kappa^{z_1}$ for which $\sigma_1 = \sigma_2 = \cdots = \sigma_{k-1} = 0$, $\sigma_k \neq 0$, and all parameters $\widetilde{s}(z) \in \mathbf{S}_\kappa^{z_1}$ with $\widetilde{s}(z_1) \neq 0$.

Case (ii): $|\sigma_0| > 1$.
There are no solutions to Problem 5.4 in $\mathbf{S}_0$, since the functions in this class are bounded by 1. The following theorem shows that for each $\kappa \geq 1$ there are infinitely many solutions $s(z) \in \mathbf{S}_\kappa^{z_1}$. If $s(z)$ is one of them then $s(z) - \sigma_0$ has a zero of order $k \geq 1$ at $z = z_1$ and it can be shown that $k \leq \kappa$. To get a compact solution formula it is natural to consider k as an additional parameter.

Theorem 5.6. *If $|\sigma_0| > 1$, then for each integer k with $1 \leq k \leq \kappa$, the formula*

$$s(z) = \frac{\sigma_0 \widetilde{s}(z) + b_c(z)^k}{\widetilde{s}(z) + \sigma_0^* b_c(z)^k} \tag{5.16}$$

gives a one-to-one correspondence between all solutions $s(z) \in \mathbf{S}_\kappa^{z_1}$ *of Problem* 5.4 *with* $\sigma_1 = \sigma_2 = \cdots = \sigma_{k-1} = 0$ *and* $\sigma_k \neq 0$ *and all parameters* $\widetilde{s}(z) \in \mathbf{S}_{\kappa-k}^{z_1}$ *with* $\widetilde{s}(z_1) \neq 0$.

Case (iii): $|\sigma_0| = 1$.
By the maximum modulus principle, the constant function $s(z) \equiv \sigma_0$ is the only solution in $\mathbf{S}_0$. Before we describe the solutions in the classes $\mathbf{S}_\kappa^{z_1}$ with $\kappa \geq 1$, we formulate the problem again in full with all the augmented parameters.

Problem 5.7. *Given* $\sigma_0 \in \mathbb{C}$ *with* $|\sigma_0| = 1$, *an integer* k *with* $1 \leq k \leq \kappa$, *and numbers* $s_0, s_1, \ldots, s_{k-1} \in \mathbb{C}$ *with* $s_0 \neq 0$. *Determine all functions* $s(z) \in \mathbf{S}_\kappa^{z_1}$ *with* $s(z_1) = \sigma_0, \sigma_1 = s_0$ *if* $k = 1$, *and* $s(z_1) = \sigma_0$, $\sigma_1 = \cdots = \sigma_{k-1} = 0$, $\sigma_{k+j} = s_j$, $j = 0, 1, \ldots, k-1$, *if* $k > 1$.

To describe the solutions to this problem we need some notation, compare with case (iii) of the definition of the Schur transformation in Subsection 5.1. We associate with any k complex numbers $s_0 \neq 0, s_1, \dots, s_{k-1}$ the polynomial

$$\begin{aligned} q(z) &= q(z; s_0, s_1, \dots, s_{k-1}) \\ &= c_0 + c_1(z - z_1) + \cdots + c_{k-1}(z - z_1)^{k-1} \\ &\qquad - \big(c_{k-1}^* z^{k+1}(1 - zz_1^*)^{k-1} + c_{k-2}^* z^{k+2}(1 - zz_1^*)^{k-2} + \cdots + c_0^* z^{2k}\big) \end{aligned}$$

of degree $2k$, where the coefficients $c_0, c_1, \dots, c_{k-1}$ are determined by the formula

$$c_0 s_\ell + \cdots + c_\ell s_0 = \sigma_0 \binom{k}{\ell} (-z_1^*)^\ell \big(1 - |z_1|^2\big)^{k-\ell}, \quad \ell = 0, 1, \dots, k-1. \tag{5.17}$$

Theorem 5.8. *If $|\sigma_0| = 1$, for all integers κ and k with $1 \leq k \leq \kappa$ and any choice of complex numbers $s_0 \neq 0, s_1, \dots, s_{k-1}$, the formula*

$$s(z) = \frac{\big(q(z) + (z - z_1)^k (1 - zz_1^*)^k\big)\widetilde{s}(z) \, - \, \sigma_0 q(z)}{\sigma_0^* q(z)\widetilde{s}(z) \, - \, \big(q(z) - (z - z_1)^k (1 - zz_1^*)^k\big)} \tag{5.18}$$

with $q(z) = q(z; s_0, s_1, \dots, s_{k-1})$ gives a one-to-one correspondence between all solutions $s(z) \in \mathbf{S}_\kappa^{z_1}$ of Problem 5.7 and all parameters $\widetilde{s}(z) \in \mathbf{S}_{\kappa-k}$ with $\widetilde{s}(z_1) \neq \sigma_0$ if $\widetilde{s}(z) \in \mathbf{S}_{\kappa-k}^{z_1}$.

If $s(z)$ is a given by (5.18), then

$$s(z) - \sigma_0 = \frac{(z - z_1)^k (1 - zz_1^*)^k (\widetilde{s}(z) - \sigma_0)}{\sigma_0^* q(z)\widetilde{s}(z) \, - \, \big(q(z) - (z - z_1)^k (1 - zz_1^*)^k\big)},$$

which shows that k is the order of the zero of $s(z) - \sigma_0$ at $z = z_1$ and hence

$$\begin{aligned} &\sigma_0^* q(z; s_1, \dots, s_{k_1})(s(z) - \sigma_0) - (z - z_1)^k (1 - zz_1^*)^k \\ &\qquad = \frac{(z - z_1)^k (1 - zz_1^*)^k (s(z) - \sigma_0)}{\widetilde{s}(z) - \sigma_0} = \mathrm{O}\big((z - z_1)^{2k}\big), \quad z \to z_1. \end{aligned}$$

By comparing this relation with (5.8), we find that

$$q(z; s_0, \dots, s_{k-1}) = q(z; \sigma_k, \dots, \sigma_{2k-1})$$

and hence, on account of (5.7) and (5.17), that $\sigma_{k+j} = s_j$, $j = 0, \dots, k-1$, that is, $s(z)$ is a solution. If the parameter $\widetilde{s}(z)$ is holomorphic at z_1 and $\widetilde{s}(z_1) \neq \sigma_0$, then σ_{2k} satisfies an inequality, because of the inequality (5.10). If $\widetilde{s}(z)$ has a pole of order $q \geq 1$, then $q \leq \kappa - k$ and the coefficients $\sigma_{2k}, \sigma_{2k+1}, \dots, \sigma_{2k+q-1}$ of $s(z)$ are determined by $s_0, s_1, \dots, s_{k-1}$, and σ_{2k+q} satisfies an inequality.

Detailed proofs of the three theorems above can be found in [10]. The connection that exists between the basic interpolation problem on the one hand and cases (i), (ii), and (iii) of the Schur transformation on the other hand can be summed up as follows. Assume $s(z)$ belongs to $\mathbf{S}^{z_1}$ and is not identically equal to a unimodular constant, then it is a solution of a basic interpolation problem with its Taylor coefficients σ_j as data and hence can be written in the form (5.15), (5.16), or (5.18) depending on $|\sigma_0| < 1, > 1$, or $= 1$. The parameter $\widetilde{s}(z)$ in these formulas is the Schur transform of $s(z)$ in the cases (i), (ii) with $k = 1$ and (iii). Indeed, the parametrization formulas (5.15), (5.16), and (5.18) are simply the inverses of the formulas in the cases (i), (ii) with $k = 1$, and (iii) of the definition of the Schur transformation. In case (ii) with $k > 1$, $\widetilde{s}(z)/b_c(z)^{k-1}$ is the Schur transform of $s(z)$ and hence $\widetilde{s}(z)$ is the k-fold composite Schur transform of $s(z)$.

In the next theorem we rewrite the parametrization formulas as linear fractional transformations in terms of the function $\Theta(z)$ described in Theorem 5.2.

Theorem 5.9. *The parametrization formulas* (5.15), (5.16), *and* (5.18) *can be written in the form*

$$s(z) = \mathcal{T}_{\Psi(z)}(\widetilde{s}(z)),$$

where in case of formula (5.15)*:* $\Psi(z) = \Theta(z)\,U$ *in which* $\Theta(z)$ *is given by* (5.11) *and* U *is the* J_c*-unitary constant*

$$U = \frac{1}{\sqrt{1-|\sigma_0|^2}}\begin{pmatrix} b_c(z_0) & \sigma_0 \\ b_c(z_0)\sigma_0^* & 1 \end{pmatrix};$$

in case of formula (5.16)*:* $\Psi(z) = \Theta(z)\,U\,\Theta'(z)^{k-1}\,V$ *in which* $\Theta(z)$ *and* $\Theta'(z)$ *are given by* (5.11) *and* (5.14)*, respectively, and* U *and* V *are the* J_c*-unitary constants*

$$U = \frac{1}{\sqrt{|\sigma_0|^2-1}}\begin{pmatrix} \sigma_0 & b_c(z_0) \\ 1 & b_c(z_0)\sigma_0^* \end{pmatrix}, \quad V = \begin{pmatrix} 1 & 0 \\ 0 & b_c(z_0)^{k-1} \end{pmatrix};$$

and, finally, in case of formula (5.18)*:* $\Psi(z) = \Theta(z)\,U$ *in which* $\Theta(z)$ *is given by* (5.12) *and* U *is the* J_c*-unitary constant*

$$U = b_c(z_0)^k I_2 + \frac{q(z_0)}{(1-z_0 z_1^*)^{2k}}\mathbf{u}\mathbf{u}^* J_c, \quad \mathbf{u} = \begin{pmatrix} 1 \\ \sigma_0^* \end{pmatrix}.$$

5.3. Factorization in the class $\mathcal{U}_c^{z_1}$

Recall that, with $z_1 \in \mathbb{D}$, $\mathcal{U}_c^{z_1}$ stands for the class of those rational 2×2 matrix functions which are J_c-unitary on $\mathbb{T}$ and which have a unique pole at $1/z_1^*$. This class is closed under taking products, and by Theorem 3.13, products are automatically minimal. In the following theorem we describe the elementary factors of this class and the factorization of an arbitrary element of $\mathcal{U}_c^{z_1}$ into elementary factors. To this end, in this subsection we fix a point $z_0 \in \mathbb{T}$ in which the matrix functions will be normalized, see Subsection 3.4. Recall that $b_c(z)$ denotes the Blaschke factor

$$b_c(z) = \frac{z - z_1}{1 - z z_1^*}.$$

Theorem 5.10. (i) *A rational matrix function* $\Theta(z) \in \mathcal{U}_c^{z_1}$*, which is normalized by* $\Theta(z_0) = I_2$ *for some* $z_0 \in \mathbb{T}$*, is elementary if and only if it is of the form*

$$\Theta(z) = I_2 + \left(\frac{b_c(z)}{b_c(z_0)} - 1\right)\frac{\mathbf{u}\mathbf{u}^* J_c}{\mathbf{u}^* J_c \mathbf{u}}$$

for 2×1 *vector* $\mathbf{u}$ *with* $\mathbf{u}^* J_c \mathbf{u} \neq 0$*, or of the form*

$$\Theta(z) = \left(\frac{b_c(z)}{b_c(z_0)}\right)^k I_2 + \frac{q_1(z)}{(1 - z z_1^*)^{2k}}\mathbf{u}\mathbf{u}^* J_c,$$

where $k \geq 1$, $\mathbf{u}$ is a J_c-neutral nonzero 2×1 vector, and $q_1(z)$ is a polynomial of degree $\leq 2k$ with the properties $q_1(z_0) = 0$ and

$$b_c(z_0)^k q_1(z) + b_c(z_0)^{*k} z^{2k} q_1\left(1/z^*\right)^* = 0.$$

(ii) *Every $\Theta(z) \in \mathcal{U}_c^{z_1}$ can be written in a unique way as*

$$\Theta(z) = \left(\frac{b_c(z)}{b_c(z_0)}\right)^n \Theta_1(z) \cdots \Theta_m(z)\, U, \tag{5.19}$$

where n, m are nonnegative integers, the $\Theta_j(z)$, $j = 1, 2, \ldots, m$, are elementary factors in $\mathcal{U}_c^{z_1}$, normalized by $\Theta_j(z_0) = I_2$, and $U = \Theta(z_0)$ is a J_c-unitary constant.

The proof of this theorem, which can be found in [9, Theorem 5.4], is based on Theorems 5.3 for part (i) and on Theorem 3.15 for part (ii). If $\Theta(z)$ is elementary and has one of the forms given in part (i) of the theorem and U is a J_c-unitary constant, then $U\,\Theta(z)\,U^*$ is elementary and has the same form with $\mathbf{u}$ replaced by $U\mathbf{u}$.

We now outline how the factorization (5.19) of a matrix function $\Theta(z) \in \mathcal{U}_c^{z_1}$ can be obtained using the Schur algorithm. For further details and proofs we refer to [9, Section 6].

(a) First we normalize $\Theta(z)$ by writing $\Theta(z) = \Theta(z)\Theta(z_0)^{-1}\Theta(z_0)$. Then we take out a scalar factor $\big(b_c(z)/b_c(z_0)\big)^n$ from $\Theta(z)\Theta(z_0)^{-1}$ so that the remaining factor is not the zero matrix at z_1. Finally we split off a factor of the form

$$\begin{pmatrix} 1 & 0 \\ 0 & \dfrac{b_c(z)}{b_c(z_0)} \end{pmatrix}^r$$

to get the factorization

$$\Theta(z) = \left(\frac{b_c(z)}{b_c(z_0)}\right)^n \begin{pmatrix} 1 & 0 \\ 0 & \dfrac{b_c(z)}{b_c(z_0)} \end{pmatrix}^r \Psi(z)\Theta(z_0) \tag{5.20}$$

with $\Psi(z) \in \mathcal{U}_c^{z_1}$ having the properties $\Psi(z_0) = I_2$, $\Psi(z_1) \neq 0$, and, if

$$\Psi(z) = \begin{pmatrix} a(z) & b(z) \\ c(z) & d(z) \end{pmatrix},$$

then $|c(z_1)| + |d(z_1)| \neq 0$. If $\Psi(z)$ is constant, then (5.20) with $\Psi(z) = I_2$ is the desired factorization.

(b) Now assume that $\Psi(z)$ is not constant: $\deg \Psi > 0$. Choose a number $\tau \in \mathbb{T}$ such that

(b$_1$) $c(z_1)\tau + d(z_1) \neq 0$ and

(b$_2$) the function $s(z) = \dfrac{a(z)\tau + b(z)}{c(z)\tau + d(z)}$ is not a constant.

Condition (b_1) implies that $s(z)$ is holomorphic at z_1. Since $|c(z_1)| + |d(z_1)| \neq 0$, there is at most one $\tau \in \mathbb{T}$ for which (b_1) does not hold. We claim that there are at most two unimodular values of τ for which condition (b_2) does not hold. To see this assume that there are three different points $\tau_1, \tau_2, \tau_3 \in \mathbb{T}$ such that $s(z)$ is constant for $\tau = \tau_1, \tau_2, \tau_3$. Then, since $\Psi(z_0) = I_2$, we have

$$\frac{a(z)\tau_j + b(z)}{c(z)\tau_j + d(z)} \equiv \tau_j, \quad j = 1, 2, 3,$$

that is, the quadratic equation $c(z)\tau^2 + (d(z) - a(z))\tau - b(z) \equiv 0$ has three different solutions. It follows that $c(z) \equiv b(z) \equiv 0$ and $a(z) \equiv d(z)$, hence

$$a(z)^2 = d(z)^2 = \det \Psi(z).$$

Since, by Theorem 3.12, for some unimodular complex number c

$$\det \Psi(z) \;=\; c\, b_c(z)^{\deg \Psi}$$

and, by assumption, $\deg \Psi > 0$ we see that $\Psi(z_1) = 0$ which is in contradiction with one of the properties of $\Psi(z)$. This proves the claim. We conclude that $s(z)$ has the properties (b_1) and (b_2) for all but three values of $\tau \in \mathbb{T}$.

Since $\Psi(z) \in \mathcal{U}_c^{z_1}$, the function $s(z)$ belongs to the class $\mathbf{S}^{z_1}$. It is not identically equal to a unimodular constant, so we can apply the Schur algorithm:

$$s_0(z) = s(z),\;\; s_1(z) = \mathcal{T}_{\Psi_1(z)^{-1}}(s_0(z)),\;\; s_2(z) = \mathcal{T}_{\Psi_2(z)^{-1}}(s_1(z)), \ldots,$$

$$s_q(z) = \mathcal{T}_{\Psi_q(z)^{-1}}(s_{q-1}(z))$$

where the $\Psi_j(z)$'s are as in Theorem 5.9 and, hence, apart from constant J_c-unitary factors, elementary factors or products of elementary factors. The algorithm stops, because after finitely many, say q, iterations the function $s_q(z)$ is a unimodular constant. Moreover, it can be shown that

$$\Psi(z) = \Psi_1(z)\Psi_2(z)\cdots\Psi_q(z)V, \tag{5.21}$$

where V is a J_c-unitary constant.

(c) Via Steps (a) and (b) we have obtained a factorization of $\Theta(z)$ of the form

$$\Theta(z) = \Omega_1(z)\Omega_2(z)\cdots\Omega_m(z)V\Theta(z_0),$$

in which each of the factors $\Omega_j(z)$ is elementary but not necessarily normalized. The desired normalized factorization can now be obtained by the formulas

$$\begin{aligned}
\Theta_1(z) &= \Omega_1(z)\Omega_1(z_0)^{-1}, \\
\Theta_2(z) &= \Omega_1(z_0)\Omega_2(z)\Omega_2(z_0)^{-1}\Omega_1(z_0)^{-1}, \\
\Theta_3(z) &= \Omega_1(z_0)\Omega_2(z_0)\Omega_3(z)\Omega_3(z_0)^{-1}\Omega_2(z_0)^{-1}\Omega_1(z_0)^{-1}
\end{aligned}$$

etc., ending with $\Omega_m(z_0)^{-1}\cdots\Omega_1(z_0)^{-1}V = I_2$.

The basic idea why the above procedure works is that, by Theorem 1.2,

$$\begin{aligned}
\big(1 \quad -s\big)\,\mathcal{P}(\Psi) &= \mathcal{P}(s)\\
&= \big(1 \quad -s\big)\,\mathcal{P}(\Psi_1) \oplus (a_1 - c_1 s)\mathcal{P}(s_1)\\
&\;\;\vdots\\
&= \big(1 \quad -s\big)\,\big(\mathcal{P}(\Psi_1) \oplus \Psi_1\mathcal{P}(\Psi_2)\big) \oplus (a_1 - c_1 s)(a_2 - c_2 s_1)\mathcal{P}(s_2)\\
&= \big(1 \quad -s\big)\,\big(\mathcal{P}(\Psi_1) \oplus \Psi_1\mathcal{P}(\Psi_2) \oplus \Psi_1\Psi_2\mathcal{P}(\Psi_3) \oplus \cdots\big)\\
&= \big(1 \quad -s\big)\,\mathcal{P}(\Psi_1\Psi_2\cdots\Psi_q),
\end{aligned}$$

where

$$\Psi_j(z) = \begin{pmatrix} a_j(z) & b_j(z)\\ c_j(z) & d_j(z)\end{pmatrix}, \quad j = 1, 2, \ldots.$$

Hence $\mathcal{P}(\Psi) = \mathcal{P}(\Psi_1\Psi_2\cdots\Psi_q)$ and this implies (5.21).

5.4. Realization

The realizations of functions $s(z) \in \mathbf{S}^{z_1}$ which we consider in this section are given by formula (2.18):

$$s(z) = \gamma + b_c(z)\big\langle(1 - b_c(z)T)^{-1}u, v\big\rangle, \quad b_c(z) = \frac{z - z_1}{1 - zz_1^*}; \tag{5.22}$$

here γ is a complex number: $\gamma = s(z_1)$, T is a bounded operator in some Pontryagin space $(\mathcal{P}, \langle\,\cdot\,,\,\cdot\,\rangle)$, u and v are elements from $\mathcal{P}$. With the entries of (5.22) we form the operator matrix (2.19)

$$\mathcal{V} = \begin{pmatrix} T & u\\ \langle\,\cdot\,, v\rangle & \gamma\end{pmatrix} : \begin{pmatrix}\mathcal{P}\\ \mathbb{C}\end{pmatrix} \to \begin{pmatrix}\mathcal{P}\\ \mathbb{C}\end{pmatrix}.$$

In the rest of this section we are interested in the effect of the Schur transformation on the realizations, that is, we describe the realizations $\widehat{\mathcal{V}}$ of the Schur transform $\widehat{s}(z)$ of $s(z)$ or the realizations $\widetilde{\mathcal{V}}$ of the composite Schur transform $\widetilde{s}(z)$ by means of the realizations $\mathcal{V}$ of the given $s(z)$. The composite Schur transform is defined in Subsection 5.1. By definition it is holomorphic at z_1, in particular the 1-fold composite Schur transform of $s(z)$ is defined if $\widehat{s}(z)$ is holomorphic at z_1 and then it is equal to $\widehat{s}(z)$.

We consider only the closely outerconnected coisometric case and formulate the results related to the cases (i), (ii), and (iii) of the definition of the Schur transformation as separate theorems. For proofs of these theorems and of the theorems for the closely innerconnected isometric and the closely connected unitary cases, see [7], [8], [11], [125], and [126]. Recall that if $s(z) \in \mathbf{S}^{z_1}$, we denote its Taylor expansion around z_1 by (4.1):

$$s(z) = \sum_{i=0}^{\infty} \sigma_i (z - z_1)^i.$$

Theorem 5.11. *Assume $s(z) \in \mathbf{S}^{z_1}$ with $|\sigma_0| < 1$ and let $\widehat{s}(z)$ be the Schur transform of $s(z)$. If* (5.22) *is the closely outerconnected coisometric realization of $s(z)$, then*

(i) $\operatorname{span}\{v\}$ *is a* 1*-dimensional positive subspace of $\mathcal{P}$, so that the space*

$$\widehat{\mathcal{P}} = \mathcal{P} \ominus \operatorname{span}\{v\}$$

and the orthogonal projection P in $\mathcal{P}$ onto $\widehat{\mathcal{P}}$ are well defined, and

(ii) *with*

$$\begin{aligned} \widehat{T} &= PTP, & \widehat{u} &= \frac{1}{\sqrt{1-|\gamma|^2}}\,Pu, \\ \widehat{v} &= \frac{1}{\sqrt{1-|\gamma|^2}}\,PT^*v, & \widehat{\gamma} &= \frac{\langle u, v\rangle}{1-|\gamma|^2} \end{aligned}$$

the formula

$$\widehat{s}(z) = \widehat{\gamma} + b_c(z)\big\langle (1 - b_c(z)\widehat{T})^{-1}\widehat{u}, \widehat{v}\big\rangle$$

is the closely outerconnected coisometric realization of $\widehat{s}(z)$.

Moreover, $\operatorname{ind}_-(\widehat{\mathcal{P}}) = \operatorname{ind}_-(\mathcal{P})$.

Theorem 5.12. *Assume $s(z) \in \mathbf{S}^{z_1}$ with $|\sigma_0| > 1$, denote by k the smallest integer ≥ 1 such that $\sigma_k \neq 0$, and let $\widetilde{s}(z)$ be the k-fold composite Schur transform of $s(z)$. If* (5.22) *is the closely outerconnected coisometric realization of $s(z)$, then*

(i) $\operatorname{span}\{v, T^*v, \ldots, T^{*(k-1)}v\}$ *is a k-dimensional negative subspace of $\mathcal{P}$, so that the space*

$$\widetilde{\mathcal{P}} = \mathcal{P} \ominus \operatorname{span}\{v, T^*v, \ldots, T^{*(k-1)}v\}$$

and the orthogonal projection P in $\mathcal{P}$ onto $\widetilde{\mathcal{P}}$ are well defined, and

(ii) *with*

$$\begin{aligned} \widetilde{T} &= PTP - \frac{\big\langle \,\cdot\,, PT^{*k}v\big\rangle}{\sigma_k}Pu, & \widetilde{u} &= \frac{\sqrt{|\gamma|^2-1}}{\sigma_k}Pu, \\ \widetilde{v} &= \frac{\sqrt{|\gamma|^2-1}}{\sigma_k^*}PT^{*k}v, & \widehat{\gamma} &= \frac{1-|\gamma|^2}{\sigma_k} \end{aligned}$$

the formula

$$\widetilde{s}(z) = \widetilde{\gamma} + b_c(z)\big\langle (1 - b_c(z)\widetilde{T})^{-1}\widetilde{u}, \widetilde{v}\big\rangle$$

is the closely outerconnected coisometric realization of $\widehat{s}(z)$.

Moreover, $\operatorname{ind}_-(\widetilde{\mathcal{P}}) = \operatorname{ind}_-(\mathcal{P}) - k$.

The complex number t_{2k+q} in the next theorem is the nonzero coefficient in the expansion (5.9); if $q = 0$, then t_{2k} is given by (5.10).

Theorem 5.13. *Assume $s(z) \in \mathbf{S}^{z_1}$ with $|\sigma_0| = 1$, denote by k the smallest integer ≥ 1 such that $\sigma_k \neq 0$ and let $\widetilde{s}(z)$ be the $q+1$-fold composite Schur transform of $s(z)$, where q is the order of the pole of the Schur transform of $s(z)$. If* (2.18) *is the closely outerconnected coisometric realization of $s(z)$, then*

(i) *the space* $\operatorname{span}\{v, T^*v, \ldots, T^{*(2k+q-1)}v\}$ *is a* $(2k+q)$*-dimensional Pontryagin subspace of* $\mathcal{P}$ *with negative index equal to* $k+q$*, so that the space*

$$\widetilde{\mathcal{P}} = \mathcal{P} \ominus \operatorname{span}\{v, T^*v, \ldots, T^{*(2k+q-1)}v\}$$

and the orthogonal projection P *in* $\mathcal{P}$ *onto* $\widetilde{\mathcal{P}}$ *are well defined, and*

(ii) *with*

$$\begin{array}{rclrcl} \widetilde{T} & = & PTP + \dfrac{1}{\sigma_k t_{2k+q}} \left\langle \,\cdot\,, PT^{*(2k+q)}v \right\rangle Pu, & \widetilde{u} & = & \dfrac{1}{t_{2k+q}} P_{2k+q} u, \\ \widetilde{v} & = & \dfrac{1}{t^*_{2k+q}} PT^{*(2k+q)}v, & \widetilde{\gamma} & = & \dfrac{\sigma_k}{t_{2k+q}} \end{array}$$

the formula

$$\widetilde{s}(z) = \widetilde{\gamma} + b_c(z)\langle (1 - b_c(z)\widetilde{T})^{-1}\widetilde{u}, \widetilde{v}\rangle$$

is the closely outerconnected coisometric realization of $\widetilde{s}(z)$*.*

Moreover, $\operatorname{ind}_-(\widetilde{\mathcal{P}}) = \operatorname{ind}_-(\mathcal{P}) - k - q$*. If* $q = 0$*, then* (i) *and* (ii) *hold with* $q = 0$ *and* $\widetilde{\gamma}$ *replaced by*

$$\widetilde{\gamma} = \sigma_0 - \frac{\sigma_k}{t_{2k}}.$$

Theorem 5.1 follows from the last statements in the previous theorems. These theorems can also be used to give a geometric proof of the following result, see [7, Section 9]. It first appeared in [42, Lemma 3.4.5] with an analytic proof and implies that after finitely many steps the Schur algorithm applied to $s(z) \in \mathbf{S}$ only yields classical Schur functions.

Theorem 5.14. *Let* $s(z)$ *be a generalized Schur function which is not a unimodular constant and set*

$$s_0(z) = s(z), \quad s_j(z) = \widehat{s}_{j-1}(z), \; j = 1, 2, \ldots.$$

Then there is an index j_0 *such that* $s_j(z) \in \mathbf{S}_0$ *for all integers* $j \geq j_0$*.*

5.5. Additional remarks and references

It is well known that there is one-to-one correspondence between the class of Schur functions $s(z)$ and the set of sequences of Schur parameters $(\rho_j)_j \geq 0$ defined via the Schur algorithm centered at $z = 0$ applied to $s(z)$ by the formulas

$$s_0(z) = s(z), \quad \rho_0 = s_0(0)$$

and for $j = 0, 1, \ldots$, by (1.2):

$$s_{j+1}(z) = \widehat{s}_j(z) = \frac{1}{z}\frac{s_j(z) - s_j(0)}{1 - s_j(z)s_j(0)^*}, \quad \rho_{j+1} = s_{j+1}(0),$$

see for example [82, Section 9]. This has been generalized to generalized Schur functions and sequences of augmented Schur parameters in [72].

The Schur algorithm is also related to the study of inverse problems, see [87], [52], and [51], and to the theory of discrete first order systems of the form

$$X_{n+1}(z) = \begin{pmatrix} 1 & -\rho_n \\ -\rho_n^* & 1 \end{pmatrix} \begin{pmatrix} z & 0 \\ 0 & 1 \end{pmatrix} X_n(z),$$

see [2], [27], and [28]. All these works should have counterparts in the indefinite settings; we leave this question to a forthcoming publication.

A real algebraic integer $\theta > 1$ is called a Pisot–Vijayaraghavan number if all its conjugates are in the open unit disk (note that the definition is for an algebraic integer, and so in the minimal polynomial equation which defines θ and its conjugates the coefficient of the highest power of the indeterminate is equal to 1). If at least one of the conjugates of θ lies on the unit circle, θ is called a Salem number. These numbers were studied first by Ch. Pisot, R. Salem and J. Dufresnoy[1]; they have various important properties which play a role, for instance, in the study of uniqueness sets for trigonometric series, see [109], [110], [112], and [111]. After the paper [77] J. Dufresnoy and Ch. Pisot introduced in [78] new methods, and, in particular, relations with meromorphic functions and generalized Schur functions.

6. Generalized Schur functions: $z_1 \in \mathbb{T}$

6.1. The Schur transformation

The Schur transformation centered at the point $z_1 \in \mathbb{T}$ will be defined for the functions from the class $\mathbf{S}^{z_1;2p}$, where p is an integer ≥ 1. First we introduce some notation and recall some facts along the way.

Assume $s(z)$ belongs to $\mathbf{S}^{z_1;2p}$ with asymptotic expansion (2.16):

$$s(z) = \tau_0 + \sum_{i=1}^{2p-1} \tau_j (z - z_1)^i + \mathrm{O}\left((z - z_1)^{2p}\right), \quad z \hat{\to} z_1,$$

where the coefficients τ_j satisfy the conditions (1)–(3) of Subsection 2.3. Denote by Γ_p the Hermitian $p \times p$ Pick matrix associated with the kernel $K_s(z, w)$ at z_1, see Theorem 4.5. Let k be the smallest integer ≥ 1 such that $\tau_k \neq 0$. Then $k \leq p$ and $k = k_0(\Gamma_p)$, that is, k is the smallest integer $j \geq 1$ for which the $j \times j$ principal submatrix $\Gamma_j := (\Gamma_p)_j$ of Γ_p is invertible, and the Hermitian $k \times k$ matrix Γ_k has the form (4.16). Whereas the Schur transformation with an interior point $z_1 \in \mathbb{D}$ in the cases (i): $\Gamma_1 > 0$, (ii) $\Gamma_1 < 0$, and (iii) $\Gamma_1 = 0$ had different forms, in the case $z_1 \in \mathbb{T}$ (and for Nevanlinna functions in the case $z_1 = \infty$, see Subsection 8.1), the transformation formula can be written in the same form in all three cases.

We define the vector function

$$R(z) = \begin{pmatrix} \dfrac{1}{1 - z z_1^*} & \dfrac{z}{(1 - z z_1^*)^2} & \cdots & \dfrac{z^{k-1}}{(1 - z z_1^*)^k} \end{pmatrix},$$

[1] They were first discovered by A. Thue and G. Hardy, see [42, Preface].

fix some normalization point $z_0 \in \mathbb{T}$, $z_0 \neq z_1$, and introduce the polynomial $p(z)$ by

$$p(z) = (1 - zz_1^*)^k \, R(z)\Gamma_k^{-1} R(z_0)^*.$$

It has the properties

$$\deg p(z) \leq k-1, \quad p(z_1) \neq 0,$$

and

$$p(z) - z_0(-z_1^*)^k z^{k-1} p(1/z^*)^* = 0.$$

The asymptotic formula

$$\tau_0 \frac{(1 - zz_1^*)^k}{(1 - zz_0^*)p(z)} = -\sum_{i=k}^{2k-1} \tau_i (z - z_1)^i + \mathrm{O}\big((z - z_1)^{2k}\big), \quad z \hat{\to} z_1,$$

shown in [18, Lemma 3.1], is the analog of (5.4). Now the Schur transform $\widehat{s}(z)$ of $s(z)$ is defined by the formula

$$\widehat{s}(z) = \frac{\big((1 - zz_1^*)^k + (1 - zz_0^*)p(z)\big)s(z) - \tau_0(1 - zz_0^*)p(z)}{\tau_0^*(1 - zz_0^*)p(z)s(z) + \big((1 - zz_1^*)^k - (1 - zz_0^*)p(z)\big)}. \tag{6.1}$$

Note that the numerator and the denominator both tend to 0 when $z \hat{\to} z_1$. The denominator cannot be identically equal to 0. Indeed, if it would be then

$$s(z) = \tau_0 \left(1 - \frac{(1 - zz_1^*)^k}{(1 - zz_0^*)p(z)}\right)$$

and hence $s(z)$ would have a pole at z_0 in contradiction with (2.15). In the particular case $k = 1$ we have that $\Gamma_1 = \gamma_{00} = \tau_0^* \tau_1 z_1$, see (4.17), is a nonzero real number and the polynomial $p(z)$ is a constant:

$$p(z) = \frac{1}{\tau_0^* \tau_1 z_1 (1 - z_0^* z_1)}.$$

Recall that the number $\kappa_-(\Gamma_k)$ of negative eigenvalues of the Hermitian matrix Γ_k is given by (4.18).

Theorem 6.1. *Assume $s(z) \in \mathbf{S}_\kappa^{z_1;2p}$ is not equal to a unimodular constant and let k be the smallest integer ≥ 1 such that $\tau_k \neq 0$. Then $\kappa_-(\Gamma_k) \leq \kappa$ and for the Schur transform $\widehat{s}(z)$ from (6.1) it holds $\widehat{s}(z) \in \mathbf{S}_{\widehat{\kappa}}$ with*

$$\widehat{\kappa} = \kappa - \kappa_-(\Gamma_k).$$

This theorem follows from Theorem 6.3 and the relation (6.5) in the next subsection. Formula (6.1) for the Schur transformation can be written as the linear fractional transformation

$$\widehat{s}(z) = \mathcal{T}_{\Phi(z)}(s(z))$$

with

$$\begin{aligned}\Phi(z) &= \frac{1}{(1 - zz_1^*)^k} \begin{pmatrix} (1 - zz_1^*)^k + (1 - zz_0^*)p(z) & \tau_0(1 - zz_0^*)p(z) \\ \tau_0^*(1 - zz_0^*)p(z) & (1 - zz_1^*)^k - (1 - zz_0^*)p(z) \end{pmatrix} \\ &= I_2 + \frac{(1 - zz_0^*)p(z)}{(1 - zz_1^*)^k} \mathbf{u}\mathbf{u}^* J_c, \quad \mathbf{u} = \begin{pmatrix} \tau_0^* \\ 1 \end{pmatrix}, \quad J_c = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.\end{aligned}$$

Hence the inverse Schur transformation of (6.1) is given by

$$s(z) = \mathcal{T}_{\Theta(z)}(\widehat{s}(z)),$$

where

$$\Theta(z) \;=\; \Phi(z)^{-1} \;=\; I_2 - \frac{(1-zz_0^*)p(z)}{(1-zz_1^*)^k}\mathbf{u}\mathbf{u}^* J_c, \quad \mathbf{u} = \begin{pmatrix} \tau_0^* \\ 1 \end{pmatrix}.$$

The connection between $\Theta(z)$ and $s(z)$ follows from Theorem 1.1 with $z_1 \in \mathbb{T}$ and $X(z)$ etc. given by (1.11), and Theorem 3.10. This implies that $\Theta(z)$ can be written in the form (3.24):

$$\Theta(z) = I_2 - (1-zz_0^*)C(I-zA)^{-1}G^{-1}(I-z_0A)^{-*}C^* J_c$$

with

$$C = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \sigma_0^* & 0 & \cdots & 0 \end{pmatrix}, \quad A = z_1^* I_k + S_k, \quad G = \Gamma_k. \tag{6.2}$$

It follows that $\Theta(z)$ is normalized and belongs to $\mathcal{U}_c^{z_1}$.

6.2. The basic boundary interpolation problem

The basic boundary interpolation problem for generalized Schur functions can be formulated as follows.

Problem 6.2. *Given $z_1 \in \mathbb{T}$, an integer $k \geq 1$, and complex numbers τ_0, τ_k, $\tau_{k+1}, \ldots, \tau_{2k-1}$ with $|\tau_0| = 1$, $\tau_k \neq 0$ and such that the $k \times k$ matrix Γ_k in* (4.16) *is Hermitian. Determine all functions $s(z) \in \mathbf{S}$ such that*

$$s(z) = \tau_0 + \sum_{i=k}^{2k-1} \tau_i (z-z_1)^i + \mathrm{O}((z-z_1)^{2k}), \quad z \hat{\to} z_1.$$

If $s(z)$ is a solution of the problem, then it belongs to some class $\mathbf{S}_\kappa^{z_1;2k}$ where κ is an integer $\geq \kappa_-(\Gamma_k)$, see (4.19). With the data of the problem and a fixed point $z_0 \in \mathbb{T} \setminus \{z_1\}$, we define the polynomial $p(z)$ as in Subsection 6.1.

Theorem 6.3. *The linear fractional transformation*

$$s(z) = \frac{\big((1-zz_1^*)^k - (1-zz_0)p(z)\big)\,\widetilde{s}(z) + \tau_0(1-zz_0^*)}{-\tau_0^*(1-zz_0^*)p(z)\widetilde{s}(z) + (1-zz_1^*)^k + (1-zz_0^*)p(z)} \tag{6.3}$$

establishes a one-to-one correspondence between all solutions $s(z) \in \mathbf{S}_\kappa^{z_1;2k}$ of Problem 6.2 *and all parameters $\widetilde{s}(z) \in \mathbf{S}_{\widetilde{\kappa}}$ with the property*

$$\liminf_{z \hat{\to} z_1} |\widetilde{s}(z) - \tau_0| > 0, \tag{6.4}$$

where

$$\widetilde{\kappa} = \kappa - \kappa_-(\Gamma_k).$$

For a proof of this theorem and a generalization of it to multipoint boundary interpolation, see [18, Theorem 3.2]. In the particular case that the parameter $\widetilde{s}(z)$ is rational the inequality (6.4) is equivalent to the fact that the denominator in (6.3):

$$-\tau_0^*(1-zz_0^*)p(z)(\widetilde{s}(z)-\tau_0)+(1-zz_1^*)^k$$

is not zero at $z=z_1$.

Note that the linear fractional transformation (6.3) is the inverse of the Schur transformation and

$$\widetilde{s}(z)=\mathcal{T}_{\Theta(z)^{-1}}(s(z))=\widehat{s}(z). \tag{6.5}$$

6.3. Factorization in the class $\mathcal{U}_c^{z_1}$

We repeat that $\mathcal{U}_c^{z_1}$ with $z_1\in\mathbb{T}$ is the class of all rational 2×2 matrix functions which are J_c-unitary on $\mathbb{T}\setminus\{z_1\}$ and have a unique pole in z_1. Since $\mathcal{U}_c^{z_1}$ is closed under taking inverses, products of elements from this class need not be minimal. To describe the elementary factors of $\mathcal{U}_c^{z_1}$ we fix a normalization point z_0 in $\mathbb{T}\setminus\{z_1\}$.

Theorem 6.4.

(i) *A normalized matrix function $\Theta(z)\in\mathcal{U}_c^{z_1}$ is elementary if and only if it is of the form*

$$\Theta(z)=I_2-\frac{(1-zz_0^*)p(z)}{(1-zz_1^*)^k}\mathbf{u}\mathbf{u}^*J_c,\quad J_c=\begin{pmatrix}1&0\\0&-1\end{pmatrix},$$

where k is an integer ≥ 1, $\mathbf{u}$ is a J_c-neutral nonzero 2×1 vector: $\mathbf{u}^*J_c\mathbf{u}=0$, and $p(z)$ is a polynomial of degree $\leq k-1$ satisfying $p(z_1)\neq 0$ and

$$p(z)=z_0(-z_1^*)^kz^{k-1}p(1/z^*)^*.$$

(ii) *Every $\Theta(z)\in\mathcal{U}_c^{z_1}$ admits a unique minimal factorization*

$$\Theta(z)=\Theta_1(z)\cdots\Theta_n(z)U,$$

in which each factor $\Theta_j(z)$ is a normalized elementary matrix function from $\mathcal{U}_c^{z_1}$ and $U=\Theta(z_0)$ is a J_c-unitary constant.

A proof of this theorem is given in [18, Theorem 5.2]. Part (ii) follows from Theorem 3.17. Part (i) is related to (3.24) with C, A and G as in (6.2). It shows that the function $\Theta(z)$ associated with the Schur transformation and the basic interpolation problem in the previous subsections is a normalized elementary factor in $\mathcal{U}_c^{z_1}$. In the positive case the factors of degree 1 are called *Brune sections* or *Potapov–Blaschke sections of the third kind*, see [67].

We sketch how to obtain the factorization of an arbitrary $\Theta(z)\in\mathcal{U}_c^{z_1}$ via the Schur algorithm. A proof that the procedure works can be found in [18, Section 6].

(a) We first normalize $\Theta(z)$ and write

$$\Theta(z)=\Psi(z)\Theta(z_0),\quad \Psi(z)=\Theta(z)\Theta(z_0)^{-1}=\begin{pmatrix}a(z)&b(z)\\c(z)&d(z)\end{pmatrix}.$$

Assume that $\Psi(z)$ is not a J_c-unitary constant, otherwise the procedure stops right here. We denote by $o_{z_1}(g)$ the order of the pole of the function $g(z)$ at z_1. We choose $\tau \in \mathbb{T}$ such that

(a$_1$) $c(0)\tau + d(0) \neq 0$,

(a$_2$) $o_{z_1}(a\tau + b) = \max\{o_{z_1}(a), o_{z_1}(b)\}$,

(a$_3$) $o_{z_1}(c\tau + d) = \max\{o_{z_1}(c), o_{z_1}(d)\}$, and

(a$_4$) the function $s(z) = \dfrac{a(z)\tau + b(z)}{c(z)\tau + d(z)}$ is not identically equal to a constant.

Each of the first three conditions holds for all but at most one value of τ. The fourth condition holds for all except two values. The argument here is similar to the one given in Subsection 5.3; now one uses that $\det \Psi(z) = 1$, see Theorem 3.12. So all in all there are at most five forbidden values for $\tau \in \mathbb{T}$. Since $\Psi(z) \in \mathcal{U}_c^{z_1}$, $s(z)$ is a rational generalized Schur function and therefore it is holomorphic on $\mathbb{T}$ and satisfies $|s(z)| = 1$ for all $z \in \mathbb{T}$, that is, $s(z)$ is the quotient of two Blaschke factors. It follows that the kernel $K_s(z,w)$ has an asymptotic expansion (4.15) for any integer $p \geq 1$. Since it is symmetric in the sense that $K_s(z,w)^* = K_s(w,z)$, the corresponding Pick matrices Γ of all sizes are Hermitian. Thus we can apply the Schur algorithm to $s(z)$ and continue as in Steps (b) and (c) in Subsection 5.3.

6.4. Additional remarks and references

The analogs of the realization theorems as in, for instance, Subsection 5.5 have yet to be worked out. The results of the present section can be found in greater details in [18]. For boundary interpolation in the setting of Schur functions we mention the book [36] and the paper [114]. The case of boundary interpolation for generalized Schur functions was studied in [35].

A nonconstant function $s(z) \in \mathbf{S_0}$ has in $z_1 \in \mathbb{T}$ a Carathéodory derivative, if the limits

$$\tau_0 = \lim_{z \hat{\to} z_1} s(z) \text{ with } |\tau_0| = 1, \quad \tau_1 = \lim_{z \hat{\to} z_1} \frac{s(z) - \tau_0}{z - z_1} \tag{6.6}$$

exist, and then

$$\lim_{z \hat{\to} z_1} s'(z) = \tau_1.$$

The relation (6.6) is equivalent to the fact that the limit

$$\lim_{z \hat{\to} z_1} \frac{1 - |s(z)|}{1 - |z|}$$

exists and is finite and positive; in this case it equals

$$\Gamma_1 = \tau_0^* \tau_1 z_1,$$

see [113, p. 48]. Thus Theorem 6.3 is a generalization of the interpolation results in [36] and [114] to an indefinite setting.

7. Generalized Nevanlinna functions: $z_1 \in \mathbb{C}^+$

7.1. The Schur transformation

Throughout this section the 2×2 matrix J_ℓ and the Blaschke factor $b_\ell(z)$ related to the real line and $z_1 \in \mathbb{C}^+$ are defined by

$$J_\ell = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad b_\ell(z) = \frac{z - z_1}{z - z_1^*}.$$

The Taylor expansion of a function $n(z) \in \mathbf{N}^{z_1}$ will be written as in (4.20):

$$n(z) = \sum_{j=0}^{\infty} \nu_j (z - z_1)^j, \qquad \text{and we set} \qquad \mu = \frac{\nu_0 - \nu_0^*}{z_1 - z_1^*}.$$

If $n(z) \in \mathbf{N}$ is not identically equal to a real constant, we define its *Schur transform* $\widehat{n}(z)$ as follows.

(i) Assume $n(z) \in \mathbf{N}^{z_1}$ and $\operatorname{Im} \nu_0 \neq 0$. Then $\widehat{n}(z) = \infty$ if $n(z)$ is linear and otherwise

$$\widehat{n}(z) = \frac{\beta(z)n(z) - |\nu_0|^2}{n(z) - \alpha(z)}, \tag{7.1}$$

where

$$\alpha(z) = \nu_0 + \mu(z - z_1), \quad \beta(z) = \nu_0^* - \mu(z - z_1).$$

(ii) Assume $n(z) \in \mathbf{N}^{z_1}$ and $\operatorname{Im} \nu_0 = 0$. Then, since, by assumption, $n(z) \not\equiv \nu_0$, the function

$$\frac{1}{n(z) - \nu_0}$$

has a poles at z_1 and z_1^*. Since $n(z)^* = n(z^*)$, the orders of the poles are the same and equal to the smallest integer $k \geq 1$ such that $\nu_k \neq 0$. Denote by $H_{z_1}(z)$ and $H_{z_1^*}(z)$ the principal parts of the Laurent expansion of the function $1/(n(z) - \nu_0)$ at z_1 and z_1^*. Then $H_{z_1^*}(z) = H_{z_1}(z^*)^*$ and

$$\frac{1}{n(z) - \nu_0} = H_{z_1}(z) + H_{z_1^*}(z) + a(z) = \frac{p(z)}{(z - z_1)^k (z - z_1^*)^k} + a(z) \tag{7.2}$$

with a function $a(z)$ which is holomorphic at z_1 and a polynomial $p(z)$ which is real: $p(z)^* = p(z^*)$, of degree $\leq 2k - 1$, and such that $p(z_1) \neq 0$. If the function $1/(n(z) - \nu_0)$ only has poles at z_1 and z_1^* and vanishes at ∞, that is, if (7.2) holds with $a(z) \equiv 0$, then $\widehat{n}(z) = \infty$. If (7.2) holds with $a(z) \not\equiv 0$, then

$$\widehat{n}(z) = \frac{\beta(z)n(z) - \nu_0^2}{n(z) - \alpha(z)}, \tag{7.3}$$

where

$$\alpha(z) = \nu_0 + \frac{(z - z_1)^k (z - z_1^*)^k}{p(z)}, \quad \beta(z) = \nu_0 - \frac{(z - z_1)^k (z - z_1^*)^k}{p(z)}.$$

(iii) If $n(z)$ has a pole at z_1 then

$$\widehat{n}(z) = n(z) - h_{z_1}(z) - h_{z_1^*}(z), \tag{7.4}$$

where $h_{z_1}(z)$ and $h_{z_1^*}(z) = h_{z_1}(z^*)^*$ are the principal parts of the Laurent expansion of $n(z)$ at the points z_1 and z_1^* respectively.

If in case (i) $n(z)$ is linear, that is, if $n(z) = a + bz$, $z \in \mathbb{C}$, with $a, b \in \mathbb{R}$, then it follows that $n(z) = \nu_0 + \mu(z - z_1) = \alpha(z)$. In this case the right-hand side of (7.1) is not defined, and the definition of the Schur transformation has been split into two parts.

The real polynomial $p(z)$ of degree $\le 2k - 1$ in case (ii) is determined by the asymptotic relation

$$p(z)(n(z) - \nu_0) = (z - z_1)^k (z - z_1^*)^k + \mathrm{O}\big((z - z_1)^{2k}\big), \quad z \to z_1, \tag{7.5}$$

and can be expressed in terms of the k Taylor coefficients $\nu_k, \dots, \nu_{2k-1}$ of $n(z)$ in the following way: If written in the form

$$p(z) = \sum_{j=0}^{k-1} a_j (z - z_1)^j + \sum_{j=k}^{2k-1} b_j (z - z_1)^j \tag{7.6}$$

then the coefficients $a_0, \dots, a_{k-1}$ are determined by the relations

$$a_j \nu_k + a_{j-1}\nu_{k+1} + \cdots + a_0 \nu_{k+j} = \binom{k}{j} (z_1 - z_1^*)^{k-j}, \quad j = 0, 1, \dots, k-1. \tag{7.7}$$

The other coefficients b_j, $j = k, k+1, \dots, 2k-1$, are uniquely determined by the fact that $p(z)$ is real: $p(z) = p(z^*)^*$. Indeed, this equality implies

$$\sum_{j=k}^{2k-1} b_j (z - z_1)^j = \sum_{j=0}^{k-1} a_j^* (z - z_1^*)^j - \sum_{j=0}^{k-1} a_j (z - z_1)^j + \sum_{j=k}^{2k-1} b_j^* (z - z_1^*)^j.$$

By taking the ith derivatives of the functions on both sides, $i = 0, 1, \dots, k-1$, and then evaluating them at z_1^* we get a system of k equations for the k unknowns b_j, $j = k, k+1, \dots, 2k-1$:

$$\sum_{j=k}^{2k-1} b_j \frac{j!}{(j-i)!} (z_1^* - z_1)^{j-i} = i! a_i^* - \sum_{j=i}^{k-1} a_j \frac{j!}{(j-i)!} (z_1^* - z_1)^{j-i}, \quad i = 0, 1, \dots, k-1.$$

Since the coefficient matrix of this system is invertible, these unknowns are uniquely determined.

Theorem 7.1. *Let $n(z) \in \mathbf{N}$ and assume it is not identically equal to a real constant. For its Schur transformation the following holds in the cases* (i), (ii), *and* (iii) *and with the integer k in* (ii) *as above:*

(i) $n(z) \in \mathbf{N}_\kappa^{z_1} \Longrightarrow \widehat{n}(z) \in \mathbf{N}_{\widehat{\kappa}}$ *with* $\widehat{\kappa} = \kappa$ *if* $\operatorname{Im} \nu_0 > 0$ *and* $\widehat{\kappa} = \kappa - 1$ *if* $\operatorname{Im} \nu_0 < 0$.

(ii) $n(z) \in \mathbf{N}_\kappa^{z_1} \Longrightarrow 1 \le k \le \kappa$, $\widehat{n}(z) \in \mathbf{N}_{\kappa-k}$.

(iii) $n(z) \in \mathbf{N}_\kappa$ *and* $n(z)$ *has a pole at* z_1 *of order* $q \ge 1 \Longrightarrow q \le \kappa$ *and* $\widehat{n}(z) \in \mathbf{N}_{\kappa-q}^{z_1}$.

This theorem is proved in [16, Theorem 7.3]; the proof uses the decomposition in Theorem 1.2 applied to $X(z)$ etc. as in (1.12). A proof can also be given by means of the realization results in Subsection 7.4.

The Schur transform $\widehat{n}(z)$ of $n(z)$ may have a pole at z_1 in cases (i) and (ii); evidently, in case (iii) it is holomorphic at z_1. In case (i) the Schur transform is holomorphic at z_1 if and only if $\nu_1 = \mu$ and it has a pole of order $q \geq$ if and only if

$$\nu_1 = \mu, \quad \nu_2 = \cdots = \nu_q = 0, \quad \nu_{q+1} \neq 0.$$

In case (ii) $\widehat{n}(z)$ is holomorphic at z_1 if and only if

$$a_k = b_k, \tag{7.8}$$

where a_k is the number in (7.7) with $j = k$ and b_k is the coefficient of $p(z)$ in (7.6); otherwise it has a pole and the order of the pole is equal to the order of the zero at z_1 of $n(z) - \alpha(z)$ minus $2k$. If in these cases $\widehat{n}(z)$ has a pole, then by applying the Schur transformation case (iii) to $\widehat{n}(z)$, we obtain a function which we shall call the *composite Schur transform of* $n(z)$. By definition it is holomorphic at z_1.

The formulas (7.1), (7.3), and (7.4) are of the form

$$\widehat{n}(z) = \mathcal{T}_{\Phi(z)}(n(z))$$

with in case (i)

$$\Phi(z) = \frac{1}{z - z_1} \begin{pmatrix} \beta(z) & -|\nu_0|^2 \\ 1 & -\alpha(z) \end{pmatrix},$$

in case (ii)

$$\begin{aligned} \Phi(z) &= \frac{p(z)}{b_\ell(z)^k (z - z_1)^k (z - z_1^*)^k} \begin{pmatrix} \beta(z) & -\nu_0^2 \\ 1 & -\alpha(z) \end{pmatrix} \\ &= \frac{1}{b_\ell(z)^k} \left(I_2 - \frac{p(z)}{(z - z_1)^k (z - z_1^*)^k} \mathbf{u}\mathbf{u}^* J_\ell \right), \quad \mathbf{u} = \begin{pmatrix} \nu_0 \\ 1 \end{pmatrix}, \end{aligned}$$

and in case (iii)

$$\begin{aligned} \Phi(z) &= \frac{1}{b_\ell(z)^q} \begin{pmatrix} 1 & -h_{z_1}(z) - h_{z_1^*}(z) \\ 0 & 1 \end{pmatrix} \\ &= \frac{1}{b_\ell(z)^q} \big(I_2 - (h_{z_1}(z) + h_{z_1^*}(z)) \mathbf{u}\mathbf{u}^* J_\ell \big), \quad \mathbf{u} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \end{aligned}$$

where q is the order of the pole of $n(z)$ at z_1. As in the case for Schur functions, the interest often lies in the normalized inverse transformation, and therefore we set

$$\Theta(z) = \Phi(z)^{-1} \Phi(\infty).$$

Theorem 7.2. *In case* (i)

$$\Theta(z) = I_2 + \big(b_\ell(z) - 1\big) \frac{\mathbf{u}\mathbf{u}^* J_\ell}{\mathbf{u}^* J_\ell \mathbf{u}}, \quad \mathbf{u} = \begin{pmatrix} \nu_0^* \\ 1 \end{pmatrix},$$

in case (ii),

$$\Theta(z)=b_\ell(z)^k I_2+\frac{p(z)}{(z-z_1^*)^{2k}}\mathbf{u}\mathbf{u}^* J_\ell,\quad \mathbf{u}=\begin{pmatrix}\nu_0\\ 1\end{pmatrix},\tag{7.9}$$

and in case (iii)

$$\Theta(z)=b_\ell(z)^q I_2+b_\ell(z)^q\big(h_{z_1}(z)+h_{z_1^*}(z)\big)\mathbf{u}\mathbf{u}^* J_\ell,\quad \mathbf{u}=\begin{pmatrix}1\\ 0\end{pmatrix},\tag{7.10}$$

where q is the order of the pole at z_1 of $n(z)$ and $h_{z_1}(z)$ and $h_{z_1^*}(z)=h_{z_1}(z^*)^*$ are the principal parts of the Laurent expansion of $n(z)$ at the points z_1 and z_1^*, respectively.

The first statement in the following theorem is a consequence of Theorem 3.2.

Theorem 7.3. *In all three cases $\Theta(z)$ can be written in the form* (3.9)*:*

$$\Theta(z)=I_2-C(zI-A)^{-1}G^{-1}C^*J_\ell.\tag{7.11}$$

The matrices A, C, and G are given by the following formulas.
In case (i)*:*

$$C=\begin{pmatrix}\nu_0^*\\ 1\end{pmatrix},\quad A=z_1^*,\quad G=\mu=\frac{\nu_0-\nu_0^*}{z_1-z_1^*}.$$

In case (ii)*:*

$$C=\begin{pmatrix}\nu_0^* & 0 & \cdots & 0 & \nu_k^* & \cdots & \nu_{2k-1}^*\\ 1 & 0 & \cdots & 0 & 0 & \cdots & 0\end{pmatrix},\quad A=z_1^*I_{2k}+S_{2k},\quad G=\Gamma_{2k},$$

where k is the smallest integer ≥ 1 such that $\nu_k\neq 0$ and Γ_{2k} is the $2k\times 2k$ principal submatrix of the Pick matrix Γ of $n(z)$ at z_1.
In case (iii)*:*

$$C=\begin{pmatrix}-1 & 0 & \cdots & 0 & 0 & \cdots & 0\\ 0 & 0 & \cdots & 0 & \rho_q^* & \cdots & \rho_{2q-1}^*\end{pmatrix},\quad A=z_1^*I_{2q}+S_{2q},\quad G=\Gamma'_{2q},$$

where, if

$$h_{z_1}(z)=\frac{\nu_{-q}}{(z-z_1)^q}+\frac{\nu_{-q+1}}{(z-z_1)^{q-1}}+\cdots+\frac{\nu_{-1}}{z-z_1}$$

is the principal part of the Laurent expansion of $n(z)$ at z_1, the complex numbers $\rho_q,\dots,\rho_{2q-1}$ are given by the relation

$$\begin{pmatrix}\rho_q & 0 & \cdots & 0 & 0\\ \rho_{q+1} & \rho_q & \cdots & 0 & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ \rho_{2q-2} & \rho_{2q-3} & \cdots & \rho_q & 0\\ \rho_{2q-1} & \rho_{2q-2} & \cdots & \rho_{q+1} & \rho_q\end{pmatrix}=\begin{pmatrix}\nu_{-q} & 0 & \cdots & 0 & 0\\ \nu_{-q+1} & \nu_{-q} & \cdots & 0 & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ \nu_{-2} & \nu_{-3} & \cdots & \nu_{-q} & 0\\ \nu_{-1} & \nu_{-2} & \cdots & \nu_{-q+1} & \nu_{-q}\end{pmatrix}^{-1}\tag{7.12}$$

and Γ'_{2q} is obtained from formula (4.24) *by replacing k by q and Δ by the matrix on the left-hand side of* (7.12).

In the cases (i) and (ii) the matrix G in the theorem is the smallest invertible principal submatrix of Γ, see Theorem 4.9. The proof of the theorem for these two cases can be found in [16, (5.5) and Theorem 6.3]. We derive the formula (7.11) for case (iii) from the one of case (ii) as follows. From (7.10) we obtain

$$-J_\ell\Theta(z)J_\ell = b_\ell(z)^q I_2 + \frac{r(z)}{(z-z_1^*)^{2k}}\mathbf{v}\mathbf{v}^* J_\ell, \quad \mathbf{v} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{7.13}$$

where $r(z)$ is the polynomial

$$\begin{aligned} r(z) &= (z-z_1)^q (z-z_1^*)^q \big(h_{z_1}(z) + h_{z_1}(z^*)^*\big) \\ &= (z-z_1^*)^q \left(\nu_{-q} + \nu_{-q+1}(z-z_1) + \cdots + \nu_{-1}(z-z_1)^{q-1}\right) \\ &\qquad + (z-z_1)^q \left(\nu_{-q}^* + \nu_{-q+1}^*(z-z_1^*) + \cdots + \nu_{-1}^*(z-z_1^*)^{q-1}\right) \\ &= (z-z_1^*)^q \left(\nu_{-q} + \nu_{-q+1}(z-z_1) + \cdots + \nu_{-1}(z-z_1)^{q-1}\right) + \mathrm{O}\big((z-z_1)^q\big), \end{aligned}$$

as $z \to z_1$. The right-hand side of (7.13) has the same form as the right-hand side of (7.9) with $\nu_0 = 0$ and polynomial $p(z)$ satisfying (7.5). The analog of (7.5) for the polynomial $r(z)$ reads as

$$r(z)\big(\rho_q + \rho_{q+1}(z-z_1) + \cdots + \rho_{2q-1}(z-z_1)^{q-1}\big) = (z-z_1^*)^q + \mathrm{O}\big((z-z_1)^q\big),$$

as $z \to z_1$. Equating coefficients we obtain the relation (7.12), and the formula for $-J_\ell\Theta(z)J_\ell$ now follows from case (ii):

$$-J_\ell\Theta(z)J_\ell = I_2 - C'(zI-A)^{-1}\left(G'\right)^{-1}\left(C'\right)^* J_\ell$$

with

$$C' = \begin{pmatrix} 0 & 0 & \cdots & 0 & \rho_q^* & \cdots & \rho_{2q-1}^* \\ 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}$$

and A and $G' = \Gamma_{2q}'$ as in the theorem. The asserted formula for $\Theta(z)$ now follows by setting $C = -J_\ell C'$.

Evidently, from the forms of the functions $\Theta(z)$ in Theorems 7.2 and 7.3 we see that they belong to the class $\mathcal{U}_\ell^{z_1}$ and are normalized by $\Theta(\infty) = I_2$.

7.2. The basic interpolation problem

The basic interpolation problem at $z = z_1 \in \mathbb{C}^+$ for generalized Nevanlinna functions reads as follows.

Problem 7.4. *Given $\nu_0 \in \mathbb{C}$ and an integer $\kappa \geq 0$. Determine all $n(z) \in \mathbf{N}_\kappa^{z_1}$ with $n(z_1) = \nu_0$.*

To describe the solutions of this basic interpolation problem we consider two cases. As in Subsection 5.2 in the second case we reformulate the problem in adaptation to our method with augmented parameters.

Case (i): $\operatorname{Im}\nu_0 \neq 0$. If $\kappa = 0$ and $\operatorname{Im}\nu_0 < 0$, then the problem does not have a solution, because $\operatorname{Im} n(z_1) \geq 0$ for all functions $n(z) \in \mathbf{N}_0$. If $\operatorname{Im}\nu_0 > 0$, then for each $\kappa \geq 0$ and if $\operatorname{Im}\nu_0 < 0$ then for each $\kappa \geq 1$ there are infinitely many solutions as the following theorem shows.

Theorem 7.5. *If* $\mathrm{Im}\ \nu_0 \neq 0$, *the formula*

$$n(z) = \frac{\alpha(z)\widetilde{n}(z) - |\nu_0|^2}{\widetilde{n}(z) - \beta(z)} \tag{7.14}$$

with

$$\alpha(z) = \nu_0 + \mu(z - z_1), \quad \beta(z) = \nu_0^* - \mu(z - z_1), \quad \mu = \frac{\nu_0 - \nu_0^*}{z_1 - z_1^*},$$

gives a one-to-one correspondence between all solutions $n(z) \in \mathbf{N}_\kappa^{z_1}$ *of Problem* 7.4 *and all parameters* $\widetilde{n}(z) \in \mathbf{N}_{\widetilde{\kappa}}$ *which, if holomorphic at* z_1, *satisfy the inequality* $\widetilde{n}(z_1) \neq \nu_0^*$, *where*

$$\widetilde{\kappa} = \begin{cases} \kappa, & \mathrm{Im}\ \nu_0 > 0, \\ \kappa - 1, & \mathrm{Im}\ \nu_0 < 0. \end{cases}$$

Note that for all parameters $\widetilde{n}(z) \in \mathbf{N}$ which have a pole at z_1 the solution satisfies

$$n'(z_1) = \frac{\nu_0 - \nu_0^*}{z_1 - z_1^*}.$$

This follows from

$$n(z_1) - \nu_0 = (z - z_1)\frac{\nu_0 - \nu_0^*}{z_1 - z_1^*}\frac{\widetilde{N}(z) - \nu_0}{\widetilde{N}(z) - \beta(z)}.$$

Case (ii): $\mathrm{Im}\,\nu_0 = 0$. By the maximum modulus principle there is a unique solution in the class $\mathbf{N}_0$, namely $n(z) \equiv \nu_0$. There are infinitely many solutions in $\mathbf{N}_\kappa^{z_1}$ for $\kappa \geq 1$. To describe them we reformulate the problem with augmented parameters.

Problem 7.6. *Given* $\nu_0 \in \mathbb{C}$ *with* $\mathrm{Im}\ \nu_0 = 0$, *integers* κ *and* k *with* $1 \leq k \leq \kappa$, *and numbers* $s_0, s_1, \ldots, s_{k-1} \in \mathbb{C}$ *with* $s_0 \neq 0$. *Determine all functions* $n(z) \in \mathbf{N}_\kappa^{z_1}$ *with* $n(z_1) = \nu_0$, *and* $\nu_{k+j} = s_j$, $j = 0, 1, \ldots, k-1$, *and, if* $k > 1$, $\nu_1 = \cdots = \nu_{k-1} = 0$.

With the data of the problem we associate the polynomial $p(z) = p(z, s_0, \ldots, s_{k-1})$ of degree $\leq 2k - 1$ with the properties:

(1) The coefficients $a_j = p^{(j)}(z_1)/j!$, $j = 0, \ldots, k-1$ satisfy the relations

$$a_j s_k + a_{j-1}s_{k+1} + \cdots + a_0 s_{k+j} = \binom{k}{j}(z_1 - z_1^*)^{k-j}, \quad j = 0, 1, \ldots, k-1.$$

(2) $p(z)$ is real, that is, $p(z) = p(z^*)^*$.

That $p(z)$ is uniquely determined follows from considerations as in Subsection 7.1.

Theorem 7.7. *If* $\mathrm{Im}\ \nu_0 = 0$, *for each integer* k *with* $1 \leq k \leq \kappa$ *and any choice of the complex numbers* $s_0 \neq 0, s_1, \ldots, s_{k-1}$ *the formula*

$$n(z) = \frac{\alpha(z)\widetilde{n}(z) - \nu_0^2}{\widetilde{n}(z) - \beta(z)} \qquad \textit{with}$$

$$\alpha(z) = \nu_0 + \frac{(z - z_1)^k(z - z_1^*)^k}{p(z)}, \quad \beta(z) = \nu_0 - \frac{(z - z_1)^k(z - z_1^*)^k}{p(z)}$$

gives a one-to-one correspondence between all solutions $n(z) \in \mathbf{N}_\kappa^{z_1}$ *of Problem* 7.6 *and all parameters* $\widetilde{n}(z) \in \mathbf{N}_{\kappa-k}$ *such that* $\widetilde{n}(z_1) \neq \nu_0$ *if* $\widetilde{n}(z)$ *is holomorphic at* z_1.

The parametrization formulas are the inverse of the Schur transformation, that is, the parameter $\widetilde{n}(z)$ corresponding to the solution $n(z)$ is the Schur transform of $n(z)$: $\widetilde{n}(z) = \widehat{n}(z)$. This can be seen from the following theorem. Recall that

$$b_\ell(z) = \frac{z - z_1}{z - z_1^*}.$$

Theorem 7.8. *The parametrization formula can be written as the linear fractional transformation*

$$n(z) = T_{\Theta(z)}(\widetilde{n}(z)),$$

where in case (i)

$$\Theta(z) = I_2 + (b_\ell(z) - 1)\frac{\mathbf{u}\mathbf{u}^* J_\ell}{\mathbf{u}^* J_\ell \mathbf{u}}, \quad \mathbf{u} = \begin{pmatrix} \nu_0^* \\ 1 \end{pmatrix}.$$

and in case (ii)

$$\Theta(z) = \left(b_\ell(z)^k I_2 - \frac{p(z)}{(z - z_1^*)^{2k}}\mathbf{u}\mathbf{u}^* J_\ell \right), \quad \mathbf{u} = \begin{pmatrix} \nu_0 \\ 1 \end{pmatrix}.$$

7.3. Factorization in the class $\mathcal{U}_\ell^{z_1}$

Recall that the class $\mathcal{U}_\ell^{z_1}$ consists of all rational J_ℓ-unitary 2×2 matrix functions, which have a pole only in $1/z_1^*$, and that $\Theta(z) \in \mathcal{U}_\ell^{z_1}$ is called normalized if $\Theta(\infty) = I_2$. By Theorem 3.13, products in the class $\mathcal{U}_\ell^{z_1}$ are always minimal. The following result is from [16, Theorems 6.2 and 6.4]. Part (i) is closely connected with Theorem 7.3 and part (ii) with Theorem 3.15.

Theorem 7.9. (i) *A normalized matrix function $\Theta(z)$ in $\mathcal{U}_\ell^{z_1}$ is elementary if and only if it has either of the following two forms*:

$$\Theta(z) = I_2 + (b_\ell(z) - 1)\frac{\mathbf{u}\mathbf{u}^* J_\ell}{\mathbf{u}^* J_\ell \mathbf{u}},$$

where $\mathbf{u}$ is a 2×1 vector such that $\mathbf{u}^* J_\ell \mathbf{u} \neq 0$, or

$$\Theta(z) = b_\ell(z)^k I_2 - \frac{p(z)}{(z - z_1^*)^{2k}}\mathbf{u}\mathbf{u}^* J_\ell\,,$$

where $\mathbf{u}$ is a J_ℓ-neutral nonzero 2×1 vector: $\mathbf{u}^* J_\ell \mathbf{u} = 0$*, k is an integer ≥ 1, and $p(z)$ is a real polynomial of degree $\leq 2k - 1$ with $p(z_1) \neq 0$.*

(ii) *Every $\Theta(z) \in \mathcal{U}_\ell^{z_1}$ has the unique minimal factorization*:

$$\Theta(z) = b_\ell(z)^n \Theta_1(z) \cdots \Theta_m(z) U,$$

where n is the largest nonnegative integer such that $b_\ell(z)^{-n}\Theta(z) \in \mathcal{U}_\ell^{z_1}$, $\Theta_j(z)$, $j = 1, \ldots, m$, is a normalized elementary factor from $\mathcal{U}_\ell^{z_1}$, and $U = \Theta(\infty)$ is a J_ℓ-unitary constant.

The theorem implies that the coefficient matrices $\Theta(z)$ of the inverse Schur transformation for generalized Nevanlinna functions are elementary factors in $\mathcal{U}_\ell^{z_1}$.

We describe how this fact can be used in a procedure to obtain the unique factorization of an element $\Theta(z)$ in $\mathcal{U}_\ell^{z_1}$ into elementary factors. Proofs of the various statements can be found in [16, Section 8].

(a) First we normalize and extract a power $b_\ell(z)^n$ of $\Theta(z)$: $\Theta(z)=b_\ell(z)^n\Psi(z)\Theta(\infty)$, so that $\Psi(z)\in\mathcal{U}_\ell^{z_1^*}$, $\Psi(z_1)\neq 0$ and $\Psi(\infty)=I_2$. If $\Psi(z)$ is a constant matrix stop the procedure. In this case the factorization is simply

$$\Theta(z)=b_\ell(z)^n\Theta(\infty).$$

So we assume from now that $\Psi(z)$ is not a constant matrix. We write

$$\Psi(z)=\begin{pmatrix} a(z) & b(z) \\ c(z) & d(z)\end{pmatrix}.$$

(b) Choose a real number $\tau\neq 0$ such that the function

$$\frac{a(z)\tau+b(z)}{c(z)\tau+d(z)} \tag{7.15}$$

is not a constant, and such that

$$(\mathrm{b}_1)\qquad c(z_1)\tau+d(z_1)\neq 0,$$

or, if (b_1) is not possible (which is the case, for example, if $c(z_1)=0$ and $d(z_1)=0$), then such that

$$(\mathrm{b}_2)\qquad a(z_1)\tau+b(z_1)\neq 0.$$

Except for at most three values of τ these conditions can be satisfied: The same argument as in Subsection 5.3, shows that there are at most two real numbers τ for which the function in (7.15) is a constant. The choice (b_1) or (b_2) is possible, because for at most one $\tau\in\mathbb{R}$ we have that

$$c(z_1)\tau+d(z_1)=0,\quad a(z_1)\tau+b(z_1)=0.$$

Indeed assume on the contrary that both $a(z_1)\tau+b(z_1)$ and $c(z_1)\tau+d(z_1)$ vanish for at least two different real numbers τ_1 and τ_2. Then

$$\Psi(z_1)\begin{pmatrix}\tau_1 & \tau_2\\ 1 & 1\end{pmatrix}=\begin{pmatrix}0&0\\0&0\end{pmatrix},$$

which implies $\Psi(z_1)=0$, contradicting the hypothesis.

(c) If (b_1) holds, form the function

$$n(z)=\mathcal{T}_{\Psi(z)}(\tau)=\frac{a(z)\tau+b(z)}{c(z)\tau+d(z)}.$$

Since $\Psi\in\mathcal{U}_\ell^{z_1}$, $n(z)$ is a rational generalized Nevanlinna function. Moreover, it is holomorphic at z_1 and not identically equal to a real constant. From

$$\lim_{z\to\infty} n(z)=\tau$$

and the definition of the Schur transformation, it follows that also

$$\lim_{z\to\infty}\widehat{n}(z)=\tau.$$

Thus $\widehat{n}(z)$ is a rational generalized Schur function which is either a real constant or has a Schur transform. Hence the Schur algorithm can be applied to $n(z)$ and like in Subsection 5.3 it leads to a factorization of $\Psi(z)$ and to the desired factorization of $\Theta(z)$.

(d) If (b_2) holds form the function

$$n(z) = \mathcal{T}_{J_\ell \Psi(z)}(\tau) = -\frac{c(z)\tau + d(z)}{a(z)\tau + b(z)}.$$

As in (c) the Schur algorithm can be applied to $n(z)$ and yields the factorization, say

$$J_\ell \Psi(z) = \Psi_1(z)\Psi_2(z)\cdots\Psi_q(z)J_\ell,$$

where the factors $\Psi_j(z)$ are elementary and normalized. Then

$$\Psi(z) = (-J_\ell\Psi_1(z)J_\ell)(-J_\ell\Psi_2(z)J_\ell)\cdots(-J_\ell\Psi_q(z)J_\ell),$$

is the factorization of $\Psi(z)$ into normalized elementary factors. Substituting this in the formula for $\Theta(z)$ we obtain the desired factorization of $\Theta(z)$.

7.4. Realization

We assume that $n(z)$ belongs to $\mathbf{N}^{z_1}$ with Taylor expansion (4.20):

$$n(z) = \sum_{i=0}^{\infty} \nu_1 (z - z_1)^i$$

and that it has the minimal self-adjoint realization

$$n(z) = n(z_1)^* + (z - z_1^*)\big\langle (1 + (z - z_0)(A - z)^{-1})u, u\big\rangle_{\mathcal{P}}, \tag{7.16}$$

The Taylor coefficients of $n(z)$ can be written as $\nu_0 = n(z_1)$,

$$\nu_i = \big\langle (A - z_1)^{-i+1}(I + (z_1 - z_1^*)(A - z_1)^{-1})u, u\big\rangle_{\mathcal{P}},\ i = 1, 2, \ldots, \tag{7.17}$$

and, moreover,

$$\langle u, u\rangle_{\mathcal{P}} = \mu\ \big(= (\nu_0 - \nu_0^*)/(z_1 - z_1^*)\big). \tag{7.18}$$

We study the effect of the Schur transformation on this realization and express the minimal self-adjoint realization of the Schur transform $\widehat{n}(z)$ or the composite Schur transform $\widetilde{n}(z)$ of $n(z)$ in terms of the realization (7.16). In the following theorems we may take $q = 0$: Then by saying $\widehat{n}(z)$ has a pole at z_1 of order 0 we mean that $\widehat{n}(z)$ is holomorphic at z_1, and the composite Schur transform is the Schur transform itself.

Case (i): $\operatorname{Im}\nu_0 \neq 0$. We recall that the Schur transform in part (i) of the definition is holomorphic at z_1 if and only if $\mu \neq \nu_1$ and that it has a pole of order q if and only if q is the smallest nonnegative integer such that $\nu_{q+1} \neq 0$ (hence $\nu_1 = \cdots = \nu_q = 0$ if $q > 0$).

Theorem 7.10. *Assume that $n(z) \in \mathbf{N}^{z_1}$ has the Taylor expansion* (4.20) *at z_1 with $\operatorname{Im}\nu_0 \neq 0$ and that $\widehat{n}(z)$ is defined and has a pole of order q at z_1. Then the minimal self-adjoint realization of the composite Schur transform $\widetilde{n}(z)$ of $n(z)$ is given by*

$$\widetilde{n}(z) = \widetilde{n}(z_1)^* + (z - z_1^*)\big\langle \big(I + (z - z_1)(\widetilde{A} - z)^{-1}\big)\widetilde{u}, \widetilde{u}\big\rangle_{\widetilde{\mathcal{P}}}$$

with

$$\widetilde{\mathcal{P}} = \mathcal{P} \ominus \mathcal{L}, \quad \widetilde{A} = \widetilde{P}A\big|_{\widetilde{\mathcal{P}}}, \quad \widetilde{u} = \widetilde{\nu}(q)\,\widetilde{P}\,(A - z_1)^{-q-1}u,$$

where $\mathcal{L}$ is the nondegenerate subspace

$$\mathcal{L} = \operatorname{span}\big\{u, (A - z_1)^{-1}u, \ldots, (A - z_1)^{-q}u, (A - z_1^*)^{-1}u, \ldots, (A - z_1^*)^{-q}u\big\}$$

of $\mathcal{P}$, $\widetilde{P}$ is the orthogonal projection in $\mathcal{P}$ onto $\widetilde{\mathcal{P}}$, and

$$\widetilde{\nu}(q) = \frac{\langle u, u\rangle_{\mathcal{P}}}{\big\langle (A - z_1)^{-q-1}u, u\big\rangle_{\mathcal{P}}} = \begin{cases} \dfrac{\nu_0 - \nu_0^*}{\nu_1 - \mu}, & q = 0, \\[2ex] \dfrac{\nu_0 - \nu_0^*}{\nu_{q+1}}, & q > 0. \end{cases} \tag{7.19}$$

The space $\widetilde{\mathcal{P}}$ is a Pontryagin space with negative index

$$\operatorname{ind}_-(\mathcal{P}) = \begin{cases} \operatorname{ind}_-(\mathcal{P}) - q, & \operatorname{Im}\nu_0 > 0, \\ \operatorname{ind}_-(\mathcal{P}) - q - 1, & \operatorname{Im}\nu_0 < 0. \end{cases}$$

The theorem is a combination of Corollaries 4.3 and 6.4 in [17]. Note that if $q = 0$, then $\mathcal{L}$ is just a 1-dimensional space spanned by u and then also

$$\widehat{n}(z_1) = \widetilde{n}(z_1) = \frac{\nu_0^*\nu_1 - \nu_0\mu}{\nu_1 - \mu}.$$

This and the second equality in (7.19) readily follows from the formulas (7.17) and (7.18).

Case (ii): $\operatorname{Im}\nu_0 = 0$. We assume that $\widehat{n}(z)$ exists according to part (ii) of the definition. We recall that k, the smallest integer ≥ 1 such that $\nu_k \neq 0$, exists because we assume that $n(z)$ is not a real constant. We also recall that $\widehat{n}(z)$ has a pole if and only if $a_k = b_k$, see (7.8). In this case the order of the pole is q if and only if $n(z)$ has the asymptotic expansion

$$n(z) - \alpha(z) = c_q(z - z_1)^{2k+q} + \mathrm{O}\big((z - z_1)^{2k+q+1}\big), \quad c_q \neq 0.$$

This q is finite, because $n(z) \not\equiv \alpha(z)$.

Theorem 7.11. *Assume that $n(z) \in \mathbf{N}^{z_1}$ has Taylor expansion* (4.20) *at z_1 with $\operatorname{Im}\nu_0 = 0$ and let $k \geq 1$ be the smallest integer such that $\nu_k \neq 0$. Assume also that $\widehat{n}(z)$ is defined and has a pole of order q at z_1. Then the minimal self-adjoint realization of the composite Schur transform $\widetilde{n}(z)$ centered at z_1 is given by*

$$\widetilde{n}(z) = \widetilde{n}(z_1)^* + (z - z_1^*)\big\langle \big(I + (z - z_1)(\widetilde{A} - z)^{-1}\big)\widetilde{u}, \widetilde{u}\big\rangle_{\widetilde{\mathcal{P}}},$$

with

$$\widetilde{\mathcal{P}} = \mathcal{P} \ominus \mathcal{L}, \quad \widetilde{A} = \widetilde{P}A\big|_{\widetilde{\mathcal{P}}}, \quad \widetilde{u} = \widetilde{\nu}(q)\widetilde{P}(A - z_1)^{-k-q}u$$

where $\mathcal{L}$ is the nondegenerate subspace

$$\mathcal{L} = \operatorname{span}\left\{u, (A-z_1)^{-1}u, \dots, (A-z_1)^{-k-q+1}u, (A-z_1^*)^{-1}u, \dots, (A-z_1^*)^{-k-q}u\right\}$$

of $\mathcal{P}$, $\widetilde{P}$ is the orthogonal projection in $\mathcal{P}$ onto $\widetilde{\mathcal{P}}$, and

$$\widetilde{\nu}(q) = \begin{cases} \dfrac{(z_1 - z_1^*)^k}{\nu_k(b_k - a_k)}, & q = 0, \\ \dfrac{\nu_k}{c_q}, & q > 0. \end{cases} \tag{7.20}$$

Moreover, $\widetilde{\mathcal{P}}$ is a Pontryagin space with negative index

$$\operatorname{ind}_-(\widetilde{\mathcal{P}}) = \operatorname{ind}_-(\mathcal{P}) - k - q. \tag{7.21}$$

This theorem is a combination of Corollaries 5.3 and 6.6 of [17]. If $q = 0$, then

$$\widehat{n}(z_1) = \widetilde{n}(z_1) = \nu_0 - \frac{(z_1 - z_1^*)^k}{(b_k - a_k)},$$

otherwise

$$\widetilde{n}(z_1) = \lim_{z \to z_1} \left(\widehat{n}(z) - \widehat{h}_{z_1}(z)\right) - \widehat{h}_{z_1^*}(z_1),$$

where $\widehat{h}_{z_1}(z)$ and $\widehat{h}_{z_1^*}(z) = \widehat{h}_{z_1}(z^*)^*$ are the principal parts of the Laurent expansions of $\widehat{n}(z)$ at z_1 and z_1^*.

The analog of Theorem 5.14 reads as follows, see [17, Theorem 7.1].

Theorem 7.12. *If $n(z) \in \mathbf{N}$ is not a real constant and the Schur algorithm applied to $n(z)$ yields the functions*

$$n_0(z) = n(z), \qquad n_j(z) = \widehat{n}_{j-1}(z), \; j = 1, 2, \dots,$$

then there exists an index j_0 such that $n_j(z) \in \mathbf{N}_0$ for all integers $j \geq j_0$.

The basic idea of the geometric proof of this theorem in [17], in terms of the minimal self-adjoint realization (7.16) of $n(z)$, is that (i) for each integer $j \geq 0$ the linear space

$$\mathcal{H}_j = \operatorname{span}\left\{u, (A-z_1)^{-1}u, (A-z_1)^{-2}u, \dots, (A-z_1)^{-j}u\right\}$$

is a subspace of the orthogonal complement of the state space in the minimal self-adjoint realization of $n_{j+1}(z)$ and (ii) that for sufficiently large j the space $\mathcal{H}_j$ contains a negative subspace of dimension $\operatorname{sq}_-(n)$. Since this negative subspace then is maximal negative, it follows that for sufficiently large j the state space in the realization of $n_{j+1}(z)$ is a Hilbert space, which means that $n_{j+1}(z)$ is classical Nevanlinna function.

7.5. Additional remarks and references

It is well known that the class of Schur functions can be transformed into the class of Nevanlinna functions by applying the Möbius transformation to the dependent and independent variables. The question arises if this idea can be carried over to the Schur transform: Do items (i)–(iv) in the definition of the Schur transformation of Schur functions in Subsection 5.1 via the Möbius transformation correspond in some way to items (i)–(iii) in the definition of the Schur transformation of Nevanlinna functions in Subsection 7.1? We did not pursue this question, partly because the more complicated formulas in Subsection 5.1 do not seem to transform easily to the ones in Subsection 7.1, and partly because this was not the way we arrived at the definition of the Schur transform for Nevanlinna functions. We obtained this definition by constructing suitable subspaces in $\mathcal{L}(n)$ in the same way as was done in the space $\mathcal{P}(s)$ and as explained in the three steps in Subsection 1.4. The basic idea is that the matrix function $\Theta(z)$ which minimizes the dimension of the space $\mathcal{P}(\Theta)$ in the decomposition in Theorem 1.2 with $X(z)$ etc. given by (1.11) and (1.12) is the matrix function that appears in the description of the inverse of the Schur transformation.

That in the positive case it is possible to use the Möbius transformation to make a reduction to the case of Schur function was remarked already by P.I. Richards in 1948. He considered functions $\varphi(s)$ which are analytic in Re $s>0$ and such that

$$\operatorname{Re}\varphi(s)\geq 0 \quad \text{for} \quad \operatorname{Re} s>0. \tag{7.22}$$

This case is of special importance in network theory. We reformulate P.I. Richard's result in the setting of Nevanlinna functions.

Theorem 7.13. *Let $n(z)$ a Nevanlinna function and assume that $n(ik)$ is purely imaginary for $k>0$. Then for every $k>0$, the function*

$$\widehat{n}(z)=\mathrm{i}\frac{z\,n(\mathrm{i}k)-\mathrm{i}k\,n(z)}{z\,n(z)-\mathrm{i}k\,n(\mathrm{i}k)} \tag{7.23}$$

is a Nevanlinna function. If $n(z)$ is rational, then

$$\deg n=1+\deg\widehat{n}.$$

It appears that equation (7.23) is just the Schur transformation after two changes of variables. This is mentioned, without proof, in the paper [123]. We give a proof for completeness.

Proof. Define ζ and functions $s(\zeta)$ and $\widehat{s}(\zeta)$ via:

$$z=\mathrm{i}\frac{1-\zeta}{1+\zeta},\quad n(z)=\mathrm{i}\frac{1-s(\zeta)}{1+s(\zeta)},\quad \widehat{n}(z)=\mathrm{i}\frac{1-\widehat{s}(\zeta)}{1+\widehat{s}(\zeta)}.$$

Then $\zeta\in\mathbb{D}$ if and only if $z\in\mathbb{C}^+$. Equation (7.23) is equivalent to:

$$\mathrm{i}\frac{1-\widehat{s}(\zeta)}{1+\widehat{s}(\zeta)}=\frac{z\,n(ik)-\mathrm{i}k\,n(z)}{z\,n(z)-\mathrm{i}k\,n(\mathrm{i}k)},$$

that is, with $a = \dfrac{1-k}{1+k}$,

$$\begin{aligned}\widehat{s}(\zeta) &= \frac{z+\mathrm{i}k}{z-\mathrm{i}k}\frac{n(z)-n(\mathrm{i}k)}{n(z)+n(\mathrm{i}k)}\\ &= \frac{\mathrm{i}\dfrac{1-\zeta}{1+\zeta}+\mathrm{i}k}{\mathrm{i}\dfrac{1-\zeta}{1+\zeta}-\mathrm{i}k}\cdot\frac{\mathrm{i}\dfrac{1-s(\zeta)}{1+s(\zeta)}-\mathrm{i}\dfrac{1-s(a)}{1+s(a)}}{\mathrm{i}\dfrac{1-s(\zeta)}{1+s(\zeta)}+\mathrm{i}\dfrac{1-s(a)}{1+s(a)}}\\ &= \frac{(1-\zeta)+(1+\zeta)k}{(1-\zeta)-(1+\zeta)k}\frac{s(\zeta)-s(a)}{1-s(\zeta)s(a)}\\ &= \frac{1-\zeta a}{\zeta-a}\frac{s(\zeta)-s(a)}{1-s(\zeta)s(a)}.\end{aligned}$$

By hypothesis, $n(\mathrm{i}k)$ is purely imaginary and so $s(a)$ is real. Thus, the last equation is the Schur transformation centered at the point a. The last claim can be found in [124, p. 173] and [127, pp. 455, 461–462]. □

Following [124, (7.40)] we rewrite (7.23) as

$$n(z) = n(\mathrm{i}k)\frac{z+k\,\widehat{n}(z)}{\mathrm{i}k-\mathrm{i}z\,\widehat{n}(z)}.$$

This linear fractional transformation provides a description of all Nevanlinna functions such that $n(\mathrm{i}k)$ is preassigned, and is a particular case of the linear fractional transformation (7.14) with $n(\mathrm{i}k) = \nu_0 \in \mathrm{i}\mathbb{R}$.

8. Generalized Nevanlinna functions with asymptotic at ∞

8.1. The Schur transformation

The Schur transformation centered at the point ∞ is defined for the generalized Nevanlinna functions from the class $\mathbf{N}^{\infty;2p}$, where p is an integer ≥ 1. We recall from Subsection 2.4 that a function $n(z) \in \mathbf{N}$ belongs to $\mathbf{N}^{\infty;2p}$ if it has an asymptotic expansion at ∞ of the form

$$n(z) = -\frac{\mu_0}{z} - \frac{\mu_1}{z^2} - \cdots - \frac{\mu_{2p-1}}{z^{2p}} + \mathrm{O}\left(\frac{1}{z^{2p+1}}\right), \quad z = \mathrm{i}y,\ y \uparrow \infty, \tag{8.1}$$

where

(1) $\mu_j \in \mathbb{R}$, $j = 0, 1, \ldots, 2p-1$, and
(2) at least one of the coefficients $\mu_0, \mu_1, \ldots, \mu_{p-1}$ is not equal to 0.

As remarked in Subsection 2.4, if this holds then there exists a real number μ_{2p} such that

$$n(z) = -\frac{\mu_0}{z} - \frac{\mu_1}{z^2} - \cdots - \frac{\mu_{2p}}{z^{2p+1}} + \mathrm{o}\left(\frac{1}{z^{2p+1}}\right), \quad z = \mathrm{i}y,\ y \uparrow \infty. \tag{8.2}$$

For $0 \le m \le p+1$ the $m \times m$ Hankel matrix Γ_m is defined by

$$\Gamma_m = \begin{pmatrix} \mu_0 & \mu_1 & \cdots & \mu_{m-1} \\ \mu_1 & \mu_2 & \cdots & \mu_m \\ \vdots & \vdots & & \vdots \\ \mu_{m-1} & \mu_m & \cdots & \mu_{2m-2} \end{pmatrix} \tag{8.3}$$

and we set

$$\gamma_m = \det \Gamma_m.$$

By k we denote the smallest integer ≥ 1 such that $\mu_{k-1} \neq 0$ and set

$$\varepsilon_{k-1} = \operatorname{sgn} \mu_{k-1}. \tag{8.4}$$

Then $1 \le k \le p$, $\gamma_k = (-1)^{[k/2]} \mu_{k-1}^k$, and $\kappa_-(\Gamma_k)$ is given by (4.26):

$$\kappa_-(\Gamma_k) = \begin{cases} [k/2], & \varepsilon_{k-1} > 0, \\ [(k+1)/2]\,, & \varepsilon_{k-1} < 0. \end{cases}$$

With the polynomial

$$e_k(z) = \frac{1}{\gamma_k} \det \begin{pmatrix} 0 & 0 & \ldots & \mu_{k-1} & \mu_k \\ 0 & 0 & \ldots & \mu_k & \mu_{k+1} \\ \vdots & \vdots & & \vdots & \vdots \\ \mu_{k-1} & \mu_k & \ldots & \mu_{2k-2} & \mu_{2k-1} \\ 1 & z & \ldots & z^{k-1} & z^k \end{pmatrix} \tag{8.5}$$

we define the *Schur transform* $\widehat{n}(z)$ of the function $n(z) \in \mathbf{N}^{\infty;2p}$ by

$$\widehat{n}(z) = -\frac{e_k(z)n(z) + \mu_{k-1}}{\varepsilon_{k-1} n(z)}. \tag{8.6}$$

Theorem 8.1. *If $n(z) \in \mathbf{N}_\kappa^{\infty;2p}$ has expansion* (8.2) *and k is the smallest integer ≥ 1 such that $\mu_{k-1} \neq 0$, then $\kappa_-(\Gamma_k) \le \kappa$ and $\widehat{n}(z) \in \mathbf{N}_{\widehat{\kappa}}$ with*

$$\widehat{\kappa} = \kappa - \kappa_-(\Gamma_k).$$

This theorem is [65, Theorem 3.2]. If, under the assumptions of Theorem 8.1, $\gamma_{p+1} \neq 0$, then $\widehat{n}(z)$ has an asymptotic expansion of the form (8.2) with p replaced by $p-k$ and explicit formulas for the coefficients in this expansion are given in [65, Lemma 2.4]. If the asymptotic expansion of $\widehat{n}(z)$ is such that $\widehat{n}(z)$ belongs to the class $\mathbf{N}^{\infty;2(k-p)}$ then the Schur transformation can be applied to $\widehat{n}(z)$, and so on, and we speak of the *Schur algorithm.*

Evidently, the inverse of the transformation (8.6) is given by

$$n(z) = -\frac{\mu_{k-1}}{\varepsilon_{k-1}\widehat{n}(z) + e_k(z)}. \tag{8.7}$$

This is a generalization of the transformation considered in [4, Lemma 3.3.6]. Indeed, if $n(z) \in \mathbf{N_0}$ has the asymptotic expansion

$$n(z) = -\frac{\mu_0}{z} - \frac{\mu_1}{z^2} + \mathrm{o}\left(\frac{1}{z^2}\right), \quad z = \mathrm{i}y,\ y \uparrow \infty, \tag{8.8}$$

and does not vanish identically, then $\mu_0 > 0$, hence $k = 1$, $\varepsilon_0 = 1$, and the relation (8.7) becomes

$$n(z) = -\frac{\mu_0}{\widehat{n}(z) + z - \dfrac{\mu_1}{\mu_0}}. \tag{8.9}$$

In [4, Lemma 3.3.6] it was shown that $\widehat{n}(z)$ is again a function of class $\mathbf{N_0}$ and that $\widehat{n}(z) = \mathrm{o}(1)$, $z = \mathrm{i}y$, $y \uparrow \infty$. If (8.8) holds with the term $\mathrm{o}(1/z^2)$ replaced by $\mathrm{O}(1/z^3)$, that is, if $n(z) \in \mathbf{N}_0^{\infty;2}$, then $\widehat{n}(z) = \mathrm{O}(1/z)$, $z = \mathrm{i}y$, $y \uparrow \infty$. The relations (8.9) and (8.7) can also be considered as the first step in a continuous fraction expansion of $n(z)$.

The transformation (8.7) can be written in the form

$$n(z) = \mathcal{T}_{\Phi(z)}(\widehat{n}(z)), \tag{8.10}$$

where

$$\Phi(z) = \frac{1}{\sqrt{|\mu_{k-1}|}} \begin{pmatrix} 0 & -\mu_{k-1} \\ \varepsilon_{k-1} & e_k(z) \end{pmatrix},$$

which belongs to $\mathcal{U}_\ell^\infty$. It we normalize the matrix function $\Phi(z)$ we obtain

$$\Theta(z) = \Phi(z)\Phi(0)^{-1},$$

where $\Phi(0)$ is a J_ℓ-unitary constant, and the transformation (8.10) becomes

$$n(z) = \mathcal{T}_{\Theta(z)U}(\widehat{n}(z)), \quad U = \Phi(0).$$

The matrix function $\Theta(z)$ also belongs to $\mathcal{U}_\ell^\infty$ and we have

$$\Theta(z) = \begin{pmatrix} 1 & 0 \\ \dfrac{e_k(0) - e_k(z)}{\mu_{k-1}} & 1 \end{pmatrix} = I_2 + p(z)\mathbf{u}\mathbf{u}^* J_\ell$$

with

$$p(z) = \frac{1}{\mu_{k-1}}\big(e_k(z) - e_k(0)\big), \quad \mathbf{u} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad J_\ell = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

The connection between the given function $n(z)$ and $\Theta(z)$ can be explained by applying the general setting of Subsection 1.4 to $X(z)$ etc. given by (1.12). Then $\mathcal{B}(X) = \mathcal{L}(n)$ and by letting $z_1 \to \infty$ we find that this space contains elements of the form

$$f_0(z) = n(z),\ f_j(z) = z^j n(z) + z^{j-1}\mu_0 + \cdots + \mu_{j-1},\ j = 1, \ldots, k-1,$$

see [15, Lemma 5.2]. If in Steps 2 and 3 in Subsection 1.4 we replace $\mathcal{M}_k$ by the span of the vector functions

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix},\ z^j \begin{pmatrix} 0 \\ 1 \end{pmatrix} + z^{j-1} \begin{pmatrix} \mu_0 \\ 0 \end{pmatrix} + \cdots + \begin{pmatrix} \mu_{j-1} \\ 0 \end{pmatrix},\ j = 1, \ldots, k-1,$$

we obtain a function $\Theta(z)$ of the above form and we find that it can be written according to formula (3.13) as

$$\Theta(z) = I_2 - zC(I - zA)^{-1}G^{-1}C^* J_\ell$$

with

$$C = \begin{pmatrix} 0 & \mu_0 & \mu_1 & \cdots & \mu_{k-2} \\ -1 & 0 & 0 & \cdots & 0 \end{pmatrix}, \quad A = S_k, \quad G = \Gamma_k,$$

where S_k is the $k \times k$ principal submatrix of the shift S and Γ_k, given by (8.3), is the $k \times k$ principal submatrix of the Pick matrix Γ of $n(z)$ at $z_1 = \infty$, see Subsection 4.4. (The formula for C differs from the one in [15] because in that paper we consider $-J_\ell$ instead of J_ℓ and for $X(z)$ the vector $\begin{pmatrix} 1 & n(z) \end{pmatrix}$ instead of $\begin{pmatrix} 1 & -n(z) \end{pmatrix}$.)

8.2. The basic boundary interpolation problem at ∞

The basic boundary interpolation problem which we consider here corresponds to the basic boundary interpolation Problem 6.2. It reads as follows.

Problem 8.2. *Given an integer $k \geq 1$, real numbers $\mu_{k-1}, \ldots, \mu_{2k-1}$ with $\mu_{k-1} \neq 0$. Determine all functions $n(z) \in \mathbf{N}$ such that*

$$n(z) = -\frac{\mu_{k-1}}{z^k} - \cdots - \frac{\mu_{2k-1}}{z^{2k}} + \mathrm{O}\left(\frac{1}{z^{2k+1}}\right), \qquad z = iy, \ y \uparrow \infty.$$

With the data of the problem we define the matrix Γ_k, the number ϵ_{k-1}, and the polynomial $e_k(z)$ by (8.3), (8.4), and (8.5). Evidently, if $n(z)$ is a solution, then it belongs to the class $\mathbf{N}_\kappa^{\infty;2k}$ with $\kappa \geq \kappa_-(\Gamma_k)$, see the inequality (4.27).

Theorem 8.3. *The formula*

$$n(z) = -\frac{\mu_{k-1}}{\varepsilon_{k-1}\widetilde{n}(z) + e_k(z)} \tag{8.11}$$

gives a bijective correspondence between all solutions $n(z) \in \mathbf{N}_\kappa^{\infty;2k}$ of Problem 8.2 and all parameters $\widetilde{n}(z)$ in the class $\mathbf{N}_{\widetilde{\kappa}}$ with $\widetilde{n}(z) = \mathrm{O}(1/z)$, $z = \mathrm{i}y$, $y \uparrow \infty$, where

$$\widetilde{\kappa} = \kappa - \kappa_-(\Gamma_k).$$

Proof. If $n(z)$ is a solution, then (8.2) holds with $p = k$ and some real number μ_{2k} and we may apply [65, Lemma 2.4]. It follows that $n(z)$ can be expressed as the linear fractional transformation (8.11) with a scalar function $\widetilde{n}(z)$ which behaves as $\mathrm{O}(1/z)$, $z = iy$, $y \uparrow \infty$. To show that this function is a generalized Nevanlinna function we use the relation

$$L_n(z,w) = n(z)\big(L_{e_k/\mu_{k-1}}(z,w) + L_{\widetilde{n}}(z,w)\big)n(w)^*, \tag{8.12}$$

which follows directly from (8.11). The polynomial $e_k(z)/\mu_{k-1}$ has real coefficients and hence is a generalized Nevanlinna function. The number of negative squares of the kernel $L_{e_k/\mu_{k-1}}(z,w)$ is equal to $\kappa_-(\Gamma_k)$ given by (4.26). We assume that $n(z)$ is a solution and hence it is a generalized Nevanlinna function. From the relation (8.12) it follows that $\widetilde{n}(z)$ also is a generalized Nevanlinna function. Since $e_k(z)$ and $\widetilde{n}(z)$ behave differently near $z = \infty$, and using well-known results from the theory of reproducing kernel spaces, see, for instance, [19, Section 1.5], we find that

$$\mathcal{L}(n) \cong \mathcal{L}(e_k/\mu_{k-1}) \oplus \mathcal{L}(\widetilde{n}),$$

that is, the spaces on the left and right are unitarily equivalent. (This can also be proved using Theorem 1.2 in the present setting.) Hence

$$\kappa = \kappa_-(\Gamma_k) + \widetilde{\kappa}.$$

As to the converse, assume $n(z)$ is given by (8.11) with parameter $\widetilde{n}(z)$ from $\mathbf{N}_{\widetilde{\kappa}}$ satisfying $\widetilde{n}(z) = \mathrm{O}\,(1/z)$, $z = \mathrm{i}y,\ y \uparrow \infty$. Then, since the degree of the polynomial $e_k(z)$ is k,

$$n(z) + \frac{\mu_{k-1}}{e_k(z)} = \frac{|\mu_{k-1}|\widetilde{n}(z)}{e_k(z)\big(\epsilon_{k-1}\widetilde{n}(z) + e_k(z)\big)} = \mathrm{O}\left(\frac{1}{z^{2k+1}}\right), \qquad z = iy,\ y \uparrow \infty.$$

By [15, Lemma 5.2],

$$\frac{\mu_{k-1}}{e_k(z)} = \frac{\mu_{k-1}}{z^k} + \cdots + \frac{\mu_{2k-1}}{z^{2k}} + \mathrm{O}\left(\frac{1}{z^{2k+1}}\right), \qquad z = iy,\ y \uparrow \infty,$$

and hence $n(z)$ has the asymptotic expansion (8.1). That it is a generalized Nevanlinna function with κ negative squares follows from (8.12) and arguments similar to the ones following it. □

8.3. Factorization in the class $\mathcal{U}_\ell^\infty$

Recall from Subsection 3.4 that the class $\mathcal{U}_\ell^\infty$, which consists of the J_ℓ-unitary 2×2 matrix polynomials, is closed under multiplication and taking inverses. The latter implies that a product need not be minimal. Nevertheless, by Theorem 3.17, each element in $\mathcal{U}_\ell^\infty$ admits a unique minimal factorization. An element $\Theta(z)$ of this class is called normalized if $\Theta(0) = I_2$.

Theorem 8.4.

(i) *A normalized $\Theta(z) \in \mathcal{U}_\ell^\infty$ is elementary if and only if it is of the form*

$$\Theta(z) = I_2 + p(z)\mathbf{u}\mathbf{u}^* J, \tag{8.13}$$

where $\mathbf{u}$ is 2×1 vector satisfying $\mathbf{u}^* J\mathbf{u} = 0$ and $p(z)$ is a real polynomial with $p(0) = 0$.

(ii) *$\Theta(z)$ admits a unique minimal factorization*

$$\Theta(z) = \Theta_1(z) \cdots \Theta_m(z) U$$

with normalized elementary factors $\Theta_j(z)$ from $\mathcal{U}_\ell^\infty$, $j = 1, 2, \ldots, m$, and the J_ℓ-unitary constant $U = \Theta(0)$.

This theorem is proved in [15, Theorem 6.4]. For part (i) see also formula (3.13). We note that if $\Theta(z)$ is of the form (8.13) and $p(z) = t_k z^k + \cdots + t_1 z$ with $t_k \neq 0$, then $k = \dim \mathcal{P}(\Theta)$ and the negative index κ of the Pontryagin space $\mathcal{P}(\Theta)$ is given by

$$\kappa = \begin{cases} [k/2], & t_k > 0, \\ [(k+1)/2]\,, & t_k < 0. \end{cases}$$

We now describe in four constructive steps how the Schur algorithm can be applied to obtain the factorization of Theorem 8.4(ii). For the details see [15, Section 6].

Assume $\Theta(z)$ belongs to $\mathcal{U}_\ell^\infty$ and is not equal to a J_ℓ-unitary constant.

(a) Determine a J_ℓ-unitary constant V_0 such that if

$$\Psi(z) = V_0\Theta(z) = \begin{pmatrix} a(z) & b(z) \\ c(z) & d(z) \end{pmatrix},$$

then

$$\max(\deg a, \deg b) < \max(\deg c, \deg d).$$

The matrix function $\Psi(z)$ also belongs to the class $\mathcal{U}_\ell^\infty$. For a proof that such a V_0 exists we refer to [15, Lemma 6.1].

(b) Choose $\tau \in \mathbb{R}$ such that

$$\deg(a\tau + b) = \max\{\deg a,\ \deg b\}, \quad \deg(c\tau + d) = \max\{\deg c,\ \deg d\},$$

and consider the function

$$n(z) = \frac{a(z)\tau + b(z)}{c(z)\tau + d(z)} = \mathcal{T}_{\Psi(z)}(\tau).$$

If $\deg c < \deg a$ we can also choose $\tau = \infty$ and

$$n(z) = a(z)/c(z) = \mathcal{T}_{\Psi(z)}(\infty).$$

Since $\Psi(z) \in \mathcal{U}_\ell^\infty$, $n(z)$ is a generalized Nevanlinna function and the kernels $K_\Psi(z,w)$ and $L_n(z,w)$ have the same number of negative squares. Evidently, in both cases $n(z)$ is rational and has the property

$$\lim_{y\to\infty} n(iy) = 0.$$

This implies that $n(z)$ belongs to $\mathbf{N}^{\infty;2p}$ for any sufficiently large integer p and that its Schur transform $\widehat{n}(z)$ is well defined and has the same properties, and so on, in other words, the Schur algorithm can be applied to $n(z)$.

(c) Apply, as in Subsection 5.3, the Schur algorithm to $n(z)$ to obtain the minimal factorization

$$\Psi(z) = \Psi_1(z)\Psi_2(z)\cdots\Psi_m(z)V_1$$

with normalized factors $\Psi_j(z)$ and $V_1 = \Psi(0)$, and hence

$$\Theta(z) = V_0^{-1}\Psi_1(z)\Psi_2(z)\cdots\Psi_m(z)\,V_1. \tag{8.14}$$

(d) Normalize the factors in (8.14) to obtain the factorization

$$\Theta(z) = \Theta_1(z)\Theta_2(z)\cdots\Theta_m(z)\Theta(0)$$

with normalized elementary factors. This factorization is obtained from (8.14) via the formulas

$$\begin{aligned}
\Theta_1(z) &= V_0^{-1}\Psi_1(z)\Psi_1(0)^{-1}V_0, \\
\Theta_2(z) &= V_0^{-1}\Psi_1(0)\Psi_2(z)\Psi_2(0)^{-1}\Psi_1(0)^{-1}V_0, \\
\Theta_3(z) &= V_0^{-1}\Psi_1(0)\Psi_2(0)\Psi_3(z)\Psi_3(0)^{-1}\Psi_2(0)^{-1}\Psi_1(0)^{-1}V_0,
\end{aligned}$$

and so on.

In [91] it is shown that the factorization in Theorem 8.4(ii) can also be obtained using purely algebraic tools, without the Schur transformation and the more geometric considerations in reproducing kernel Pontryagin spaces used in this paper.

8.4. Realization

With the function $n(z) \in \mathbf{N}^{\infty;2p}$ the following Pontryagin space $\Pi(n)$ can be associated. We consider the linear span of the functions $\mathbf{r}_z$, $z \in \text{hol}\,(n)$, $z \neq z^*$, defined by

$$\mathbf{r}_z(t) = \frac{1}{t-z}, \quad t \in \mathbb{C}.$$

Equipped with the inner product

$$\langle \mathbf{r}_z, \mathbf{r}_\zeta \rangle = \frac{n(z) - n(\zeta)^*}{z - \zeta^*}, \quad z, \zeta \in \text{hol}\,(n),\ z \neq \zeta^*,$$

this linear span becomes a pre-Pontryagin space, the completion of which is by definition the space $\Pi(n)$. It follows from the asymptotic expansion (8.2) of $n(z)$ that $\Pi(n)$ contains the functions

$$\mathbf{t}_j(t) := t^j, \quad j = 0, 1, \ldots, p,$$

and that

$$\langle \mathbf{t}_j, \mathbf{t}_k \rangle = \mu_{j+k}, \quad 0 \leq j, k \leq p,\ j + k \leq 2p\,, \tag{8.15}$$

see [94, Satz 1.10]. In $\Pi(n)$ the operator of multiplication by the independent variable t can be defined, which is self-adjoint and possibly unbounded; we denote it by A. Let $u \equiv e_0(t) := \mathbf{t}_0(t) \equiv 1$, $t \in \mathbb{C}$. Then $u \in \text{dom}\,(A^j)$ and $\mathbf{t}_j = A^j u$, $j = 0, 1, \ldots, p$, and the function $n(z)$ admits the representation

$$n(z) = \big\langle (A - z)^{-1} u, u \big\rangle\,, \quad z \in \text{hol}\,(n).$$

Now let $k\,(\leq p)$ be again the smallest positive integer such that $\mu_{k-1} \neq 0$. We introduce the subspace

$$\mathcal{H}_k = \text{span}\{\mathbf{t}_0, \mathbf{t}_1, \ldots, \mathbf{t}_{k-1}\} \tag{8.16}$$

of $\Pi(n)$. It is nondegenerate since $\mu_{k-1} \neq 0$ and its negative index equals

$$\text{ind}_-(\mathcal{H}_k) = \begin{cases} [k/2], & \mu_{k-1} > 0, \\ [(k+1)/2]\,, & \mu_{k-1} < 0. \end{cases}$$

Denote by $\widehat{\mathcal{H}}_k$ the orthogonal complement of $\mathcal{H}_k$ in $\Pi(n)$:

$$\Pi(n) = \mathcal{H}_k \oplus \widehat{\mathcal{H}}_k, \tag{8.17}$$

and let $\widehat{P}$ be the orthogonal projection onto $\widehat{\mathcal{H}}_k$ in $\Pi(n)$.

Theorem 8.5. *Let $n(z) \in \mathbf{N}_\kappa^{\infty;2p}$ have the asymptotic expansion* (8.2) *and let k be the smallest integer ≥ 1 such that $\mu_{k-1} \neq 0$. If*

$$n(z) = \big\langle (A - z)^{-1} u, u \big\rangle$$

with a densely defined self-adjoint operator A and an element u in the Pontryagin space $\Pi(n)$ is a minimal realization of $n(z)$, then a minimal realization of the Schur transform $\widehat{n}(z)$ from (8.6) *is*

$$\widehat{n}(z) = \left\langle (\widehat{A} - z)^{-1}\widehat{u}, \widehat{u} \right\rangle, \quad z \in \rho(\widehat{A}),$$

where $\widehat{A}$ is the densely defined self-adjoint operator $\widehat{A} = \widehat{P}A\widehat{P}$ in $\widehat{\mathcal{H}}_k$ and $\widehat{u} = \widehat{P}A^k u$.

Proof. Clearly, the function $e_k(t)$ from (8.4) belongs to the space $\Pi(n)$, and it is easy to see that it belongs even to $\widehat{\mathcal{H}}_k$. We write $e_k(t)$ in the form

$$e_k(t) = t^k + \eta_{k-1}t^{k-1} + \cdots + \eta_1 t + \eta_0$$

with coefficients η_j given by the corresponding submatrices from (8.4). The relation

$$\mathbf{t}_k = -\big(\eta_{k-1}\mathbf{t}_{k-1} + \cdots + \eta_1\mathbf{t} + \eta_0\mathbf{t}_0\big) + e_k$$

gives the decomposition of the element $\mathbf{t}_k \in \Pi(n)$ according to (8.17). The elements of the space $\mathcal{H}_k$ in the decomposition (8.17) we write as vectors with respect to the basis $\mathbf{t}_0, \mathbf{t}_1, \ldots, \mathbf{t}_{k-1}$. If we observe that $A\mathbf{t}_{k-1} = \mathbf{t}_k$, the operator A in the realization of $n(z)$ in $\Pi(n)$ becomes

$$A = \left(\begin{array}{ccccc|c} 0 & 0 & \cdots & 0 & \eta_0 & \varepsilon_0[\cdot, e_k] \\ 1 & 0 & \cdots & 0 & \eta_1 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \eta_{k-1} & 0 \\ 0 & 0 & \cdots & 0 & e_k(t) & \widehat{A} \end{array}\right).$$

Next we find the component ξ_{k-1} of the solution vector x of the equation $(A-z)x = u$, that is,

$$(A - z)\begin{pmatrix} \xi_0 \\ \xi_1 \\ \vdots \\ \xi_{k-1} \\ \widehat{x} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}.$$

An easy calculation yields

$$\xi_{k-1} = -\frac{1}{\varepsilon_0 \left\langle (\widehat{A} - z)^{-1}e_k, e_k \right\rangle + e_k(z)},$$

and we obtain finally

$$\big\langle (A - z)^{-1}u, u \big\rangle = \xi_{k-1}\mu_{k-1} = -\frac{\mu_{k-1}}{\varepsilon_0 \left\langle (\widehat{A} - z)^{-1}e_k, e_k \right\rangle + e_k(z)}. \qquad \square$$

8.5. Additional remarks and references

The Akhiezer transformation (8.9) is the analog of the classical Schur transformation in the positive case and is proved, as already mentioned, in [4, Lemma 3.3.6]. The proof that the Schur transform of a Schur function is again a Schur function can be proved using the maximum modulus principle, whereas the analog for Nevanlinna functions follows easily from the integral representations of Nevanlinna functions.

The self-adjoint realization in Subsection 8.4 is more concrete than the realizations considered in the corresponding Subsections 5.4 and 7.4. The approach here seems simpler, because we could exhibit explicitly elements that belong to the domain of the self-adjoint operator in the realization. The realizations and the effect of the Schur transformation on them, exhibited in the Subsections 5.4, 7.4, and 8.4 can also be formulated in terms of backward-shift operators in the reproducing kernel Pontryagin spaces $\mathcal{P}(s)$ and $\mathcal{L}(n)$, see, for example, [125], [126], and [16, Section 8].

In this section the main role was played by Hankel matrices. Such matrices, but with coefficients in a finite field, appear in a completely different area, namely in the theory of error correcting codes. A recursive fast algorithm to invert a Hankel matrix with coefficients in a finite field was developed by E.R. Bekerlamp and J.L. Massey in the decoding of Bose–Chauduri–Hocquenghem codes, see [43, chapter 7, §7.4 and §7.5]. Since the above formulas for elementary factors do not depend on the field and make sense if the field of complex numbers is replaced by any finite field, there should be connections between the Bekerlamp–Massey algorithm and the present section.

References

[1] R. Ackner, H. Lev-Ari, and T. Kailath, *The Schur algorithm for matrix-valued meromorphic functions*, SIAM J. Math. Anal. **15** (1994), 140–150.

[2] V.M. Adamyan and S.E. Nechayev, *Nuclear Hankel matrices and orthogonal trigonometric polynomials*, Contemporary Mathematics **189** (1995), 1–15.

[3] N.I. Akhiezer, *On a minimum problem in function theory and the number of roots of an algebraic equation inside the unit disc*, Izv. Akad. Nauk SSSR Mat. Estestv. Nauk No. 9 (1930), 1169–1189 (Russian); English translation: Operator Theory: Adv. Appl. **95**, Birkhäuser Verlag, Basel, 1997, 19–35.

[4] N.I. Akhiezer, *The classical moment problem*, Hafner, New-York, 1965.

[5] D. Alpay, *Reproducing kernel Krein spaces of analytic functions and inverse scattering*, PhD thesis, Weizmann Institute of Sciences, Rehovot, Israel, 1986.

[6] D. Alpay, *The Schur algorithm, reproducing kernel spaces and system theory*, Panoramas et Synthèses, 1998 (French); English translation: Amer. Math. Soc., Providence, RI, 2001.

[7] D. Alpay, T. Azizov, A. Dijksma, and H. Langer, *The Schur algorithm for generalized Schur functions I: Coisometric realizations*, Operator Theory: Adv. Appl. **129**, Birkhäuser Verlag, Basel, 2001, 1–36.

[8] D. Alpay, T.Ya. Azizov, A. Dijksma, and H. Langer, *The Schur algorithm for generalized Schur functions II: Jordan chains and transformations of characteristic functions*, Monatsh. Math. **138**(1) (2003), 1–29.

[9] D. Alpay, T. Azizov, A. Dijksma, and H. Langer, *The Schur algorithm for generalized Schur functions III: J-unitary matrix polynomials on the circle*, Linear Algebra Appl. **369** (2003), 113–144.

[10] D. Alpay, T.Ya. Azizov, A. Dijksma, H. Langer, and G. Wanjala, *A basic interpolation problem for generalized Schur functions and coisometric realizations*, Operator Theory: Adv. Appl. **143**, Birkhäuser Verlag, Basel, 2003, 39–76.

[11] D. Alpay, T.Ya. Azizov, A. Dijksma, H. Langer, and G. Wanjala, *The Schur algorithm for generalized Schur functions IV: Unitary realizations*, Operator Theory: Adv. Appl. **149**, Birkhäuser Verlag, Basel, 2004, 23–45.

[12] D. Alpay, T.Ya. Azizov, A. Dijksma, and J. Rovnyak, *Colligations in Pontryagin spaces with a symmetric characteristic function*, Operator Theory: Adv. Appl. **130**, Birkhäuser Verlag, Basel, 2002, 55–82.

[13] D. Alpay, P. Bruinsma, A. Dijksma, and H.S.V. de Snoo, *Interpolation problems, extensions of symmetric operators and reproducing kernel spaces II*, Integral Equations and Operator Theory **14** (1991), 465–500.

[14] D. Alpay, P. Dewilde, and H. Dym, *On the existence and construction of solutions to the partial lossless inverse scattering problem with applications to estimation theory*, IEEE Trans. Inform. Theory **35** (1989), 1184–1205.

[15] D. Alpay, A. Dijksma, and H. Langer, *Factorization of J-unitary matrix polynomials on the line and a Schur algorithm for generalized Nevanlinna functions*, Linear Algebra Appl. **387** (2004), 313–342.

[16] D. Alpay, A. Dijksma, and H. Langer, *J_ℓ-unitary factorization and the Schur algorithm for Nevanlinna functions in an indefinite setting*, Linear Algebra Appl. **419** (2006), 675–709.

[17] D. Alpay, A. Dijksma, H. Langer, and Y. Shondin, *The Schur transformation for generalized Nevanlinna functions: interpolation and self-adjoint operator realizations*, to appear in Complex Analysis and Operator Theory **1**(2) (2006).

[18] D. Alpay, A. Dijksma, H. Langer, and G. Wanjala, *Basic boundary interpolation for generalized Schur functions and factorization of rational J-unitary matrix functions*, Operator Theory: Adv. Appl. **165**, Birkhäuser Verlag, Basel, 2006, 1–29.

[19] D. Alpay, A. Dijksma, J. Rovnyak, and H. de Snoo, *Schur functions, operator colligations, and reproducing kernel Pontryagin spaces*, Operator Theory: Adv. Appl. **96**, Birkhäuser Verlag, Basel, 1997.

[20] D. Alpay and H. Dym, *Hilbert spaces of analytic functions, inverse scattering and operator models, I*, Integral Equations and Operator Theory **7** (1984), 589–641.

[21] D. Alpay and H. Dym, *On applications of reproducing kernel spaces to the Schur algorithm and rational J-unitary factorization*, Operator Theory: Adv. Appl. **18**, Birkhäuser Verlag, Basel, 1986, 89–159.

[22] D. Alpay and H. Dym, *Structured invariant spaces of vector valued functions, sesquilinear forms, and a generalization of the Iohvidov laws*, Linear Algebra Appl. **137/138** (1990), 413–451.

[23] D. Alpay and H. Dym, *On reproducing kernel spaces, the Schur algorithm, and interpolation in a general class of domains*, Operator Theory: Adv. Appl. **59**, Birkhäuser Verlag, Basel, 1992, 30–77.

[24] D. Alpay and H. Dym, *On a new class of reproducing kernel spaces and a new generalization of the Iohvidov laws*, Linear Algebra Appl. **178** (1993), 109–183.

[25] D. Alpay and H. Dym, *On a new class of realization formulas and their applications*, Linear Algebra Appl. **241/243** (1996), 3–84.

[26] D. Alpay and I. Gohberg, *Unitary rational matrix functions*, Operator Theory: Adv. Appl. **33**, Birkhäuser Verlag, Basel, 1988, 175–222.

[27] D. Alpay and I. Gohberg, *Discrete analogs of canonical systems with pseudo–exponential potential. Definitions and formulas for the spectral matrix functions*, Operator Theory: Adv. Appl. **161**, Birkhäuser Verlag, Basel, 2006, 1–47.

[28] D. Alpay and I. Gohberg, *Discrete analogs of canonical systems with pseudo–exponential potential. Inverse problems*, Operator Theory: Adv. Appl. **165**, Birkhäuser Verlag, Basel, 2006, 31–65.

[29] N. Aronszajn, *La théorie des noyaux reproduisants et ses applications. I*, Proc. Cambridge Philos. Soc. **39** (1943), 133–153.

[30] N. Aronszajn, *Theory of reproducing kernels*, Trans. Amer. Math. Soc. **68** (1950), 337–404.

[31] D. Arov and H. Dym, *Some remarks on the inverse monodromy problem for 2×2 canonical differential systems*, Operator Theory: Adv. Appl. **122**, Birkhäuser Verlag, Basel, 2001, 53–87.

[32] D. Arov and H. Dym, *J-inner matrix functions, interpolation and inverse problems for canonical systems. IV. Direct and inverse bitangential input scattering problems*, Integral Equations and Operator Theory **43**(1) (2002), 1–67.

[33] T.Ya. Azizov and I.S. Iohvidov, *Foundations of the theory of linear operators in spaces with indefinite metric*, Nauka, Moscow, 1986 (Russian); English translation: *Linear operators in spaces with an indefinite metric*, John Wiley, New York, 1989.

[34] M. Bakonyi and T. Constantinescu, *Schur's algorithm and several applications*, Pitman Research Notes in Mathematics Series **261**, Longman Scientific & Technical, Harlow, 1992.

[35] J. Ball, *Interpolation problems of Pick-Nevanlinna and Loewner types for meromorphic matrix functions*, Integral Equations and Operator Theory **6**(6) (1983), 804–840.

[36] J. Ball, I. Gohberg, and L. Rodman, *Interpolation of rational matrix functions*, Operator Theory: Adv. Appl. **45**, Birkhäuser Verlag, Basel, 1990

[37] H. Bart, I. Gohberg, and M.A. Kaashoek, *Minimal factorization of matrix and operator functions*, Operator Theory: Adv. Appl. **1**, Birkhäuser Verlag, Basel, 1979.

[38] A. Benveniste, R. Nikoukhah, and A. Willsky, *Multiscale system theory*, Rapport de Recherche **1194**, INRIA, Mars 1990.

[39] A. Benveniste, R. Nikoukhah, and A. Willsky, *Multiscale system theory*, IEEE Trans. Circuits Systems I Fund. Theory Appl. **41**(1) (1994), 2–15.

[40] C. Berg, J. Christensen, and P. Ressel, *Harmonic analysis on semigroups*, Graduate Texts in Mathematics **100**, Springer-Verlag, New York, 1984.

[41] M.-J. Bertin, *K-nombres de Pisot et de Salem*, Acta Arith. **68**(2) (1994), 113–131.

[42] M.-J. Bertin, A. Decomps-Guilloux, M. Grandet-Hugot, M. Pathiaux-Delefosse, and J.-P. Schreiber, *Pisot and Salem numbers*, Birkhäuser Verlag, Basel, 1992.

[43] R.E. Blahut, *Theory and practice of error control codes*, Addison–Wesley, Reading, MA, 1983.

[44] V. Bolotnikov, A. Kheifets, and L. Rodman, *Functions with Pick matrices having bounded number of negative eigenvalues*, Contemp. Math. **323**, Amer. Math. Soc., Providence, RI, 2003, 393–417.

[45] V. Bolotnikov, A. Kheifets, and L. Rodman, *Jet functions having indefinite Carathéodory-Pick matrices*, Linear Algebra Appl. **385** (2004), 215–286.

[46] V. Bolotnikov, A. Kheifets, and L. Rodman, *Operator-valued jet functions with positive Carathéodory-Pick operators*, Integral Equations and Operator Theory **50**(3) (2004), 291–304.

[47] L. de Branges, *Some Hilbert spaces of analytic functions I*, Trans. Amer. Math. Soc. **106** (1963), 445–468.

[48] L. de Branges, *Espaces Hilbertiens de fonctions entières*, Masson, Paris, 1972.

[49] L. de Branges and J. Rovnyak, *Canonical models in quantum scattering theory*, Wiley, New York, 1966, 295–392.

[50] L. de Branges and J. Rovnyak, *Square summable power series*, Holt, Rinehart and Winston, New York, 1966.

[51] A. Bruckstein, B. Levy, and T. Kailath, *Differential methods in inverse scattering*, SIAM J. Appl. Math. **45** (1985), 312–335.

[52] A. Bruckstein and T. Kailath, *Inverse scattering for discrete transmission-line models*, SIAM Rev. **29**(3) (1987), 359–389.

[53] C. Carathéodory, *Über den Variabilitätsbereich der Koeffizienten von Potenzreihen, die gegebene Werte nicht annehmen*, Math. Ann. **64** (1907), 95–115.

[54] C. Carathéodory, *Über den Variabilitätsbereich der Fourierschen Konstanten von positiven harmonischen Funktionen*, Palermo Rend. **32** (1911), 193–217.

[55] Ch. Chamfy, *Fonctions méromorphes sur le cercle unité et leurs séries de Taylor*, Ann. Inst. Fourier **8** (1958), 211–251.

[56] T. Constantinescu, *Schur analysis with a finite number of negative squares*, Operator Theory: Adv. Appl. **17**, Birkhäuser Verlag, Basel, 1986, 87–108.

[57] T. Constantinescu and A. Gheondea, *The Schur algorithm and coefficient characterizations for generalized Schur functions*, Proc. Amer. Math. Soc. **128**(9) (2000), 2705–2713.

[58] T. Constantinescu and J. L. Johnson, *Tensor algebras and displacement structure. I. The Schur algorithm*, Z. Anal. Anwendungen **21**(1) (2002), 3–20.

[59] T. Constantinescu and J. L. Johnson, *Tensor algebras and displacement structure. II. Non-commutative Szegő polynomials*, Z. Anal. Anwendungen **21**(3) (2002), 611–626.

[60] T. Constantinescu, A. Sayed, and T. Kailath, *Inverse scattering experiments, structured matrix inequalities, and tensor algebra*, Linear Algebra Appl. **343/344** (2002), 147–169.

[61] Ph. Delsarte, Y. Genin, and Y. Kamp, *Schur parametrization of positive definite block-Toeplitz systems*, SIAM J. Appl. Math. **36** (1979), 34–46.

[62] Ph. Delsarte, Y. Genin, and Y. Kamp, *A generalization of the Levinson algorithm for Hermitian Toeplitz matrices with any rank profile*, IEEE Trans. Acoust. Speech Signal Process **33**(4) (1985), 964–971.

[63] Ph. Delsarte, Y. Genin, and Y. Kamp, *Pseudo-Carathéodory functions and Hermitian Toeplitz matrices*, Philips J. Res. **41**(1), 1–54.

[64] E. Deprettere and P. Dewilde, *The generalized Schur algorithm*, Operator Theory: Adv. Appl. **29**, Birkhäuser Verlag, Basel, 1988, 97–115.

[65] M. Derevyagin, *On the Schur algorithm for indefinite moment problem*, Methods Funct. Anal. Topology **9**(2) (2003), 133–145.

[66] P. Dewilde and H. Dym, *Lossless chain scattering matrices and optimum linear prediction: the vector case*, Int. J. in Circuit Theory and Appl. **9** (1981), 135–175.

[67] P. Dewilde and H. Dym, *Schur recursions, error formulas and convergence of rational estimators for stationary stochastic processes*, IEEE Trans. Inform. Theory **27** (1981), 446–461.

[68] P. Dewilde and A.-J. van der Veen, *Time-varying systems and computations*, Kluwer Academic Publishers, Boston, MA, 1998.

[69] A. Dijksma, H. Langer, and H. de Snoo, *Eigenvalues and pole functions of Hamiltonian systems with eigenvalue depending boundary conditions*, Math. Nachr. **161** (1993), 107–154.

[70] A. Dijksma, H. Langer, A. Luger, and Yu. Shondin, *A factorization result for generalized Nevanlinna functions of the class $\mathcal{N}_\kappa$*, Integral Equations and Operator Theory **36**(1) (2000), 121–125.

[71] A. Dijksma, H. Langer, A. Luger, and Y. Shondin, *Minimal realizations of scalar generalized Nevanlinna functions related to their basic factorization*, Operator Theory: Adv. Appl. **154**, Birkhäuser Verlag, Basel, 2004, 69–90.

[72] A. Dijksma and G. Wanjala, *Generalized Schur functions and augmented Schur parameters*, Operator Theory: Adv. Appl. **162**, Birkhäuser Verlag, Basel, 2005, 135–144.

[73] W.F. Donoghue, *Monotone matrix functions and analytic continuation*, Grundlehren der mathematischen Wissenschaften **207**, Springer-Verlag, Berlin, 1974.

[74] V. Dubovoj, B. Fritzsche, and B. Kirstein, *Matricial version of the classical Schur problem*, Teubner Texte zur Mathematik **129**, B.G. Teubner Verlagsgesellschaft, Stuttgart–Leipzig, 1992.

[75] V. Dubovoy, *Shift operators contained in contractions, Schur parameters and pseudocontinuable Schur functions*, Operator Theory: Adv. Appl. **165**, Birkhäuser Verlag, Basel, 2006, 175–250.

[76] J. Dufresnoy, *Le problème des coefficients pour certaines fonctions méromorphes dans le cercle unité*, Ann. Acad. Sci. Fenn. Ser. A., I. Mathematica **250**/9 (1958), 1–7.

[77] J. Dufresnoy and Ch. Pisot, *Sur un ensemble fermé d'entiers algébriques*, Ann. Sci. Ecole Norm. Sup. **70** (1953), 105–133.

[78] J. Dufresnoy and Ch. Pisot, *Étude de certaines fonctions méromorphes bornées sur le cercle unité. Application à un ensemble fermé d'entiers algébriques*, Ann. Sci. Ecole Norm. Sup. **72** (1955), 69–92.

[79] H. Dym, *J-contractive matrix functions, reproducing kernel Hilbert spaces and interpolation*, Conference Board of the Mathematical Sciences, regional conference series in mathematics **71**, Amer. Math. Soc., Providence, RI, 1989.

[80] H. Dym, *On reproducing kernel spaces, J-unitary matrix functions, interpolation and displacement rank*, Operator Theory: Adv. Appl. **41**, Birkhäuser Verlag, Basel, 1989, 173–239.

[81] H. Dym and A. Iacob, *Positive definite extensions, canonical equations and inverse problems*, Operator Theory: Adv. Appl. **12**, Birkhäuser Verlag, Basel, 1984, 141–240.

[82] H. Dym and V. Katsnelson, *Contributions of Issai Schur to analysis*, Progr. Math. **210**, Birkhäuser Verlag, Boston, MA, 2003, xci–clxxxviii.

[83] B. Fritzsche and B. Kirstein (editors), *Ausgewählte Arbeiten zu den Ursprüngen der Schur-Analysis*, Teubner Archiv zur Mathematik **16**, B.G. Teubner Verlagsgesellschaft, Stuttgart–Leipzig, 1991.

[84] B. Fritzsche, B. Kirstein, and A. Lasarow, *The matricial Carathéodory problem in both nondegenerate and degenerate cases*, Operator Theory: Adv. Appl. **165**, Birkhäuser Verlag, Basel, 2006, 251–290.

[85] P.A. Fuhrmann, *Linear systems and operators in Hilbert space*, McGraw-Hill, New York, 1981.

[86] L. Geronimus, *Orthogonal polynomials*, Consultants Bureau, New York, 1961.

[87] G.M.L. Gladwell, *The application of Schur's algorithm to an inverse eigenvalue problem*, Inverse Problems **7**(4) (1991), 557–565.

[88] I. Gohberg (editor), *I. Schur methods in operator theory and signal processing*, Operator Theory: Adv. Appl. **18**, Birkhäuser Verlag, Basel, 1986.

[89] I.S. Iohvidov, M.G. Krein, and H. Langer, *Introduction to the spectral theory of operators in spaces with an indefinite metric*, Akademie-Verlag, Berlin, 1982.

[90] M. Kaashoek, *State space theory of rational matrix functions and applications*, Fields Institute Monographs **3**, Amer. Math. Soc., 1996, 235–333.

[91] M. Kaltenbäck and H. Woracek, *Unique prime factorization in a partial semigroup of matrix-polynomials*, Discussiones Mathematicae General Algebra and Applications **26** (2006), 21–43.

[92] S. Khrushchev, *Schur's algorithm, orthogonal polynomials, and convergence of Wall's continued fractions in* $L^2(\mathbb{T})$, J. Approx. Theory **108**(2) (2001), 161–248.

[93] M.G. Krein, *Integral representation of a continuous Hermitian-indefinite function with a finite number of negative squares*, Dokl. Akad. Nauk. SSSR **125** (1959), 31–34.

[94] M.G. Krein and H. Langer, *Über die verallgemeinerten Resolventen und die charakteristische Funktion eines isometrischen Operators im Raume* Π_k, Colloquia Math. Soc. Janos Bolyai, Tihany (Hungary), **5**, Hilbert space operators and operator algebras, North-Holland, Amsterdam, 1972, 353–399.

[95] M.G. Krein and H. Langer, *Über die Q-Funktion eines* π*-hermiteschen Operators im Raume* π_κ, Acta Sci. Math. Szeged **34** (1973),191–230.

[96] M.G. Krein and H. Langer, *Über einige Fortsetzungsprobleme, die eng mit der Theorie hermitescher Operatoren im Raume* π_k *zusammenhängen. I. Einige Funktionenklassen und ihre Darstellungen*, Math. Nachr. **77** (1977), 187–236.

[97] M.G. Krein and H. Langer, *On some extension problems which are closely connected with the theory of Hermitian operators in a space* π_κ*. III. Indefinite analogues of the Hamburger and Stieltjes moment problems. Part I*, Beiträge Anal. **14** (1979), 25–40.

[98] M.G. Krein and H. Langer, *On some extension problems which are closely connected with the theory of Hermitian operators in a space* π_κ*. III. Indefinite analogues of the Hamburger and Stieltjes moment problems. Part II*, Beiträge Anal. **15** (1980), 27–45.

[99] M.G. Krein and H. Langer, *Some propositions on analytic matrix functions related to the theory of operators in the space* Π_k, Acta Sci. Math. Szeged **43** (1981), 181–205.

[100] H. Langer, *A characterization of generalized zeros of negative type of the class* $\mathcal{N}_\kappa$, Operator Theory: Adv. Appl. **17**, Birkhäuser Verlag, Basel, 1986, 201–212.

[101] H. Lev-Ari, *Nonstationary lattice filter modeling*, PhD thesis, Stanford University, 1983.

[102] H. Lev-Ari and T. Kailath, *Lattice filter parametrization and modeling of nonstationary processes*, IEEE Trans. Inform. Theory **30**(1) (1984), 2–16.

[103] H. Lev-Ari and T. Kailath, *Triangular factorisation of structured hermitian matrices*, Operator Theory: Adv. Appl. **18**, Birkhäuser Verlag, Basel, 1986, 301–324.

[104] J. Mercer, *Functions of positive and negative type and their connection with the theory of integral equations*, Philos. Trans. Roy. Soc. London Ser. A **209** (1909), 415–446.

[105] Ch. Pisot, *Familles compactes de fractions rationnelles et ensembles fermés de nombres algébriques*, Ann. Sci. École Norm. Sup. **81** (1964), 165–188.

[106] P.I. Richards, *A special class of functions with positive real part in half-plane*, Duke Math. Jour. **14** (1947), 777–786.

[107] J. Rovnyak, *Characterization of spaces* $\mathfrak{H}(M)$, unpublished paper, 1968. Available at the URL `http://www.people.virginia.edu/ jlr5m/home.html`.

[108] S. Saitoh, *Theory of reproducing kernels and its applications*, Pitman Research Notes in Mathematics Series **189**, Longman Scientific & Technical, Harlow, 1988.

[109] R. Salem, *A remarkable class of algebraic integers. Proof of a conjecture of Vijayaraghavan*, Duke Math. J. **11** (1944), 103–108.

[110] R. Salem, *Sets of uniqueness and sets of multiplicity. II*, Trans. Amer. Math. Soc. **56** (1944), 32–49.

[111] R. Salem and A. Zygmund, *Sur les ensembles parfaits dissymétriques à rapport constant*, C. R. Acad. Sci. Paris **240** (1955), 2281–2283.

[112] R. Salem and A. Zygmund, *Sur un théorème de Piatetçki-Shapiro*, C. R. Acad. Sci. Paris **240** (1955), 2040–2042.

[113] D. Sarason, *Sub-Hardy Hilbert spaces in the unit disk*, University of Arkansas lecture notes in the mathematical sciences **10**, Wiley, New York, 1994.

[114] D. Sarason, *Nevanlinna-Pick interpolation with boundary data*, Integral Equations and Operator Theory **30**(2) (1998), 231–250.

[115] Z. Sasvári, *The extension problem for positive definite functions. A short historical survey*, Operator Theory: Adv. Appl. **163**, Birkhäuser Verlag, Basel, 2006, 365–379.

[116] I. Schur, *Über die Potenzreihen, die im Innern des Einheitkreises beschränkt sind, I*, J. Reine Angew. Math. **147** (1917), 205–232; English translation: Operator Theory: Adv. Appl. **18**, Birkhäuser Verlag, Basel, 1986, 31–59.

[117] I. Schur, *Über die Potenzreihen, die im Innern des Einheitkreises beschränkt sind, II*, J. Reine Angew. Math. **148** (1918), 122–145; English translation: Operator Theory: Adv. Appl. **18**, Birkhäuser Verlag, Basel, 1986, 61–88.

[118] L. Schwartz, *Sous espaces hilbertiens d'espaces vectoriels topologiques et noyaux associés (noyaux reproduisants)*, J. Analyse Math. **13** (1964), 115–256.

[119] J.A. Shohat and J.D. Tamarkin, *The Problem of Moments*, Amer. Math. Soc. Math. surveys **II**, Amer. Math. Soc., New York, 1943.

[120] P. Sorjonen, *Pontryagin Räume mit einem reproduzierenden Kern*, Ann. Acad. Sci. Fenn. Ser. A I Math. **594** (1975), 1–30.

[121] T. Takagi, *On an algebraic problem related to an analytic theorem of Carathéodory and Fejér and on an allied theorem of Landau*, Japanese J. Math. **1** (1924), 83–93.

[122] O. Toeplitz, *Über die Fouriersche Entwicklung positiver Funktionen*, Rend. Circ. Mat. Palermo **32** (1911), 191–192.

[123] P.P. Vaidyanathan and S.K. Mitra, *Discrete version of Richard's theorem and applications to cascaded lattice realization of digital filter transfer matrices and functions*, IEEE Trans. Circuits Systems **33** (1986), 26–34.

[124] M.E. van Valkenburg, *Modern Network Analysis*, John Wiley and Sons, New York, 1960.

[125] G. Wanjala, *Closely connected unitary realizations of the solutions to the basic interpolation problem for generalized Schur functions*, Operator Theory: Adv. Appl. **160**, Birkhäuser Verlag, Basel, 2005, 441–468.

[126] G. Wanjala, *The Schur transform of a generalized Schur function and operator realizations*, PhD thesis, University of Groningen, Groningen, The Netherlands, 2005.

[127] L. Weinberg, *Network analysis and synthesis*, Krieger Publishing Company, Melbourne, FL., 1976.

D. Alpay
Department of Mathematics
Ben-Gurion University of the Negev
P.O. Box 653
84105 Beer-Sheva, Israel
e-mail: `dany@math.bgu.ac.il`

A. Dijksma
Department of Mathematics
University of Groningen
P.O. Box 800
9700 AV Groningen, The Netherlands
e-mail: `dijksma@math.rug.nl`

H. Langer
Institute of Analysis and Computational Mathematics
Vienna University of Technology
Wiedner Hauptstrasse 8–10
1040 Vienna, Austria
e-mail: `hlanger@mail.zserv.tuwien.ac.at`

Operator Theory:
Advances and Applications, Vol. 176, 99–164

A Truncated Matricial Moment Problem on a Finite Interval. The Case of an Odd Number of Prescribed Moments

Abdon E. Choque Rivero, Yuriy M. Dyukarev,
Bernd Fritzsche and Bernd Kirstein

Abstract. The main goal of this paper is to study the truncated matricial moment problem on a finite closed interval in the case of an odd number of prescribed moments by using of the FMI method of V.P. Potapov. The solvability of this problem is characterized by the fact that two block Hankel matrices built from the data of the problem are nonnegative Hermitian (Theorem 1.3). An essential step to solve the problem under consideration is to derive an effective coupling identity between both block Hankel matrices (Proposition 2.5). In the case that these Hankel matrices are both positive Hermitian we parametrize the set of solutions via a linear fractional transformation the generating matrix-valued function of which is a matrix polynomial whereas the set of parameters consists of distinguished pairs of meromorphic matrix-valued functions.

Mathematics Subject Classification (2000). Primary 44A60, 47A57, 30E05.

Keywords. Matricial moment problem, system of fundamental matrix inequalities of Potapov-type, Stieltjes transform.

0. Introduction

This paper continues the authors' investigations on the truncated matricial moment problem on a finite closed interval of the real axis. The scalar version of this problem was treated by M.G. Krein (see [K], [KN, Chapter III]) by different methods. A closer look at this work shows that the cases of an even or odd number of prescribed moments were handled separately. These cases turned out to be intimately related with different classes of functions holomorphic outside the fixed interval $[a, b]$.

The same situation can be met in the matricial version of the moment problem under consideration. After studying the case of an even number of prescribed moments in [CDFK], we handle in this paper the case that the number of prescribed moments is odd. It is not quite unexpected that there are several features which are common to our treatment of both cases. This concern particularly our basic strategy. As in [CDFK] we will use V.P. Potapov's so-called Fundamental Matrix Inequality (FMI) approach in combination with L.A. Sakhnovich's method of operator identities. (According to applications of V.P. Potapov's approach to matrix versions of classical moment and interpolation problems we refer, e.g., to the papers Dubovoj [Du]; Dyukarev/Katsnelson [DK]; Dyukarev [Dy1]; Golinskii [G1], [G2]; Katsnelson [Ka1]–[Ka3]; Kovalishina [Ko1], [Ko2]. Concerning the operator identity method, e.g., the works [IS], [S1], [S2], [BS], and [AB] are mentioned. Roughly speaking, at a first view the comparison between the even and odd cases shows that the principal steps are similar whereas the detailed realizations of these steps are rather different. More precisely, the odd case turned out to be much more difficult. Some reasons for this will be listed in the following more concrete description of the contents of this paper. Similarly as in [CDFK], we will again meet the situation that our matrix moment problem has solutions if and only if two block Hankel matrices built from the given data are nonnegative Hermitian (see Theorem 1.3). Each of these block Hankel matrices satisfies a certain Ljapunov type identity (see Proposition 2.1). An essential point in the paper is to find an effective algebraic coupling between both block Hankel matrices. The desired coupling identity will be realized in Proposition 2.5. The comparison with the analogous coupling identity in the even case (see [CDFK, Proposition 2.2]) shows that the coupling identity in Proposition 2.5 is by far more complicated. This is caused by the fact that in contrast with [CDFK] now two block Hankel matrices of different sizes and much more involved structure have to be coupled. A first main result (see Theorem 1.2) indicates that (after Stieltjes transform) the original matrix moment problem is equivalent to a system of two fundamental matrix inequalities (FMI) of Potapov type. Our proof of Theorem 1.2 is completely different from the proof of the corresponding result in the even case (see [CDFK, Theorem 1.2]). Whereas the proof of Theorem 1.2 in [CDFK] is mainly based on an appropriate application of generalized inversion formula of Stieltjes-Perron type, in this paper we prefer a much more algebraically orientated approach. Hereby, in particular Lemma 4.5 should be mentioned. It occupies a key role in the proof of Theorem 1.2.

Assuming positive Hermitian block Pick matrices we will parametrize the set of all solutions of the system of FMI's of Potapov type. In the first step we will treat separately each of the two inequalities of the system by the factorization method of V.P. Potapov. The main difficulty is hidden in the second step. One has to find a suitable coupling between the solutions of the two single inequalities. This goal will be realized in Proposition 6.10. The proof of Proposition 6.10 is doubtless one of the crucial points of the whole paper. This is the place where we make essential use of the algebraic coupling identity stated in Proposition 2.5. Since this coupling identity is much more involved as its analogue in the case of an even number of prescribed moments (see [CDFK, Proposition 6.10]) it is not astonishing that the proof of Proposition 6.10 is much more complicated than the proof of the analogous result in [CDFK].

A common feature in our treatment of both cases is that the set of solutions will be parametrized via a linear fractional transformation the generating matrix-valued function of which is a matrix polynomial, whereas the set of parameters consists of different classes of distinguished pairs of matrix-valued functions. More precisely, depending on the even or odd cases this class of pairs of meromorphic matrix-valued can be considered as a 'projective extension' of the classes $\mathcal{R}_q[a,b]$ and $\mathcal{S}_q[a,b]$ of holomorphic matrix-valued functions introduced in Sections 1 and 5, respectively.

1. Notation and preliminaries

Throughout this paper, let p, q, and r be positive integers. We will use $\mathbb{C}, \mathbb{R}, \mathbb{N}_0$, and $\mathbb{N}$ to denote the set of all complex numbers, the set of all real numbers, the set of all nonnegative integers, and the set of all positive integers, respectively. For every nonnegative integers m and n, let $\mathbb{N}_{m,n}$ designate the set of all integers k which satisfy $m \leq k \leq n$. The notation $\mathbb{C}^{p\times q}$ stands for the set of all complex $p \times q$ matrices. Further, let $\Pi_+ := \{w \in \mathbb{C} : \operatorname{Im} w \in (0, +\infty)\}$, let $\Pi_- := \{w \in \mathbb{C} : \operatorname{Im} w \in (-\infty, 0)\}$, and we will write $\mathfrak{B}$ for the Borel σ-algebra on $\mathbb{R}$ (respectively, $\tilde{\mathfrak{B}}$ for the Borel σ-algebra on $\mathbb{C}$). The Borel σ-algebra on $\mathbb{C}^{p\times q}$ will be denoted by $\tilde{\mathfrak{B}}_{p\times q}$. If $\mathfrak{X}$ and $\mathcal{Y}$ are nonempty sets, if $f : \mathfrak{X} \to \mathcal{Y}$ is a mapping, and if $\mathcal{Z}$ is a nonempty subset of $\mathfrak{X}$, then $\operatorname{Rstr.}_{\mathcal{Z}} f$ stands for the restriction of f onto $\mathcal{Z}$.

The matricial generalization of M.G. Krein's classical moment problem considered in this paper is formulated using the notion of nonnegative Hermitian $q\times q$ measure. Let Λ be a nonempty set and let $\mathfrak{A}$ be a σ-algebra on Λ. A matrix-valued function μ whose domain is the σ-algebra $\mathfrak{A}$ and whose values belong to the set $\mathbb{C}^{q\times q}_{\geq}$ of all nonnegative Hermitian complex $q \times q$ matrices is called nonnegative Hermitian $q \times q$ measure on $(\Lambda, \mathfrak{A})$ if it is countably additive, i.e., if μ satisfies

$$\mu\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} \mu(A_j)$$

for each infinite sequence $(A_j)_{j=1}^{\infty}$ of pairwise disjoint sets that belong to $\mathfrak{A}$. We will use $\mathcal{M}^q_{\geq}(\Lambda, \mathfrak{A})$ to denote the set of all nonnegative Hermitian $q \times q$ measures on $(\Lambda, \mathfrak{A})$. Let $\mu = (\mu_{jk})_{j,k=1}^{q}$ belong to $\mathcal{M}^q_{\geq}(\Lambda, \mathfrak{A})$. Then every entry function μ_{jk} of μ is a complex-valued measure on $(\Lambda, \mathfrak{A})$. For each complex-valued function f defined on Λ which is, for all $j \in \mathbb{N}_{1,q}$ and all $k \in \mathbb{N}_{1,q}$, integrable with respect to the variation $|\mu_{jk}|$ of μ_{jk}, the integral

$$\int_\Lambda f d\mu := \left(\int_\Lambda f d\mu_{jk}\right)_{j,k=1}^{q} \tag{1.1}$$

is defined. We will also write $\int_\Lambda f(\lambda)\mu(d\lambda)$ for this integral.

Now let us formulate the matricial version of M.G. Krein's moment problem.

Let a and b be real numbers with $a < b$, let l be a nonnegative integer, and let $(s_j)_{j=0}^{l}$ be a sequence of complex $q \times q$ matrices. Describe the set $\mathcal{M}_{\geq}^{q}\left[[a,b], \mathfrak{B} \cap [a,b]; (s_j)_{j=0}^{l}\right]$ of all nonnegative Hermitian $q \times q$ measures σ which are defined on the Borel σ-algebra $\mathfrak{B} \cap [a,b]$ on the interval $[a,b]$ and which satisfy

$$\int_{[a,b]} t^j \sigma(dt) = s_j$$

for each integer j with $0 \leq j \leq l$.

We continue to assume that a and b are real numbers such that $a < b$. In this paper, we turn our attention to the case of an odd number of given moments, i.e., to the situation that $l = 2n$ holds with some nonnegative integer n. (The case of an even number of given moments is discussed in [CDFK].) If $n = 0$, then the situation is obvious: The set $\mathcal{M}_{\geq}^{q}\left[[a,b], \mathfrak{B} \cap [a,b]; (s_j)_{j=0}^{0}\right]$ is nonempty if and only if the matrix s_0 is nonnegative Hermitian. If s_0 is nonnegative Hermitian then $\mathcal{M}_{\geq}^{q}\left[[a,b], \mathfrak{B} \cap [a,b]; (s_j)_{j=0}^{0}\right]$ consists exactly of all nonnegative Hermitian $q \times q$ measures σ defined on $\mathfrak{B} \cap [a,b]$ such that $\sigma([a,b]) = s_0$. Thus we can focus our considerations on the case of a given sequence $(s_j)_{j=0}^{2n}$ of complex $q \times q$ matrices where n is a positive integer. According to the idea which was used by M.G. Krein and A.A. Nudelman in the scalar case $q = 1$ (see [KN, IV, §7]), by Stieltjes transformation we will translate the moment problem into the language of the class $\mathcal{R}_q[a,b]$ of matrix-valued functions $S : \mathbb{C} \setminus [a,b] \to \mathbb{C}^{q \times q}$ which satisfy the following four conditions:

(i) S is holomorphic in $\mathbb{C} \setminus [a,b]$.
(ii) For each $w \in \Pi_+$, the matrix $\operatorname{Im} S(w)$ is nonnegative Hermitian.
(iii) For each $t \in (-\infty, a)$, the matrix $S(t)$ is nonnegative Hermitian.
(iv) For each $t \in (b, +\infty)$, the matrix $-S(t)$ is nonnegative Hermitian.

Let us observe that, according to the investigations of M.G. Krein and A.A. Nudelman, one can show that the class $\tilde{\mathcal{R}}_q[a,b]$ of all matrix-valued functions $S : \Pi_+ \cup (\mathbb{R} \setminus [a,b]) \to \mathbb{C}^{q \times q}$ which satisfy (ii), (iii), (iv), and

(i′) : S is holomorphic in Π_+ and continuous in $H := \Pi_+ \cup (\mathbb{R} \setminus [a,b])$.

admits the representation $\tilde{\mathcal{R}}_q[a,b] = \{\mathrm{Rstr.}_H S : S \in \mathcal{R}_q[a,b]\}$.

The following theorem describes the interrelation between the set

$$\mathcal{M}_{\geq}^{q}([a,b], \mathfrak{B} \cap [a,b])$$

of all nonnegative Hermitian $q \times q$ measures defined on $\mathfrak{B} \cap [a,b]$ and the set $\mathcal{R}_q[a,b]$.

Theorem 1.1. (a) *For each $\sigma \in \mathcal{M}_{\geq}^{q}([a,b], \mathfrak{B} \cap [a,b])$, the matrix-valued function $S^{[\sigma]} : \mathbb{C} \setminus [a,b] \to \mathbb{C}^{q \times q}$ defined by*

$$S^{[\sigma]}(z) := \int_{[a,b]} \frac{1}{t-z} \sigma(dt) \tag{1.2}$$

belongs to $\mathcal{R}_q[a,b]$.

(b) *For each $S \in \mathcal{R}_q[a,b]$, there exists a unique nonnegative Hermitian measure $\sigma \in \mathcal{M}_{\geq}^{q}([a,b],\mathfrak{B}\cap[a,b])$ such that*

$$S(z) = \int_{[a,b]} \frac{1}{t-z}\sigma(dt) \tag{1.3}$$

is satisfied for all $z \in \mathbb{C}\setminus[a,b]$.

Theorem 1.1 can be proved by modifying the proof in the case $q = 1$. This scalar case is considered in [KN, Appendix, Ch. 3]. According to Theorem 1.1, the mapping $f : \mathcal{M}_{\geq}^{q}([a,b],\mathfrak{B}\cap[a,b]) \to \mathcal{R}_q[a,b]$ given by $f(\sigma) := S^{[\sigma]}$ is bijective. For every nonnegative Hermitian measure $\sigma \in \mathcal{M}_{\geq}^{q}([a,b],\mathfrak{B}\cap[a,b])$, the matrix-valued function $S^{[\sigma]} : \mathbb{C}\setminus[a,b] \to \mathbb{C}^{q\times q}$ defined by (1.2) is called the *Stieltjes transform of* σ. Conversely, if a matrix-valued function $S \in \mathcal{R}_q[a,b]$ is given, then the unique $\sigma \in \mathcal{M}_{\geq}^{q}([a,b],\mathfrak{B}\cap[a,b])$ which satisfies (1.3) for all $z \in \mathbb{C}\setminus[a,b]$ is said to be the *Stieltjes measure of* S.

With these notations the matricial version of M.G. Krein's moment problem can be reformulated:

Let a and b be real numbers with $a < b$, let l be a nonnegative integer, and let $(s_j)_{j=0}^{l}$ be a sequence of complex $q \times q$ matrices. Describe then the set $\mathcal{R}_q\left[[a,b];(s_j)_{j=0}^{l}\right]$ of the Stieltjes transforms of all nonnegative Hermitian measures which belong to $\mathcal{M}_{\geq}^{q}\left[[a,b],\mathfrak{B}\cap[a,b];(s_j)_{j=0}^{l}\right]$.

The consideration of this reformulated version of the moment problem has the advantage that one can apply function-theoretic methods. Because of the inclusion $\mathcal{R}_q\left[[a,b];(s_j)_{j=0}^{l}\right] \subseteq \mathcal{R}_q[a,b]$ it is an interpolation problem in the class $\mathcal{R}_q[a,b]$. Note that $\mathcal{M}_{\geq}^{q}\left[[a,b],\ \mathfrak{B}\cap[a,b];(s_j)_{j=0}^{l}\right] \neq \emptyset$ if and only if $\mathcal{R}_q\left[[a,b];(s_j)_{j=0}^{l}\right] \neq \emptyset$. As already mentioned in this paper we will consider the case that an odd number of moments is given. We will show that, for every nonnegative integer n, the set $\mathcal{R}_q\left[[a,b];(s_j)_{j=0}^{2n}\right]$ can be characterized as the set of solutions of an appropriately constructed system of two fundamental matrix inequalities of Potapov-type. To state this result we give some further notation. We will use I_q to designate the identity matrix which belongs to $\mathbb{C}^{q\times q}$. The notation $0_{p\times q}$ stands for the null matrix which belongs to $\mathbb{C}^{p\times q}$. If the size of an identity matrix or a null matrix is obvious, we will omit the indices. For each $A \in \mathbb{C}^{q\times q}$, the notation $\det A$ indicates the determinant of A. If $A \in \mathbb{C}^{q\times q}$, then let $\operatorname{Re} A$ and $\operatorname{Im} A$ be the real part of A and the imaginary part of A, respectively: $\operatorname{Re} A := \frac{1}{2}(A+A^*)$ and $\operatorname{Im} A := \frac{1}{2i}(A-A^*)$. For all $A \in \mathbb{C}^{p\times q}$, we will use A^+ to denote the Moore-Penrose inverse of A. Further, for each $A \in \mathbb{C}^{p\times q}$, let $\|A\|_E$ (respectively, $\|A\|$) be the Euclidean norm (respectively, operator norm) of A. The notation $\mathbb{C}_H^{q\times q}$ stands for the set of all Hermitian complex $q \times q$ matrices. If A and B are complex $q \times q$ matrices and if we write $A \geq B$ or $B \leq A$, then we mean that A and B are Hermitian complex matrices for which the matrix $A - B$ is nonnegative Hermitian. Further, if $\mathcal{Z}$ is a nonempty subset of $\mathbb{C}$ and if f is a matrix-valued function defined on $\mathcal{Z}$, then for each $z \in \mathcal{Z}$ the notation $f^*(z)$ is short for $(f(z))^*$. For all $j \in \mathbb{N}_0$ and all $k \in \mathbb{N}_0$, let δ_{jk} be the Kronecker symbol, i.e., let $\delta_{jk} := 1$ if $j = k$ and $\delta_{jk} := 0$ if $j \neq k$. For each $n \in \mathbb{N}_0$, let

$$T_n := (\delta_{j,k+1} I_q)_{j,k=0}^{n} \tag{1.4}$$

and let $R_{T_n} : \mathbb{C} \to \mathbb{C}^{(n+1)q \times (n+1)q}$ be defined by

$$R_{T_n}(z) := (I - zT_n)^{-1}. \tag{1.5}$$

Observe that, for each $n \in \mathbb{N}_0$, the matrix-valued function R_{T_n} can be represented via

$$R_{T_n}(z) = \begin{pmatrix} I_q & 0 & 0 & \dots & 0 & 0 \\ zI_q & I_q & 0 & \dots & 0 & 0 \\ z^2 I_q & zI & I_q & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ z^n I_q & z^{n-1} I_q & z^{n-2} I_q & \dots & zI_q & I_q \end{pmatrix} \tag{1.6}$$

and

$$R_{T_n}(z) = \sum_{j=0}^{n} z^j T_n^j \tag{1.7}$$

for each $z \in \mathbb{C}$. For each $n \in \mathbb{N}$, we will use the notations

$$\mathbf{R}_{1,n} := R_{T_n}, \quad \mathbf{R}_{2,n} := R_{T_{n-1}}, \quad \mathbf{T}_{1,n} := T_n, \quad \text{and} \quad \mathbf{T}_{2,n} := T_{n-1}. \tag{1.8}$$

Let $v_0 := I_q$ and, for each $n \in \mathbb{N}$, let

$$v_n := \begin{pmatrix} I_q \\ 0_{nq \times q} \end{pmatrix}, \quad \mathbf{v}_{1,n} := v_n, \quad \text{and} \quad \mathbf{v}_{2,n} := v_{n-1}. \tag{1.9}$$

For each $n \in \mathbb{N}_0$ and each sequence $(s_j)_{j=0}^{2n}$ of complex $q \times q$ matrices, let

$$\tilde{H}_{0,n} := \begin{pmatrix} s_0 & s_1 & s_2 & \dots & s_n \\ s_1 & s_2 & s_3 & \dots & s_{n+1} \\ s_2 & s_3 & s_4 & \dots & s_{n+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_n & s_{n+1} & s_{n+2} & \dots & s_{2n} \end{pmatrix}.$$

If $n \in \mathbb{N}_0$ and if $(s_j)_{j=1}^{2n+1}$ is a sequence of complex $q \times q$ matrices, then let

$$\tilde{H}_{1,n} := \begin{pmatrix} s_1 & s_2 & s_3 & \dots & s_{n+1} \\ s_2 & s_3 & s_4 & \dots & s_{n+2} \\ s_3 & s_4 & s_5 & \dots & s_{n+3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_{n+1} & s_{n+2} & s_{n+3} & \dots & s_{2n+1} \end{pmatrix}.$$

Further, for each $n \in \mathbb{N}_0$ and each sequence $(s_j)_{j=2}^{2n+2}$ of complex $q \times q$ matrices, let

$$\tilde{H}_{2,n} := \begin{pmatrix} s_2 & s_3 & s_4 & \dots & s_{2n+2} \\ s_3 & s_4 & s_5 & \dots & s_{2n+3} \\ s_4 & s_5 & s_6 & \dots & s_{2n+4} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_{n+2} & s_{n+3} & s_{n+4} & \dots & s_{2n+2} \end{pmatrix}.$$

Moreover, for all real numbers a and b such that $a < b$, for each positive integer n, and for all sequences $(s_j)_{j=0}^{2n}$, let

$$\mathbf{H}_{1,n} := \tilde{H}_{0,n} \quad \text{and} \quad \mathbf{H}_{2,n} := -ab\tilde{H}_{0,n-1} + (a+b)\tilde{H}_{1,n-1} - \tilde{H}_{2,n-1}. \tag{1.10}$$

If n is a positive integer and if $(s_j)_{j=0}^{2n}$ is a sequence of complex $q \times q$ matrices, then we will also use the following notations: For all $j, k \in \mathbb{N}_{0,2n}$ with $j < k$, let

$$y_{[j,k]} := \begin{pmatrix} s_j \\ s_{j+1} \\ \vdots \\ s_k \end{pmatrix} \quad \text{and} \quad z_{[j,k]} := (s_j, s_{j+1}, \dots, s_k). \tag{1.11}$$

Further, let $\mathbf{u}_{1,0} := 0_{q \times q}$ and, for each $n \in \mathbb{N}$,

$$\mathbf{u}_{1,n} := \begin{pmatrix} 0_{q \times q} \\ -y_{[0,n-1]} \end{pmatrix} \tag{1.12}$$

and

$$\mathbf{u}_{2,n} := -ab\mathbf{u}_{1,n-1} - (a+b)y_{[0,n-1]} + y_{[1,n]}. \tag{1.13}$$

Now we are able to formulate the first main result of this paper.

Theorem 1.2. *Let a and b be real numbers such that $a < b$, let n be a positive integer, and let $(s_j)_{j=0}^{2n}$ be a sequence of complex $q \times q$ matrices. Let $S : \mathbb{C} \setminus [a,b] \to \mathbb{C}^{q \times q}$ be a $q \times q$ matrix-valued function, let $S_1 := S$, and let $S_2 : \mathbb{C} \setminus [a,b] \to \mathbb{C}^{q \times q}$ be defined by*

$$S_2(z) := (z-a)(b-z)S(z) - s_0 z. \tag{1.14}$$

Then S belongs to $\mathcal{R}_q\left[[a,b]; (s_j)_{j=0}^{2n}\right]$ if and only if the following conditions are satisfied:

(i) *S is holomorphic in $\mathbb{C} \setminus [a,b]$.*

(ii) *For each $r \in \mathbb{N}_{1,2}$ and each $z \in \mathbb{C} \setminus \mathbb{R}$, the matrix*

$$\mathbf{K}_{r,n}^{[S]}(z) := \begin{pmatrix} \mathbf{H}_{r,n} & \mathbf{R}_{r,n}(z)(\mathbf{v}_{r,n}S_r(z) - \mathbf{u}_{r,n}) \\ (\mathbf{R}_{r,n}(z)(\mathbf{v}_{r,n}S_r(z) - \mathbf{u}_{r,n}))^* & \frac{S_r(z) - S_r^*(z)}{z - \overline{z}} \end{pmatrix} \tag{1.15}$$

is nonnegative Hermitian.

A proof of Theorem 1.2 will be given at the end of Section 4. We will use Theorem 1.2 in order to describe the case $\mathcal{M}_{\geq}^{q}\left[[a,b], \mathfrak{B} \cap [a,b]; (s_j)_{j=0}^{2n}\right] \neq \emptyset$. In Section 7, we will prove the following result which in the scalar case $q = 1$ is due to M.G. Krein (see [K], [KN, Theorem 2.3, S. 90]).

Theorem 1.3. *Let a and b be real numbers such that $a < b$, let n be a nonnegative integer, and let $(s_j)_{j=0}^{2n}$ be a sequence of complex $q \times q$ matrices. Then $\mathcal{M}_{\geq}^{q}\left[[a,b], \mathfrak{B} \cap [a,b]; (s_j)_{j=0}^{2n}\right]$ is nonempty if and only if the block Hankel matrices $\mathbf{H}_{1,n}$ and $\mathbf{H}_{2,n}$ are both nonnegative Hermitian.*

Let the assumption of Theorem 1.2 be satisfied. Then we will say that the matrix-valued function $S : \mathbb{C} \setminus [a,b] \to \mathbb{C}^{q \times q}$ is a *solution of the system of the fundamental matrix inequalities of Potapov-type associated with* $[a,b]$ *and* $(s_j)_{j=0}^{2n}$ if S is holomorphic in $\mathbb{C} \setminus [a,b]$ and if the matrix inequalities $\mathbf{K}_{1,n}^{[S]}(z) \geq 0$ and $\mathbf{K}_{2,n}^{[S]}(z) \geq 0$ hold for every choice of z in $\mathbb{C} \setminus \mathbb{R}$.

2. Main algebraic identities

In this section we will single out essential identities concerning the block matrices introduced in Section 1.

Proposition 2.1. (**Ljapunov type identities**) *Let $n \in \mathbb{N}_0$ and let $(s_j)_{j=0}^{2n}$ be a sequence of Hermitian complex $q \times q$ matrices. For each $r \in \mathbb{N}_{1,2}$, then*

$$\mathbf{H}_{r,n}\mathbf{T}_{r,n}^{*} - \mathbf{T}_{r,n}\mathbf{H}_{r,n} = \mathbf{u}_{r,n}\mathbf{v}_{r,n}^{*} - \mathbf{v}_{r,n}\mathbf{u}_{r,n}^{*}.$$

The proof of Proposition 2.1 is straightforward. We omit the lengthy calculations.

Now we are going to derive the main coupling identity between the block Hankel matrices $\mathbf{H}_{1,n}$ and $\mathbf{H}_{2,n}$. On the way to this identity we will obtain some auxiliary identities of own interest.

For each positive integer n, let

$$\mathbf{L}_{1,n} := (\delta_{j,k+1} I_q)_{\substack{j=0,\dots,n \\ k=0,\dots,n-1}} \quad \text{and} \quad \mathbf{L}_{2,n} := (\delta_{j,k} I_q)_{\substack{j=0,\dots,n \\ k=0,\dots,n-1}} \tag{2.1}$$

where $\delta_{j,k}$ is the Kronecker symbol: $\delta_{jk} := 1$ if $j = k$ and $\delta_{jk} := 0$ if $j \neq k$.

Remark 2.2. *Let $n \in \mathbb{N}$. Then the following identities hold:*

$$\mathbf{L}_{2,n}^{*}\mathbf{v}_{1,n} = \mathbf{v}_{2,n}, \quad \mathbf{L}_{2,n}\mathbf{v}_{2,n} = \mathbf{v}_{1,n}, \tag{2.2}$$

$$\mathbf{T}_{1,n}^{*}\mathbf{L}_{1,n} = \mathbf{L}_{2,n}, \tag{2.3}$$

$$\mathbf{T}_{1,n}\mathbf{L}_{2,n} = \mathbf{L}_{1,n}, \tag{2.4}$$

$$\mathbf{L}_{2,n}\mathbf{L}_{2,n}^{*}\mathbf{T}_{1,n}^{*} = \mathbf{T}_{1,n}^{*}, \tag{2.5}$$

$$\mathbf{L}_{1,n}\mathbf{L}_{2,n}^{*} = \mathbf{T}_{1,n}, \tag{2.6}$$

$$\mathbf{L}_{2,n}^{*}\mathbf{L}_{2,n} = I, \tag{2.7}$$

$$\mathbf{T}_{1,n}^{*}\mathbf{L}_{2,n} = \mathbf{L}_{2,n}\mathbf{T}_{2,n}^{*}, \tag{2.8}$$

and

$$\mathbf{T}_{1,n}^{*}\mathbf{L}_{1,n} - \mathbf{L}_{1,n}\mathbf{T}_{2,n}^{*} = \mathbf{v}_{1,n}\mathbf{v}_{2,n}^{*}. \tag{2.9}$$

Lemma 2.3. *Let $n \in \mathbb{N}$ and let $(s_j)_{j=0}^{2n}$ be a sequence of complex $q \times q$ matrices. Then*

$$-\mathbf{T}_{1,n}\mathbf{H}_{1,n}\mathbf{v}_{1,n} = \mathbf{u}_{1,n}, \tag{2.10}$$

$$\tilde{H}_{0,n-1}\mathbf{L}_{1,n}^* = \mathbf{L}_{2,n}^*\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*, \tag{2.11}$$

$$\mathbf{L}_{1,n}^*\mathbf{H}_{1,n}\mathbf{T}_{1,n}^* = \tilde{H}_{1,n-1}\mathbf{L}_{1,n}^*, \tag{2.12}$$

$$\mathbf{u}_{2,n} = -(a+b)\mathbf{L}_{2,n}^*\mathbf{H}_{1,n}\mathbf{v}_{1,n} + ab\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}\mathbf{H}_{1,n}\mathbf{v}_{1,n} + \mathbf{L}_{1,n}^*\mathbf{H}_{1,n}\mathbf{v}_{1,n}, \tag{2.13}$$

$$s_0\mathbf{v}_{2,n}^* + \mathbf{v}_{1,n}^*\mathbf{H}_{1,n}\mathbf{L}_{1,n}\mathbf{T}_{2,n}^* = \mathbf{v}_{1,n}^*\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*\mathbf{L}_{1,n}, \tag{2.14}$$

$$s_0\mathbf{v}_{1,n}^* - \mathbf{v}_{1,n}^*\mathbf{H}_{1,n}(I - a\mathbf{T}_{1,n}^*) - b\mathbf{u}_{1,n}^*(I - a\mathbf{T}_{1,n}^*) = -\mathbf{u}_{2,n}^*\mathbf{L}_{1,n}^*, \tag{2.15}$$

and

$$s_0\mathbf{v}_{2,n}^* + \mathbf{u}_{2,n}^*\mathbf{T}_{2,n}^* + \mathbf{u}_{1,n}^*\left(\mathbf{L}_{1,n}(I_{nq} - a\mathbf{T}_{2,n}^*) - b(I_{(n+1)q} - a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}\right) = 0. \tag{2.16}$$

Proof. Equation (2.13) implies that the left-hand side of (2.16) is equal to

$$\begin{aligned} &s_0\mathbf{v}_{2,n}^* + \mathbf{v}_{1,n}^*\mathbf{H}_{1,n}(\mathbf{L}_{1,n}\mathbf{T}_{2,n}^* - \mathbf{T}_{1,n}^*\mathbf{L}_{1,n}) - a\mathbf{v}_{1,n}^*\mathbf{H}_{1,n}(\mathbf{L}_{2,n} - \mathbf{T}_{1,n}^*\mathbf{L}_{1,n})\mathbf{T}_{2,n}^* \\ &\quad - b\mathbf{v}_{1,n}^*\mathbf{H}_{1,n}(\mathbf{L}_{2,n}\mathbf{T}_{2,n}^* - \mathbf{T}_{1,n}^*\mathbf{L}_{2,n}) + ab\mathbf{v}_{1,n}^*\mathbf{H}_{1,n}(\mathbf{L}_{2,n}\mathbf{T}_{2,n}^* - \mathbf{T}_{1,n}^*\mathbf{L}_{2,n}). \end{aligned}$$

Using Remark 2.2 thus the asserted equations follow by straightforward calculation. □

Observe that identity (2.10) also holds in the case $n = 0$.

Remark 2.4. *Let $n \in \mathbb{N}$. Then for every complex number z the identities*

$$(\mathbf{L}_{1,n}^* - z\mathbf{L}_{2,n}^*)\mathbf{R}_{1,n}(z)\mathbf{v}_{1,n} = 0, \tag{2.17}$$

$$\mathbf{L}_{2,n}(\mathbf{R}_{2,n}(\overline{z}))^* = (\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{L}_{2,n}, \tag{2.18}$$

$$(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{T}_{1,n}^* = \mathbf{L}_{2,n}(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*, \tag{2.19}$$

$$\mathbf{R}_{1,n}(z)\mathbf{T}_{1,n} = \mathbf{T}_{1,n}\mathbf{R}_{1,n}(z), \quad \text{and} \quad \mathbf{R}_{2,n}(z)\mathbf{T}_{2,n} = \mathbf{T}_{2,n}\mathbf{R}_{2,n}(z) \tag{2.20}$$

hold. Moreover,

$$\mathbf{R}_{k,n}(w)\mathbf{R}_{k,n}(z) = \mathbf{R}_{k,n}(z)\mathbf{R}_{k,n}(w) \tag{2.21}$$

and

$$\mathbf{R}_{k,n}(w) - \mathbf{R}_{k,n}(z) = (w-z)\mathbf{R}_{k,n}(z)\mathbf{T}_{k,n}\mathbf{R}_{k,n}(w) \tag{2.22}$$

are valid for each $k \in \{1, 2\}$ and all complex numbers w and z.

Note that the identities (2.21) and (2.22) also hold if $n = 0$ and $k = 1$. Further, the first identity in (2.20) is true for $n = 0$.

Now we state an essential coupling identity between the block Hankel matrices $\mathbf{H}_{1,n}$ and $\mathbf{H}_{2,n}$.

Proposition 2.5. (**Coupling Identity**) *Let $(s_j)_{j=0}^{2n}$ be a sequence of complex $q \times q$ matrices. Then*

$$\begin{aligned} &\mathbf{R}_{2,n}(a) \cdot (\mathbf{u}_{2,n}\mathbf{v}_{1,n}^* + a\mathbf{v}_{2,n}s_0\mathbf{v}_{1,n}^*)(\mathbf{R}_{1,n}(a))^* \\ &\quad = \mathbf{H}_{2,n}\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*(\mathbf{R}_{1,n}(a))^* + \mathbf{R}_{2,n}(a) \cdot \big(-(b-a)\mathbf{L}_{2,n}^*(I_{(n+1)q} - a\mathbf{T}_{1,n}) \\ &\qquad + (I_{nq} - a\mathbf{T}_{2,n})(\mathbf{L}_{1,n}^* - a\mathbf{L}_{2,n}^*)\big)\,\mathbf{H}_{1,n}. \end{aligned} \tag{2.23}$$

Proof. We denote the right-hand side of (2.23) by Φ_n. Then we can conclude

$$\Phi_n = \mathbf{R}_{2,n}(a)\left((I - a\mathbf{T}_{2,n})\mathbf{H}_{2,n}\mathbf{L}^*_{2,n}\mathbf{T}^*_{1,n} + \left[-(b-a)\mathbf{L}^*_{2,n}(I - a\mathbf{T}_{1,n})\right.\right.$$
$$\left.\left. + \ (I - a\mathbf{T}_{2,n})(\mathbf{L}^*_{1,n} - a\mathbf{L}^*_{2,n})\right]\mathbf{H}_{1,n}(I - a\mathbf{T}^*_{1,n})\right)(\mathbf{R}_{1,n}(a))^*$$

and hence using (2.8) and (1.10) then

$$\begin{aligned}\Phi_n &= \mathbf{R}_{2,n}(a)\Big(\Big[-ab\tilde{H}_{0,n-1} + (a+b)\tilde{H}_{1,n-1} - \tilde{H}_{2,n-1}\Big]\mathbf{L}^*_{2,n}\mathbf{T}^*_{1,n}\\
&-a\mathbf{T}_{2,n}\Big[-ab\tilde{H}_{0,n-1} + (a+b)\tilde{H}_{1,n-1} - \tilde{H}_{2,n-1}\Big]\mathbf{L}^*_{2,n}\mathbf{T}^*_{1,n}\\
&+\big[-b\mathbf{L}^*_{2,n} + ab\mathbf{L}^*_{2,n}\mathbf{T}_{1,n} + \mathbf{L}^*_{1,n} - a\mathbf{T}_{2,n}\mathbf{L}^*_{1,n}\big]\left(\mathbf{H}_{1,n} - a\mathbf{H}_{1,n}\mathbf{T}^*_{1,n}\right)\Big)(\mathbf{R}_{1,n}(a))^*\\
&= \mathbf{R}_{2,n}(a)\big(\mathbf{A}_{00} + a\mathbf{A}_{10} + b\mathbf{A}_{01} + a^2\mathbf{A}_{20} + ab\mathbf{A}_{11} + a^2b\mathbf{A}_{21}\big)(\mathbf{R}_{1,n}(a))^* \qquad (2.24)\end{aligned}$$

where

$$\mathbf{A}_{00} := -\tilde{H}_{2,n-1}\mathbf{L}^*_{2,n}\mathbf{T}^*_{1,n} + \mathbf{L}^*_{1,n}\mathbf{H}_{1,n},$$

$$\mathbf{A}_{10} := \tilde{H}_{1,n-1}\mathbf{L}^*_{2,n}\mathbf{T}^*_{1,n} + \mathbf{T}_{2,n}\tilde{H}_{2,n-1}\mathbf{L}^*_{2,n}\mathbf{T}^*_{1,n} - \mathbf{T}_{2,n}\mathbf{L}^*_{1,n}\mathbf{H}_{1,n} - \mathbf{L}^*_{1,n}\mathbf{H}_{1,n}\mathbf{T}^*_{1,n},$$

$$\mathbf{A}_{01} := \tilde{H}_{1,n-1}\mathbf{L}^*_{2,n}\mathbf{T}^*_{1,n} - \mathbf{L}^*_{2,n}\mathbf{H}_{1,n},$$

$$\mathbf{A}_{20} := -\mathbf{T}_{2,n}\tilde{H}_{1,n-1}\mathbf{L}^*_{2,n}\mathbf{T}^*_{1,n} + \mathbf{T}_{2,n}\mathbf{L}^*_{1,n}\mathbf{H}_{1,n}\mathbf{T}^*_{1,n},$$

$$\mathbf{A}_{11} := -\tilde{H}_{0,n-1}\mathbf{L}^*_{2,n}\mathbf{T}^*_{1,n} - \mathbf{T}_{2,n}\tilde{H}_{1,n-1}\mathbf{L}^*_{2,n}\mathbf{T}^*_{1,n} + \mathbf{L}^*_{2,n}\mathbf{T}_{1,n}\mathbf{H}_{1,n} + \mathbf{L}^*_{2,n}\mathbf{H}_{1,n}\mathbf{T}^*_{1,n},$$

and

$$\mathbf{A}_{21} := \mathbf{T}_{2,n}\tilde{H}_{0,n-1}\mathbf{L}^*_{2,n}\mathbf{T}^*_{1,n} - \mathbf{L}^*_{2,n}\mathbf{T}_{1,n}\mathbf{H}_{1,n}\mathbf{T}^*_{1,n}.$$

It is readily checked that

$$\mathbf{A}_{00} = y_{[1,n]}\mathbf{v}^*_{1,n}.$$

Taking into account (2.4) and (2.11) it follows

$$\mathbf{A}_{10} = \mathbf{T}_{2,n}\tilde{H}_{2,n-1}\mathbf{L}^*_{1,n} - \mathbf{T}_{2,n}\mathbf{L}^*_{1,n}\mathbf{H}_{1,n} = \mathbf{v}_{2,n}s_0\mathbf{v}^*_{1,n} - y_{[0,n-1]}\mathbf{v}^*_{1,n}$$

and

$$\begin{aligned}\mathbf{A}_{11} &= -\tilde{H}_{0,n-1}\mathbf{L}^*_{1,n} - \mathbf{T}_{2,n}\tilde{H}_{1,n-1}\mathbf{L}^*_{2,n}\mathbf{T}^*_{1,n} + \mathbf{L}^*_{2,n}\mathbf{T}_{1,n}\mathbf{H}_{1,n} + \mathbf{L}^*_{2,n}\mathbf{H}_{1,n}\mathbf{T}^*_{1,n}\\
&= -\mathbf{T}_{2,n}\tilde{H}_{1,n-1}\mathbf{L}^*_{2,n}\mathbf{T}^*_{1,n} + \mathbf{L}^*_{2,n}\mathbf{T}_{1,n}\mathbf{H}_{1,n} = -\mathbf{u}_{1,n-1}\mathbf{v}^*_{1,n}.\end{aligned}$$

Equation (2.4) also yields

$$\mathbf{A}_{01} = -y_{[0,n-1]}\mathbf{v}^*_{1,n}.$$

Further, one obtains easily $\mathbf{A}_{20} = 0$ and, using (2.8), $\mathbf{A}_{21} = 0$. Thus from (2.24) we get

$$\begin{aligned}\Phi_n &= \mathbf{R}_{2,n}(a)\left(y_{[1,n]}\mathbf{v}^*_{1,n} + a(\mathbf{v}_{2,n}s_0\mathbf{v}^*_{1,n} - y_{[0,n-1]}\mathbf{v}^*_{1,n}) - by_{[0,n-1]}\mathbf{v}^*_{1,n}\right.\\
&\quad\left.-ab\mathbf{u}_{1,n-1}\mathbf{v}^*_{1,n}\right)(\mathbf{R}_{1,n}(a))^*\\
&= \mathbf{R}_{2,n}(a)\left(\mathbf{u}_{2,n}\mathbf{v}^*_{1,n} + a\mathbf{v}_{2,n}s_0\mathbf{v}^*_{1,n}\right)(\mathbf{R}_{1,n}(a))^*. \qquad \square\end{aligned}$$

3. From the moment problem to the system of fundamental matrix inequalities of Potapov-type

In this section, we will show that every matrix-valued function which belongs to the set $\mathcal{R}_q\left[[a,b];(s_j)_{j=0}^{2n}\right]$ is a solution of the system of the fundamental matrix inequalities associated with $[a,b]$ and $(s_j)_{j=0}^{2n}$. Hereby, we will use integral representations for special matrices and functions. For this reason, first we recall some results of the integration theory of nonnegative Hermitian measures (for details, see [Kt], [R1],[R2], [R3]).

Let $(\Lambda,\mathfrak{A})$ be a measurable space. For each subset A of Λ, we will write 1_A for the indicator function of the set A (defined on Λ). If ν is a nonnegative real-valued measure on $(\Lambda,\mathfrak{A})$, then let $p\times q-\mathcal{L}^1(\Lambda,\mathfrak{A},\nu;\mathbb{C})$ denote the class of all $\mathfrak{A}-\mathfrak{B}_{p\times q}$-measurable complex $p\times q$ matrix-valued functions $\Phi=(\varphi_{jk})_{\substack{j=1,\dots,p\\k=1,\dots,q}}$ defined on Λ for which every entry function φ_{jk} is integrable with respect to ν.

Now let $\mu\in\mathcal{M}_{\geq}^q(\Lambda,\mathfrak{A})$. Then every entry function μ_{jk} of $\mu=(\mu_{jk})_{j,k=1}^n$ is a complex-valued measure on $(\Lambda,\mathfrak{A})$. In particular, $\mu_{11},\mu_{22},\dots,\mu_{qq}$ are finite nonnegative real-valued measures on $(\Lambda,\mathfrak{A})$. Moreover, μ is absolutely continuous with respect to the so-called trace measure $\tau:=\sum_{j=1}^q\mu_{jj}$ of μ, i.e., for each $A\in\mathfrak{A}$ which satisfies $\tau(A)=0$ it follows $\mu(A)=0_{q\times q}$. The corresponding Radon-Nikodym derivatives $\frac{d\mu_{jk}}{d\tau}$ are thus well defined up to sets of zero τ-measure. Setting

$$\mu'_\tau:=\left(\frac{d\mu_{jk}}{d\tau}\right)_{j,k=1}^q$$

we have then

$$\mu(A)=\left(\int_A\frac{d\mu_{jk}}{d\tau}d\tau\right)_{j,k=1}^q=\int_A\mu'_\tau d\tau$$

for each $A\in\mathfrak{A}$. An ordered pair $[\Phi,\Psi]$ consisting of an $\mathfrak{A}-\tilde{\mathfrak{B}}_{p\times q}$-measurable complex $p\times q$ matrix-valued function $\Phi=(\varphi_{jk})_{\substack{j=1,\dots,p\\k=1,\dots,q}}$ defined on Λ and an $\mathfrak{A}-\tilde{\mathfrak{B}}_{r\times q}$-measurable complex $r\times q$ matrix-valued function $\Psi=(\psi_{lk})_{\substack{l=1,\dots,r\\k=1,\dots,q}}$ defined on Λ is said to be left-integrable with respect to μ if $\Phi\mu'_\tau\Psi^*$ belongs to $p\times q-\mathcal{L}^1(\Lambda,\mathfrak{A},\mu;\mathbb{C})$. In this case, for each $A\in\mathfrak{A}$, the ordered pair $[1_A\Phi,1_A\Psi]$ is also left-integrable with respect to μ and the integral of $[\Phi,\Psi]$ over A is defined by

$$\int_A\Phi d\mu\Psi^*:=\int_\Lambda(1_A\Phi)\mu'_\tau(1_A\Psi)^*d\tau.$$

We will also write $\int_A\Phi(\lambda)\mu(d\lambda)\Psi^*(\lambda)$ for this integral. Let us consider an arbitrary σ-finite nonnegative real-valued measure ν on $(\Lambda,\mathfrak{A})$ such that μ is absolutely continuous with respect to ν and let $\mu'_\nu:=\left(\frac{d\mu_{jk}}{d\nu}\right)_{j,k=1}^q$ be a version of the matrix-valued function of the corresponding Radon-Nikodym derivatives. For each ordered pair $[\Phi,\Psi]$ of an $\mathfrak{A}-\tilde{\mathfrak{B}}_{p\times q}$-measurable matrix-valued function $\Phi:\Lambda\to\mathbb{C}^{p\times q}$ and an $\mathfrak{A} - \tilde{\mathfrak{B}}_{r\times q}$-measurable matrix-valued function $\Psi : \Lambda \to \mathbb{C}^{r\times q}$ which is left-integrable with respect to μ and each $A \in \mathfrak{A}$, then

$$\int_A \Phi d\mu \Psi^* = \int_A (1_A\Phi)\mu'_\nu(1_A\Psi)^* d\nu$$

holds. We will use $p\times q - \mathcal{L}^2(\Lambda, \mathfrak{A}, \mu)$ to denote the set of all $\mathfrak{A} - \tilde{\mathfrak{B}}_{p\times q}$-measurable mappings $\Phi : \Lambda \to \mathbb{C}^{p\times q}$ for which the pair $[\Phi, \Phi]$ is left-integrable with respect to μ. Note that if $\Phi \in p \times q - \mathcal{L}^2(\Lambda, \mathfrak{A}, \mu)$ and if $\Psi \in r \times q - \mathcal{L}^2(\Lambda, \mathfrak{A}, \mu)$, then the pair $[\Phi, \Psi]$ is left-integrable with respect to μ. If $\Phi : \Lambda \to \mathbb{C}^{p\times q}$ is an $\mathfrak{A} - \tilde{\mathfrak{B}}_{p\times q}$-measurable mapping for which a set $N \in \mathfrak{A}$ with $\mu(N) = 0$ and a nonnegative real number C exist such that $\|\Phi(\lambda)\| \le C$ holds for each $\lambda \in \Lambda \setminus N$, then Φ belongs to $p\times q - \mathcal{L}^2(\Lambda, \mathfrak{A}, \mu)$. For all complex-valued functions f and g which are defined on Λ and for which the function $h := f\overline{g}$ is integrable with respect to $|\mu_{jk}|$ for every choice of j and k in $\mathbb{N}_{1,q}$, the pair $[fI_q, gI_q]$ is left-integrable with respect to μ and, in view of (1.1),

$$\int_A (fI_q)d\mu(gI_q)^* = \int_A (f\overline{g})d\mu \tag{3.1}$$

holds for all $A \in \mathfrak{A}$.

Remark 3.1. *Let $\mu \in \mathcal{M}^q_{\geq}(\Lambda, \mathfrak{A})$ and let $\Phi \in p \times q - \mathcal{L}^2(\Lambda, \mathfrak{A}, \mu)$. Then $\mu_{[\Phi]} : \mathfrak{A} \to \mathbb{C}^{p\times p}$ given by*

$$\mu_{[\Phi]}(A) := \int_A \Phi d\mu \Phi^*$$

belongs to $\mathcal{M}^p_{\geq}(\Lambda, \mathfrak{A})$. If $\Psi : \Lambda \to \mathbb{C}^{r\times p}$ is $\mathfrak{A} - \tilde{\mathfrak{B}}_{r\times p}$-measurable and if $\Theta : \Lambda \to \mathbb{C}^{t\times p}$ is $\mathfrak{A} - \tilde{\mathfrak{B}}_{t\times p}$-measurable, then $[\Psi, \Theta]$ is left-integrable with respect to $\mu_{[\Phi]}$ if and only if $[\Psi\Phi, \Theta\Phi]$ is left-integrable with respect to μ. Moreover, in this case,

$$\int_\Lambda \Psi d\mu_{[\Phi]}\Theta^* = \int_\Lambda \Psi\Phi d\mu(\Theta\Phi)^*.$$

Remark 3.2. *Let $\mu \in \mathcal{M}^q_{\geq}(\Lambda, \mathfrak{A})$ and let $\mathfrak{C} := \{\lambda \in \Lambda : \{\lambda\} \in \mathfrak{A}\}$. Then one can easily see that $\mathfrak{C}_\mu := \{\lambda \in \mathfrak{C} : \mu(\{\lambda\}) \neq 0\}$ is a countable subset of Λ.*

Throughout this paper, we assume now that a and b are real numbers which satisfy $a < b$. Let us turn our attention to nonnegative Hermitian $q\times q$ measures on the Borel σ-algebra $\mathfrak{B}\cap[a, b]$ on the closed finite interval $[a, b]$. For each nonnegative Hermitian measure $\sigma \in \mathcal{M}^q_{\geq}([a, b], \mathfrak{B} \cap [a, b])$ and each $j \in \mathbb{N}_0$, let

$$s_j^{[\sigma]} := \int_{[a,b]} t^j \sigma(dt). \tag{3.2}$$

Further, for all $\sigma \in \mathcal{M}^q_{\geq}([a, b], \mathfrak{B} \cap [a, b])$ and all $m \in \mathbb{N}_0$, let

$$\tilde{H}^{[\sigma]}_{0,m} := (s^{[\sigma]}_{j+k})^m_{j,k=0}, \quad \tilde{H}^{[\sigma]}_{1,m} := (s^{[\sigma]}_{j+k+1})^m_{j,k=0}, \quad \tilde{H}^{[\sigma]}_{2,m} := (s^{[\sigma]}_{j+k+2})^m_{j,k=0}, \tag{3.3}$$

$$\mathbf{H}^{[\sigma]}_{1,m} := \tilde{H}^{[\sigma]}_{0,m}, \tag{3.4}$$

and in the case $m \geq 1$ moreover

$$\mathbf{H}_{2,m}^{[\sigma]} := -ab\tilde{H}_{0,m-1}^{[\sigma]} + (a+b)\tilde{H}_{1,m-1}^{[\sigma]} - \tilde{H}_{2,m-1}^{[\sigma]}. \tag{3.5}$$

For each $m \in \mathbb{N}_0$, let the $(m+1)q \times q$ matrix polynomial E_m be defined by

$$E_m(z) := \begin{pmatrix} I_q \\ zI_q \\ z^2 I_q \\ \vdots \\ z^m I_q \end{pmatrix}. \tag{3.6}$$

Obviously, from (1.9) we get $E_m(0) = v_m$ for each $m \in \mathbb{N}_0$. Further, for each $m \in \mathbb{N}_0$ and each $z \in \mathbb{C}$, from (1.6) and (1.9) it follows immediately

$$E_m(z) = R_{T_m}(z)v_m. \tag{3.7}$$

Now we state important integral representations for the block Hankel matrices introduced in (3.3), (3.4), and (3.5).

Lemma 3.3. *Let $\sigma \in \mathcal{M}_{\geq}^q([a,b], \mathfrak{B} \cap [a,b])$. For each $m \in \mathbb{N}_0$, then*

$$\mathbf{H}_{1,m}^{[\sigma]} = \int_{[a,b]} E_m(t)\sigma(dt)E_m^*(t), \quad \tilde{H}_{1,m}^{[\sigma]} = \int_{[a,b]} tE_m(t)\sigma(dt)E_m^*(t),$$

$$\tilde{H}_{2,m}^{[\sigma]} = \int_{[a,b]} t^2 E_m(t)\sigma(dt)E_m^*(t),$$

and in the case $m \geq 1$ moreover

$$\mathbf{H}_{2,m}^{[\sigma]} = \int_{[a,b]} \sqrt{(t-a)(b-t)}E_{m-1}(t)\sigma(dt)\left(\sqrt{(t-a)(b-t)}E_{m-1}(t)\right)^*.$$

In particular, for each $m \in \mathbb{N}_0$, the matrices $\mathbf{H}_{1,m}^{[\sigma]}$ and $\tilde{H}_{2,m}^{[\sigma]}$ are all nonnegative Hermitian and, moreover, the matrix $\tilde{H}_{1,m}^{[\sigma]}$ is Hermitian. Furthermore, for each $m \in \mathbb{N}$ the matrix $\mathbf{H}_{2,m}^{[\sigma]}$ is nonnegative Hermitian.

The proof of Lemma 3.3 is straightforward. We omit the details.

Remark 3.4. *Let $n \in \mathbb{N}$ and let $(s_j)_{j=0}^{2n}$ be a sequence of complex $q \times q$ matrices such that $\mathcal{M}_{\geq}^q\left[[a,b], \mathfrak{B} \cap [a,b]; (s_j)_{j=0}^{2n}\right] \neq \emptyset$. From Lemma 3.3 one can easily see then that all the matrices $\mathbf{H}_{1,n}, \tilde{H}_{2,n-1}$, and $\mathbf{H}_{2,n}$ are nonnegative Hermitian and that the matrix $\tilde{H}_{1,n-1}$ is Hermitian. In particular, $s_j^* = s_j$ for each $j \in \mathbb{N}_{0,2n}$ and $s_{2k} \in \mathbb{C}_{\geq}^{q\times q}$ for each $k \in \mathbb{N}_{0,n}$.*

Now we turn our attention to the right lower blocks in the block matrices given by (1.15). Let $\mathcal{R}_q$ be the set of all matrix-valued functions $F : \Pi_+ \to \mathbb{C}^{q\times q}$ which are holomorphic in Π_+ and which satisfy $\operatorname{Im} F(w) \geq 0$ for each $w \in \Pi_+$.

Lemma 3.5. *Let $S \in \mathcal{R}_q[a,b]$ and σ be the Stieltjes measure of S.*

(a) *The matrix-valued function* $S_1 := S$ *is holomorphic in* $\mathbb{C} \setminus [a,b]$ *and*

$$\frac{S_1(z) - S_1^*(z)}{z - \overline{z}} = \int_{[a,b]} \left(\frac{1}{\overline{t-z}} I_q \right) \sigma(dt) \left(\frac{1}{\overline{t-z}} I_q \right)^* \tag{3.8}$$

holds for each $z \in \mathbb{C} \setminus \mathbb{R}$. *Moreover,* $\operatorname{Rstr.}_{\Pi_+} S_1$ *belongs to* $\mathcal{R}_q$.

(b) *The matrix-valued function* $S_2 : \mathbb{C} \setminus [a,b] \to \mathbb{C}^{q \times q}$ *defined by* (1.14) *is holomorphic in* $\mathbb{C} \setminus [a,b]$ *and*

$$\frac{S_2(z) - S_2^*(z)}{z - \overline{z}} = \int_{[a,b]} \left(\frac{\sqrt{(t-a)(b-t)}}{\overline{t-z}} I_q \right) \sigma(dt) \left(\frac{\sqrt{(t-a)(b-t)}}{\overline{t-z}} I_q \right)^* \tag{3.9}$$

holds for each $z \in \mathbb{C} \setminus \mathbb{R}$. *Moreover,* $\operatorname{Rstr.}_{\Pi_+} S_2$ *belongs to* $\mathcal{R}_q$.

Proof. We are going to prove part (b). Part (a) can be verified analogously. Obviously, S_2 is holomorphic in $\mathbb{C} \setminus [a,b]$. Let $z \in \mathbb{C} \setminus [a,b]$. Then we get

$$\begin{aligned}
S_2(z) - S_2^*(z) &= (z-a)(b-z) \int_{[a,b]} \frac{1}{t-z} \sigma(dt) - z \int_{[a,b]} \sigma(dt) \\
&\quad -(\overline{z}-a)(b-\overline{z}) \int_{[a,b]} \frac{1}{\overline{t-z}} \sigma(dt) + \overline{z} \int_{[a,b]} \sigma(dt) \\
&= \int_{[a,b]} \frac{(z-a)(b-z)(t-\overline{z}) - (\overline{z}-a)(b-\overline{z})(t-z)}{(t-z)\overline{(t-z)}} \sigma(dt) \\
&\quad -(z-\overline{z}) \int_{[a,b]} \sigma(dt) \\
&= (z-\overline{z}) \int_{[a,b]} \frac{(t-a)(b-t)}{(t-z)\overline{(t-z)}} \sigma(dt) \\
&= (z-\overline{z}) \int_{[a,b]} \left(\frac{\sqrt{(t-a)(b-t)}}{t-z} I_q \right) \sigma(dt) \left(\frac{\sqrt{(t-a)(b-t)}}{t-z} I_q \right)^*
\end{aligned}$$

and consequently (3.9) follows. The right-hand side of (3.9) is nonnegative Hermitian. Thus for each $z \in \Pi_+$ we have

$$\operatorname{Im} S_2(z) = (\operatorname{Im} z) \cdot \frac{S_2(z) - S_2^*(z)}{z - \overline{z}} \in \mathbb{C}_{\geq}^{q \times q}.$$

Hence $\operatorname{Rstr.}_{\Pi_+} S_2$ belongs to $\mathcal{R}_q$. □

Now we give integral representations for the block matrices defined by (1.15). These formulas provide the key for the proof of the main result of this section.

Lemma 3.6. *Let* $n \in \mathbb{N}$ *and let* $(s_j)_{j=0}^{2n}$ *be a sequence of complex* $q \times q$ *matrices such that* $\mathcal{M}_{\geq}^q \left[[a,b], \mathfrak{B} \cap [a,b]; (s_j)_{j=0}^{2n} \right] \neq \emptyset$. *Let* $S \in \mathcal{R}_q \left[[a,b]; (s_j)_{j=0}^{2n} \right]$ *and let* σ *be the Stieltjes measure of* S. *Let* $S_1 := S$ *and let* $S_2 : \mathbb{C} \setminus [a,b] \to \mathbb{C}^{q \times q}$ *be defined* *by* (1.14). *For each* $z \in \mathbb{C} \setminus \mathbb{R}$, *then the matrices* $\mathbf{K}_{1,n}^{[S]}(z)$ *and* $\mathbf{K}_{2,n}^{[S]}(z)$ *given by* (1.15) *admit the representations*

$$\mathbf{K}_{1,n}^{[S]}(z) = \int_{[a,b]} \begin{pmatrix} E_n(z) \\ \frac{1}{\overline{t-z}} I_q \end{pmatrix} \sigma(dt) \begin{pmatrix} E_n(z) \\ \frac{1}{\overline{t-z}} I_q \end{pmatrix}^* \tag{3.10}$$

and

$$\mathbf{K}_{2,n}^{[S]}(z) = \int_{[a,b]} \sqrt{(t-a)(b-t)} \begin{pmatrix} E_{n-1}(t) \\ \frac{1}{\overline{t-z}} I_q \end{pmatrix} \sigma(dt) \left(\sqrt{(t-a)(b-t)} \begin{pmatrix} E_{n-1}(t) \\ \frac{1}{\overline{t-z}} I_q \end{pmatrix} \right)^*, \tag{3.11}$$

and are, in particular, both nonnegative Hermitian.

Proof. Since S belongs to $\mathcal{R}_q\left[[a,b];(s_j)_{j=0}^{2n}\right]$ we have $s_j^{[\sigma]} = s_j$ for each $j \in \mathbb{N}_{0,2n}$. Consequently $\mathbf{H}_{1,n}^{[\sigma]} = \mathbf{H}_{1,n}$ and $\mathbf{H}_{2,n}^{[\sigma]} = \mathbf{H}_{2,n}$. Thus Lemma 3.3 yields

$$\mathbf{H}_{1,n} = \int_{[a,b]} E_n(t)\sigma(dt)E_n^*(t) \tag{3.12}$$

and

$$\mathbf{H}_{2,n} = \int_{[a,b]} \sqrt{(t-a)(b-t)}E_{n-1}(t)\sigma(dt)\left(\sqrt{(t-a)(b-t)}E_{n-1}(t)\right)^*. \tag{3.13}$$

Because of Lemma 3.5 the identities (3.8) and (3.9) hold for every choice of z in $\mathbb{C} \setminus \mathbb{R}$. Now we are going to check that

$$\mathbf{R}_{1,n}(z)(\mathbf{v}_{1,n}S_1(z) - \mathbf{u}_{1,n}) = \int_{[a,b]} E_n(t)\sigma(dt)\left(\frac{1}{\overline{t-z}} I_q\right)^* \tag{3.14}$$

and

$$\begin{aligned} &\mathbf{R}_{2,n}(z)(\mathbf{v}_{2,n}S_2(z) - \mathbf{u}_{2,n}) \\ &= \int_{[a,b]} \sqrt{(t-a)(b-t)}E_{n-1}(t)\sigma(dt)\left(\frac{\sqrt{(t-a)(b-t)}}{\overline{t-z}} I_q\right)^* \end{aligned} \tag{3.15}$$

hold for each $z \in \mathbb{C} \setminus \mathbb{R}$. Using (3.12) we obtain

$$\begin{aligned} \mathbf{u}_{1,n} &= -T_n y_{[0,n]} = -T_n\mathbf{H}_{1,n}v_n = -\int_{[a,b]} T_nE_n(t)\sigma(dt)(v_n^*E_n(t))^* \\ &= -\int_{[a,b]} T_nR_{T_n}(t)v_n\sigma(dt)I_q^*. \end{aligned} \tag{3.16}$$

For every choice of z in $\mathbb{C} \setminus \mathbb{R}$ and t in $\mathbb{R}$, from (2.22) and (1.8) we get

$$\frac{R_{T_n}(z)}{t-z} + R_{T_n}(z)T_nR_{T_n}(t) = \frac{R_{T_n}(t)}{t-z}.$$

From (3.16) we can conclude then

$$\begin{aligned}
&\mathbf{R}_{1,n}(z)(\mathbf{v}_{1,n}S_1(z)-\mathbf{u}_{1,n})\\
&= R_{T_n}(z)\left(v_n\int_{[a,b]}\frac{1}{t-z}I_q\sigma(dt)I_q^* + \int_{[a,b]}T_nR_{T_n}(t)v_n\sigma(dt)I_q^*\right)\\
&= \int_{[a,b]}\left(\frac{R_{T_n}(z)}{t-z}+R_{T_n}(z)T_nR_{T_n}(t)\right)v_n\sigma(dt)I_q^*\\
&= \int_{[a,b]}\frac{R_{T_n}(t)}{t-z}v_n\sigma(dt)I_q^* = \int_{[a,b]}E_n(t)\sigma(dt)\left(\frac{1}{\overline{t-z}}I_q\right)^*.
\end{aligned}$$

Thus (3.14) is proved. Using

$$\begin{aligned}
\mathbf{u}_{2,n} &= ab\int_{[a,b]}\mathbf{T}_{2,n}\mathbf{R}_{2,n}(t)\mathbf{v}_{2,n}\sigma(dt)I_q^* - (a+b)\int_{[a,b]}\mathbf{R}_{2,n}(t)\mathbf{v}_{2,n}\sigma(dt)I_q^*\\
&+\int_{[a,b]}\mathbf{R}_{2,n}(t)\mathbf{v}_{2,n}\sigma(dt)(tI_q)^*,
\end{aligned}$$

the identity

$$\begin{aligned}
&\frac{(z-a)(b-z)}{t-z}\mathbf{R}_{2,n}(z) - z\mathbf{R}_{2,n}(z) + (a+b)\mathbf{R}_{2,n}(z)\mathbf{R}_{2,n}(t)\\
&-ab\mathbf{R}_{2,n}(z)\mathbf{T}_{2,n}\mathbf{R}_{2,n}(t) - t\mathbf{R}_{2,n}(z)\mathbf{R}_{2,n}(t) = \frac{(t-a)(b-t)}{t-z}\mathbf{R}_{2,n}(t),
\end{aligned}$$

which is true for each $z\in\mathbb{C}\setminus\mathbb{R}$, and the equation

$$\mathbf{R}_{2,n}(z)\mathbf{v}_{2,n}S_2(z) = \int_{[a,b]}\left(\frac{(z-a)(b-z)}{t-z}\mathbf{R}_{2,n}(z) - z\mathbf{R}_{2,n}(z)\right)\mathbf{v}_{2,n}\sigma(dt)I_q^*$$

it follows analogously (3.15) for every choice of z in $\mathbb{C}\setminus\mathbb{R}$. From (3.12), (3.8), and (3.14) we obtain (3.10). The representation (3.11) is a consequence of (3.13), (3.15) and (3.9). □

Now we are going to state the main result of this section. If $n\in\mathbb{N}_0$ and if $(s_j)_{j=0}^{2n}$ is a sequence of complex $q\times q$ matrices, then we will use the notation $\mathcal{P}_q\left[[a,b];(s_j)_{j=0}^{2n}\right]$ to denote the set of all solutions of the system of the fundamental matrix inequalities of Potapov-type associated with $[a,b]$ and $(s_j)_{j=0}^{2n}$, i.e., the set of all matrix-valued functions $S:\mathbb{C}\setminus[a,b]\to\mathbb{C}^{q\times q}$ which are holomorphic in $\mathbb{C}\setminus[a,b]$ and for which the matrices $\mathbf{K}_{1,n}^{[S]}(z)$ and $\mathbf{K}_{2,n}^{[S]}(z)$ are both nonnegative Hermitian for every choice of z in $\mathbb{C}\setminus\mathbb{R}$.

Proposition 3.7. *Let $n\in\mathbb{N}$ and let $(s_j)_{j=0}^{2n}$ be a sequence of complex $q\times q$ matrices. Then $\mathcal{R}_q\left[[a,b];(s_j)_{j=0}^{2n}\right]$ is a subset of the set $\mathcal{P}_q\left[[a,b];(s_j)_{j=0}^{2n}\right]$ of all solutions of the system of the fundamental matrix inequalities of Potapov-type associated with $[a,b]$ and $(s_j)_{j=0}^{2n}$.*

Proof. Apply Lemma 3.6. □

4. From the system of fundamental matrix inequalities to the moment problem

Throughout this section, we again assume that a and b are real numbers which satisfy $a < b$. Further, let n be a positive integer and let $(s_j)_{j=0}^{2n}$ be a sequence of complex $q \times q$ matrices. We will continue to work with the notations given above. In particular, if a matrix-valued function $S : \mathbb{C} \setminus [a,b] \to \mathbb{C}^{q\times q}$ is given, then let $S_1 := S$ and let $S_2 : \mathbb{C} \setminus [a,b] \to \mathbb{C}^{q\times q}$ be defined by (1.14). We will again use the notation $\mathcal{P}_q\left[[a,b];(s_j)_{j=0}^{2n}\right]$ to denote the set of all solutions of the system of the fundamental matrix inequalities of Potapov-type associated with $[a,b]$ and $(s_j)_{j=0}^{2n}$.

Remark 4.1. *If $\mathcal{P}_q\left[[a,b];(s_j)_{j=0}^{2n}\right]$ is nonempty, then $\mathbf{H}_{1,n} \geq 0$, $\mathbf{H}_{2,n} \geq 0$, and in particular $s_j^* = s_j$ for each $j \in \mathbb{N}_{0,2n}$ and $s_{2k} \in \mathbb{C}_{\geq}^{q\times q}$ for each $k \in \mathbb{N}_{0,n}$. Moreover, $\mathcal{P}_q\left[[a,b];(s_j)_{j=0}^{2n}\right] \neq \emptyset$ implies $\tilde{H}_{1,n-1}^* = \tilde{H}_{1,n-1}$ and $\tilde{H}_{2,n-1} \geq 0$.*

Remark 4.2. *Suppose $\mathcal{P}_q\left[[a,b];(s_j)_{j=0}^{2n}\right] \neq \emptyset$. Let $S \in \mathcal{P}_q\left[[a,b];(s_j)_{j=0}^{2n}\right]$. Then S_1 and S_2 are both holomorphic in $\mathbb{C} \setminus [a,b]$. Let $r \in \{1,2\}$. For each $z \in \mathbb{C} \setminus \mathbb{R}$, we have*

$$\operatorname{Im} S_r(z) = \frac{S_r(z) - S_r^*(z)}{z - \overline{z}} \cdot \operatorname{Im} z.$$

Consequently, since the matrix $\mathbf{K}_{r,n}^{[S]}(z)$ is nonnegative Hermitian for each $z \in \mathbb{C}\setminus\mathbb{R}$, the matrix $\operatorname{Im} S_r(w)$ is nonnegative Hermitian for each $w \in \Pi_+$ and nonpositive Hermitian for each $w \in \Pi_-$.

Lemma 4.3. $\mathcal{P}_q\left[[a,b];(s_j)_{j=0}^{2n}\right] \subseteq \mathcal{R}_q[a,b]$.

Proof. It is sufficient to consider the case $\mathcal{P}_q\left[[a,b];(s_j)_{j=0}^{2n}\right] \neq \emptyset$.

Let $S \in \mathcal{P}_q\left[[a,b];(s_j)_{j=0}^{2n}\right]$. Then S_1 and S_2 are holomorphic in $\mathbb{C} \setminus [a,b]$ and Remarks 4.1 and 4.2 yield that

$$\operatorname{Im} S(w) = \operatorname{Im} S_1(w) \geq 0 \tag{4.1}$$

and

$$\operatorname{Im}\ [(w-a)(b-w)S(w)] \geq \operatorname{Im} S_2(w) \geq 0 \tag{4.2}$$

hold for every choice of w in Π_+. On the other hand, Remarks 4.1 and 4.2 imply for each $w \in \Pi_-$ the inequality

$$\operatorname{Im} S(w) = \operatorname{Im} S_1(w) \leq 0. \tag{4.3}$$

Let $t \in \mathbb{R} \setminus [a,b]$. From (4.1) it follows

$$\operatorname{Im} S(t) = \lim_{\varepsilon \to +0} \operatorname{Im} S(t + i\varepsilon) \geq 0.$$

Inequality (4.3) yields

$$\operatorname{Im} S(t) = \lim_{\varepsilon \to +0} \operatorname{Im} S(t - i\varepsilon) \leq 0.$$

Hence

$$\operatorname{Im} S(t) = 0. \tag{4.4}$$

For every choice of $\varepsilon \in (0,+\infty)$ we have

$$((t+i\varepsilon)-a)(b-(t+i\varepsilon)) = (t-a)(b-t)+\varepsilon^2+i(b-2t+a)\varepsilon$$

and, because of (4.2), consequently

$$\begin{aligned} 0 &\leq \operatorname{Im}\ [((t+i\varepsilon)-a)(b-(t+i\varepsilon))S(t+i\varepsilon)] \\ &= (t-a)(b-t)\operatorname{Im} S(t+i\varepsilon)+\varepsilon^2 \operatorname{Im} S\ (t+i\varepsilon) \\ &\quad +\frac{\varepsilon}{2}(b-2t-a)\big(S(t+i\varepsilon)+S^*(t+i\varepsilon)\big), \end{aligned} \tag{4.5}$$

Now let $t \in \mathbb{R}\setminus[a,b]$. For each $\varepsilon \in (0,+\infty)$ from (4.1) then we get

$$-(t-a)(b-t)\operatorname{Im} S(t+i\varepsilon) \geq 0.$$

For every positive real number ε, then

$$0 \leq \varepsilon^2 \operatorname{Im} S(t+i\varepsilon)+\frac{\varepsilon}{2}(b-2t-a)\big(S(t+i\varepsilon)+S^*(t+i\varepsilon)\big)$$

and consequently

$$0 \leq \varepsilon \operatorname{Im} S(t+i\varepsilon)+\frac{1}{2}(b-2t+a)\big(S(t+i\varepsilon)+S^*(t+i\varepsilon)\big).$$

Setting $\varepsilon \to +0$ and using (4.4) implies

$$0 \leq \frac{1}{2}(b-2t+a)\big(S(t)+S^*(t)\big) = (b-2t+a)S(t) = \big((b-t)+(a-t)\big)S(t).$$

Thus, if $t \in (-\infty,a)$, then $S(t) \geq 0$ holds. If $t \in (b,+\infty)$, then $-S(t) \geq 0$ is valid. Taking into account that S is holomorphic in $\mathbb{C}\setminus[a,b]$ and that (4.1) holds we see then that S belongs to $\mathcal{R}_q[a,b]$. The proof is complete. □

Lemma 4.4. *Let $S \in \mathcal{R}_q[a,b]$ and let σ be the Stieltjes measure of S. For every complex number z which satisfies $|z| >$ max $(|a|,|b|)$, then*

$$S(z) = -\sum_{j=0}^{\infty} s_j^{[\sigma]} z^{-j-1} \tag{4.6}$$

where $s_j^{[\sigma]}$ is defined in (3.2).

Proof. First we note that for each $x \in \mathbb{C}^q$ the mapping $\sigma_{[x]}$ defined as in Remark 3.1 is a (nonnegative) measure defined on $\mathfrak{B}\cap[a,b]$. Denote $\kappa := \max(|a|,|b|)$. Let $z \in \mathbb{C}$ be such that $|z| > \kappa$. For each $t \in [a,b]$, we have then

$$\left|\left(\frac{t}{z}\right)^k\right| \leq \left(\frac{\kappa}{|z|}\right)^k < 1.$$

Using this, Remark 3.1, (3.1), the equations

$$\frac{1}{t-z} = -\frac{1}{z}\sum_{j=0}^{\infty}\left(\frac{t}{z}\right)^j, \quad \frac{1}{\kappa-|z|} = -\frac{1}{|z|}\sum_{j=0}^{\infty}\left(\frac{\kappa}{|z|}\right)^j,$$

and Lebesgue's dominated convergence theorem, for each $x \in \mathbb{C}^q$, we obtain

$$\begin{aligned}
x^* S(z) x &= x^* \left(\int_{[a,b]} \left(-\frac{1}{z} \sum_{j=0}^{\infty} \left(\frac{t}{z} \right)^j \right) \sigma(dt) \right) x \\
&= \int_{[a,b]} \left(-\frac{1}{z} \sum_{j=0}^{\infty} \left(\frac{t}{z} \right)^j \sigma_{[x]}(dt) \right) = -\frac{1}{z} \sum_{j=0}^{\infty} \frac{1}{z^j} \int_{[a,b]} t^j \sigma_{[x]}(dt) \\
&= x^* \left(-\frac{1}{z} \sum_{j=0}^{\infty} \frac{1}{z^j} s_j^{[\sigma]} \right) x
\end{aligned}$$

and consequently (4.6). □

For each $m \in \mathbb{N}_0$ let $\varepsilon_{m,q} : \mathbb{C} \to \mathbb{C}^{(m+1)q \times q}$ be defined by

$$\varepsilon_{m,q}(z) := \begin{pmatrix} z^m I_q \\ z^{m-1} I_q \\ \vdots \\ z I_q \\ I_q \end{pmatrix}.$$

Lemma 4.5. *Let $n \in \mathbb{N}$ and let $(s_j)_{j=0}^{2n}$ be a sequence of complex $q \times q$ matrices such that $\mathcal{P}_q \left[[a,b]; (s_j)_{j=0}^{2n} \right] \neq \emptyset$. For each $j \in \mathbb{N}_{0,2n-2}$ let*

$$\check{s}_j := -ab s_j + (a+b) s_{j+1} - s_{j+2}.$$

Let $S \in \mathcal{P}_q \left[[a,b]; (s_j)_{j=0}^{2n} \right]$. Furthermore, let $\check{S}_{1,n} : \mathbb{C} \backslash \mathbb{R} \to \mathbb{C}^{q \times q}$ and $\check{S}_{2,n} : \mathbb{C} \backslash \mathbb{R} \to \mathbb{C}^{q \times q}$ be defined by

$$\check{S}_{1,n}(z) := z^{2n} \left(S(z) + \sum_{j=0}^{2n-1} s_j z^{-j-1} \right) \tag{4.7}$$

and

$$\check{S}_{2,n}(z) := \begin{cases} (z-a)(b-z)S(z) - s_0 z + (a+b)s_0 - s_1 & \text{if } n = 1 \\ z^{2n-2} \left((z-a)(b-z)S(z) - s_0 z + (a+b)s_0 - s_1 + \sum\limits_{j=0}^{2n-3} \check{s}_j z^{-j-1} \right) & \text{if } n \geq 2 \end{cases}. \tag{4.8}$$

For each $z \in \mathbb{C} \setminus \mathbb{R}$ then

$$\begin{pmatrix} s_{2n} & \check{S}_{1,n}(z) \\ \check{S}_{1,n}^*(z) & \frac{\check{S}_{1,n}(z) - \check{S}_{1,n}^*(z)}{z - \overline{z}} \end{pmatrix} \geq 0 \tag{4.9}$$

and

$$\begin{pmatrix} \check{s}_{2n-2} & \check{S}_{2,n}(z) \\ \check{S}^*_{2,n}(z) & \frac{\check{S}_{2,n}(z)-\check{S}^*_{2,n}(z)}{z-\overline{z}} \end{pmatrix} \geq 0. \tag{4.10}$$

Proof. For each $k \in \mathbb{N}$ and each $z \in \mathbb{C}$, let

$$\mathbf{M}_k(z) := \begin{pmatrix} 0_{q\times q} & 0_{q\times q} & \dots & 0_{q\times q} & I_q & 0_{q\times q} \\ \overline{z}^{k-1} I_q & \overline{z}^{k-2} I_q & \dots & \overline{z}^0 I_q & 0_{q\times q} & \overline{z}^k I_q \end{pmatrix}.$$

Since S belongs to $\mathcal{P}_q\left[[a,b];(s_j)_{j=0}^{2n}\right]$ for each $z \in \mathbb{C}\setminus\mathbb{R}$ the matrices $\mathbf{K}^{[S]}_{1,n}(z)$ and $\mathbf{K}^{[S]}_{2,n}(z)$ are nonnegative Hermitian. This implies

$$\mathbf{M}_n(z)\mathbf{K}^{[S]}_{1,n}(z)\mathbf{M}^*_n(z) \geq 0 \quad \text{and} \quad \mathbf{M}_n(z)\mathbf{K}^{[S]}_{2,n}(z)\mathbf{M}^*_n(z) \geq 0 \tag{4.11}$$

for each $z \in \mathbb{C}\setminus\mathbb{R}$. Thus it is sufficient to check that

$$\mathbf{M}_n(z)\mathbf{K}^{[S]}_{1,n}(z)\mathbf{M}^*_n(z) = \begin{pmatrix} s_{2n} & \check{S}_{1,n}(z) \\ \check{S}^*_{1,n}(z) & \frac{\check{S}_{1,n}(z)-\check{S}^*_{1,n}(z)}{z-\overline{z}} \end{pmatrix} \tag{4.12}$$

and

$$\mathbf{M}_n(z)\mathbf{K}^{[S]}_{2,n}(z)\mathbf{M}^*_n(z) = \begin{pmatrix} \check{s}_{2n-2} & \check{S}_{2,n}(z) \\ \check{S}^*_{2,n}(z) & \frac{\check{S}_{2,n}(z)-\check{S}^*_{2,n}(z)}{z-\overline{z}} \end{pmatrix} \tag{4.13}$$

hold for each $z \in \mathbb{C}\setminus\mathbb{R}$. First we are going to prove that (4.12) is valid. For each $z \in \mathbb{C}\setminus\mathbb{R}$, let

$$\mathbf{M}_n(z)\mathbf{K}^{[S]}_{1,n}(z)\mathbf{M}^*_n(z) = \begin{pmatrix} A_1(z) & B_1(z) \\ C_1(z) & D_1(z) \end{pmatrix} \tag{4.14}$$

be the $q \times q$ block representation of the matrix $\mathbf{M}_n(z)\mathbf{K}^{[S]}_{1,n}(z)\mathbf{M}^*_n(z)$. Obviously, $C_1(z) = B^*_1(z)$ for each $z \in \mathbb{C}\setminus\mathbb{R}$. Moreover, for each $z \in \mathbb{C}\setminus\mathbb{R}$ we get $A_1(z) = s_{2n}$,

$$B_1(z) = \left(z_{[n,2n]}, \left(z^n I_q, z^{n-1} I_q, \dots, z^0 I_q\right)\left(\begin{pmatrix} I_q \\ 0_{nq\times q} \end{pmatrix} S_1(z) - \begin{pmatrix} 0_{q\times q} \\ -y_{[0,n-1]} \end{pmatrix}\right)\right) \begin{pmatrix} z^{n-1} I_q \\ z^{n-2} I_q \\ \vdots \\ z^0 I_q \\ 0_{q\times q} \\ z^n I_q \end{pmatrix}$$

$$= \left(s_n, s_{n+1}, \dots, s_{2n}, z^n S_1(z) + \sum_{j=0}^{n-1} s_j z^{n-1-j}\right) \begin{pmatrix} z^{n-1} I_q \\ z^{n-2} I_q \\ \vdots \\ z^0 I_q \\ 0_{q\times q} \\ z^n I_q \end{pmatrix}$$

$$= z^{2n}\left(\sum_{j=0}^{n-1} s_j z^{-1-j} + \sum_{j=n}^{2n-1} s_j z^{-1-j} + S(z)\right) = \check{S}_{1,n}(z),$$

and

$$D_1(z) = \left(\overline{z}^{n-1} I_q, \overline{z}^{n-2} I_q, \dots, \overline{z}^0 I_q, 0_{q\times q}, \overline{z}^n I_q\right) \mathbf{K}_{1,n}^{[S]}(z) \begin{pmatrix} z^{n-1} I_q \\ z^{n-2} I_q \\ \vdots \\ z^0 I_q \\ 0_{q\times q} \\ z^n I_q \end{pmatrix}$$
$$= \triangle_1(z) + \triangle_2(z) + \triangle_2^*(z) + \triangle_3(z),$$

where

$$\triangle_1(z) := \left(\varepsilon_{n-1,q}^*(z), 0_{q\times 2q}\right) \mathbf{K}_{1,n}^{[S]}(z) \begin{pmatrix} \varepsilon_{n-1,q}(z) \\ 0_{2q\times q} \end{pmatrix},$$
$$\triangle_2(z) := \left(\varepsilon_{n-1,q}^*(z), 0_{q\times 2q}\right) \mathbf{K}_{1,n}^{[S]}(z) \begin{pmatrix} 0_{(n+1)q\times q} \\ z^n I_q \end{pmatrix},$$

and

$$\triangle_3(z) := \left(0_{q\times(n+1)q}, \overline{z}^n I_q\right) \mathbf{K}_{1,n}^{[S]}(z) \begin{pmatrix} 0_{(n+1)q\times q} \\ z^n I_q \end{pmatrix}.$$

Now we compute the functions $\triangle_1, \triangle_2$, and $\triangle_3$. For each $z \in \mathbb{C} \setminus \mathbb{R}$ we obtain

$$\begin{aligned}\triangle_1(z) &= \varepsilon_{n-1,q}^*(z) \mathbf{H}_{1,n-1} \varepsilon_{n-1,q}(z) = \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} s_{j+k} \overline{z}^{n-1-j} z^{n-1-k} \\ &= \sum_{m=0}^{n-1} s_m \sum_{l=0}^{m} \overline{z}^{n-1-l} z^{n-1-m+l} + \sum_{m=0}^{n-2} s_{n+m} \sum_{l=0}^{n-2-m} \overline{z}^{n-2-m-l} z^l \\ &= \sum_{m=0}^{n-1} s_m z^{n-m-1} \overline{z}^{n-m-1} \sum_{l=0}^{m} z^l \overline{z}^{m-l} + \sum_{m=0}^{n-2} s_{n+m} \sum_{l=0}^{n-2-m} z^l \overline{z}^{n-2-m-l}.\end{aligned}$$

For each $z \in \mathbb{C} \setminus \mathbb{R}$ and each $m \in \mathbb{N}_0$ we have

$$\sum_{l=0}^{m} z^l \overline{z}^{m-l} = \frac{z^{m+1} - \overline{z}^{m+1}}{z - \overline{z}}$$

and consequently

$$\triangle_1(z) = \sum_{m=0}^{n-1} s_m z^{n-m-1} \overline{z}^{n-m-1} \frac{z^{m+1} - \overline{z}^{m+1}}{z - \overline{z}} + \sum_{m=0}^{n-2} s_{n+m} \frac{z^{n-m-1} - \overline{z}^{n-m-1}}{z - \overline{z}}.$$

For every choice of z in $\mathbb{C}\setminus\mathbb{R}$ we also get

$$\begin{aligned}
\triangle_2(z) &= \left(\varepsilon_{n-1,q}^*(z), 0_{q\times 2q}\right)\begin{pmatrix} z^n R_{T_n}(z)(v_n S_1(z) - \mathbf{u}_{1,n}) \\ \\ \frac{z^n\left(S_1(z)-S_1^*(z)\right)}{z-\overline{z}}\end{pmatrix}\\
&= \left(\overline{z}^{n-1}I_q, \overline{z}^{n-2}I_q, \dots, \overline{z}^0 I_q, 0_{q\times q}\right)\left(z^n\begin{pmatrix} z^0 I_q \\ z^1 I_q \\ \vdots \\ z^n I_q\end{pmatrix} S_1(z) - z^n R_{T_n}(z)\begin{pmatrix} 0_{q\times q} \\ -y_{[0,n-1]}\end{pmatrix}\right)\\
&= S_1(z)\cdot z^n \sum_{j=0}^{n-1} z^j \overline{z}^{n-1-j} - z^n\left(\overline{z}^{n-1}I_q, \overline{z}^{n-2}I_q, \dots, \overline{z}^0 I_q,\right)\\
&\quad\cdot\begin{pmatrix} z^0 I_q & 0_{q\times q} & \dots & 0_{q\times q} & 0_{q\times q} \\ zI_q & z^0 I_q & \dots & 0_{q\times q} & 0_{q\times q} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ z^{n-1}I_q & z^{n-2}I_q & \dots & z^0 I_q & 0_{q\times q}\end{pmatrix}\begin{pmatrix} 0_{q\times q} \\ -y_{[0,n-1]}\end{pmatrix}\\
&= S_1(z) z^n \frac{z^n - \overline{z}^n}{z-\overline{z}} + z^n \sum_{k=0}^{n-2} s_k \frac{z^{n-k-1} - \overline{z}^{n-k-1}}{z-\overline{z}}
\end{aligned}$$

and

$$\triangle_3(z) = \frac{z^n\overline{z}^n}{z-\overline{z}}\left(S_1(z) - S_1^*(z)\right).$$

For each $z\in\mathbb{C}\setminus\mathbb{R}$ thus it follows

$$\begin{aligned}
D_1(z) &= \sum_{m=0}^{n-1} s_m z^{n-m-1}\overline{z}^{n-m-1}\frac{z^{m+1}-\overline{z}^{m+1}}{z-\overline{z}} + \sum_{m=0}^{n-2} s_{n+m}\frac{z^{n-m-1}-\overline{z}^{n-m-1}}{z-\overline{z}}\\
&\quad + S_1(z) z^n \frac{z^n-\overline{z}^n}{z-\overline{z}} + z^n\sum_{k=0}^{n-2} s_k \frac{z^{n-k-1}-\overline{z}^{n-k-1}}{z-\overline{z}} + S_1^*(z)\overline{z}^n\frac{z^n-\overline{z}^n}{z-\overline{z}}\\
&\quad + \overline{z}^n \sum_{k=0}^{n-2} s_k\frac{z^{n-k-1}-\overline{z}^{n-k-1}}{z-\overline{z}} + z^n\overline{z}^n\frac{S_1(z)-S_1^*(z)}{z-\overline{z}}\\
&= \frac{1}{z-\overline{z}}\Bigg(z^{2n}S_1(z) - \overline{z}^{2n}S_1^*(z) + \sum_{m=0}^{n-1} s_m z^n \overline{z}^{n-m-1} - \sum_{m=0}^{n-1} s_m z^{n-m-1}\overline{z}^n\\
&\qquad + \sum_{j=n}^{2n-2} s_j z^{2n-j-1} - \sum_{j=n}^{2n-2} s_j \overline{z}^{2n-j-1} + z^n\sum_{k=0}^{n-2} s_k z^{n-k-1}\\
&\qquad - z^n\sum_{k=0}^{n-2} s_k\overline{z}^{n-k-1} + \overline{z}^n\sum_{k=0}^{n-2} s_k z^{n-k-1} - \overline{z}^n\sum_{k=0}^{n-2} s_k\overline{z}^{n-k-1}\Bigg)
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{z-\overline{z}} \Bigg(z^{2n} S_1(z) - \overline{z}^{2n} S_1^*(z) + z^{2n} \sum_{j=n}^{2n-2} s_j z^{-j-1} - \overline{z}^{2n} \sum_{j=n}^{2n-2} s_j \overline{z}^{-j-1} \\
&\qquad + z^{2n} \sum_{k=0}^{n-2} s_k z^{-k-1} - \overline{z}^{2n} \sum_{k=0}^{n-2} s_k \overline{z}^{-k-1} + s_{n-1} z^n \overline{z}^0 - s_{n-1} z^0 \overline{z}^n \Bigg) \\
&= \frac{1}{z-\overline{z}} \left(z^{2n} \left(S_1(z) + \sum_{j=0}^{2n-2} s_j z^{-j-1} \right) - \overline{z}^{2n} \left(S_1^*(z) + \sum_{j=0}^{2n-2} s_j \overline{z}^{-j-1} \right) \right) \\
&= \frac{1}{z-\overline{z}} \Bigg(z^{2n} \left(S_1(z) + \sum_{j=0}^{2n-2} s_j z^{-j-1} + s_{2n-1} z^{-(2n-1)-1} \right) \\
&\qquad - \overline{z}^{2n} \left(S_1^*(z) + \sum_{j=0}^{2n-2} s_j \overline{z}^{-j-1} + s_{2n-1} \overline{z}^{-(2n-1)-1} \right) \Bigg) \\
&= \frac{1}{z-\overline{z}} \left(\check{S}_{1,n}(z) - \check{S}_{1,n}^*(z) \right) \quad .
\end{aligned}
$$

Thus (4.12) is verified. Because of the first inequality in (4.11) then (4.9) is proved. Now we are going to check that (4.13) holds for each $z \in \mathbb{C} \setminus \mathbb{R}$. For this reason, we observe that

$$
\begin{aligned}
\mathbf{H}_{2,n} &= -ab\tilde{H}_{0,n-1} + (a+b)\tilde{H}_{1,n-1} - \tilde{H}_{2,n-1} \\
&= (-abs_{j+k} + (a+b)s_{j+k+1} - s_{j+k+2})_{j,k=0}^{n-1} = (\check{s}_{j+k})_{j,k=0}^{n-1}
\end{aligned} \tag{4.15}
$$

holds. For all $z \in \mathbb{C}$ let $\mathbf{M}_0(z) := I_{2q}$. Further, for each $z \in \mathbb{C} \setminus \mathbb{R}$, let

$$
\mathbf{M}_{n-1}(z)\mathbf{K}_{2,n}^{[S]}(z)\mathbf{M}_{n-1}^*(z) = \begin{pmatrix} A_2(z) & B_2(z) \\ C_2(z) & D_2(z) \end{pmatrix} \tag{4.16}
$$

be the $q \times q$ block representation of $\mathbf{M}_{n-1}(z)\mathbf{K}_{2,n}^{[S]}(z)\mathbf{M}_{n-1}^*(z)$. Obviously,

$$
C_2(z) = B_2^*(z) \tag{4.17}
$$

is valid for each $z \in \mathbb{C} \setminus \mathbb{R}$. Because of (4.16) we easily see that

$$
A_2(z) = \check{s}_{2n-2} \tag{4.18}
$$

holds for each $z \in \mathbb{C} \setminus \mathbb{R}$. Now we suppose that $n \geq 2$. For each $z \in \mathbb{C} \setminus \mathbb{R}$ we have then

$$
\begin{aligned}
&B_2(z) \\
&= \left(\check{s}_{n-1}, \check{s}_n, \ldots, \check{s}_{2n-2}, \left(z^{n-1} I_q, z^{n-2} I_2, \ldots, z^0 I_q \right) \left[\begin{pmatrix} I_q \\ 0_{(n-1)q \times q} \end{pmatrix} S_2(z) - \mathbf{u}_{2,n} \right] \right)
\end{aligned}
$$

$$
\begin{aligned}
&\cdot \left(\begin{pmatrix} z^{n-2}I_q \\ z^{n-3}I_q \\ \vdots \\ z^0 I_q \\ 0_{q\times q} \end{pmatrix} \atop z^{n-1}I_q \right) \\
&= \sum_{j=n-1}^{2n-3} \check{s}_j z^{2n-3-j} + z^{2n-2}S_2(z) - z^{n-1}\left(z^{n-1}I_q, z^{n-2}I_q, z^{n-3}I_q, \dots, z^0 I_q\right)\mathbf{u}_{2,n} \\
&= z^{2n-2}\left(S_2(z) + \sum_{j=n-1}^{2n-3} \check{s}_j z^{-j-1}\right) + z^{n-1}(a+b)\sum_{k=0}^{n-1} s_k z^{n-1-k} \\
&\quad - z^{n-1}ab\sum_{k=0}^{n-2} s_k z^{n-2-k} - z^{n-1}\sum_{k=1}^{n} s_k z^{n-k} \\
&= z^{2n-2}\Bigg((z-a)(b-z)S(z) - s_0 z + \sum_{j=n-1}^{2n-3} \check{s}_j z^{-j-1} + \sum_{j=0}^{n-2}(-abs_j)z^{-j-1} \\
&\quad + \sum_{j=0}^{n-2}(a+b)s_{j+1}z^{-j-1} + (a+b)s_0 - \sum_{j=0}^{n-2} s_{j+2}z^{-j-1} - s_1\Bigg) \\
&= z^{2n-2}\left((z-a)(b-z)S(z) - s_0 z + (a+b)s_0 - s_1 + \sum_{j=0}^{2n-3} \check{s}_j z^{-j-1}\right) \\
&= \check{S}_{2,n}(z).
\end{aligned}
\tag{4.19}
$$

For each $z \in \mathbb{C} \setminus \mathbb{R}$ we also obtain

$$
\begin{aligned}
D_2(z) &= \left(\overline{z}^{n-2}I_q, \overline{z}^{n-3}I_q, \dots, \overline{z}^0 I_q, 0_{q\times q}, \overline{z}^{n-1}I_q\right)\mathbf{K}_{2,n}^{[S]}(z) \begin{pmatrix} z^{n-2}I_q \\ z^{n-3}I_q \\ \vdots \\ z^0 I_q \\ 0_{q\times q} \\ z^{n-1}I_q \end{pmatrix} \\
&= \Gamma_1(z) + \Gamma_2(z) + \Gamma_2^*(z) + \Gamma_3(z),
\end{aligned}
$$

where

$$
\begin{aligned}
\Gamma_1(z) &:= (\varepsilon_{n-2,q}^*(z), 0_{q\times 2q})\mathbf{K}_{2,n}^{[S]}(z) \begin{pmatrix} \varepsilon_{n-2,q}(z) \\ 0_{2q\times q} \end{pmatrix}, \\
\Gamma_2(z) &:= (\varepsilon_{n-2,q}^*(z), 0_{q\times 2q})\mathbf{K}_{2,n}^{[S]}(z) \begin{pmatrix} 0_{nq\times q} \\ z^{n-1}I_q \end{pmatrix},
\end{aligned}
$$

and

$$\Gamma_3(z) := (0_{q\times nq}, \overline{z}^{n-1} I_q)\mathbf{K}_{2,n}^{[S]}(z) \begin{pmatrix} 0_{nq\times q} \\ z^{n-1} I_q \end{pmatrix}.$$

In the next step we compute the functions Γ_1, Γ_2, and Γ_3. For each $z \in \mathbb{C} \setminus \mathbb{R}$ we have

$$\begin{aligned}
\Gamma_1(z) &= \varepsilon_{n-2,q}^*(z)\mathbf{H}_{2,n-1}\varepsilon_{n-2,q}(z) = \sum_{j=0}^{n-2}\sum_{k=0}^{n-2} \check{s}_{j+k}\overline{z}^{n-2-j} z^{n-2-k} \\
&= \sum_{m=0}^{n-2} \check{s}_m \sum_{l=0}^{m} \overline{z}^{n-2-l} z^{n-2-m+l} + \sum_{m=0}^{n-3} \check{s}_{n-1+m} \sum_{l=0}^{n-3-m} \overline{z}^{n-3-m-l} z^l \\
&= \sum_{m=0}^{n-2} \check{s}_m z^{n-m-2}\overline{z}^{n-m-2} \sum_{l=0}^{m} z^l \overline{z}^{m-l} + \sum_{m=0}^{n-3} \check{s}_{n-1+m} \frac{z^{n-m-2} - \overline{z}^{n-m-2}}{z - \overline{z}} \\
&= \sum_{m=0}^{n-2} \check{s}_m z^{n-m-2}\overline{z}^{n-m-2} \frac{z^{m+1} - \overline{z}^{m+1}}{z - \overline{z}} + \sum_{l=n-1}^{2n-4} \check{s}_l \frac{z^{2n-3-l} - \overline{z}^{2n-3-l}}{z - \overline{z}} \\
&= \frac{1}{z - \overline{z}} \left(z^{n-1} \sum_{m=0}^{n-2} \check{s}_m \overline{z}^{n-m-2} - \overline{z}^{n-1} \sum_{m=0}^{n-2} \check{s}_m z^{n-m-2} + z^{2n-2} \sum_{l=n-1}^{2n-4} \check{s}_l z^{-l-1} \right. \\
&\qquad \left. - \overline{z}^{2n-2} \sum_{l=n-1}^{2n-4} \overline{z}^{-l-1} \right),
\end{aligned}$$

$$\begin{aligned}
\Gamma_2(z) &= (\varepsilon_{n-2,q}^*(z), 0_{q\times 2q}) \begin{pmatrix} z^{n-1} R_{T_{n-1}}(z)(v_{n-1}S_2(z) - \mathbf{u}_{2,n}) \\ z^{n-1}\frac{S_2(z) - S_2^*(z)}{z - \overline{z}} \end{pmatrix} \\
&= \left(\overline{z}^{\,n-2} I_q, \overline{z}^{n-3} I_q, \ldots, \overline{z}^0 I_q, 0_{q\times q}\right) \left(z^{n-1} \begin{pmatrix} I_q \\ z I_q \\ \vdots \\ z^{n-2} I_q \\ z^{n-1} I_q \end{pmatrix} S_2(z) - z^{n-1} R_{T_{n-1}}(z)\mathbf{u}_{2,n} \right) \\
&= z^{n-1} S_2(z) \sum_{j=0}^{n-2} \overline{z}^{n-2-j} z^j \\
&\quad - z^{n-1} \left(\sum_{k=0}^{n-2} \overline{z}^{n-2-k} z^k, \sum_{k=0}^{n-3} \overline{z}^{n-3-k} z^k, \ldots, \sum_{k=0}^{0} \overline{z}^{0-k} z^k, 0_{q\times q} \right) \mathbf{u}_{2,n} \\
&= S_2(z) z^{n-1} \frac{z^{n-1} - \overline{z}^{n-1}}{z - \overline{z}} \\
&\quad - z^{n-1} \left(\sum_{k=0}^{n-2} \overline{z}^{n-2-k} z^k, \sum_{k=0}^{n-3} \overline{z}^{n-3-k} z^k, \ldots, \sum_{k=0}^{0} \overline{z}^{0-k} z^k \right)
\end{aligned}$$

$$\cdot \left(-(a+b)y_{[0,n-2]} - ab \begin{pmatrix} 0_{q\times q} \\ -y_{[0,n-3]} \end{pmatrix} + y_{[1,n-1]} \right)$$

$$= S_2(z) z^{n-1} \frac{z^{n-1} - \overline{z}^{n-1}}{z - \overline{z}}$$

$$- z^{n-1} \left(\frac{z^{n-1} - \overline{z}^{n-1}}{z - \overline{z}}, \frac{z^{n-2} - \overline{z}^{n-2}}{z - \overline{z}}, \ldots, \frac{z^{1} - \overline{z}^{1}}{z - \overline{z}} \right)$$

$$\cdot \left(-(a+b)y_{[0,n-2]} - ab \begin{pmatrix} 0_{q\times q} \\ -y_{[0,n-3]} \end{pmatrix} + y_{[1,n-1]} \right)$$

$$= S_2(z) z^{n-1} \frac{z^{n-1} - \overline{z}^{n-1}}{z - \overline{z}} - \frac{z^{n-1}}{z - \overline{z}} \left(- \sum_{j=0}^{n-2} (a+b) s_j (z^{n-1-j} - \overline{z}^{n-1-j}) \right.$$

$$\left. + \sum_{j=0}^{n-3} ab s_j (z^{n-2-j} - \overline{z}^{n-2-j}) + \sum_{j=1}^{n-1} s_j (z^{n-j} - \overline{z}^{n-j}) \right)$$

$$= \frac{z^{n-1}}{z - \overline{z}} \left(S_2(z)(z^{n-1} - \overline{z}^{n-1}) - \sum_{j=0}^{n-3} ab s_j (z^{n-2-j} - \overline{z}^{n-2-j}) \right.$$

$$+ \sum_{j=0}^{n-3} (a+b) s_{j+1} (z^{n-2-j} - \overline{z}^{n-2-j}) + (a+b) s_0 (z^{n-1} - \overline{z}^{n-1})$$

$$\left. - \sum_{j=0}^{n-3} s_{j+2} (z^{n-2-j} - \overline{z}^{n-2-j}) - s_1 (z^{n-1} - \overline{z}^{n-1}) \right)$$

$$= \frac{z^{n-1}}{z - \overline{z}} \Big(S_2(z)(z^{n-1} - \overline{z}^{n-1}) + \sum_{j=0}^{n-3} \check{s}_j (z^{n-2-j} - \overline{z}^{n-2-j})$$

$$+ (a+b) s_0 (z^{n-1} - \overline{z}^{n-1}) - s_1 (z^{n-1} - \overline{z}^{n-1}) \Big),$$

and

$$\Gamma_3(z) = \frac{z^{n-1} \overline{z}^{n-1}}{z - \overline{z}} \big(S_2(z) - S_2^*(z) \big).$$

Consequently, for each $z \in \mathbb{C} \setminus \mathbb{R}$ it follows

$$D_2(z) = \frac{1}{z - \overline{z}} \Big(z^{n-1} \sum_{m=0}^{n-2} \check{s}_m \overline{z}^{n-m-2} - \overline{z}^{n-1} \sum_{m=0}^{n-2} \check{s}_m z^{n-m-2} + z^{2n-2} \sum_{l=n-1}^{2n-4} \check{s}_l z^{-l-1}$$

$$- \overline{z}^{2n-2} \sum_{l=n-1}^{2n-4} \check{s}_l \overline{z}^{-l-1} + z^{n-1} S_2(z)(z^{n-1} - \overline{z}^{n-1})$$

$$+ z^{n-1} \sum_{j=0}^{n-3} \check{s}_j (z^{n-2-j} - \overline{z}^{n-2-j}) + z^{n-1}(a+b) s_0 (z^{n-1} - \overline{z}^{n-1})$$

$$- z^{n-1} s_1 (z^{n-1} - \overline{z}^{n-1}) + \overline{z}^{n-1} S_2^*(z)(z^{n-1} - \overline{z}^{n-1})$$

$$
\begin{aligned}
&+\overline{z}^{n-1}\sum_{j=0}^{n-3}\check{s}_j(z^{n-2-j}-\overline{z}^{n-2-j})+\overline{z}^{n-1}(a+b)s_0(z^{n-1}-\overline{z}^{n-1})\\
&-\overline{z}^{n-1}s_1(z^{n-1}-\overline{z}^{n-1})+z^{n-1}\overline{z}^{n-1}\big(S_2(z)-S_2^*(z)\big)\Big)\\
&=\frac{1}{z-\overline{z}}\Big(z^{n-1}\sum_{m=0}^{n-2}\check{s}_m\overline{z}^{n-m-2}-\overline{z}^{n-1}\sum_{m=0}^{n-2}\check{s}_m z^{n-m-2}\\
&+z^{2n-2}\sum_{l=n-1}^{2n-4}\check{s}_l z^{-l-1}-\overline{z}^{2n-2}\sum_{l=n-1}^{2n-4}\check{s}_l\overline{z}^{-l-1}+z^{2n-2}S_2(z)-z^{n-1}\overline{z}^{n-1}S_2(z)\\
&+z^{n-1}\sum_{j=0}^{n-3}\check{s}_j z^{n-2-j}-z^{n-1}\sum_{j=0}^{n-3}\check{s}_j\overline{z}^{n-2-j}+z^{2n-2}(a+b)s_0-z^{n-1}\overline{z}^{n-1}(a+b)s_0\\
&-z^{2n-2}s_1+z^{n-1}\overline{z}^{n-1}s_1+z^{n-1}\overline{z}^{n-1}S_2^*(z)-\overline{z}^{2n-2}S_2^*(z)\\
&+\overline{z}^{n-1}\sum_{j=0}^{n-3}\check{s}_j z^{n-2-j}-\overline{z}^{n-1}\sum_{j=0}^{n-3}\check{s}_j\overline{z}^{n-2-j}+z^{n-1}\overline{z}^{n-1}(a+b)s_0\\
&-\overline{z}^{2n-2}(a+b)s_0-z^{n-1}\overline{z}^{n-1}s_1+\overline{z}^{2n-2}s_1+z^{n-1}\overline{z}^{n-1}S_2(z)-z^{n-1}\overline{z}^{n-1}S_2^*(z)\Big)\\
&=\frac{1}{z-\overline{z}}\Big(z^{n-1}\overline{z}^{n-1}\sum_{m=0}^{n-2}\check{s}_m\overline{z}^{-m-1}-z^{n-1}\overline{z}^{n-1}\sum_{m=0}^{n-2}\check{s}_m z^{-m-1}\\
&+z^{2n-2}\sum_{l=n-1}^{2n-4}\check{s}_l z^{-l-1}-\overline{z}^{2n-2}\sum_{l=n-1}^{2n-4}\check{s}_l\overline{z}^{-l-1}+z^{2n-2}S_2(z)\\
&+z^{2n-2}\sum_{m=0}^{n-3}\check{s}_m z^{-m-1}-z^{n-1}\overline{z}^{n-1}\sum_{m=0}^{n-3}\check{s}_m\overline{z}^{-m-1}+z^{2n-2}(a+b)s_0\\
&-z^{2n-2}s_1-\overline{z}^{2n-2}S_2^*(z)+z^{n-1}\overline{z}^{n-1}\sum_{m=0}^{n-3}\check{s}_m z^{-m-1}-\overline{z}^{2n-2}\sum_{m=0}^{n-3}\check{s}_m\overline{z}^{-m-1}\\
&-\overline{z}^{2n-2}(a+b)s_0+\overline{z}^{2n-2}s_1\Big)\\
&=\frac{1}{z-\overline{z}}\Big(z^{2n-2}\Big(S_2(z)+(a+b)s_0-s_1+\sum_{j=0}^{n-3}\check{s}_j z^{-j-1}+\sum_{l=n-1}^{2n-4}\check{s}_l z^{-l-1}\Big)\\
&-\overline{z}^{2n-2}\Big(S_2^*(z)+(a+b)s_0-s_1+\sum_{j=0}^{n-3}\check{s}_j\overline{z}^{-j-1}+\sum_{l=n-1}^{2n-4}\check{s}_l\overline{z}^{-l-1}\Big)\\
&+z^{n-1}\overline{z}^{n-1}\check{s}_{n-2}\overline{z}^{-n+1}-z^{n-1}\overline{z}^{n-1}\check{s}_{n-2}z^{-n+1}\Big).
\end{aligned}
$$

Taking into account that

$$
\begin{aligned}
&z^{n-1}\overline{z}^{n-1}\check{s}_{n-2}\overline{z}^{-n+1}-z^{n-1}\overline{z}^{n-1}\check{s}_{n-2}z^{-n+1}=z^{n-1}\check{s}_{n-2}-\overline{z}^{n-1}\check{s}_{n-2}\\
&=z^{2n-2}\check{s}_{n-2}z^{-(n-2)-1}-\overline{z}^{2n-2}\check{s}_{n-2}\overline{z}^{-(n-2)-1}\\
&=z^{2n-2}\check{s}_{n-2}z^{-(n-2)-1}-\overline{z}^{2n-2}\check{s}_{n-2}\overline{z}^{-(n-2)-1}+z^{2n-2}\check{s}_{2n-3}z^{-(2n-3)-1}\\
&\quad-\overline{z}^{2n-2}\check{s}_{2n-3}\overline{z}^{-(2n-3)-1}
\end{aligned}
$$

holds for each $z \in \mathbb{C} \setminus \mathbb{R}$, we can conclude that

$$\begin{aligned} D_2(z) &= \frac{1}{z-\overline{z}}\Big(z^{2n-2}\big(S_2(z)+(a+b)s_0-s_1+\sum_{j=0}^{2n-3}\check{s}_j z^{-j-1}\big) \\ &\qquad -\overline{z}^{2n-2}\big(S_2^*(z)+(a+b)s_0-s_1+\sum_{j=0}^{2n-3}\check{s}_j \overline{z}^{-j-1}\big)\Big) \\ &= \frac{1}{z-\overline{z}}\big(\check{S}_{2,n}(z)-\check{S}_{2,n}^*(z)\big) \end{aligned} \tag{4.20}$$

holds for each $z \in \mathbb{C} \setminus \mathbb{R}$. By virtue of (4.16), (4.17), (4.18), (4.19), and (4.20), equation (4.13) is valid. Because of the second inequality in (4.11) we can conclude that (4.10) is proved for each $z \in \mathbb{C} \setminus \mathbb{R}$ in the case $n \geq 2$. If $n = 1$, then for each $z \in \mathbb{C} \setminus \mathbb{R}$ we obtain

$$\begin{aligned} &\begin{pmatrix} \check{s}_{2n-2} & \check{S}_{2,n}(z) \\ \check{S}_{2,n}^*(z) & \frac{\check{S}_{2,n}(z)-\check{S}_{2,n}(z)}{z-\overline{z}} \end{pmatrix} = \begin{pmatrix} \check{s}_0 & \check{S}_{2,1}(z) \\ \check{S}_{2,1}^*(z) & \frac{\check{S}_{2,1}(z)-\check{S}_{2,1}(z)}{z-\overline{z}} \end{pmatrix} \\ &= \begin{pmatrix} -abs_0+(a+b)s_1-s_2 & \check{S}_{2,1}(z) \\ \check{S}_{2,1}^*(z) & \frac{\check{S}_{2,1}(z)-\check{S}_{2,1}(z)}{z-\overline{z}} \end{pmatrix} = \begin{pmatrix} \mathbf{H}_{2,n} & \check{S}_{2,1}(z) \\ \check{S}_{2,1}^*(z) & \frac{\check{S}_{2,1}(z)-\check{S}_{2,1}(z)}{z-\overline{z}} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{H}_{2,1} & S_2(z)+(a+b)s_0-s_1 \\ S_2^*(z)+(a+b)s_0-s_1 & \frac{S_2(z)-S_2^*(z)}{z-\overline{z}} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{H}_{2,1} & R_{T_0}(z)\left(\mathbf{v}_{2,1}S_2(z)-\mathbf{u}_{2,1}\right) \\ \left(R_{T_0}(z)\left(\mathbf{v}_{2,1}S_2(z)-\mathbf{u}_{2,1}\right)\right)^* & \frac{S_2(z)-S_2^*(z)}{z-\overline{z}} \end{pmatrix} = \mathbf{K}_{2,1}^{[S]}(z) \\ &= \mathbf{K}_{2,n}^{[S]}(z) \geq 0. \end{aligned}$$

Thus (4.10) also holds in the case $n = 1$. □

Now we prove the main result of this section.

Proposition 4.6. *Let $n \in \mathbb{N}$ and let $(s_j)_{j=0}^{2n}$ be a sequence of complex $q \times q$ matrices. Then the set $\mathcal{P}_q\left[[a,b];(s_j)_{j=0}^{2n}\right]$ of all solutions of the system of the fundamental matrix inequalities of Potapov-type associated with $[a,b]$ and $(s_j)_{j=0}^{2n}$ is a subset of $\mathcal{R}_q\left[[a,b];(s_j)_{j=0}^{2n}\right]$.*

Proof. Let $S \in \mathcal{P}_q\left[[a,b];(s_j)_{j=0}^{2n}\right]$. Lemma 4.3 yields then that S belongs to $\mathcal{R}_q[a,b]$. Denote by σ the Stieltjes measure of S. From Lemma 4.4 we get then that for every complex number z with $|z| > \max(|a|,|b|)$ the matrix-valued function S can be represented via (4.6). In particular, (4.6) holds for $z = \kappa + 1$ where $\kappa := \max(|a|,|b|)$. Thus, we see that there is a positive real number K such that $\|s_j^{[\sigma]}\| \leq K(\kappa+1)^{j+1}$ is satisfied for every nonnegative integer j. For each real number y which fulfills $y > \kappa + 1$ and each $m \in \mathbb{N}_0$, we have

$$\left\| \sum_{j=m}^{\infty} s_j^{[\sigma]} (iy)^{m-1-j} \right\| \leq K \sum_{j=m}^{\infty} (\kappa+1)^{j+1} y^{m-1-j}$$
$$= \frac{K(\kappa+1)^{m+1}}{y} \sum_{k=0}^{\infty} \left(\frac{\kappa+1}{y} \right)^k = \frac{K(\kappa+1)^{m+1}}{y-\kappa-1}$$

and consequently

$$\lim_{y\to+\infty} \sum_{j=m}^{\infty} s_j^{[\sigma]} (iy)^{m-1-j} = 0. \tag{4.21}$$

From Lemma 4.5 we can conclude that the matrix-valued functions $\check{S}_{1,n}$ and $\check{S}_{2,n}$ defined on $\mathbb{C} \setminus \mathbb{R}$ and given by (4.7) and (4.8) fulfill the matrix inequalities (4.9) and (4.10) for each $z \in \mathbb{C} \setminus \mathbb{R}$. Thus, applying [CDFK, Lemma 8.9] and [CDFK, part (a) of Theorem 8.7] we obtain that there exist the limits

$$\Sigma_{1,n} := \lim_{y\to+\infty} (-iy\check{S}_{1,n}(iy)) \quad \text{and} \quad \Sigma_{2,n} := \lim_{y\to+\infty} (-iy\check{S}_{2,n}(iy)) \tag{4.22}$$

and that the inequalities

$$\Sigma_{1,n} \leq s_{2n} \quad \text{and} \quad \Sigma_{2,n} \leq -ab\, s_{2n-2} + (a+b)s_{2n-1} - s_{2n} \tag{4.23}$$

hold. Using (4.22), (4.7), and (4.6) we get

$$\Sigma_{1,n} = \lim_{y\to+\infty} \left(\sum_{j=0}^{2n-1} \left(s_j^{[\sigma]} - s_j \right) (iy)^{2n-j} + s_{2n}^{[\sigma]} + \sum_{j=2n+1}^{\infty} s_j^{[\sigma]} (iy)^{2n-j} \right)$$

and, because of (4.21) (with $m = 2n+1$) and the first inequality in (4.23) then

$$\lim_{y\to+\infty} \sum_{j=0}^{2n-1} \left(s_j^{[\sigma]} - s_j \right) (iy)^{2n-j} + s_{2n}^{[\sigma]} = \Sigma_{1,n} \leq s_{2n}. \tag{4.24}$$

This implies

$$s_j^{[\sigma]} = s_j \quad \text{for each} \quad j \in \mathbb{N}_{0,2n-1} \tag{4.25}$$

and hence

$$s_{2n}^{[\sigma]} \leq s_{2n}. \tag{4.26}$$

By virtue of (4.22), (4.25), (4.8), (4.6), and (4.21) we get

$$\Sigma_{2,n} = \lim_{y\to+\infty} \left(- \sum_{j=2n}^{\infty} s_j^{[\sigma]} (iy)^{2n-j} + (a+b) \sum_{j=2n-1}^{\infty} s_j^{[\sigma]} (iy)^{2n-j-1} \right.$$
$$\left. -ab \sum_{j=2n-2}^{\infty} s_j^{[\sigma]} (iy)^{2n-j-2} \right)$$
$$= -s_{2n}^{[\sigma]} + (a+b)s_{2n-1}^{[\sigma]} - ab s_{2n-2}^{[\sigma]}.$$

Consequently, (4.25) and the second inequality in (4.23) imply $s_{2n} \geq s_{2n}^{[\sigma]}$. Thus from (4.26) it follows $s_{2n}^{[\sigma]} = s_{2n}$. This and (4.25) show that σ belongs to $\mathcal{M}_{\geq}^{q}\left[[a,b], \mathfrak{B} \cap [a,b]; (s_j)_{j=0}^{2n}\right]$, i.e., S belongs to $\mathcal{R}_q\left[[a,b]; (s_j)_{j=0}^{2n}\right]$. The proof is complete. □

Now we obtain immediately a proof of Theorem 1.2.

Proof of Theorem 1.2. Apply Proposition 3.7 and 4.6. □

5. Nonnegative column pairs

The main theme of this section is to introduce the class of pairs of meromorphic matrix-valued functions which will play the role of parameters in the description of the solution set $\mathcal{R}\left[[a,b]; (s_j)_{j=0}^{2n}\right]$ of our reformulated truncated matrix moment problem in the positive definite case.

Let J be a $p \times p$ signature matrix, i.e., J is a complex $p \times p$ matrix which satisfies $J^* = J$ and $J^2 = I$. A complex $p \times p$ matrix A is said to be J-*contractive* (respectively, J-*expansive*) if $J - A^*JA \geq 0$ (respectively, $A^*JA - J \geq 0$). If A is a complex $p \times p$ matrix, then A is J-contractive (respectively, J-expansive) if and only if A^* is J-contractive (respectively, J-expansive) (see, e.g., [DFK, Theorem 1.3.3]). Moreover, if A is a nonsingular complex $p \times p$ matrix, then A is J-contractive if and only if A^{-1} is J-expansive (see, e.g., [DFK, Lemma 1.3.15]). A complex $p \times p$ matrix is said to be J-*unitary* if $J - A^*JA = 0$. If A is a J-unitary complex $p \times p$ matrix, then A is nonsingular and the matrices A^* and A^{-1} are J-unitary as well.

For our further considerations, the $2q \times 2q$ signature matrix

$$\tilde{J}_q := \begin{pmatrix} 0 & -iI_q \\ iI_q & 0 \end{pmatrix} \tag{5.1}$$

is of particular interest. Indeed, on the one hand, we work with the class $\mathcal{R}_q[a,b]$ and, on the other hand, for all complex $q \times q$ matrices C we have

$$\begin{pmatrix} C \\ I_q \end{pmatrix}^* (-\tilde{J}_q) \begin{pmatrix} C \\ I_q \end{pmatrix} = 2 \operatorname{Im} C. \tag{5.2}$$

Remark 5.1. *Let $n \in \mathbb{N}$ and let $(s_j)_{j=0}^{2n}$ be a sequence of complex Hermitian $q \times q$ matrices. From Proposition* 2.1 *one can easily see that for each $r \in \mathbb{N}_{1,2}$ the identity*

$$(\mathbf{u}_{r,n}, \mathbf{v}_{r,n})\tilde{J}_q(\mathbf{u}_{r,n}, \mathbf{v}_{r,n})^* = -i\left(\mathbf{H}_{r,n}\mathbf{T}_{r,n}^* - \mathbf{T}_{r,n}\mathbf{H}_{r,n}\right)$$

holds.

For each Hermitian complex $(p+q) \times (p+q)$ matrix J in [FKK, Definition 51] the notion of a J-nonnegative pair is introduced. We are going to modify this definition for our purpose in this paper. In the following, we continue to suppose that a and b are real numbers which satisfy $a < b$.

Definition 5.2. *Let P and Q be $q \times q$ complex matrix-valued functions which are meromorphic in $\mathbb{C} \setminus [a,b]$. Then $\binom{P}{Q}$ is called a column pair of second type which is nonnegative with respect to $-\tilde{J}_q$ and $[a,b]$ if there exists a discrete subset $\mathcal{D}$ of $\mathbb{C} \setminus [a,b]$ such that the following four conditions are satisfied:*

(i) *The matrix-valued functions P and Q are holomorphic in $\mathbb{C} \setminus ([a,b] \cup \mathcal{D})$.*

(ii) *For all $z \in \mathbb{C} \setminus ([a,b] \cup \mathcal{D})$, rank $\binom{P(z)}{Q(z)} = q$.*

(iii) *For all $z \in \mathbb{C} \setminus (\mathbb{R} \cup \mathcal{D})$,*

$$\frac{1}{2 \operatorname{Im} z} \begin{pmatrix} P(z) \\ Q(z) \end{pmatrix}^* (-\tilde{J}_q) \begin{pmatrix} P(z) \\ Q(z) \end{pmatrix} \geq 0. \tag{5.3}$$

(iv) *For all $z \in \mathbb{C} \setminus (\mathbb{R} \cup \mathcal{D})$,*

$$\frac{1}{2 \operatorname{Im} z} \begin{pmatrix} (z-a)P(z) \\ (b-z)Q(z) \end{pmatrix}^* (-\tilde{J}_q) \begin{pmatrix} (z-a)P(z) \\ (b-z)Q(z) \end{pmatrix} \geq 0. \tag{5.4}$$

In the following, let $\mathfrak{P}(-\tilde{J}_q, [a,b])$ denote the set of all column pairs of second type which are nonnegative with respect to $-\tilde{J}_q$ and $[a,b]$.

Now we turn our attention to a distinguished class of holomorphic $q \times q$ matrix-valued functions which will turn out to be intimately connected to the class $\mathfrak{P}\left(-\tilde{J}_q, [a,b]\right)$. The class $\mathcal{S}_q[a,b]$ consists of all matrix-valued functions S for which the following three conditions hold:

(i) S is holomorphic in $\mathbb{C} \setminus [a,b]$.

(ii) For each $w \in \Pi_+$, the matrix $\operatorname{Im} S(w)$ is nonnegative Hermitian.

(iii) For each $t \in \mathbb{R} \setminus [a,b]$, the matrix $S(t)$ is nonnegative Hermitian.

Theorem 5.3. *Let $S : \mathbb{C} \setminus [a,b] \to \mathbb{C}^{q \times q}$ be a matrix-valued function. Then S belongs to $\mathcal{S}_q[a,b]$ if and only if there is a $\rho \in \mathcal{M}^q_{\geq}([a,b], \mathfrak{B} \cap [a,b])$ such that*

$$S(z) = (b-z) \int_{[a,b]} \frac{1}{t-z} \rho(dt) \tag{5.5}$$

holds for each $z \in \mathbb{C} \setminus [a,b]$. In this case, the nonnegative Hermitian measure ρ is unique.

Theorem 5.3 can be proved by modifying the proof in the case $q = 1$. This scalar case is considered in [KN, Appendix, Ch. 4].

Corollary 5.4. *Let $S : \mathbb{C} \setminus [a,b] \to \mathbb{C}^{q \times q}$ be a matrix-valued function and let $F : \mathbb{C} \setminus [a,b] \to \mathbb{C}^{q \times q}$ be defined by $F(z) := \frac{1}{b-z} S(z)$. Then S belongs to $\mathcal{S}_q[a,b]$ if and only if F belongs to $\mathcal{R}_q[a,b]$.*

Proof. Use Theorems 1.1 and 5.3. $\square$

Lemma 5.5. *Let $S \in \mathcal{S}_q[a,b]$ and let ρ be the unique nonnegative Hermitian $q \times q$ measure on $\mathfrak{B} \cap [a,b]$ such that* (5.5) *holds for each $z \in \mathbb{C} \setminus [a,b]$. Let $\tilde{S}_1 := S$ and let $\tilde{S}_2 : \mathbb{C} \setminus [a,b] \to \mathbb{C}^{q\times q}$ be defined by*

$$\tilde{S}_2(z) := \frac{z-a}{b-z} S(z). \tag{5.6}$$

(a) *For each $z \in \mathbb{C} \setminus \mathbb{R}$,*

$$\frac{\tilde{S}_1(z) - \tilde{S}_1^*(z)}{z - \overline{z}} = \int_{[a,b]} \left(\frac{\sqrt{b-t}}{t-z} I_q \right) \rho(dt) \left(\frac{\sqrt{b-t}}{t-z} I_q \right)^* \tag{5.7}$$

and

$$\frac{\tilde{S}_2(z) - \tilde{S}_2^*(z)}{z - \overline{z}} = \int_{[a,b]} \left(\frac{\sqrt{t-a}}{t-z} I_q \right) \rho(dt) \left(\frac{\sqrt{t-a}}{t-z} I_q \right)^* . \tag{5.8}$$

(b) *The matrix-valued functions $\tilde{S}_1$ and $\tilde{S}_2$ are both holomorphic in $\mathbb{C} \setminus [a,b]$ and, for each $w \in \Pi_+$, the matrices $\operatorname{Im} \tilde{S}_1(w)$ and $\operatorname{Im} \tilde{S}_2(w)$ are both nonnegative Hermitian.*

Proof. Let $z \in \mathbb{C} \setminus \mathbb{R}$. Using the equations

$$\frac{1}{z-\overline{z}}\left(\frac{b-z}{t-z} - \frac{b-\overline{z}}{t-\overline{z}}\right) = \frac{b-t}{(t-z)(\overline{t-z})} \quad \text{and} \quad \frac{1}{z-\overline{z}}\left(\frac{z-a}{t-z} - \frac{\overline{z}-a}{t-\overline{z}}\right) = \frac{t-a}{(t-z)(\overline{t-z})}$$

application of (5.5) yields (5.7) and (5.8). Thus part (a) is proved. Part (b) is an immediate consequence of part (a). □

Remark 5.6. *For each $S \in \mathcal{S}_q[a,b]$, one can show that $S(z) = S^*(\overline{z})$ holds for every choice of z in $\mathbb{C} \setminus \mathbb{R}$. Consequently, if S belongs to $\mathcal{S}_q[a,b]$, then for each $z \in \mathbb{C} \setminus \mathbb{R}$ we have*

$$\frac{1}{2 \operatorname{Im} z} \begin{pmatrix} S(z) \\ I_q \end{pmatrix}^* (-\tilde{J}_q) \begin{pmatrix} S(z) \\ I_q \end{pmatrix} = \frac{1}{\operatorname{Im} z} \operatorname{Im} S(z) \geq 0$$

and

$$\frac{1}{2 \operatorname{Im} z} \begin{pmatrix} (z-a)S(z) \\ (b-z)I_q \end{pmatrix}^* (-\tilde{J}_q) \begin{pmatrix} (z-a)S(z) \\ (b-z)I_q \end{pmatrix} = \frac{1}{\operatorname{Im} z} \operatorname{Im} \left(\frac{z-a}{b-z} S(z) \right) \geq 0.$$

In particular, $\begin{pmatrix} S \\ I_q \end{pmatrix}$ belongs to $\mathfrak{P}(-\tilde{J}_q, [a,b])$ for each $S \in \mathcal{S}_q[a,b]$.

Now we characterize the membership of a given matrix-valued function S to the class $\mathcal{S}_q[a,b]$ in terms of the matrix-valued functions $\tilde{S}_1$ and $\tilde{S}_2$ introduced in Lemma 5.5.

Theorem 5.7. *Let $S : \mathbb{C} \setminus [a,b]$ be a matrix-valued function. Let $\tilde{S}_1 = S$ and let $\tilde{S}_2 : \mathbb{C} \setminus [a,b] \to \mathbb{C}^{q\times q}$ be defined by* (5.6)*. Then S belongs to $\mathcal{S}_q[a,b]$ if and only if the following statements hold:*

(i) *$\tilde{S}_1$ and $\tilde{S}_2$ are both holomorphic in $\mathbb{C} \setminus [a,b]$.*
(ii) $\operatorname{Rstr.}_{\Pi_+} \tilde{S}_1$ *and* $\operatorname{Rstr.}_{\Pi_+} \tilde{S}_2$ *both belong to $\mathcal{R}_q$.*

Proof. If S belongs to $\mathcal{S}_q[a,b]$, then (i) and (ii) follow by application of Lemma 5.5. Now we suppose that (i) and (ii) hold. To verify that S belongs to $\mathcal{S}_q[a,b]$ it is sufficient to prove that for each $t \in \mathbb{R} \setminus [a,b]$ the matrix $S(t)$ is nonnegative Hermitian. Let $t \in \mathbb{R} \setminus [a,b]$. Because of (i) and (ii) we get then

$$\operatorname{Im} S(t) = \lim_{\varepsilon \to 0+0} \operatorname{Im} S(t+i\varepsilon) \geq 0 \tag{5.9}$$

and

$$\frac{t-a}{b-t} \operatorname{Im} S(t) = \lim_{\varepsilon \to 0+0} \operatorname{Im} \tilde{S}_2(t+i\varepsilon) \geq 0.$$

Since either $t < a$ or $t > b$ holds the last relation implies $\operatorname{Im} S(t) \leq 0$. Consequently $\operatorname{Im} S(t) = 0$ is true. For every positive real number ε, we have moreover

$$\begin{aligned} 0 &\leq \operatorname{Im} \tilde{S}_2(t+i\varepsilon) \\ &= \left((b-t)^2+\varepsilon^2\right)^{-1} \left(\left[(t-a)(b-t)-\varepsilon^2\right] \operatorname{Im} S(t+i\varepsilon) + \varepsilon(b-a) \operatorname{Re} S(t+i\varepsilon)\right) \\ &\leq \varepsilon \left((b-t)^2+\varepsilon^2\right)^{-1} (b-a) \operatorname{Re} S(t+i\varepsilon) \end{aligned}$$

and consequently $0 \leq \operatorname{Re} S(t+i\varepsilon)$. Letting $\varepsilon \to 0+0$ we get $S(t) = \operatorname{Re} S(t) \geq 0$. $\square$

Theorem 5.7 should be compared with [CDFK, Lemma 3.6] where an analogous result for the class $\mathcal{R}_q[a,b]$ is proved.

If $\binom{P}{Q}$ belongs to $\mathfrak{P}(-\tilde{J}_q,[a,b])$ and if F is a $q \times q$ complex matrix-valued function which is meromorphic in $\mathbb{C} \setminus [a,b]$ and for which the complex-valued function $\det F$ does not vanish identically, then it is readily checked that $\binom{PF}{QF}$ also belongs to $\mathfrak{P}(-\tilde{J}_q,[a,b])$. Pairs $\binom{P_1}{Q_1}$ and $\binom{P_2}{Q_2}$ which belong to $\mathfrak{P}(-\tilde{J}_q,[a,b])$ are said to be *equivalent* if there exists a $q \times q$ complex matrix-valued function F which is meromorphic in $\mathbb{C} \setminus [a,b]$ such that the following conditions are satisfied:

(i) The function $\det F$ does not vanish identically.
(ii) The identities $P_2 = P_1 F$ and $Q_2 = Q_1 F$ hold.

One can easily see that this relation is really an equivalence relation on $\mathfrak{P}(-\tilde{J}_q,[a,b])$. If $\binom{P}{Q} \in \mathfrak{P}(-\tilde{J}_q,[a,b])$, then we will write $\left\langle \binom{P}{Q} \right\rangle$ for the equivalence class of all column pairs $\binom{R}{S} \in \mathfrak{P}(-\tilde{J}_q,[a,b])$ which are equivalent to $\binom{P}{Q}$.

If f is a meromorphic matrix-valued function, then let $\mathbb{H}_f$ be the set of all points at which f is holomorphic.

Proposition 5.8. *Let P and Q be $q \times q$ matrix-valued functions which are meromorphic in $\mathbb{C} \setminus [a,b]$. Suppose that $\binom{P}{Q}$ belongs to $\mathfrak{P}(-\tilde{J}_q,[a,b])$ and that the function $\det Q$ does not vanish identically in $\mathbb{C} \setminus [a,b]$. Then $S := PQ^{-1}$ belongs to $\mathcal{S}_q[a,b]$.*

Proof. Since $\binom{P}{Q}$ belongs to $\mathfrak{P}(-J_q,[a,b])$ there is a discrete subset $\mathcal{D}$ of $\mathbb{C} \setminus [a,b]$ such that the conditions (i), (ii), (iii), and (iv) of Definition 5.2 are fulfilled. Let $P_1 := P$ and let $Q_1 : \mathbb{C} \setminus ([a,b] \cup \mathcal{D}) \to \mathbb{C}^{q \times q}$ be defined by $Q_1(z) := (b-z)Q(z)$.

Obviously, P_1 and Q_1 are both holomorphic in $\mathbb{C} \setminus ([a,b] \cup \mathcal{D})$, and $\det Q_1$ does not vanish identically. In particular, $\tilde{\mathcal{D}} := \{z \in \mathbb{H}_{\det Q_1} : \det Q_1(z) = 0\}$ and $\mathcal{D}_1 := \mathcal{D} \cup \tilde{\mathcal{D}}$ are discrete subsets of $\mathbb{C} \setminus [a,b]$. Moreover, we get

$$\operatorname{rank} \begin{pmatrix} P_1(z) \\ Q_1(z) \end{pmatrix} = \operatorname{rank} \begin{pmatrix} P(z) \\ Q(z) \end{pmatrix} = q$$

for each $z \in \mathbb{C} \setminus ([a,b] \cup \mathcal{D})$. Condition (iii) of Definition 5.2 yields

$$\begin{aligned} &\frac{1}{2\operatorname{Im} z} \begin{pmatrix} (b-z)P_1(z) \\ Q_1(z) \end{pmatrix}^* (-\tilde{J}_q) \begin{pmatrix} (b-z)P_1(z) \\ Q_1(z) \end{pmatrix} \\ &= \frac{|b-z|^2}{2\operatorname{Im} z} \begin{pmatrix} P(z) \\ Q(z) \end{pmatrix}^* (-\tilde{J}_q) \begin{pmatrix} P(z) \\ Q(z) \end{pmatrix} \geq 0 \end{aligned}$$

for each $z \in \mathbb{C} \setminus ([a,b] \cup \mathcal{D})$. From condition (iv) of Definition 5.2 it follows

$$\begin{aligned} &\frac{1}{2\operatorname{Im} z} \begin{pmatrix} (z-a)P_1(z) \\ Q_1(z) \end{pmatrix}^* (-\tilde{J}_q) \begin{pmatrix} (z-a)P_1(z) \\ Q_1(z) \end{pmatrix} \\ &= \frac{1}{2\operatorname{Im} z} \begin{pmatrix} (z-a)P(z) \\ (b-z)Q(z) \end{pmatrix}^* (-\tilde{J}_q) \begin{pmatrix} (z-a)P(z) \\ (b-z)Q(z) \end{pmatrix} \geq 0 \end{aligned}$$

for each $z \in \mathbb{C} \setminus ([a,b] \cup \mathcal{D})$. Applying [CDFK, Proposition 5.7] we obtain that $F := P_1 Q_1^{-1}$ belongs to the class $\mathcal{R}_q[a,b]$. Thus Corollary 5.4 provides us that $G : \mathbb{C} \setminus [a,b] \to \mathbb{C}^{q \times q}$ given by $G(z) := (b-z)F(z)$ belongs to $\mathcal{S}_q[a,b]$. Because of

$$G(z) = (b-z)P(z)((b-z)Q(z))^{-1} = P(z)Q^{-1}(z)$$

for each $z \in \mathbb{C} \setminus ([a,b] \cup \mathcal{D}_1)$ the proof is complete. □

The combination of Proposition 5.8 and Remark 5.6 shows that there is a one-to-one correspondence between the class $\mathcal{S}_q[a,b]$ and the set of equivalence classes of elements $\binom{P}{Q} \in \mathcal{P}(-\tilde{J}_q, [a,b])$ for which $\det Q$ does not identically vanish. For this reason, the set $\mathcal{P}(-\tilde{J}_q, [a,b])$ can be considered as a projective extension of $\mathcal{S}_q[a,b]$.

6. Description of the solution set in the positive definite case

In this section, we suppose again that a and b are real numbers such that $a < b$. Further, let n be a positive integer. Let $\mathcal{M}_{\geq}^q([a,b], \mathfrak{B} \cap [a,b])$ denote the set of all nonnegative Hermitian $q \times q$ measures defined on $\mathfrak{B} \cap [a,b]$. For all $\sigma \in \mathcal{M}_{\geq}^q([a,b], \mathfrak{B} \cap [a,b])$ and all nonnegative integers j, let $s_j^{[\sigma]}$ be given by (3.2). From Remark 3.4 we know that, for each $\sigma \in \mathcal{M}_{\geq}^q([a,b], \mathfrak{B} \cap [a,b])$ and for every nonnegative integer m, the matrices $\mathbf{H}_{1,m}^{[\sigma]}$ and $\mathbf{H}_{2,m}^{[\sigma]}$ given by (3.4) and (3.5) are both nonnegative Hermitian. Hence, in view of the considerations in Section 1, if $(s_j)_{j=0}^{2n}$ is a sequence of complex $q \times q$ matrices such that the solution set $\mathcal{R}_q\left[[a,b]; (s_j)_{j=0}^{2n}\right]$ of the (reformulated) matricial version of M.G. Krein's moment problem is nonempty, then $\mathbf{H}_{1,n}$ and $\mathbf{H}_{2,n}$ are both nonnegative Hermitian. In this section, we will give a parametrization of the set $\mathcal{R}_q\left[[a,b]; (s_j)_{j=0}^{2n}\right]$ under the assumption that the block Hankel matrices $\mathbf{H}_{1,n}$ and $\mathbf{H}_{2,n}$ are both positive Hermitian. However, first we are going now to present a class of measures $\sigma \in \mathcal{M}_{\geq}^{q}([a,b], \mathfrak{B} \cap [a,b])$ for which the block Hankel matrices $\mathbf{H}_{1,m}^{[\sigma]}$ and $\mathbf{H}_{2,m}^{[\sigma]}$ are positive Hermitian for every nonnegative integer m. Let λ denote the Lebesgue measure defined on $\mathfrak{B} \cap [a,b]$ and let $\mathcal{L}^1([a,b], \mathfrak{B} \cap [a,b], \lambda; \mathbb{C})$ designate the set of all $(\mathfrak{B} \cap [a,b]) - \tilde{\mathfrak{B}}$-measurable complex-valued functions which are defined on $[a,b]$ and which are integrable with respect to λ. Recall that $\mathbb{C}_{\geq}^{q\times q}$ denotes the set of all nonnegative Hermitian complex $q \times q$ matrices. Further, let $\mathbb{C}_{>}^{q\times q}$ be the set of all positive Hermitian complex $q \times q$ matrices.

Lemma 6.1. *Let $X = (X_{jk})_{j,k=1}^{q} : [a,b] \to \mathbb{C}^{q\times q}$ be a $q \times q$ matrix-valued function every entry function X_{jk} of which belongs to $\mathcal{L}^1([a,b], \mathfrak{B} \cap [a,b], \lambda; \mathbb{C})$ and which satisfies $\lambda\left(\{t \in [a,b] : X(t) \in \mathbb{C}^{q\times q} \setminus \mathbb{C}_{>}^{q\times q}\}\right) = 0$. Then $\mu : \mathfrak{B} \cap [a,b] \to \mathbb{C}^{q\times q}$ defined by*

$$\mu(B) := \int_B X d\lambda$$

belongs to $\mathcal{M}_{\geq}^{q}([a,b], \mathfrak{B} \cap [a,b])$ and, for every nonnegative integer m, the block Hankel matrices $\mathbf{H}_{1,m}^{[\mu]}$ and $\mathbf{H}_{2,m}^{[\mu]}$ are both positive Hermitian.

Proof. Let m be a nonnegative integer. From Lemma 3.3 we see that the representations

$$\mathbf{H}_{1,m}^{[\mu]} = \int_{[a,b]} E_m(t)\mu(dt)E_m(t) = \int_{[a,b]} E_m(t)X(t)E_m^*(t)\lambda(dt)$$

and

$$\mathbf{H}_{2,m}^{[\mu]} = \int_{[a,b]} (t-a)(b-t)E_m(t)X(t)E_m^*(t)\lambda(dt)$$

hold where E_m is the matrix polynomial which is for each $z \in \mathbb{C}$ given by (3.7). Let $x \in \mathbb{C}^{(m+1)q\times 1} \setminus \{0\}$. Then one can easily see that the set $M_x := \{t \in [a,b] : E_m^*(t)x = 0\}$ is finite. In particular, $\lambda(M_x \cup \{a,b\}) = 0$. Hence we obtain

$$\lambda\left(\{t \in [a,b] : (t-a)(b-t)x^* E_m(t)X(t)E_m^*(t)x \in (-\infty, 0]\}\right) = 0$$

and consequently

$$x^*\mathbf{H}_{2,m}^{[\mu]}x = \int_{[a,b]} (t-a)(b-t)\left(E_m^*(t)x\right)^* X(t)E_m^*(t)x\,\lambda(dt) \in (0, +\infty).$$

Analogously, one can see that $x^*\mathbf{H}_{1,m}^{[\mu]}x \in (0, +\infty)$ holds. $\square$

Observe that the constant matrix-valued function $X : [a,b] \to \mathbb{C}^{q\times q}$ with value $\frac{1}{b-a} I_q$ is a simple example for a matrix-valued function which satisfies the assumptions of Lemma 6.1. In particular, there exists a sequence $(r_j)_{j=0}^{2n}$ of complex $q \times q$ matrices such that the block Hankel matrices

$$(r_{j+k})_{j,k=0}^{n} \quad \text{and} \quad (-abr_{j+k} + (a+b)r_{j+k+1} - r_{j+k+2})_{j,k=0}^{n-1}$$

are both positive Hermitian.

For our following considerations we will apply the description of the set $\mathcal{R}_q\left[[a,b];(s_j)_{j=0}^{2n}\right]$ given in Theorem 1.2 where the matrix-valued functions $\mathbf{K}_{1,n}^{[S]}$ and $\mathbf{K}_{2,n}^{[S]}$ given by (1.15) are used. Recall that Theorem 1.2 shows that a given matrix-valued function $S : \mathbb{C}\setminus[a,b] \to \mathbb{C}^{q\times q}$ belongs to $\mathcal{R}_q\left[[a,b];(s_j)_{j=0}^{2n}\right]$ if and only if S is a solution of the system of the fundamental matrix inequalities of Potapov-type associated with the interval $[a,b]$ and the sequence $(s_j)_{j=0}^{2n}$ of complex $q\times q$ matrices, i.e., if S is a holomorphic function for which the matrices $\mathbf{K}_{1,n}^{[S]}(z)$ and $\mathbf{K}_{2,n}^{[S]}(z)$ given by (1.15) are both nonnegative Hermitian for all $z\in\mathbb{C}\setminus\mathbb{R}$.

Assuming now that the left upper blocks in the block matrices (1.15) are moreover positive Hermitian we will construct a linear fractional description of the set $\mathcal{R}_q\left[[a,b];(s_j)_{j=0}^{2n}\right]$. In the first step using the factorization method of V.P. Potapov we will handle separately each of the two fundamental matrix inequalities of our system.

Remark 6.2. *Suppose that $(s_j)_{j=0}^{2n}$ is a sequence of complex $q\times q$ matrices such that the matrices $\mathbf{H}_{1,n}$ and $\mathbf{H}_{2,n}$ are both positive Hermitian. Let $S:\mathbb{C}\setminus[a,b]\to\mathbb{C}^{q\times q}$ be a matrix-valued function. Using a well-known criterion for the characterization of nonnegative Hermitian block matrices* (*see, e.g.,* [DFK, Lemma 1.1.9]) *it is easily checked that the matrices $\mathbf{K}_{1,n}^{[S]}(z)$ and $\mathbf{K}_{1,n}^{[S]}(z)$ are both nonnegative Hermitian for all $z\in\mathbb{C}\setminus\mathbb{R}$ if and only if for each $k\in\{1,2\}$ and each $z\in\mathbb{C}\setminus\mathbb{R}$ the matrix*

$$C_{k,n}^{[S]}(z) := \frac{S_k(z)-S_k^*(z)}{z-\overline{z}} \\ -(\mathbf{R}_{k,n}(z)(\mathbf{v}_{k,n}S_k(z)-\mathbf{u_{k,n}}))^*\mathbf{H}_{k,n}^{-1}(\mathbf{R}_{k,n}(z)(\mathbf{v}_{k,n}S_k(z)-\mathbf{u}_{k,n})) \tag{6.1}$$

is nonnegative Hermitian where $S_1 := S$ and where $S_2:\mathbb{C}\setminus[a,b]\to\mathbb{C}^{q\times q}$ is defined by (1.14).

A matrix-valued entire function W is said to belong to the *Potapov class* $\mathfrak{P}_{\tilde{J}_q}(\Pi_+)$ if for each $z\in\Pi_+$ the matrix $W(z)$ is $\tilde{J}_q$-contractive. If W belongs to $\mathfrak{P}_{\tilde{J}_q}(\Pi_+)$, then W is called an $\tilde{J}_q$-*inner function* if $\tilde{J}_q - W^*(x)\tilde{J}_qW(x)=0$ is valid for all $x\in\mathbb{R}$.

Proposition 6.3. *Let $(s_j)_{j=0}^{2n}$ be a sequence of complex $q\times q$ matrices such that the block Hankel matrices $\mathbf{H}_{1,n}$ and $\mathbf{H}_{2,n}$ are both positive Hermitian. For each $k\in\{1,2\}$, then $\tilde{U}_{k,n}:\mathbb{C}\to\mathbb{C}^{2q\times 2q}$ defined by*

$$\tilde{U}_{k,n}(z) := I_{2q}+i(z-a)(\mathbf{u}_{k,n},\mathbf{v}_{k,n})^*[\mathbf{R}_{k,n}(\overline{z})]^*\mathbf{H}_{k,n}^{-1}[\mathbf{R}_{k,n}(a)]\cdot(\mathbf{u}_{k,n},\mathbf{v}_{k,n})\tilde{J}_q \tag{6.2}$$

is a $2q\times 2q$ matrix polynomial of degree not greater that $n+2-k$. Furthermore, for each $k\in\{1,2\}$, the following statements hold:

(a) *For all $z\in\mathbb{C}$,*

$$\tilde{J}_q-\tilde{U}_{k,n}(z)\tilde{J}_q[\tilde{U}_{k,n}(z)]^* \\ = i(\overline{z}-z)(\mathbf{u}_{k,n},\mathbf{v}_{k,n})^*[\mathbf{R}_{k,n}(\overline{z})]^*\mathbf{H}_{k,n}^{-1}[\mathbf{R}_{k,n}(\overline{z})](\mathbf{u}_{k,n},\mathbf{v}_{k,n}). \tag{6.3}$$

In particular for each $w \in \Pi_+$,

$$\tilde{J}_q - \tilde{U}_{k,n}(w)\tilde{J}_q[\tilde{U}_{k,n}(w)]^* \geq 0. \tag{6.4}$$

Moreover, for each real number x,

$$\tilde{J}_q - \tilde{U}_{k,n}(x)\tilde{J}_q[\tilde{U}_{k,n}(x)]^* = 0. \tag{6.5}$$

(b) *The matrix-valued function* $\tilde{U}_{k,n}$ *is a* $\tilde{J}_q$*-inner function belonging to the Potapov class* $\mathfrak{P}_{\tilde{J}_q}(\Pi_+)$.

(c) *For all* $z \in \mathbb{C}$, *the matrix* $\tilde{U}_{k,n}(z)$ *is nonsingular and the identities*

$$[\tilde{U}_{k,n}(z)]^{-1}$$
$$= I_{2q} - i(z-a)(\mathbf{u}_{k,n}, \mathbf{v}_{k,n})^*[\mathbf{R}_{k,n}(a)]^*\mathbf{H}_{k,n}^{-1}[\mathbf{R}_{k,n}(z)](\mathbf{u}_{k,n}, \mathbf{v}_{k,n})\tilde{J}_q$$

and

$$\begin{aligned}&\tilde{J}_q - [\tilde{U}_{k,n}(z)]^{-*}\tilde{J}_q[\tilde{U}_{k,n}(z)]^{-1} \\ &= i(z-\overline{z})\tilde{J}_q(\mathbf{u}_{k,n}, \mathbf{v}_{k,n})^*[\mathbf{R}_{k,n}(z)]^*\mathbf{H}_{k,n}^{-1}\mathbf{R}_{k,n}(z)(\mathbf{u}_{k,n}, \mathbf{v}_{k,n})\tilde{J}_q\end{aligned} \tag{6.6}$$

hold.

Proof. Let $k \in \{1,2\}$. From (1.6) and (1.8) one can see that $\mathbf{R}_{k,n}$ is a matrix polynomial of degree $n+1-k$. Consequently, $\tilde{U}_{k,n}$ is a matrix polynomial of degree not greater that $n+2-k$. Obviously, for each $z \in \mathbb{C}$ from (1.5) and (1.8) the identities

$$\mathbf{R}_{k,n}(z) \cdot (I_{(n+2-k)q} - z\mathbf{T}_{k,n}) = I_{(n+2-k)q} \tag{6.7}$$

and

$$(I_{(n+2-k)q} - z\mathbf{T}_{k,n}) \cdot \mathbf{R}_{k,n}(z) = I_{(n+2-k)q} \tag{6.8}$$

hold. Using these equations and Remark 5.1 straightforward calculations yield

$$\begin{aligned}&\tilde{J}_q - \tilde{U}_{k,n}(z)\tilde{J}_q[\tilde{U}_{k,n}(z)]^* \\ &= i(\mathbf{u}_{k,n}, \mathbf{v}_{k,n})^*[\mathbf{R}_{k,n}(\overline{z})]^*\mathbf{H}_{k,n}^{-1}\mathbf{R}_{k,n}(a) \\ &\quad\cdot\Big(-(z-a)[\mathbf{R}_{k,n}(\overline{z})]^{-1}\mathbf{H}_{k,n}[\mathbf{R}_{k,n}(a)]^{-*} \\ &\qquad+(\overline{z}-a)[\mathbf{R}_{k,n}(a)]^{-1}\mathbf{H}_{k,n}[\mathbf{R}_{k,n}(\overline{z})]^{-*} \\ &\qquad\quad+|z-a|^2(\mathbf{H}_{k,n}\mathbf{T}_{k,n}^* - \mathbf{T}_{k,n}\mathbf{H}_{k,n})\Big) \\ &\quad\cdot[\mathbf{R}_{k,n}(a)]^*\mathbf{H}_{k,n}^{-1}[\mathbf{R}_{k,n}(\overline{z})] \cdot (\mathbf{u}_{k,n}, \mathbf{v}_{k,n})\end{aligned}$$

for every complex number z. For each $z \in \mathbb{C}$ we also have

$$\begin{aligned}a(z-a) - z(\overline{z}-a) + |z-a|^2 &= -a(\overline{z}-z), \\ \overline{z}(z-a) - a(\overline{z}-a) - |z-a|^2 &= -a(\overline{z}-z), \\ -(z-a)a\overline{z} + (\overline{z}-a)az &= a^2(\overline{z}-z),\end{aligned}$$

and because of (6.7) consequently

$$\begin{aligned}
&-(z-a)[\mathbf{R}_{k,n}(\overline{z})]^{-1}\mathbf{H}_{k,n}[\mathbf{R}_{k,n}(a)]^{-*}+(\overline{z}-a)[\mathbf{R}_{k,n}(a)]^{-1}\mathbf{H}_{k,n}[\mathbf{R}_{k,n}(\overline{z})]^{-*}\\
&\quad+|z-a|^2(\mathbf{H}_{k,n}\mathbf{T}_{k,n}^*-\mathbf{T}_{k,n}\mathbf{H}_{k,n})\\
&=-(z-a)(I-\overline{z}\mathbf{T}_{k,n})\mathbf{H}_{k,n}(I-a\mathbf{T}_{k,n}^*)+(\overline{z}-a)(I-a\mathbf{T}_{k,n})\mathbf{H}_{k,n}(I-z\mathbf{T}_{k,n}^*)\\
&\quad+|z-a|^2(\mathbf{H}_{k,n}\mathbf{T}_{k,n}^*-\mathbf{T}_{k,n}\mathbf{H}_{k,n})\\
&=(\overline{z}-z)(\mathbf{H}_{k,n}-a\mathbf{H}_{k,n}\mathbf{T}_{k,n}^*-a\mathbf{T}_{k,n}\mathbf{H}_{k,n}+a^2\mathbf{T}_{k,n}\mathbf{H}_{k,n}\mathbf{T}_{k,n}^*)\\
&=(\overline{z}-z)(I-a\mathbf{T}_{k,n})\mathbf{H}_{k,n}(I-a\mathbf{T}_{k,n}^*)\\
&=(\overline{z}-z)(\mathbf{R}_{k,n}(a))^{-1}\mathbf{H}_{k,n}(\mathbf{R}_{k,n}(a))^{-*}.
\end{aligned}$$

For every complex number z, thus we obtain identity (6.3). Since $\mathbf{H}_{k,n}$ is positive Hermitian (6.4) and (6.5) follow. Part (a) is proved. Parts (b) and (c) are immediate consequences of part (a), [DFK, Theorem 1.3.3], and [CDFK, Lemma 5.1]. □

Proposition 6.3 provides us now useful expressions of the Schur complements $C_{k,n}^{[s]}(z)$ in terms of the $2q\times 2q$ matrix polynomials $\tilde{U}_{k,n}$.

Lemma 6.4. *Let $(s_j)_{j=0}^{2n}$ be a sequence of complex $q\times q$ matrices such that the block Hankel matrices $\mathbf{H}_{1,n}$ and $\mathbf{H}_{2,n}$ are both positive Hermitian. For each $k\in\{1,2\}$, let $\tilde{U}_{k,n}:\mathbb{C}\to\mathbb{C}^{2q\times 2q}$ be given by* (6.2)*. Further, let $\mathcal{D}$ be a discrete subset of $\mathbb{C}\setminus[a,b]$, let $S:\mathbb{C}\setminus([a,b]\cup\mathcal{D})\to\mathbb{C}^{q\times q}$ be a matrix-valued function, let $S_1:=S$, and let $S_2:\mathbb{C}\setminus([a,b]\cup\mathcal{D})\to\mathbb{C}^{q\times q}$ be defined by* (1.14)*. For each $k\in\{1,2\}$ and each $z\in\mathbb{C}\setminus(\mathbb{R}\cup\mathcal{D})$, then the matrix $C_{k,n}^{[S]}(z)$ given by* (6.1) *admits the representation*

$$C_{k,n}^{[S]}(z)=\frac{1}{i(z-\overline{z})}\begin{pmatrix}S_k(z)\\ I\end{pmatrix}^*[\tilde{U}_{k,n}(z)]^{-*}\tilde{J}_q[\tilde{U}_{k,n}(z)]^{-1}\begin{pmatrix}S_k(z)\\ I\end{pmatrix}. \tag{6.9}$$

Proof. Let $k\in\{1,2\}$ and let $z\in\mathbb{C}\setminus(\mathbb{R}\cup\mathcal{D})$. From parts (c) and (d) of Proposition 6.3 we get that $\tilde{U}_{k,n}(z)$ is a nonsingular matrix and that (6.6) holds. Moreover, we have

$$\frac{1}{z-\overline{z}}\left(S_k(z)-S_k^*(z)\right)=\frac{1}{i(z-\overline{z})}\begin{pmatrix}S_k(z)\\ I\end{pmatrix}^*\tilde{J}_q\begin{pmatrix}S_k(z)\\ I\end{pmatrix}$$

and

$$\mathbf{v}_{k,n}S_k(z)-\mathbf{u}_{k,n}=\frac{1}{i}(\mathbf{u}_{k,n},\mathbf{v}_{k,n})\tilde{J}_q\begin{pmatrix}S_k(z)\\ I\end{pmatrix}$$

Thus it follows (6.9). □

Our next considerations are aimed to look for an appropriate coupling between the $2q\times 2q$ matrix polynomials $\tilde{U}_{1,n}$ and $\tilde{U}_{2,n}$. Taking into account the formulas (6.4) and (6.5) we see that a right multiplication of $\tilde{U}_{1,n}$ and $\tilde{U}_{2,n}$ by appropriately chosen $\tilde{J}_q$-unitary matrices A_1 and A_2 preserves the $\tilde{J}_q$-properties of these functions. So we are looking for particular $\tilde{J}_q$-unitary matrices $\tilde{A}_1$ and $\tilde{A}_2$ which generate the desired coupling. For this reason, we start with some observations on $\tilde{J}_q$-unitary matrices with a special product structure.

Remark 6.5. *Let M and N be complex $q \times q$ matrices, and let*

$$A := \begin{pmatrix} I & 0 \\ M & I \end{pmatrix} \begin{pmatrix} I & N \\ 0 & I \end{pmatrix} \quad \text{and} \quad B := \begin{pmatrix} I & N \\ 0 & I \end{pmatrix} \begin{pmatrix} I & 0 \\ M & I \end{pmatrix}.$$

Then

$$A^* \tilde{J}_q A = \begin{pmatrix} i(M - M^*) & -i[(M - M^*)N + I] \\ i[N^*(M^* - M) + I] & i[N^*(M^* - M)N + N - N^*] \end{pmatrix}$$

and

$$B^* \tilde{J}_q B = \begin{pmatrix} i[M^* - M + M^*(N - N^*)M] & -i[I + M^*(N - N^*)] \\ i[I + (N - N^*)M] & i(N - N^*) \end{pmatrix}.$$

In particular, the following statements are equivalent:

(i) $A^* \tilde{J}_q A = \tilde{J}_q$.

(ii) $B^* \tilde{J}_q B = \tilde{J}_q$.

(iii) $M = M^*$ *and* $N = N^*$.

Using Remark 6.5 we introduce in the following the concrete choice of the matrices M and N which will be used in the sequel.

Lemma 6.6. *Let $(s_j)_{j=0}^{2n}$ be a sequence of complex $q \times q$ matrices such that the block Hankel matrices $\mathbf{H}_{1,n}$ and $\mathbf{H}_{2,n}$ are both positive Hermitian. Then the matrix*

$$B_1 := \mathbf{v}_{1,n}^* (\mathbf{R}_{1,n}(a))^* \mathbf{H}_{1,n}^{-1} \mathbf{R}_{1,n}(a) \mathbf{v}_{1,n}, \tag{6.10}$$

is positive Hermitian and the matrix

$$\begin{aligned} M_1 := {} & \left((b-a)\mathbf{v}_{1,n}^* (\mathbf{R}_{1,n}(a))^* \mathbf{H}_{1,n}^{-1} \mathbf{R}_{1,n}(a) \mathbf{v}_{1,n}\right)^{-1} \\ & \big(I + (b-a)\left[\mathbf{v}_{1,n}^* (\mathbf{R}_{1,n}(a))^* \mathbf{H}_{1,n}^{-1} \mathbf{R}_{1,n}(a) \mathbf{u}_{1,n}\right. \\ & - \mathbf{v}_{2,n}^* (\mathbf{R}_{2,n}(a))^* \mathbf{H}_{2,n}^{-1} \mathbf{R}_{2,n}(a) \mathbf{u}_{2,n} \\ & \left.- a\mathbf{v}_{2,n}^* (\mathbf{R}_{2,n}(a))^* \mathbf{H}_{2,n}^{-1} \mathbf{R}_{2,n}(a) \mathbf{v}_{2,n} s_0\right]\big) \end{aligned} \tag{6.11}$$

is Hermitian.

Proof. It is readily checked that B_1 is a positive Hermitian matrix. A straightforward calculation yields

$$\begin{aligned} & M_1 - M_1^* \\ & \quad = (b-a)^{-1} B_1^{-1} \left(C_1^* B_1 - C_2^* B_1 - a B_2 s_0 B_1 - B_1 C_1 + B_1 C_2 + a B_1 s_0 B_2\right) B_1^{-1} \end{aligned}$$

where

$$B_2 := \mathbf{v}_{2,n}^* (\mathbf{R}_{2,n}(a))^* \mathbf{H}_{2,n}^{-1} \mathbf{R}_{2,n}(a) \mathbf{v}_{2,n}, \tag{6.12}$$

$$C_1 := \mathbf{u}_{1,n}^* (\mathbf{R}_{1,n}(a))^* \mathbf{H}_{1,n}^{-1} \mathbf{R}_{1,n}(a) \mathbf{v}_{1,n}, \tag{6.13}$$

and

$$C_2 := \mathbf{u}_{2,n}^* (\mathbf{R}_{2,n}(a))^* \mathbf{H}_{2,n}^{-1} \mathbf{R}_{2,n}(a) \mathbf{v}_{2,n}. \tag{6.14}$$

Because of Proposition 2.1, Lemma 2.5, (2.2), (2.17), and (2.18) we have

$$\begin{aligned} & C_1^* B_1 - C_2^* B_1 - a B_2 s_0 B_1 - B_1 C_1 + B_1 C_2 + a B_1 s_0 B_2 \\ & \quad = \mathbf{v}_{1,n}^* (\mathbf{R}_{1,n}(a))^* \mathbf{H}_{1,n}^{-1} \mathbf{R}_{1,n}(a) (\mathbf{u}_{1,n} \mathbf{v}_{1,n}^* - \mathbf{v}_{1,n} \mathbf{u}_{1,n}^*) (\mathbf{R}_{1,n}(a))^* \mathbf{H}_{1,n}^{-1} \mathbf{R}_{1,n}(a) \mathbf{v}_{1,n} \end{aligned}$$

$$
\begin{aligned}
&-\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(a))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)(\mathbf{u}_{2,n}\mathbf{v}_{1,n}^*+a\mathbf{v}_{2,n}s_0\mathbf{v}_{1,n}^*)(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&+\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)(\mathbf{v}_{1,n}\mathbf{u}_{2,n}^*+a\mathbf{v}_{1,n}s_0\mathbf{v}_{2,n}^*)(\mathbf{R}_{2,n}(a))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&=\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)(\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*-\mathbf{T}_{1,n}\mathbf{H}_{1,n})(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&-\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(a))^*\mathbf{H}_{2,n}^{-1}\Big(\mathbf{H}_{2,n}\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\\
&\qquad+\mathbf{R}_{2,n}(a)\big[-(b-a)\mathbf{L}_{2,n}^*(I_{(n+1)q}-a\mathbf{T}_{1,n})\\
&+(I_{nq}-a\mathbf{T}_{2,n})(\mathbf{L}_{1,n}^*-a\mathbf{L}_{2,n}^*)\big]\mathbf{H}_{1,n}\Big)\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&+\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\Big(\mathbf{R}_{1,n}(a)\mathbf{T}_{1,n}\mathbf{L}_{2,n}\mathbf{H}_{2,n}\\
&+\mathbf{H}_{1,n}\big[-(b-a)(I_{(n+1)q}-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}\\
&+(\mathbf{L}_{1,n}-a\mathbf{L}_{2,n})(I_{nq}-a\mathbf{T}_{2,n}^*)\big](\mathbf{R}_{2,n}(a))^*\Big)\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&=\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)(\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*-\mathbf{T}_{1,n}\mathbf{H}_{1,n})(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&-\mathbf{v}_{1,n}^*\mathbf{L}_{2,n}(\mathbf{R}_{2,n}(a))^*\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&+\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{T}_{1,n}\mathbf{L}_{2,n}\mathbf{R}_{2,n}(a)\mathbf{L}_{2,n}^*\mathbf{v}_{1,n}.
\end{aligned}
$$

Applying (2.18) and (2.5) it follows

$$
\begin{aligned}
&C_1^*B_1-C_2^*B_1-aB_2s_0B_1-B_1C_1+B_1C_2+aB_1s_0B_2\\
&=\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)(\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*-\mathbf{T}_{1,n}\mathbf{H}_{1,n})(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad-\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\mathbf{L}_{2,n}\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad+\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{T}_{1,n}\mathbf{L}_{2,n}\mathbf{L}_{2,n}^*\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&=\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)(\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*-\mathbf{T}_{1,n}\mathbf{H}_{1,n})(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad-\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\mathbf{T}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad+\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{T}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&=\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\left(\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*-\mathbf{T}_{1,n}\mathbf{H}_{1,n}\right.\\
&\quad-(I_{(n+1)q}-a\mathbf{T}_{1,n})\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*\\
&\quad\left.+\mathbf{T}_{1,n}\mathbf{H}_{1,n}(I_{(n+1)q}-a\mathbf{T}_{1,n}^*)\right)(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}=0.
\end{aligned}
$$

The proof is complete. □

Lemma 6.7. *Let $(s_j)_{j=0}^{2n}$ be a sequence of complex $q\times q$ matrices such that the block Hankel matrices $\mathbf{H}_{1,n}$ and $\mathbf{H}_{2,n}$ are both positive Hermitian. Let M_1 be defined by* (6.11)*, let*

$$M_2:=-as_0,\ N_2:=-(b-a)^{-1}\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\cdot\mathbf{v}_{1,n}, \tag{6.15}$$

$$A_1:=\begin{pmatrix} I_q & M_1\\ 0_{q\times q} & I_q\end{pmatrix},\quad \text{and}\quad A_2:=\begin{pmatrix} I_q & M_2\\ 0_{q\times q} & I_q\end{pmatrix}\begin{pmatrix} I_q & 0_{q\times q}\\ N_2 & I_q\end{pmatrix}. \tag{6.16}$$

For each $k\in\{1,2\}$, then the matrix A_k is $\tilde{J}_q$-unitary and $U_{k,n}:\mathbb{C}\to\mathbb{C}^{2q\times 2q}$ defined by $U_{k,n}:=\tilde{U}_{k,n}A_k$ and (6.2) *is a $2q\times 2q$ matrix polynomial of degree not greater than $n+2-k$ which belongs to the Potapov class $\mathfrak{P}_{\tilde{J}_q}(\Pi_+)$. Moreover, for every choice of z in $\mathbb{C}$ and k in $\{1,2\}$, the matrix $U_{k,n}$ is nonsingular and the identities*

$$U_{k,n}(z)\tilde{J}_q[U_{k,n}(z)]^*=\tilde{U}_{k,n}(z)\tilde{J}_q[\tilde{U}_{k,n}(z)]^* \tag{6.17}$$

and

$$[U_{k,n}(z)]^{-*}\tilde{J}_q[U_{k,n}(z)]^{-1} = [\tilde{U}_{k,n}(z)]^{-*}\tilde{J}_q[\tilde{U}_{k,n}(z)]^{-1} \tag{6.18}$$

hold.

Proof. Lemma 6.6 yields $M_1^* = M_1$. Furthermore, the matrices M_2 and N_2 are also Hermitian. According to Remark 6.5 we see then that A_1 and A_2 are $\tilde{J}_q$-unitary matrices. In particular, A_1 and A_2 are nonsingular and A_1^{-1} and A_2^{-1} are $\tilde{J}_q$-unitary as well. From Proposition 6.3 the identities (6.17) and (6.18) follow. □

Using the $2q \times 2q$ matrix polynomials $U_{1,n}$ and $U_{2,n}$ we obtain the following description of the set $\mathcal{R}_q\left[[a,b];(s_j)_{j=0}^{2n}\right]$.

Proposition 6.8. *Let $(s_j)_{j=0}^{2n}$ be a sequence of complex $q \times q$ matrices such that the block Hankel matrices $\mathbf{H}_{1,n}$ and $\mathbf{H}_{2,n}$ are both positive Hermitian. Let $S : \mathbb{C} \setminus [a,b] \to \mathbb{C}^{q\times q}$ be a $q \times q$ matrix-valued function, let $S_1 := S$, and let $S_2 : \mathbb{C} \setminus [a,b] \to \mathbb{C}^{q\times q}$ be defined by* (1.14)*. Further, for each $k \in \{1,2\}$, let $U_{k,n}$ be the matrix polynomial defined by* (6.3)*,* (6.16)*, and $U_{k,n} := \tilde{U}_{k,n}A_k$. Then S belongs to $\mathcal{R}_q[[a,b];(s_j)_{j=0}^{2n}]$ if and only if S is holomorphic in $\mathbb{C} \setminus [a,b]$ and the matrix inequality*

$$\frac{1}{i(z-\overline{z})}\begin{pmatrix} S_k(z) \\ I \end{pmatrix}^* [U_{k,n}(z)]^{-*}\tilde{J}_q[U_{k,n}(z)]^{-1}\begin{pmatrix} S_k(z) \\ I \end{pmatrix} \geq 0 \tag{6.19}$$

holds for each $k \in \{1,2\}$ and each $z \in \mathbb{C} \setminus \mathbb{R}$.

Proof. Use Theorem 1.2, Remark 6.2, Lemma 6.4, and Lemma 6.7. □

Now we are going to verify that the solution set of the system (6.19) of matrix inequalities can be described by an appropriately constructed linear fractional transformation. This will be done in several steps. In the first step we analyze the $q \times q$ block structure of the $2q \times 2q$ matrix polynomials $U_{1,n}$ and $U_{2,n}$ which are introduced in Lemma 6.7.

Remark 6.9. *Let $(s_j)_{j=0}^{2n}$ be a sequence of complex $q\times q$ matrices such that the block Hankel matrices $\mathbf{H}_{1,n}$ and $\mathbf{H}_{2,n}$ are positive Hermitian. Let $k \in \{1,2\}$, let $\tilde{U}_{k,n}$ be the matrix polynomial defined by* (6.2)*, let A_k be the matrix given by* (6.16)*, and let $U_{k,n} := \tilde{U}_{k,n}A_k$. Then straightforward calculations show that $U_{1,n}$ and $U_{2,n}$ admit for every complex number z the $q \times q$ block representations*

$$U_{1,n}(z) = \begin{pmatrix} U_{11,n}^{(1)}(z) & U_{12,n}^{(1)}(z) \\ U_{21,n}^{(1)}(z) & U_{22,n}^{(1)}(z) \end{pmatrix} \quad \text{and} \quad U_{2,n}(z) = \begin{pmatrix} U_{11,n}^{(2)}(z) & U_{12,n}^{(2)}(z) \\ U_{21,n}^{(2)}(z) & U_{22,n}^{(2)}(z) \end{pmatrix} \tag{6.20}$$

where

$$\begin{aligned} U_{11,n}^{(1)}(z) &= I - (z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}, \\ U_{12,n}^{(1)}(z) &= \left(I - (z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\right)M_1 \\ &\quad +(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{u}_{1,n}, \end{aligned}$$

$$U^{(1)}_{21,n}(z) = -(z-a)\mathbf{v}^*_{1,n}(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n},$$
$$U^{(1)}_{22,n}(z) = -(z-a)\mathbf{v}^*_{1,n}(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}M_1 + I + (z-a)\mathbf{v}^*_{1,n}(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{u}_{1,n},$$
$$U^{(2)}_{11,n}(z) = \left(I - (z-a)\mathbf{u}^*_{2,n}(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}^{-1}_{2,n}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\right)(I + M_2N_2) + (z-a)\mathbf{u}^*_{2,n}(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}^{-1}_{2,n}\mathbf{R}_{2,n}(a)\mathbf{u}_{2,n}N_2,$$
$$U^{(2)}_{12,n}(z) = \left(I - (z-a)\mathbf{u}^*_{2,n}(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}^{-1}_{2,n}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\right)M_2 + (z-a)\mathbf{u}^*_{2,n}(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}^{-1}_{2,n}\mathbf{R}_{2,n}(a)\mathbf{u}_{2,n},$$
$$U^{(2)}_{21,n}(z) = -(z-a)\mathbf{v}^*_{2,n}(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}^{-1}_{2,n}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}(I + M_2N_2) + \left(I + (z-a)\mathbf{v}^*_{2,n}(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}^{-1}_{2,n}\mathbf{R}_{2,n}(a)\mathbf{u}_{2,n}\right)N_2,$$

and

$$U^{(2)}_{22,n}(z) = -(z-a)\mathbf{v}^*_{2,n}(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}^{-1}_{2,n}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}M_2 + I + (z-a)\mathbf{v}^*_{2,n}(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}^{-1}_{2,n}\mathbf{R}_{2,n}(a)\mathbf{u}_{2,n}.$$

The following result marks one of the crucial points of the whole paper. It contains the desired coupling between the two single matrix inequalities of our system. We introduce a $2q \times 2q$ matrix polynomial V_n which describes precisely the coupling between the $2q \times 2q$ matrix polynomials $U_{1,n}$ and $U_{2,n}$.

Proposition 6.10. *Let $(s_j)_{j=0}^{2n}$ be a sequence of complex $q \times q$ matrices such that the block Hankel matrices $\mathbf{H}_{1,n}$ and $\mathbf{H}_{2,n}$ are both positive Hermitian. Then $V_n : \mathbb{C} \to \mathbb{C}^{2q\times 2q}$ defined by*

$$V_n(z) := \begin{pmatrix} V_{11,n}(z) & V_{12,n}(z) \\ V_{21,n}(z) & V_{22,n}(z) \end{pmatrix} \tag{6.21}$$

where

$$V_{11,n}(z) := (z-a)(b-z)\left(I - (z-a)\mathbf{u}^*_{1,n}[\mathbf{R}_{1,n}(\overline{z})]^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\right) + zs_0(z-a)\mathbf{v}^*_{1,n}[\mathbf{R}_{1,n}(\overline{z})]^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}, \tag{6.22}$$

$$V_{12,n}(z) := (z-a)(b-z)\left(M_1 - (z-a)\mathbf{u}^*_{1,n}[\mathbf{R}_{1,n}(\overline{z})]^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}M_1 + (z-a)\mathbf{u}^*_{1,n}[\mathbf{R}_{1,n}(\overline{z})]^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{u}_{1,n}\right) - zs_0\left(-(z-a)\mathbf{v}^*_{1,n}[\mathbf{R}_{1,n}(\overline{z})]^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}M_1 + I + (z-a)\mathbf{v}^*_{1,n}[\mathbf{R}_{1,n}(\overline{z})]^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{u}_{1,n}\right), \tag{6.23}$$

$$V_{21,n}(z) := -(z-a)\mathbf{v}^*_{1,n}[\mathbf{R}_{1,n}(\overline{z})]^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}, \tag{6.24}$$

and

$$V_{22,n}(z) := -(z-a)\mathbf{v}^*_{1,n}[\mathbf{R}_{1,n}(\overline{z})]^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}M_1 + I + (z-a)\mathbf{v}^*_{1,n}[\mathbf{R}_{1,n}(\overline{z})]^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{u}_{1,n} \tag{6.25}$$

is a $2q \times 2q$ matrix polynomial of degree not greater than $n+3$. Moreover, the following statements hold:

(a) *For every complex number z, the identities*

$$V_n(z) = \begin{pmatrix} I & -s_0 z \\ 0 & I \end{pmatrix} \begin{pmatrix} (z-a)(b-z)I & 0 \\ 0 & I \end{pmatrix} U_{1,n}(z) \tag{6.26}$$

and

$$V_n(z) = U_{2,n}(z) \begin{pmatrix} (b-a)(z-a)I & 0 \\ 0 & \frac{b-z}{b-a} I \end{pmatrix} \tag{6.27}$$

hold where $U_{1,n} := \tilde{U}_{1,n} A_1$ and $U_{2,n} := \tilde{U}_{2,n} A_2$ and where $\tilde{U}_{1,n}, \tilde{U}_{2,n}, A_1$, and A_2 are defined by (6.2) *and* (6.16)*.*

(b) *For each $z \in \mathbb{C} \setminus \{a, b\}$, the matrix $V_n(z)$ is nonsingular.*

Proof. We use the notations given above. Let $z \in \mathbb{C}$. From Remark 6.9 we get

$$V_{11,n}(z) = (z-a)(b-z)U^{(1)}_{11,n}(z) - s_0 z U^{(1)}_{21,n}(z),$$
$$V_{12,n}(z) = (z-a)(b-z)U^{(1)}_{12,n}(z) - s_0 z U^{(1)}_{22,n}(z),$$
$$V_{21,n}(z) = U^{(1)}_{21,n}(z), \quad \text{and} \quad V_{22,n}(z) = U^{(1)}_{22,n}(z).$$

Thus (6.26) follows. If $z \notin \{a, b\}$, then (6.26) implies $\det V_n(z) \neq 0$. It remains to prove identity (6.27). For this reason, setting

$$\begin{aligned}
&\triangle_{11,n}(z) := V_{11,n}(z) \\
&-(b-a)(z-a)\big((I - (z-a)\mathbf{u}^*_{2,n}(\mathbf{R}_{2,n}(\overline{z}))^* \mathbf{H}^{-1}_{2,n} \mathbf{R}_{2,n}(a)\mathbf{v}_{2,n})(I + M_2 N_2) \\
&\quad +(z-a)\mathbf{u}^*_{2,n}(\mathbf{R}_{2,n}(\overline{z}))^* \mathbf{H}^{-1}_{2,n} \mathbf{R}_{2,n}(a)\mathbf{u}_{2,n} N_2\big),
\end{aligned} \tag{6.28}$$

$$\begin{aligned}
\triangle_{12,n}(z) := V_{12,n}(z) - \frac{b-z}{b-a}&\big(\big(I-(z-a)\mathbf{u}^*_{2,n}(\mathbf{R}_{2,n}(\overline{z}))^* \mathbf{H}^{-1}_{2,n} \mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\big) M_2 \\
+(z-a)\mathbf{u}^*_{2,n}&(\mathbf{R}_{2,n}(\overline{z}))^* \mathbf{H}^{-1}_{2,n} \mathbf{R}_{2,n}(a)\mathbf{u}_{2,n}\big),
\end{aligned} \tag{6.29}$$

$$\begin{aligned}
&\triangle_{21,n}(z) := V_{21,n}(z) \\
&-(b-a)(z-a)\big(-(z-a)\mathbf{v}^*_{2,n}(\mathbf{R}_{2,n}(\overline{z}))^* \mathbf{H}^{-1}_{2,n} \mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}(I + M_2 N_2) \\
&\quad + \big(I + (z-a)\mathbf{v}^*_{2,n}(\mathbf{R}_{2,n}(\overline{z}))^* \mathbf{H}^{-1}_{2,n} \mathbf{R}_{2,n}(a)\mathbf{u}_{2,n}\big) N_2\big),
\end{aligned} \tag{6.30}$$

and

$$\begin{aligned}
\triangle_{22,n}(z) := V_{22,n}(z) - \frac{b-z}{b-a}&\big(-(z-a)\mathbf{v}^*_{2,n}(\mathbf{R}_{2,n}(\overline{z}))^* \mathbf{H}^{-1}_{2,n} \mathbf{R}_{2,n}(a)\mathbf{v}_{2,n} M_2 \\
+I + (z-a)\mathbf{v}^*_{2,n}&(\mathbf{R}_{2,n}(\overline{z}))^* \mathbf{H}^{-1}_{2,n} \mathbf{R}_{2,n}(a)\mathbf{u}_{2,n}\big),
\end{aligned} \tag{6.31}$$

because of (6.27) and Remark 6.9 it is sufficient to verify that the identities $\triangle_{11,n}(z) = 0, \triangle_{12,n}(z) = 0, \triangle_{21,n}(z) = 0$, and $\triangle_{22,n}(z) = 0$ hold. The proof of these four identities is one of the most delicate moments of the whole paper. It requires longer computations in which we make essential use of the coupling identity of Proposition 2.5. Let B_1, B_2, C_1, and C_2 be defined by (6.10), (6.11), (6.12), and (6.14). Using (6.28), (6.22), and (6.15) we get

$$\begin{aligned}
\triangle_{11,n}(z) = (z-a)\big(&-(z-a)I - (b-z)(z-a)\mathbf{u}^*_{1,n}(\mathbf{R}_{1,n}(\overline{z}))^* \mathbf{H}^{-1}_{1,n} \mathbf{R}_{1,n}(a)\mathbf{v}_{1,n} \\
&+z s_0 \mathbf{v}^*_{1,n}(\mathbf{R}_{1,n}(\overline{z}))^* \mathbf{H}^{-1}_{1,n} \mathbf{R}_{1,n}(a)\mathbf{v}_{1,n} \\
&-(b-a)M_2 N_2 + (b-a)(z-a)\mathbf{u}^*_{2,n}(\mathbf{R}_{2,n}(\overline{z}))^* \mathbf{H}^{-1}_{2,n} \mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}
\end{aligned}$$

$$
\begin{aligned}
&\quad +(b-a)(z-a)\mathbf{u}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}M_2N_2\\
&\quad -(b-a)(z-a)\mathbf{u}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{u}_{2,n}N_2\big)\\
&= (z-a)\big(-(z-a)I-(b-z)(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad +zs_0\mathbf{v}_{1,n}^*(I-a\mathbf{T}_{1,n}^*)(\mathbf{R}_{1,n}(a))^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad -as_0\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(a))^*(I-z\mathbf{T}_{1,n}^*)(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad +(b-a)(z-a)\mathbf{u}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&\quad +(z-a)\mathbf{u}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)(a\mathbf{v}_{2,n}s_0\mathbf{v}_{1,n}^*+\mathbf{u}_{2,n}\mathbf{v}_{1,n}^*)\\
&\qquad \cdot(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\big).
\end{aligned}
\tag{6.32}
$$

Because of (2.21) and (2.20) we have

$$
\begin{aligned}
&zs_0\mathbf{v}_{1,n}^*(I-a\mathbf{T}_{1,n}^*)(\mathbf{R}_{1,n}(a))^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad -as_0\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(a))^*(I-z\mathbf{T}_{1,n}^*)(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&= s_0\mathbf{v}_{1,n}^*\big[zI-az\mathbf{T}_{1,n}^*-(aI-az\mathbf{T}_{1,n}^*)\big](\mathbf{R}_{1,n}(\overline{z}))^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&= (z-a)s_0\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}.
\end{aligned}
\tag{6.33}
$$

Moreover, application of Proposition 2.5 yields

$$
\begin{aligned}
&\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)(a\mathbf{v}_{2,n}s_0\mathbf{v}_{1,n}^*+\mathbf{u}_{2,n}\mathbf{v}_{1,n}^*)(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\\
&= \mathbf{H}_{2,n}^{-1}\big[\mathbf{H}_{2,n}\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\\
&\quad +\mathbf{R}_{2,n}(a)\big(-(b-a)\mathbf{L}_{2,n}^*(I-a\mathbf{T}_{1,n})+(I-a\mathbf{T}_{2,n})(\mathbf{L}_{1,n}^*-a\mathbf{L}_{2,n}^*)\big)\mathbf{H}_{1,n}\big]\mathbf{H}_{1,n}^{-1}\\
&= \mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}-(b-a)\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{L}_{2,n}^*(I-a\mathbf{T}_{1,n})\\
&\quad +\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)(I-a\mathbf{T}_{2,n})(\mathbf{L}_{1,n}^*-a\mathbf{L}_{2,n}^*).
\end{aligned}
\tag{6.34}
$$

From (6.33) and (6.34) we get then that the right-hand side of (6.32) (and hence $\triangle_{11,n}(z)$) is equal to

$$
\begin{aligned}
&(z-a)^2\big((-I-(b-z)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad +s_0\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad +(b-a)\mathbf{u}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&\quad +\mathbf{u}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\big[\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\\
&\quad -(b-a)\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{L}_{2,n}^*(I-a\mathbf{T}_{1,n})\\
&\quad +\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)(I-a\mathbf{T}_{2,n})(\mathbf{L}_{1,n}^*-a\mathbf{L}_{2,n}^*)\big]\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\big).
\end{aligned}
\tag{6.35}
$$

According to (2.2), (6.8), and (2.17) we have

$$
\begin{aligned}
&(b-a)\mathbf{u}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&\quad +\mathbf{u}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\big[-(b-a)\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{L}_{2,n}^*(I-a\mathbf{T}_{1,n})\\
&\quad +\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)(I-a\mathbf{T}_{2,n})(\mathbf{L}_{1,n}^*-a\mathbf{L}_{2,n}^*)\big]\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&= 0.
\end{aligned}
$$

Thus it follows from (6.35), (2.7), and (2.19) that

$$
\begin{aligned}
\triangle_{11,n}(z) &= (z-a)^2\big(-I-(b-z)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad +s_0\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad +\mathbf{u}_{2,n}^*\mathbf{L}_{2,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{T}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\big).
\end{aligned}
\tag{6.36}
$$

Using

$$\mathbf{v}_{1,n}^* \left(I - a\mathbf{T}_{1,n}\right) \mathbf{H}_{1,n}(I - a\mathbf{T}_{1,n}^*)(I - z\mathbf{T}_{1,n}^*)(\mathbf{R}_{1,n}(\overline{z}))^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1} \cdot\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n} = I, \tag{6.37}$$

(2.10), and (2.13) we obtain then

$$\begin{aligned}\triangle_{11,n}(z) = (z-a)^2 \big(&-\mathbf{v}_{1,n}^*(I - a\mathbf{T}_{1,n})\mathbf{H}_{1,n}(I - a\mathbf{T}_{1,n}^*)(I - z\mathbf{T}_{1,n}^*)\\ &\cdot(\mathbf{R}_{1,n}(\overline{z}))^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\ &+s_0\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\ &+(b-z)\mathbf{v}_{1,n}^*\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*(I - a\mathbf{T}_{1,n}^*)(\mathbf{R}_{1,n}(a))^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\ &\cdot\big[-(a+b)\mathbf{v}_{1,n}^*\mathbf{H}_{1,n}\mathbf{L}_{2,n} + ab\mathbf{v}_{1,n}^*\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*\mathbf{L}_{2,n} + \mathbf{v}_{1,n}^*\mathbf{H}_{1,n}\mathbf{L}_{1,n}\big]\\ &\cdot\mathbf{L}_{2,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{T}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\big).\end{aligned}$$

By virtue of (2.20) it follows

$$\begin{aligned}\triangle_{11,n}(z) &= (z-a)^2 \big(-\mathbf{v}_{1,n}^*(I - a\mathbf{T}_{1,n})\mathbf{H}_{1,n}(I - a\mathbf{T}_{1,n}^*)(I - z\mathbf{T}_{1,n}^*)\\ &\quad+(b-z)\mathbf{v}_{1,n}^*\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*(I - a\mathbf{T}_{1,n}^*) + s_0\mathbf{v}_{1,n}^*\\ &\quad+\big[-(a+b)\mathbf{v}_{1,n}^*\mathbf{H}_{1,n}\mathbf{L}_{2,n} + ab\mathbf{v}_{1,n}^*\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*\mathbf{L}_{2,n} + \mathbf{v}_{1,n}^*\mathbf{H}_{1,n}\mathbf{L}_{1,n}\big]\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*\big)\\ &\quad\cdot(\mathbf{R}_{1,n}(\overline{z}))^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\ &= (z-a)^2\big[\mathbf{v}_{1,n}^*\big(-\mathbf{H}_{1,n} + a\mathbf{H}_{1,n}\mathbf{T}_{1,n}^* + z\mathbf{H}_{1,n}\mathbf{T}_{1,n}^* - az\mathbf{H}_{1,n}(\mathbf{T}_{1,n}^*)^2\\ &\quad+a\mathbf{T}_{1,n}\mathbf{H}_{1,n} - az\mathbf{T}_{1,n}\mathbf{H}_{1,n}\mathbf{T}_{1,n}^* - a^2\mathbf{T}_{1,n}\mathbf{H}_{1,n}\mathbf{T}_{1,n}^* + a^2z\mathbf{T}_{1,n}\mathbf{H}_{1,n}(\mathbf{T}_{1,n}^*)^2\\ &\quad+b\mathbf{H}_{1,n}\mathbf{T}_{1,n}^* - ab\mathbf{H}_{1,n}(\mathbf{T}_{1,n}^*)^2 - z\mathbf{H}_{1,n}\mathbf{T}_{1,n}^* + az\mathbf{H}_{1,n}(\mathbf{T}_{1,n}^*)^2\\ &\quad-a\mathbf{H}_{1,n}\mathbf{L}_{2,n}\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^* - b\mathbf{H}_{1,n}\mathbf{L}_{2,n}\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*\\ &\quad+ab\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*\mathbf{L}_{2,n}\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^* + \mathbf{H}_{1,n}\mathbf{L}_{1,n}\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*\big) + s_0\mathbf{v}_{1,n}^*\big]\\ &\quad\cdot(\mathbf{R}_{1,n}(\overline{z}))^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}.\end{aligned}$$

Using (2.5) and (2.6) we get then

$$\begin{aligned}\triangle_{11,n}(z) = (z-a)^2\big(&s_0\mathbf{v}_{1,n}^* + \mathbf{v}_{1,n}^*\mathbf{H}_{1,n}(\mathbf{T}_{1,n}\mathbf{T}_{1,n}^* - I)\\ &+\mathbf{v}_{1,n}^*\mathbf{T}_{1,n}\big(a\mathbf{H}_{1,n} - az\mathbf{H}_{1,n}\mathbf{T}_{1,n}^* - a^2\mathbf{H}_{1,n}\mathbf{T}_{1,n}^* + a^2z\mathbf{H}_{1,n}(\mathbf{T}_{1,n}^*)^2\big)\big)\\ &\cdot(\mathbf{R}_{1,n}(\overline{z}))^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}.\end{aligned} \tag{6.38}$$

Because of

$$s_0\mathbf{v}_{1,n}^* + \mathbf{v}_{1,n}^*\mathbf{H}_{1,n}(\mathbf{T}_{1,n}\mathbf{T}_{1,n}^* - I) = 0 \tag{6.39}$$

and $\mathbf{v}_{1,n}^*\mathbf{T}_{1,n} = 0$ the right-hand side of (6.38) is equal to zero. Consequently, $\triangle_{11,n}(z) = 0$ is proved. Because of (6.29), (6.23), and Lemma 6.6 we have

$$\begin{aligned}\triangle_{12,n}(z) &= (z-a)(b-z)\left(M_1^* - (z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}M_1^*\right.\\ &\quad\left.+(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{u}_{1,n}\right)\\ &\quad-zs_0\left[-(z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}M_1^* + I\right.\\ &\quad\quad\left.+(z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{u}_{1,n}\right]\\ &\quad-\frac{b-z}{b-a}\left(-as_0 + a(z-a)\mathbf{u}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}s_0\right.\\ &\quad\left.+\,(z-a)\mathbf{u}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{u}_{2,n}\right)\\ &= \frac{1}{b-a}\Big((z-a)(b-z)\big[I + (b-a)C_1 - (b-a)C_2 - a(b-a)s_0B_2\end{aligned}$$

$$
\begin{aligned}
&-(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\big(I+(b-a)C_1\\
&-(b-a)C_2-a(b-a)s_0B_2\big)\\
&+(z-a)(b-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{u}_{1,n}B_1\big]\\
&-zs_0\big[-(z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\,(I+(b-a)C_1-(b-a)C_2\\
&-a(b-a)s_0B_2)+(b-a)B_1\\
&+(b-a)(z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{u}_{1,n}B_1\big]\\
&-(b-z)\big[-as_0B_1+a(z-a)\mathbf{u}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}s_0B_1\\
&+(z-a)\mathbf{u}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{u}_{2,n}B_1\big]\Big)B_1^{-1}\\
=\;&\frac{1}{b-a}\Big((z-a)(b-z)\big[I+(b-a)C_1-(b-a)C_2-a(b-a)s_0B_2\\
&-(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&+(b-a)(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)(\mathbf{u}_{1,n}\mathbf{v}_{1,n}^*-\mathbf{v}_{1,n}\mathbf{u}_{1,n}^*)\\
&\quad\cdot(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&+(b-a)(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)(\mathbf{v}_{1,n}\mathbf{u}_{2,n}^*+a\mathbf{v}_{1,n}s_0\mathbf{v}_{2,n}^*)\\
&\quad\cdot(\mathbf{R}_{2,n}(a))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\big]-s_0z\big[-(z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&+(b-a)(z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)(\mathbf{u}_{1,n}\mathbf{v}_{1,n}^*-\mathbf{v}_{1,n}\mathbf{u}_{1,n}^*)\\
&\quad\cdot(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&+(b-a)(z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)(\mathbf{v}_{1,n}\mathbf{u}_{2,n}^*+a\mathbf{v}_{1,n}s_0\mathbf{v}_{2,n}^*)\\
&\quad\cdot(\mathbf{R}_{2,n}(a))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}+(b-a)B_1\big]\\
&-(b-z)\big[-as_0B_1+(z-a)\mathbf{u}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\\
&\quad\cdot(\mathbf{u}_{2,n}\mathbf{v}_{1,n}^*+a\mathbf{v}_{2,n}s_0\mathbf{v}_{1,n}^*)\cdot(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\big]\Big)B_1^{-1}.
\end{aligned}
$$

Proposition 2.1 and Proposition 2.5 yield

$$
\begin{aligned}
\triangle_{12,n}(z)=\;&\frac{1}{b-a}\Big[(z-a)(b-z)\big[I+(b-a)C_1-(b-a)C_2-a(b-a)s_0B_2\\
&-(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&+(b-a)(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\left(\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*-\mathbf{T}_{1,n}\mathbf{H}_{1,n}\right)\\
&\quad\cdot(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&+(b-a)(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\big[\mathbf{R}_{1,n}(a)\mathbf{T}_{1,n}\mathbf{L}_{2,n}\mathbf{H}_{2,n}+\mathbf{H}_{1,n}\\
&\quad\cdot\big(-(b-a)(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}+(\mathbf{L}_{1,n}-a\mathbf{L}_{2,n})(I-a\mathbf{T}_{2,n}^*)\big)(\mathbf{R}_{2,n}(a))^*\big]\\
&\quad\cdot\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\big]\\
&-zs_0\big[-(z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&+(b-a)(z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)(\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*-\mathbf{T}_{1,n}\mathbf{H}_{1,n})\\
&\quad\cdot(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&+(b-a)(z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\big[\mathbf{R}_{1,n}(a)\mathbf{T}_{1,n}\mathbf{L}_{2,n}\mathbf{H}_{2,n}+\mathbf{H}_{1,n}\\
&\quad\cdot\big(-(b-a)(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}+(\mathbf{L}_{1,n}-a\mathbf{L}_{2,n})(I-a\mathbf{T}_{2,n}^*)\big](\mathbf{R}_{2,n}(a))^*\big)\\
&\quad\cdot\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}+(b-a)B_1\big]\\
&-(b-z)\big[-as_0B_1+(z-a)\mathbf{u}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}
\end{aligned}
$$

$$
\begin{aligned}
&\quad \cdot\big(\mathbf{H}_{2,n}\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*(\mathbf{R}_{1,n}(a))^* + \mathbf{R}_{2,n}(a)\big[-(b-a)\mathbf{L}_{2,n}^*(I-a\mathbf{T}_{1,n})\\
&\quad +(I-a\mathbf{T}_{2,n})(\mathbf{L}_{1,n}^*-a\mathbf{L}_{2,n}^*)\big]\mathbf{H}_{1,n}\big)\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\big]\Big)B_1^{-1}\\
&= \frac{1}{b-a}\Big((z-a)(b-z)\big[I+(b-a)C_1-(b-a)C_2-a(b-a)s_0B_2\\
&\quad -(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad +(b-a)(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\left(\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*-\mathbf{T}_{1,n}\mathbf{H}_{1,n}\right)\\
&\qquad \cdot(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad +(b-a)(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{T}_{1,n}\mathbf{L}_{2,n}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&\quad -(b-a)^2(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}(\mathbf{R}_{2,n}(a))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&\quad +(b-a)(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(\mathbf{L}_{1,n}-a\mathbf{L}_{2,n})(I-a\mathbf{T}_{2,n}^*)(\mathbf{R}_{2,n}(a))^*\\
&\qquad \cdot\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\big]\\
&\quad -zs_0\big[(b-a)B_1-(z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad +(b-a)(z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)(\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*-\mathbf{T}_{1,n}\mathbf{H}_{1,n})\\
&\qquad\quad \cdot(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad +(b-a)(z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{T}_{1,n}\mathbf{L}_{2,n}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&\quad -(b-a)^2(z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}(\mathbf{R}_{2,n}(a))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&\quad +(b-a)(z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(\mathbf{L}_{1,n}-a\mathbf{L}_{2,n})(I-a\mathbf{T}_{2,n}^*)(\mathbf{R}_{2,n}(a))^*\\
&\qquad\quad \cdot\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\big]\\
&\quad -(b-z)\big[-as_0B_1+(z-a)\mathbf{u}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\\
&\qquad\quad \cdot\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad -(b-a)(z-a)\mathbf{u}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{L}_{2,n}^*(I-a\mathbf{T}_{1,n})\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad +(z-a)\mathbf{u}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)(I-a\mathbf{T}_{2,n})(\mathbf{L}_{1,n}^*-a\mathbf{L}_{2,n}^*)\\
&\qquad \cdot\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\big]\Big)B_1^{-1}.
\end{aligned}
$$

Thus from (2.2) it follows

$$
\triangle_{12,n}(z) = \frac{1}{b-a}\big(G_1(z)+G_2(z)\big)B_1^{-1} \tag{6.40}
$$

where

$$
\begin{aligned}
G_1(z) &:= (z-a)(b-z)\big[I+(b-a)C_1-(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad +(b-a)(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\left[\mathbf{H}_{1,n}^{-1}\mathbf{T}_{1,n}^*-\mathbf{T}_{1,n}\mathbf{H}_{1,n}\right]\\
&\qquad\quad \cdot(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad +(b-a)(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{T}_{1,n}\mathbf{L}_{2,n}\mathbf{R}_{2,n}(a)\mathbf{L}_{2,n}^*\mathbf{v}_{1,n}\big]\\
&\quad -zs_0\Big[(b-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\qquad -(z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad +(b-a)(z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\left[\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*-\mathbf{T}_{1,n}\mathbf{H}_{1,n}\right]\\
&\qquad\quad \cdot(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad +(b-a)(z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{T}_{1,n}\mathbf{L}_{2,n}\mathbf{R}_{2,n}(a)\mathbf{L}_{2,n}^*\mathbf{v}_{1,n}\Big]\\
&\quad -(b-z)\big[-as_0B_1
\end{aligned}
$$

$$+(z-a)\mathbf{u}^*_{2,n}(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{L}^*_{2,n}\mathbf{T}^*_{1,n}(\mathbf{R}_{1,n}(a))^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\big]$$

and

$$\begin{aligned}
&G_2(z) := (z-a)(b-z)\big[-(b-a)C_2 - a(b-a)s_0B_2\\
&-(b-a)^2(z-a)\mathbf{u}^*_{1,n}(\mathbf{R}_{1,n}(\overline{z}))^*(I-a\mathbf{T}^*_{1,n})\mathbf{L}_{2,n}(\mathbf{R}_{2,n}(a))^*\mathbf{H}^{-1}_{2,n}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&+(b-a)(z-a)\mathbf{u}^*_{1,n}(\mathbf{R}_{1,n}(\overline{z}))^*(\mathbf{L}_{1,n}-a\mathbf{L}_{2,n})(I-a\mathbf{T}^*_{2,n})(\mathbf{R}_{2,n}(a))^*\\
&\quad\cdot\mathbf{H}^{-1}_{2,n}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\big]\\
&-zs_0\big[-(b-a)^2(z-a)\mathbf{v}^*_{1,n}(\mathbf{R}_{1,n}(\overline{z}))^*(I-a\mathbf{T}^*_{1,n})\mathbf{L}_{2,n}(\mathbf{R}_{2,n}(a))^*\mathbf{H}^{-1}_{2,n}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&+(b-a)(z-a)\mathbf{v}^*_{1,n}(\mathbf{R}_{1,n}(\overline{z}))^*(\mathbf{L}_{1,n}-z\mathbf{L}_{2,n}+z\mathbf{L}_{2,n}-a\mathbf{L}_{2,n})(I-a\mathbf{T}^*_{2,n})\\
&\quad\cdot(\mathbf{R}_{2,n}(a))^*\mathbf{H}^{-1}_{2,n}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\big]\\
&-(b-z)\big[-(b-a)(z-a)\mathbf{u}^*_{2,n}(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}^{-1}_{2,n}\mathbf{R}_{2,n}(a)\mathbf{L}^*_{2,n}(I-a\mathbf{T}_{1,n})\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&+(z-a)\mathbf{u}^*_{2,n}(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}^{-1}_{2,n}\mathbf{R}_{2,n}(a)(I-a\mathbf{T}_{2,n})(\mathbf{L}^*_{1,n}-a\mathbf{L}^*_{2,n})\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\big].
\end{aligned}$$

From (2.19) we obtain

$$\begin{aligned}
G_1(z) &= (z-a)(b-z)\Big[\mathbf{v}^*_{1,n}(I-a\mathbf{T}_{1,n})\mathbf{H}_{1,n}\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad+(b-a)\mathbf{u}^*_{1,n}(\mathbf{R}_{1,n}(a))^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad-(z-a)\mathbf{u}^*_{1,n}(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad+(b-a)(z-a)\mathbf{u}^*_{1,n}(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\left[\mathbf{H}_{1,n}\mathbf{T}^*_{1,n}-\mathbf{T}_{1,n}\mathbf{H}_{1,n}\right]\\
&\qquad\cdot(\mathbf{R}_{1,n}(a))^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad+(b-a)(z-a)\mathbf{u}^*_{1,n}(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{T}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\Big]\\
&\quad-zs_0\Big[(b-a)\mathbf{v}^*_{1,n}(\mathbf{R}_{1,n}(a))^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\qquad-(z-a)\mathbf{v}^*_{1,n}(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad+(b-a)(z-a)\mathbf{v}^*_{1,n}(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\left[\mathbf{H}_{1,n}\mathbf{T}^*_{1,n}-\mathbf{T}_{1,n}\mathbf{H}_{1,n}\right]\\
&\qquad\cdot(\mathbf{R}_{1,n}(a))^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad+(b-a)(z-a)\mathbf{v}^*_{1,n}(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{T}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\Big]\\
&\quad-(b-z)\big[-as_0\mathbf{v}^*_{1,n}(\mathbf{R}_{1,n}(a))^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad+(z-a)\mathbf{u}^*_{2,n}(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{L}^*_{2,n}\mathbf{T}^*_{1,n}(\mathbf{R}_{1,n}(a))^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\big]\\
&= \Big((z-a)(b-z)\big[\mathbf{v}^*_{1,n}(I-a\mathbf{T}_{1,n})\mathbf{H}_{1,n}+(b-a)\mathbf{u}^*_{1,n}(\mathbf{R}_{1,n}(a))^*\\
&\quad-(z-a)\mathbf{u}^*_{1,n}(\mathbf{R}_{1,n}(\overline{z}))^*+(b-a)(z-a)\mathbf{u}^*_{1,n}(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\\
&\qquad\cdot\left[\mathbf{H}_{1,n}\mathbf{T}^*_{1,n}-\mathbf{T}_{1,n}\mathbf{H}_{1,n}\right](\mathbf{R}_{1,n}(a))^*\\
&\quad+(b-a)(z-a)\mathbf{u}^*_{1,n}(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{T}_{1,n}\mathbf{H}_{1,n}\big]\\
&\quad-zs_0\Big[(b-a)\mathbf{v}^*_{1,n}(\mathbf{R}_{1,n}(a))^*-(z-a)\mathbf{v}^*_{1,n}(\mathbf{R}_{1,n}(\overline{z}))^*\\
&\quad+(b-a)(z-a)\mathbf{v}^*_{1,n}(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\\
&\qquad\cdot\left[\mathbf{H}_{1,n}\mathbf{T}^*_{1,n}-\mathbf{T}_{1,n}\mathbf{H}_{1,n}\right](\mathbf{R}_{1,n}(a))^*\\
&\quad+(b-a)(z-a)\mathbf{v}^*_{1,n}(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}^{-1}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{T}_{1,n}\mathbf{H}_{1,n}\Big]\\
&\quad-(b-z)\big[-as_0\mathbf{v}^*_{1,n}(\mathbf{R}_{1,n}(a))^*
\end{aligned}$$

$$
\begin{aligned}
&+ (z-a)\mathbf{u}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\big]\Big)\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&= \Big((z-a)(b-z)\big[\mathbf{v}_{1,n}^*(I-a\mathbf{T}_{1,n})\mathbf{H}_{1,n}(I-a\mathbf{T}_{1,n}^*) + (b-a)\mathbf{u}_{1,n}^*\\
&\quad -(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(I-a\mathbf{T}_{1,n}^*) + (b-a)(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\\
&\qquad \cdot\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\cdot\big[\mathbf{H}_{1,n}\mathbf{T}_{1,n}^* - \mathbf{T}_{1,n}\mathbf{H}_{1,n}\big]\\
&\quad +(b-a)(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{T}_{1,n}\mathbf{H}_{1,n}(I-a\mathbf{T}_{1,n}^*)\big]\\
&\quad -zs_0\Big[(b-a)\mathbf{v}_{1,n}^* - (z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(I-a\mathbf{T}_{1,n}^*)\\
&\quad +(b-a)(z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\,\big[\mathbf{H}_{1,n}\mathbf{T}_{1,n}^* - \mathbf{T}_{1,n}\mathbf{H}_{1,n}\big]\\
&\quad +(b-a)(z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{T}_{1,n}\mathbf{H}_{1,n}(I-a\mathbf{T}_{1,n}^*)\big]\\
&\quad -(b-z)\big[-as_0\mathbf{v}_{1,n}^* + (z-a)\mathbf{u}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*\big]\Big)\\
&\qquad \cdot(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}.
\end{aligned}
$$

Then (2.20) provides us

$$
\begin{aligned}
G_1(z) &= \Big((z-a)(b-z)\big[\mathbf{v}_{1,n}^*(I-a\mathbf{T}_{1,n})\mathbf{H}_{1,n}(I-a\mathbf{T}_{1,n}^*) + (b-a)\mathbf{u}_{1,n}^*\\
&\quad -(z-a)\mathbf{u}_{1,n}^*(I-a\mathbf{T}_{1,n}^*)(\mathbf{R}_{1,n}(\overline{z}))^* + (b-a)(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\\
&\qquad \cdot\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*\\
&\quad -a(b-a)(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{T}_{1,n}\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*\big]\\
&\quad -zs_0\big[(b-a)\mathbf{v}_{1,n}^* - (z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(I-a\mathbf{T}_{1,n}^*)\\
&\quad +(b-a)(z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*\\
&\quad -a(b-a)(z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{T}_{1,n}\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*\big]\\
&\quad -(b-z)\big[-as_0\mathbf{v}_{1,n}^* + (z-a)\mathbf{u}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*\big]\Big)(\mathbf{R}_{1,n}(a))^*\\
&\qquad \cdot\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}.
\end{aligned}
$$

Thus because of (2.7), (2.18), and (2.5) we can conclude then

$$
\begin{aligned}
G_1(z) &= \Big((z-a)(b-z)\big[\mathbf{v}_{1,n}^*(I-a\mathbf{T}_{1,n})\mathbf{H}_{1,n}(I-a\mathbf{T}_{1,n}^*) + (b-a)\mathbf{u}_{1,n}^*\\
&\quad -(z-a)\mathbf{u}_{1,n}^*(I-a\mathbf{T}_{1,n}^*)(\mathbf{R}_{1,n}(\overline{z}))^* + (b-a)(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\\
&\qquad \cdot\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)(I-a\mathbf{T}_{1,n})\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*\big]\\
&\quad -zs_0\big[(b-a)\mathbf{v}_{1,n}^* - (z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(I-a\mathbf{T}_{1,n}^*)\\
&\quad +(b-a)(z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)(I-a\mathbf{T}_{1,n})\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*\big]\\
&\quad -(b-z)\big[-as_0\mathbf{v}_{1,n}^* + (z-a)\mathbf{u}_{2,n}^*\mathbf{L}_{2,n}^*\mathbf{L}_{2,n}(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*\big]\Big)\\
&\qquad \cdot(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&= \Big((z-a)(b-z)\big[\mathbf{v}_{1,n}^*(I-a\mathbf{T}_{1,n})\mathbf{H}_{1,n}(I-a\mathbf{T}_{1,n}^*) + (b-a)\mathbf{u}_{1,n}^*\\
&\quad -(z-a)\mathbf{u}_{1,n}^*(I-a\mathbf{T}_{1,n}^*)(\mathbf{R}_{1,n}(\overline{z}))^* + (b-a)(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{T}_{1,n}^*\big]\\
&\quad -zs_0\big[(b-a)\mathbf{v}_{1,n}^* - (z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(I-a\mathbf{T}_{1,n}^*)\\
&\qquad + (b-a)(z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{T}_{1,n}^*\big]\\
&\quad -(b-z)\big[-as_0\mathbf{v}_{1,n}^* + (z-a)\mathbf{u}_{2,n}^*\mathbf{L}_{2,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{T}_{1,n}^*\big]\Big)\\
&\qquad \cdot(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}. \qquad (6.41)
\end{aligned}
$$

Using (2.20) and (2.21) it follows

$$\begin{aligned}
G_1(z) &= \Big((z-a)(b-z)\big[\mathbf{v}_{1,n}^*(I-a\mathbf{T}_{1,n})\mathbf{H}_{1,n}(I-a\mathbf{T}_{1,n}^*)\\
&\quad+\mathbf{u}_{1,n}^*\big((b-a)(I-z\mathbf{T}_{1,n}^*)-(z-a)(I-a\mathbf{T}_{1,n}^*)+(b-a)(z-a)\mathbf{T}_{1,n}^*\big)\big](\mathbf{R}_{1,n}(\overline{z}))^*\\
&\quad-zs_0\mathbf{v}_{1,n}^*\big[(b-a)(I-z\mathbf{T}_{1,n}^*)-(z-a)(I-a\mathbf{T}_{1,n}^*)+(b-a)(z-a)\mathbf{T}_{1,n}^*\big](\mathbf{R}_{1,n}(\overline{z}))^*\\
&\quad-(b-z)\big(-as_0\mathbf{v}_{1,n}^*+(z-a)\mathbf{u}_{2,n}^*\mathbf{L}_{2,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{T}_{1,n}^*\big)\Big)(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&= \Big((z-a)(b-z)\big[\mathbf{v}_{1,n}^*(I-a\mathbf{T}_{1,n})\mathbf{H}_{1,n}(I-a\mathbf{T}_{1,n}^*)\\
&\quad+(b-z)\mathbf{u}_{1,n}^*(I-a\mathbf{T}_{1,n}^*)(\mathbf{R}_{1,n}(\overline{z}))^*\big]-zs_0(b-z)\mathbf{v}_{1,n}^*(I-a\mathbf{T}_{1,n})(\mathbf{R}_{1,n}(\overline{z}))^*\\
&\quad-(b-z)\big[-as_0\mathbf{v}_{1,n}^*+(z-a)\mathbf{u}_{2,n}^*\mathbf{L}_{2,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{T}_{1,n}^*\big]\Big)\\
&\qquad\cdot(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}.
\end{aligned}$$

By virtue of (2.20) we have then

$$\begin{aligned}
G_1(z) &= (b-z)\Big((z-a)\mathbf{v}_{1,n}^*(I-a\mathbf{T}_{1,n})\mathbf{H}_{1,n}(I-a\mathbf{T}_{1,n}^*)(I-z\mathbf{T}_{1,n}^*)\\
&\quad+(z-a)(b-z)\mathbf{u}_{1,n}^*(I-a\mathbf{T}_{1,n}^*)\\
&\quad-zs_0\mathbf{v}_{1,n}^*(I-a\mathbf{T}_{1,n}^*)+as_0\mathbf{v}_{1,n}^*(I-z\mathbf{T}_{1,n}^*)-(z-a)\mathbf{u}_{2,n}^*\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*\Big)\\
&\qquad\cdot(\mathbf{R}_{1,n}(\overline{z}))^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&= (z-a)(b-z)\big[\mathbf{v}_{1,n}^*(I-a\mathbf{T}_{1,n}^*)\mathbf{H}_{1,n}(I-a\mathbf{T}_{1,n}^*)(I-z\mathbf{T}_{1,n}^*)\\
&\quad+(b-z)\mathbf{u}_{1,n}^*(I-a\mathbf{T}_{1,n}^*)-s_0\mathbf{v}_{1,n}^*-\mathbf{u}_{2,n}^*\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*\big]\\
&\qquad\cdot(\mathbf{R}_{1,n}(\overline{z}))^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}.
\end{aligned}$$

Applying (2.10), the identity

$$\mathbf{T}_{1,n}^*\mathbf{v}_{1,n}=0,$$

and (2.13), we get then

$$\begin{aligned}
G_1(z) &= (z-a)(b-z)\big[\mathbf{v}_{1,n}^*\mathbf{H}_{1,n}(I-a\mathbf{T}_{1,n}^*)(I-z\mathbf{T}_{1,n}^*)\\
&\quad-(b-z)\mathbf{v}_{1,n}^*\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*(I-a\mathbf{T}_{1,n}^*)-s_0\mathbf{v}_{1,n}^*\\
&\quad-\big(-(a+b)\mathbf{v}_{1,n}^*\mathbf{H}_{1,n}\mathbf{L}_{2,n}+ab\mathbf{v}_{1,n}^*\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*\mathbf{L}_{2,n}+\mathbf{v}_{1,n}^*\mathbf{H}_{1,n}\mathbf{L}_{1,n}\big)\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*\big]\\
&\qquad\cdot(\mathbf{R}_{1,n}(\overline{z}))^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}.
\end{aligned}$$

Thus (2.5) implies

$$\begin{aligned}
G_1(z) &= (z-a)(b-z)\big[\mathbf{v}_{1,n}^*\mathbf{H}_{1,n}-a\mathbf{v}_{1,n}^*\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*-z\mathbf{v}_{1,n}^*\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*\\
&\quad+az\mathbf{v}_{1,n}^*\mathbf{H}_{1,n}(\mathbf{T}_{1,n}^*)^2\\
&\quad-b\mathbf{v}_{1,n}^*\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*+ab\mathbf{v}_{1,n}^*\mathbf{H}_{1,n}(\mathbf{T}_{1,n}^*)^2+z\mathbf{v}_{1,n}^*\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*\\
&\quad-az\mathbf{v}_{1,n}^*\mathbf{H}_{1,n}(\mathbf{T}_{1,n}^*)^2-s_0\mathbf{v}_{1,n}^*\\
&\quad+a\mathbf{v}_{1,n}^*\mathbf{H}_{1,n}\mathbf{L}_{2,n}\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*+b\mathbf{v}_{1,n}^*\mathbf{H}_{1,n}\mathbf{L}_{2,n}\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*\\
&\quad-ab\mathbf{v}_{1,n}^*\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*\mathbf{L}_{2,n}\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*-\mathbf{v}_{1,n}^*\mathbf{H}_{1,n}\mathbf{L}_{1,n}\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*\big]\\
&\qquad\cdot(\mathbf{R}_{1,n}(\overline{z}))^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&= (z-a)(b-z)\big[\mathbf{v}_{1,n}^*\mathbf{H}_{1,n}-s_0\mathbf{v}_{1,n}^*-\mathbf{v}_{1,n}^*\mathbf{H}_{1,n}\mathbf{L}_{1,n}\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*\big]\\
&\qquad\cdot(\mathbf{R}_{1,n}(\overline{z}))^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}.
\end{aligned}$$

Applying (2.6) and (6.39) it follows

$$G_1(z) = 0. \tag{6.42}$$

From (6.7), (2.17), and (2.2) we infer

$$\begin{aligned}
G_2(z) &= (z-a)(b-z)\big[-(b-a)C_2 - a(b-a)s_0B_2\\
&\quad -(b-a)^2(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}(\mathbf{R}_{2,n}(a))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&\quad +(b-a)(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(\mathbf{L}_{1,n}-a\mathbf{L}_{2,n})\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\big]\\
&\quad -zs_0\big[-(b-a)^2(z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}(\mathbf{R}_{2,n}(a))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&\quad +(b-a)(z-a)^2\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{L}_{2,n}(I-a\mathbf{T}_{2,n}^*)(\mathbf{R}_{2,n}(a))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\big]\\
&\quad -(b-z)\big[-(b-a)(z-a)\mathbf{u}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{L}_{2,n}^*\mathbf{v}_{1,n}\big]\\
&= (z-a)(b-a)\big[-(b-z)\mathbf{u}_{2,n}^*(\mathbf{R}_{2,n}(a))^* - a(b-z)s_0\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(a))^*\\
&\quad -(b-a)(z-a)(b-z)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}(\mathbf{R}_{2,n}(a))^*\\
&\quad +(z-a)(b-z)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(\mathbf{L}_{1,n}-a\mathbf{L}_{2,n})\\
&\quad +(b-a)zs_0\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}(\mathbf{R}_{2,n}(a))^*\\
&\quad -(z-a)zs_0\mathbf{v}_{1,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{L}_{2,n} + (b-z)\mathbf{u}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\big]\mathbf{H}_{2,n}^{-1}(\mathbf{R}_{2,n}(a))^*\mathbf{v}_{2,n}\\
&= (z-a)(b-a)\big[-(b-z)\mathbf{u}_{2,n}^*\left((\mathbf{R}_{2,n}(a))^*-(\mathbf{R}_{2,n}(\overline{z}))^*\right)-a(b-z)s_0\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(a))^*\\
&\quad -(b-a)(z-a)(b-z)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}(\mathbf{R}_{2,n}(a))^*\\
&\quad +(b-z)(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(\mathbf{L}_{1,n}-a\mathbf{L}_{2,n})\\
&\quad +zs_0\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\left((b-a)(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}(\mathbf{R}_{2,n}(a))^*-(z-a)\mathbf{L}_{2,n}\right)\big]\\
&\quad \cdot\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}.
\end{aligned}$$

Using (6.7) and (2.8) we obtain

$$\begin{aligned}
&(b-a)(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}(\mathbf{R}_{2,n}(a))^*-(z-a)\mathbf{L}_{2,n}\\
&\quad = \big[(b-z)\mathbf{L}_{2,n} - ab\mathbf{T}_{1,n}^*\mathbf{L}_{2,n} - a\mathbf{L}_{2,n} + a^2\mathbf{T}_{1,n}^*\mathbf{L}_{2,n} + az\mathbf{L}_{2,n}\mathbf{T}_{2,n}^*\\
&\qquad + a\mathbf{L}_{2,n} - a^2\mathbf{L}_{2,n}\mathbf{T}_{2,n}^*\big]\cdot(\mathbf{R}_{2,n}(a))^*\\
&\quad = (b-z)(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}(\mathbf{R}_{2,n}(a))^*.
\end{aligned}\tag{6.43}$$

From (2.22) (2.20), and (2.21) we get

$$(\mathbf{R}_{2,n}(\overline{z}))^* - (\mathbf{R}_{2,n}(a))^* = (z-a)(\mathbf{R}_{2,n}(a))^*\mathbf{T}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*. \tag{6.44}$$

Because of (6.44) and (6.43) we have then

$$\begin{aligned}
G_2(z) &= (z-a)(b-a)\big[(z-a)(b-z)\mathbf{u}_{2,n}^*(\mathbf{R}_{2,n}(a))^*\mathbf{T}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\\
&\quad -a(b-z)s_0\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(a))^*\\
&\quad -(b-a)(z-a)(b-z)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}(\mathbf{R}_{2,n}(a))^*\\
&\quad +(b-z)(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(\mathbf{L}_{1,n}-a\mathbf{L}_{2,n})\\
&\quad +(b-z)zs_0\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}(\mathbf{R}_{2,n}(a))^*\big]\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}.
\end{aligned}$$

By virtue of (2.7) and (2.2) it follows

$$\begin{aligned}
G_2(z) &= (z-a)(b-a)(b-z)\big[(z-a)\mathbf{u}_{2,n}^*(\mathbf{R}_{2,n}(a))^*\mathbf{T}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\\
&\quad -as_0\mathbf{v}_{2,n}^*\mathbf{L}_{2,n}^*\mathbf{L}_{2,n}(\mathbf{R}_{2,n}(a))^*\\
&\quad -(b-a)(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}(\mathbf{R}_{2,n}(a))^*\\
&\quad +(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(\mathbf{L}_{1,n}-a\mathbf{L}_{2,n})
\end{aligned}$$

$$
\begin{aligned}
&+zs_0\mathbf{v}_{2,n}^*\mathbf{L}_{2,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}(\mathbf{R}_{2,n}(a))^*\big]\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&=(z-a)(b-a)(b-z)\big[(z-a)\mathbf{u}_{2,n}^*(\mathbf{R}_{2,n}(a))^*\mathbf{T}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\\
&\quad+s_0\mathbf{v}_{2,n}^*\mathbf{L}_{2,n}^*\left[-aI+z(\mathbf{R}_{1,n}(\overline{z}))^*(I-a\mathbf{T}_{1,n}^*)\right]\mathbf{L}_{2,n}(\mathbf{R}_{2,n}(a))^*\\
&\quad+(z-a)\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\left[(\mathbf{L}_{1,n}-a\mathbf{L}_{2,n})(I-a\mathbf{T}_{2,n}^*)\right.\\
&\qquad\left.-(b-a)(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}\right](\mathbf{R}_{2,n}(a))^*\big]\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}.
\end{aligned}
$$

Thus

$$
-aI+z(\mathbf{R}_{1,n}(\overline{z}))^*(I-a\mathbf{T}_{1,n}^*)=(z-a)(\mathbf{R}_{1,n}(\overline{z}))^*, \tag{6.45}
$$

(2.20), (2.18), (2.21), (2.8), and (2.7) imply

$$
\begin{aligned}
G_2(z)&=(b-a)(z-a)^2(b-z)\big[\mathbf{u}_{2,n}^*\mathbf{T}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*+s_0\mathbf{v}_{2,n}^*\mathbf{L}_{2,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{L}_{2,n}\\
&\quad+\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\big[\mathbf{L}_{1,n}-a\mathbf{L}_{1,n}\mathbf{T}_{2,n}^*-a\mathbf{L}_{2,n}+a^2\mathbf{L}_{2,n}\mathbf{T}_{2,n}^*-b\mathbf{L}_{2,n}\\
&\quad+ab\mathbf{T}_{1,n}^*\mathbf{L}_{2,n}+a\mathbf{L}_{2,n}-a^2\mathbf{T}_{1,n}^*\mathbf{L}_{2,n}\big]\big](\mathbf{R}_{2,n}(a))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&=(b-a)(z-a)^2(b-z)\big[\mathbf{u}_{2,n}^*\mathbf{T}_{2,n}^*+s_0\mathbf{v}_{2,n}^*\\
&\quad+\mathbf{u}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\left(\mathbf{L}_{1,n}(I-a\mathbf{T}_{2,n}^*)-b(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}\right)(I-z\mathbf{T}_{2,n}^*)\big]\\
&\quad\cdot(\mathbf{R}_{2,n}(\overline{z}))^*(\mathbf{R}_{2,n}(a))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}.
\end{aligned}
$$

From (1.7) we obtain then

$$
\begin{aligned}
G_2(z)&=(b-a)(z-a)^2(b-z)\big[s_0\mathbf{v}_{2,n}^*+\mathbf{u}_{2,n}^*\mathbf{T}_{2,n}^*+\mathbf{u}_{1,n}^*\big[\mathbf{L}_{1,n}(I-a\mathbf{T}_{2,n}^*)\\
&\quad-b(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}\big]\\
&\quad+\mathbf{u}_{1,n}^*\left(\Sigma_{k=1}^n z^k(\mathbf{T}_{1,n}^*)^k\right)\left[\mathbf{L}_{1,n}(I-a\mathbf{T}_{2,n}^*)-b(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}\right]\\
&\quad-z\mathbf{u}_{1,n}^*\left(\Sigma_{k=0}^n z^k(\mathbf{T}_{1,n}^*)^k\right)\left[\mathbf{L}_{1,n}(I-a\mathbf{T}_{2,n}^*)-b(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}\right]\mathbf{T}_{2,n}^*\big]\\
&\quad\cdot(\mathbf{R}_{2,n}(\overline{z}))^*(\mathbf{R}_{2,n}(a))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}.
\end{aligned} \tag{6.46}
$$

By virtue of (2.8) and (2.9) we can conclude that

$$
\begin{aligned}
&\mathbf{T}_{1,n}^*\left[\mathbf{L}_{1,n}(I-a\mathbf{T}_{2,n}^*)-b(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}\right]\\
&\quad-\left[\mathbf{L}_{1,n}(I-a\mathbf{T}_{2,n}^*)-b(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}\right]\mathbf{T}_{2,n}^*\\
&=\mathbf{T}_{1,n}^*\mathbf{L}_{1,n}-a\mathbf{T}_{1,n}^*\mathbf{L}_{1,n}\mathbf{T}_{2,n}^*-b\mathbf{T}_{1,n}^*\mathbf{L}_{2,n}+ab(\mathbf{T}_{1,n}^*)^2\mathbf{L}_{2,n}\\
&\qquad-\mathbf{L}_{1,n}\mathbf{T}_{2,n}^*+a\mathbf{L}_{1,n}(\mathbf{T}_{2,n}^*)^2+b\mathbf{L}_{2,n}\mathbf{T}_{2,n}^*-ab\mathbf{T}_{1,n}^*\mathbf{L}_{2,n}\mathbf{T}_{2,n}^*\\
&=\mathbf{T}_{1,n}^*\mathbf{L}_{1,n}-\mathbf{L}_{1,n}\mathbf{T}_{2,n}^*-a\left(\mathbf{T}_{1,n}^*\mathbf{L}_{1,n}-\mathbf{L}_{1,n}\mathbf{T}_{2,n}^*\right)\mathbf{T}_{2,n}^*\\
&\qquad-b\left(\mathbf{T}_{1,n}^*\mathbf{L}_{2,n}-\mathbf{L}_{2,n}\mathbf{T}_{2,n}^*\right)+ab\mathbf{T}_{1,n}^*\left(\mathbf{T}_{1,n}^*\mathbf{L}_{2,n}-\mathbf{L}_{2,n}\mathbf{T}_{2,n}^*\right)\\
&=\mathbf{v}_{1,n}\mathbf{v}_{2,n}^*(I-a\mathbf{T}_{2,n}^*)
\end{aligned} \tag{6.47}
$$

is fulfilled. Using (6.46) and the equations (2.16) and (6.47) we get then

$$
\begin{aligned}
G_2(z)&=(b-a)(z-a)^2(b-z)\Big[\mathbf{u}_{1,n}^*\sum_{k=1}^n z^k(\mathbf{T}_{1,n}^*)^{k-1}\\
&\quad\cdot\big[\mathbf{T}_{1,n}^*\big(\mathbf{L}_{1,n}(I-a\mathbf{T}_{2,n}^*)-b(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}\big)\\
&\quad-\big(\mathbf{L}_{1,n}(I-a\mathbf{T}_{2,n}^*)-b(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}\big)\mathbf{T}_{2,n}^*\big]\\
&\quad-z^{n+1}\mathbf{u}_{1,n}^*(\mathbf{T}_{1,n}^*)^n\left[\mathbf{L}_{1,n}(I-a\mathbf{T}_{2,n}^*)-b(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}\right]\mathbf{T}_{2,n}^*\Big]\\
&\quad\cdot(\mathbf{R}_{2,n}(\overline{z}))^*(\mathbf{R}_{2,n}(a))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}
\end{aligned}
$$

$$= (b-a)(z-a)^2(b-z)\left(\mathbf{u}_{1,n}^*\sum_{k=1}^{n} z^k(\mathbf{T}_{1,n}^*)^{k-1}\right)\mathbf{v}_{1,n}\mathbf{v}_{2,n}^*(I-a\mathbf{T}_{2,n}^*)$$
$$\cdot(\mathbf{R}_{2,n}(\overline{z}))^*(\mathbf{R}_{2,n}(a))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n} = 0. \tag{6.48}$$

From (6.40), (6.42), and (6.48) equation $\triangle_{12,n}(z) = 0$ follows. Now we are going to verify $\triangle_{21,n}(z) = 0$. Observe that from (6.15) and (6.10) it follows

$$N_2 = -(b-a)^{-1}B_1. \tag{6.49}$$

Using (6.24), (6.30), (6.49), and (6.10) we obtain

$$\begin{aligned}
\triangle_{21,n}(z) &= (z-a)\big[-\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad+(b-a)(z-a)\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&\quad+(z-a)\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}(as_0B_1)+B_1\\
&\quad+(z-a)\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{u}_{2,n}B_1\big]\\
&= (z-a)\big[-\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad+(b-a)(z-a)\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&\quad+(z-a)\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\big[\mathbf{u}_{2,n}\mathbf{v}_{1,n}^*+a\mathbf{v}_{2,n}s_0\mathbf{v}_{1,n}^*\big]\\
&\qquad\cdot(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}+B_1\big].
\end{aligned}$$

Application of Proposition 2.5 yields then

$$\begin{aligned}
\triangle_{21,n}(z) &= (z-a)\big[-\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad+(b-a)(z-a)\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&\quad+(z-a)\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad-(b-a)(z-a)\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{L}_{2,n}^*(I-a\mathbf{T}_{1,n})\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad+(z-a)\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)(I-a\mathbf{T}_{2,n})(\mathbf{L}_{1,n}^*-a\mathbf{L}_{2,n}^*)\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}+B_1\big].
\end{aligned}$$

Thus (2.2), (2.17), and (2.19) imply

$$\begin{aligned}
\triangle_{21,n}(z) &= (z-a)\big[-\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad+(b-a)(z-a)\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&\quad+(z-a)\mathbf{v}_{1,n}^*\mathbf{L}_{2,n}(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad-(b-a)(z-a)\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{L}_{2,n}^*\mathbf{v}_{1,n}+B_1\big]\\
&= (z-a)\big[-\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad+(b-a)(z-a)\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&\quad+(z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{T}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad-(b-a)(z-a)\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}+B_1\big]\\
&= (z-a)\big[-\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\big[(I-a\mathbf{T}_{1,n}^*)-(z-a)\mathbf{T}_{1,n}^*\big]\\
&\qquad\cdot(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}+B_1\big].\\
&= (z-a)\big[-\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(I-z\mathbf{T}_{1,n}^*)(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}+B_1\big]\\
&= 0.
\end{aligned}$$

Thus, the equation $\triangle_{21,n}(z)=0$ is proved. It remains to check $\triangle_{22,n}(z)=0$. From (6.25), (6.31), Lemma 6.6, and (6.15) we get

$$\begin{aligned}
\triangle_{22,n}(z) &= \frac{1}{b-a}\Big(-(z-a)(b-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}M_1+bI-aI\\
&\quad+(z-a)(b-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{u}_{1,n}\\
&\quad+(b-z)(z-a)\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}M_2-bI+zI\\
&\quad-(z-a)(b-z)\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{u}_{2,n}\Big)\\
&= \frac{z-a}{b-a}\Big(-(b-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}M_1^*+I\\
&\quad+(b-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{u}_{1,n}\\
&\quad+(b-z)\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}(-as_0)\\
&\quad-(b-z)\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{u}_{2,n}\Big).
\end{aligned}$$

From (6.10) and (6.11) we can conclude then

$$\begin{aligned}
&\triangle_{22,n}(z)\\
&\quad= \frac{z-a}{b-a}\Big(-\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}(I+(b-a)[C_1-C_2-as_0B_2])\\
&\quad+B_1+(b-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{u}_{1,n}B_1\\
&\quad-(b-z)a\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}s_0B_1\\
&\quad-(b-z)\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{u}_{2,n}B_1\Big)B_1^{-1}\\
&= \frac{z-a}{b-a}\Big(\mathbf{v}_{1,n}^*\big[(\mathbf{R}_{1,n}(a))^*-(\mathbf{R}_{1,n}(\overline{z}))^*\big]\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad+(b-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)[\mathbf{u}_{1,n}\mathbf{v}_{1,n}^*-\mathbf{v}_{1,n}\mathbf{u}_{1,n}^*]\\
&\qquad\cdot(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad+(b-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)[\mathbf{v}_{1,n}\mathbf{u}_{2,n}^*+a\mathbf{v}_{1,n}s_0\mathbf{v}_{2,n}^*]\\
&\qquad\cdot(\mathbf{R}_{2,n}(a))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&\quad-(b-z)\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)[\mathbf{u}_{2,n}\mathbf{v}_{1,n}^*+a\mathbf{v}_{2,n}s_0\mathbf{v}_{1,n}^*]\\
&\qquad\cdot(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\Big)B_1^{-1}.
\end{aligned}$$

Thus, Proposition 2.1 and Proposition 2.5 yield

$$\begin{aligned}
\triangle_{22,n}(z) &= \frac{z-a}{b-a}\Big(\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(a))^*[I-z\mathbf{T}_{1,n}^*-(I-a\mathbf{T}_{1,n}^*)]\\
&\qquad\cdot(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad+(b-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)[\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*-\mathbf{T}_{1,n}\mathbf{H}_{1,n}]\\
&\qquad\cdot(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad+(b-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\big[\mathbf{R}_{1,n}(a)\mathbf{T}_{1,n}\mathbf{L}_{2,n}\mathbf{H}_{2,n}\\
&\quad+\mathbf{H}_{1,n}\big(-(b-a)(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}+(\mathbf{L}_{1,n}-a\mathbf{L}_{2,n})(I-a\mathbf{T}_{2,n}^*)\big)(\mathbf{R}_{2,n}(a))^*\big]\\
&\qquad\cdot\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}
\end{aligned}$$

$$
\begin{aligned}
&-(b-z)\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\big[\mathbf{H}_{2,n}\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\\
&+\mathbf{R}_{2,n}(a)[-(b-a)\mathbf{L}_{2,n}^*(I-a\mathbf{T}_{1,n})+(I-a\mathbf{T}_{2,n})(\mathbf{L}_{1,n}^*-a\mathbf{L}_{2,n}^*)]\mathbf{H}_{1,n}\big]\\
&\quad\cdot\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\Big)B_1^{-1}\\
=&\ \frac{z-a}{b-a}\,(G_3(z)+G_4(z))B_1^{-1}
\end{aligned}
\tag{6.50}
$$

where

$$
\begin{aligned}
G_3(z) :=&-(z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\mathbf{T}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&+(b-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)[\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*-\mathbf{T}_{1,n}\mathbf{H}_{1,n}]\\
&\quad\cdot(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&+(b-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{T}_{1,n}\mathbf{L}_{2,n}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&-(b-z)\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}
\end{aligned}
$$

and

$$
\begin{aligned}
G_4(z) :=&-(b-a)^2\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}(\mathbf{R}_{2,n}(a))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&+(b-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(\mathbf{L}_{1,n}-a\mathbf{L}_{2,n})(I-a\mathbf{T}_{2,n}^*)(\mathbf{R}_{2,n}(a))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&+(b-a)(b-z)\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{L}_{2,n}^*(I-a\mathbf{T}_{1,n})\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&-(b-z)\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)(I-a\mathbf{T}_{2,n})(\mathbf{L}_{1,n}^*-a\mathbf{L}_{2,n}^*)\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}.
\end{aligned}
\tag{6.51}
$$

According to (2.2) and (2.18) we have

$$
\begin{aligned}
G_3(z) =&-(z-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{T}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&+(b-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)[\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*-\mathbf{T}_{1,n}\mathbf{H}_{1,n}]\\
&\quad(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&+(b-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{T}_{1,n}\mathbf{L}_{2,n}\mathbf{L}_{2,n}^*\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&-(b-z)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{L}_{2,n}\mathbf{L}_{2,n}^*\mathbf{T}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}.
\end{aligned}
$$

Using (2.5), and (2.19) then we obtain

$$
\begin{aligned}
G_3(z) =&\ \mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\big[-(z-a)\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\\
&+(b-a)\mathbf{R}_{1,n}(a)(\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*-\mathbf{T}_{1,n}\mathbf{H}_{1,n})(\mathbf{R}_{1,n}(a))^*\\
&+(b-a)\mathbf{T}_{1,n}\mathbf{R}_{1,n}(a)\mathbf{H}_{1,n}-(b-z)\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*(\mathbf{R}_{1,n}(a))^*\big]\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
=&\ \mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\big[-(z-a)(I-a\mathbf{T}_{1,n})\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*\\
&+(b-a)(\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*-\mathbf{T}_{1,n}\mathbf{H}_{1,n})+(b-a)\mathbf{T}_{1,n}\mathbf{H}_{1,n}(I-a\mathbf{T}_{1,n}^*)\\
&-(b-z)(I-a\mathbf{T}_{1,n})\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*\big](\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
=&\ \mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\big[-z(I-a\mathbf{T}_{1,n})\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*+a\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*\\
&\quad-a^2\mathbf{T}_{1,n}\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*\\
&+b\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*-b\mathbf{T}_{1,n}\mathbf{H}_{1,n}-a\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*+a\mathbf{T}_{1,n}\mathbf{H}_{1,n}+b\mathbf{T}_{1,n}\mathbf{H}_{1,n}\\
&\quad-ab\mathbf{T}_{1,n}\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*\\
&-a\mathbf{T}_{1,n}\mathbf{H}_{1,n}+a^2\mathbf{T}_{1,n}\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*-b\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*+ab\mathbf{T}_{1,n}\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*\\
&+z(I-a\mathbf{T}_{1,n})\mathbf{H}_{1,n}\mathbf{T}_{1,n}^*\big](\mathbf{R}_{1,n}(a))^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
=&\ 0.
\end{aligned}
\tag{6.52}
$$

Applying (6.51), (2.17), (2.2), (2.18), and (2.21) we get

$$\begin{aligned}
&G_4(z)\\
&= (b-a)\big[-(b-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}(\mathbf{R}_{2,n}(a))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&\quad +\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*(\mathbf{L}_{1,n}-z\mathbf{L}_{2,n}+z\mathbf{L}_{2,n}-a\mathbf{L}_{2,n})\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&\quad +(b-z)\mathbf{v}_{2,n}^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{L}_{2,n}^*\mathbf{v}_{1,n}\\
&= (b-a)\mathbf{v}_{1,n}^*\big[-(b-a)(\mathbf{R}_{1,n}(\overline{z}))^*(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}(\mathbf{R}_{2,n}(a))^*\\
&\quad +(z-a)(\mathbf{R}_{1,n}(\overline{z}))^*\mathbf{L}_{2,n}+(b-z)\mathbf{L}_{2,n}(\mathbf{R}_{2,n}(\overline{z}))^*\big]\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&= (b-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\big[-(b-a)(I-a\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}(I-z\mathbf{T}_{2,n}^*)\\
&\quad +(z-a)\mathbf{L}_{2,n}(I-a\mathbf{T}_{2,n}^*)(I-z\mathbf{T}_{2,n}^*)+(b-z)(I-z\mathbf{T}_{1,n}^*)\mathbf{L}_{2,n}(I-a\mathbf{T}_{2,n}^*)\big]\\
&\qquad \cdot(\mathbf{R}_{2,n}(a))^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}.
\end{aligned}$$

Thus (2.8) yields

$$\begin{aligned}
G_4(z) &= (b-a)\mathbf{v}_{1,n}^*(\mathbf{R}_{1,n}(\overline{z}))^*\big[\big(-(b-a)+(z-a)+(b-z)\big)\\
&\quad \cdot(\mathbf{L}_{2,n}-a\mathbf{T}_{1,n}^*\mathbf{L}_{2,n}-z\mathbf{T}_{1,n}^*\mathbf{L}_{2,n}+az\mathbf{T}_{1,n}^*\mathbf{L}_{2,n}\mathbf{T}_{2,n}^*)\big]\\
&\qquad \cdot(\mathbf{R}_{2,n}(a))^*(\mathbf{R}_{2,n}(\overline{z}))^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&\qquad = 0.
\end{aligned} \tag{6.53}$$

From (6.50), (6.52), and (6.53) it follows $\triangle_{22,n}(z)=0$. The proof is complete. □

In the following, for each $k\in\{1,2\}$, let $\tilde{U}_{k,n}$ be the $2q\times 2q$ matrix polynomial defined by (6.2), let A_k be the matrix given by (6.16), and $U_{k,n}:=\tilde{U}_{k,n}A_k$. Further, for every choice of $z\in\mathbb{C}$, let the $q\times q$ block partitions of $U_{1,n}(z)$ and $U_{2,n}(z)$ be given by (6.20).

Remark 6.11. *Let $(s_j)_{j=0}^{2n}$ be a sequence of complex $q\times q$ matrices such that the matrices $H_{1,n}$ and $H_{2,n}$ are both positive Hermitian. Let $U_{11,n}:\mathbb{C}\setminus[a,b]\to\mathbb{C}^{q\times q}$, $U_{12,n}:\mathbb{C}\setminus[a,b]\to\mathbb{C}^{q\times q}$, $U_{21,n}:\mathbb{C}\setminus[a,b]\to\mathbb{C}^{q\times q}$, and $U_{22,n}:\mathbb{C}\setminus[a,b]\to\mathbb{C}^{q\times q}$ be defined by*

$$U_{11,n}(z):=I-(z-a)\mathbf{u}_{1,n}^*[\mathbf{R}_{1,n}(\overline{z})]^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}, \tag{6.54}$$

$$\begin{aligned}
U_{12,n}(z) &:= \big(I-(z-a)\mathbf{u}_{1,n}^*[\mathbf{R}_{1,n}(\overline{z})]^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\big)\\
&\quad \cdot\big(I+(b-a)\big[\mathbf{u}_{1,n}^*[\mathbf{R}_{1,n}(a)]^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\qquad -\mathbf{u}_{2,n}^*[\mathbf{R}_{2,n}(a)]^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\\
&\qquad -as_0\mathbf{v}_{2,n}^*[\mathbf{R}_{2,n}(a)]^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}\big]\big)\\
&\quad \cdot\big((b-a)\mathbf{v}_{1,n}^*[\mathbf{R}_{1,n}(a)]^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\big)^{-1}\\
&\quad +(z-a)\mathbf{u}_{1,n}^*[\mathbf{R}_{1,n}(\overline{z})]^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{u}_{1,n},
\end{aligned} \tag{6.55}$$

$$U_{21,n}(z):=-(z-a)\mathbf{v}_{1,n}^*[\mathbf{R}_{1,n}(\overline{z})]^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}, \tag{6.56}$$

and

$$\begin{aligned}
U_{22,n}(z) &:= -(z-a)\mathbf{v}_{1,n}^*[\mathbf{R}_{1,n}(\overline{z})]^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\quad \cdot\big(I+(b-a)\big[\mathbf{u}_{1,n}^*[\mathbf{R}_{1,n}(a)]^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\\
&\qquad -\mathbf{u}_{2,n}^*[\mathbf{R}_{2,n}(a)]^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}
\end{aligned}$$

$$
\begin{aligned}
&-as_0\mathbf{v}_{2,n}^*[\mathbf{R}_{2,n}(a)]^*\mathbf{H}_{2,n}^{-1}\mathbf{R}_{2,n}(a)\mathbf{v}_{2,n}]\big) \\
&\cdot\big((b-a)\mathbf{v}_{1,n}^*[\mathbf{R}_{1,n}(a)]^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{v}_{1,n}\big)^{-1} \\
&+I+(z-a)\mathbf{v}_{1,n}^*[\mathbf{R}_{1,n}(\overline{z})]^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(a)\mathbf{u}_{1,n}.
\end{aligned}
\tag{6.57}
$$

For each $j,k\in\{1,2\}$ *and each* $z\in\mathbb{C}\setminus[a,b]$*, then by virtue of Lemma* 6.6 *it is readily checked that* $U_{jk,n}(z)=U_{jk,n}^{(1)}(z)$*.*

The four $q\times q$ matrix-valued functions introduced in Remark 6.11 will be used as the $q\times q$ blocks of the generating $2q\times 2q$ matrix function of our linear fractional transformation. In the next step we will show that this linear fractional transformation is well defined.

Lemma 6.12. *Let* $(s_j)_{j=0}^{2n}$ *be a sequence of complex* $q\times q$ *matrices such that the matrices* $\mathbf{H}_{1,n}$ *and* $\mathbf{H}_{2,n}$ *are both positive Hermitian. Let* $U_{11,n}$*,* $U_{12,n}$*,* $U_{21,n}$*, and* $U_{22,n}$ *be the matrix-valued functions which are defined on* $\mathbb{C}\setminus[a,b]$ *and which are given by* (6.54)*,* (6.55)*,* (6.56)*, and* (6.57)*. Let* $\binom{P}{Q}\in\mathfrak{P}(-\tilde{J}_q,[a,b])$*, let* $P_1:=U_{11,n}P+U_{12,n}Q$*, and let* $Q_1:=U_{21,n}P+U_{22,n}Q$*. Then* $\det P_1$ *and* $\det Q_1$ *are complex-valued functions which are meromorphic in* $\mathbb{C}\setminus[a,b]$ *and which do not vanish identically.*

Proof. First we observe that we know from Remark 6.11 that $U_n:=(U_{jk,n})_{j,k=1}^2$ is the restriction of $U_{1,n}$ onto $\mathbb{C}\setminus[a,b]$. According to Definition 5.2, there is a discrete subset $\mathcal{D}$ of $\mathbb{C}\setminus[a,b]$ such that the conditions (i), (ii), (iii), and (iv) of Definition 5.2 are fulfilled. Because of (i), P_1 and Q_1 are meromorphic in $\mathbb{C}\setminus[a,b]$ and holomorphic in $\mathbb{C}\setminus([a,b]\cup\mathcal{D})$. Moreover, we have $\binom{P_1(z)}{Q_1(z)}=U_1(z)\binom{P(z)}{Q(z)}$ for each $z\in\mathbb{C}\setminus([a,b]\cup\mathcal{D})$. From Lemma 6.7 we know that the matrices $U_{1,n}(z)$ and $U_{2,n}(z)$ are nonsingular for every choice of z in $\mathbb{C}$. Hence

$$
\begin{pmatrix} P(z)\\ Q(z)\end{pmatrix}=(U_{1,n}(z))^{-1}\begin{pmatrix} P_1(z)\\ Q_1(z)\end{pmatrix}
\tag{6.58}
$$

holds for each $z\in\mathbb{C}\setminus([a,b]\cup\mathcal{D})$. Furthermore, (ii) yields then

$$
\operatorname{rank}\begin{pmatrix} P_1(z)\\ Q_1(z)\end{pmatrix}=\operatorname{rank}\begin{pmatrix} P(z)\\ Q(z)\end{pmatrix}=q
\tag{6.59}
$$

for each $z\in\mathbb{C}\setminus([a,b]\cup\mathcal{D})$. Applying Proposition 6.3, Lemma 6.7, and [DFK, Theorem 1.3.3], we get that for each $k\in\{1,2\}$ the matrix polynomial $U_{k,n}$ is a $\tilde{J}_q$-inner function belonging to the Potapov class $\mathfrak{P}_{J_q}(\Pi_+)$. Consequently, Lemma 5.1 in [CDFK] yields

$$
\frac{U_{k,n}^*(z)\tilde{J}_qU_{k,n}(z)}{i(z-\overline{z})}\geq\frac{\tilde{J}_q}{i(z-\overline{z})}
$$

for each $z\in\mathbb{C}\setminus\mathbb{R}$. Now let $z\in\mathbb{C}\setminus(\mathbb{R}\cup\mathcal{D})$. From (iii) and (6.58) we can conclude

$$
\frac{1}{i(z-\overline{z})}\begin{pmatrix} P_1(z)\\ Q_1(z)\end{pmatrix}^*[U_{1,n}(z)]^{-*}\tilde{J}_q[U_{1,n}(z)]^{-1}\begin{pmatrix} P_1(z)\\ Q_1(z)\end{pmatrix}\geq 0.
$$

For each $g \in \mathcal{N}[P_1(z)] := \{h \in \mathbb{C}^q : P_1(z)h = 0\}$ this implies

$$\frac{1}{i(z-\overline{z})}\begin{pmatrix} 0 \\ Q_1(z)g \end{pmatrix}^* [U_{1,n}(z)]^{-*}\tilde{J}_q[U_{1,n}(z)]^{-1}\begin{pmatrix} 0 \\ Q_1(z)g \end{pmatrix} \geq 0.$$

Since

$$\frac{1}{i(z-\overline{z})}\begin{pmatrix} 0 \\ Q_1(z)g \end{pmatrix}^* \tilde{J}_q \begin{pmatrix} 0 \\ Q_1(z)g \end{pmatrix} = 0$$

is valid for each $g \in \mathbb{C}^q$, we get then

$$\frac{1}{i(z-\overline{z})}\begin{pmatrix} 0 \\ Q_1(z)g \end{pmatrix}^* \left(\tilde{J}_q - [U_{1,n}(z)]^{-*}\tilde{J}_q[U_{1,n}(z)]^{-1}\right)\begin{pmatrix} 0 \\ Q_1(z)g \end{pmatrix} \leq 0. \qquad (6.60)$$

for each $g \in \mathcal{N}[P_1(z)]$. On the other hand, Proposition 6.3 and Lemma 6.7 provide us

$$\begin{aligned}&\frac{1}{i(z-\overline{z})}\left(\tilde{J}_q - [U_{1,n}(z)]^{-*}\tilde{J}_q[U_{1,n}(z)]^{-1}\right) \\ &\qquad = \tilde{J}_q(\mathbf{u}_{1,n}, \mathbf{v}_{1,n})^*[\mathbf{R}_{1,n}(z)]^*\mathbf{H}_{1,n}^{-1}\mathbf{R}_{1,n}(z)(\mathbf{u}_{1,n}, \mathbf{v}_{1,n})\tilde{J}_q. \qquad (6.61)\end{aligned}$$

Since the matrix $\mathbf{H}_{1,n}$ is positive Hermitian the right-hand side of (6.61) is nonnegative Hermitian. Thus from (6.60) it follows

$$\begin{pmatrix} 0 \\ Q_1(z)g \end{pmatrix}^* \left(\tilde{J}_q - [U_{1,n}(z)]^{-*}\tilde{J}_q[U_{1,n}(z)]^{-1}\right)\begin{pmatrix} 0 \\ Q_1(z)g \end{pmatrix} = 0$$

for all $g \in \mathcal{N}[P_1(z)]$. Using (6.61) again and the nonsingularity of the matrix $\mathbf{R}_{1,n}(z)$ we get then

$$0 = (\mathbf{u}_{1,n}, \mathbf{v}_{1,n})\tilde{J}_q\begin{pmatrix} 0 \\ Q_1(z)g \end{pmatrix} = -i\mathbf{u}_{1,n}Q_1(z)g.$$

Because of (1.11) and (1.12) this implies $s_0Q_1(z)g = 0$ for each $g \in \mathcal{N}[P_1(z)]$. Since the matrix $\mathbf{H}_{1,n}$ is positive Hermitian the matrix s_0 is nonsingular. Thus,

$$\begin{pmatrix} P_1(z) \\ Q_1(z) \end{pmatrix} g = 0$$

holds for all $g \in \mathcal{N}[P_1(z)]$. Hence (6.59) shows that $\mathcal{N}[P_1(z)] = \{0\}$ is valid, i.e., the matrix $P_1(z)$ is nonsingular. Analogously one can check that $Q_1(z)$ is nonsingular as well. □

Lemma 6.13. *Let φ be a $q \times q$ matrix-valued function which is meromorphic in $\mathbb{C} \setminus [a,b]$ and which fulfills $\operatorname{Im}\varphi(z) \geq 0$ for all $z \in \Pi_+ \cap \mathbb{H}_\varphi$. Let s_0 be a Hermitian complex $q \times q$ matrix. Suppose that the matrix-valued function $\psi : \mathbb{H}_\varphi \to \mathbb{C}^{q\times q}$ defined by $\psi(w) := (b-w)(w-a)\varphi(w) - s_0w$ satisfies $\operatorname{Im}\psi(z) \geq 0$ for all $z \in \Pi_+ \cap \mathbb{H}_\varphi$. Then:*

(a) *For each $x \in (-\infty, a) \cap \mathbb{H}_\varphi$, the matrix $\varphi(x)$ is nonnegative Hermitian.*
(b) *For each $x \in (b, +\infty) \cap \mathbb{H}_\varphi$, the matrix $-\varphi(x)$ is nonnegative Hermitian.*

Proof. Let $x \in (\mathbb{R} \setminus [a,b]) \cap \mathbb{H}_\varphi$. Then we get

$$\operatorname{Im} \varphi(x) = \lim_{\varepsilon \to 0+0} \operatorname{Im} \varphi(x + i\varepsilon) \geq 0 \tag{6.62}$$

and

$$(b - x)(x - a) \operatorname{Im} \varphi(x) = \operatorname{Im} \psi(x) = \lim_{\varepsilon \to 0+0} \operatorname{Im} \psi(x + i\varepsilon) \geq 0. \tag{6.63}$$

Since $(b - x)(x - a) < 0$ holds, from (6.62) and (6.63) we obtain $\operatorname{Im} \varphi(x) = 0$. Further, for each $\varepsilon \in (0, +\infty)$ such that $x + i\varepsilon$ belongs to $\mathbb{H}_\varphi$ we have

$$\begin{aligned} &0 \leq \operatorname{Im} \psi(x + i\varepsilon) \\ &= (b - x)(x - a) \operatorname{Im} \varphi(x + i\varepsilon) + \varepsilon(b + a - 2x) \operatorname{Re} \varphi(x + i\varepsilon) + \varepsilon^2 \operatorname{Im} \varphi(x + i\varepsilon) \\ &\leq \varepsilon(b + a - 2x) \operatorname{Re} \varphi(x + i\varepsilon) + \varepsilon^2 \operatorname{Im} \varphi(x + i\varepsilon) \end{aligned}$$

and consequently

$$0 \leq (b + a - 2x) \operatorname{Re} \varphi(x + i\varepsilon) + \varepsilon \operatorname{Im} \varphi(x + i\varepsilon).$$

Letting $\varepsilon \to 0 + 0$ it follows

$$0 \leq (b + a - 2x) \operatorname{Re} \varphi(x) = (b + a - 2x)\varphi(x).$$

Thus $\varphi(x) \geq 0$ if $x < a$ and $-\varphi(x) \geq 0$ if $x > b$. $\square$

The following description of the set $\mathcal{R}_q\left[[a,b],(s_j)_{j=0}^{2n}\right]$ in the positive definite case is one of the central results of the paper.

Theorem 6.14. *Let $(s_j)_{j=0}^{2n}$ be a sequence of complex $q \times q$ matrices such that the matrices $\mathbf{H}_{1,n}$ and $\mathbf{H}_{2,n}$ are both positive Hermitian. Let $U_{11,n}$, $U_{12,n}$, $U_{21,n}$, and $U_{22,n}$ be the matrix-valued functions which are defined on $\mathbb{C} \setminus [a,b]$ and which are given by* (6.54), (6.55), (6.56), *and* (6.57). *Then:*

(a) *For each $\binom{P}{Q} \in \mathfrak{P}(-\tilde{J}_q, [a,b])$ the matrix-valued function*

$$S := (U_{11,n}P + U_{12,n}Q)(U_{21,n}P + U_{22,n}Q)^{-1}$$

belongs to $\mathcal{R}_q\left[[a,b];(s_j)_{j=0}^{2n}\right]$.

(b) *For each $S \in \mathcal{R}_q\left[[a,b];(s_j)_{j=0}^{2n}\right]$, there is a column pair $\binom{P}{Q} \in \mathfrak{P}(-\tilde{J}_q, [a,b])$ of matrix-valued functions P and Q which are holomorphic in $\mathbb{C} \setminus [a,b]$ such that S admits the representation*

$$S = (U_{11,n}P + U_{12,n}Q)(U_{21,n}P + U_{22,n}Q)^{-1}. \tag{6.64}$$

(c) *If $\binom{P_1}{Q_1}$ and $\binom{P_2}{Q_2}$ belong to $\mathfrak{P}(-\tilde{J}_q, [a,b])$, then*

$$\begin{aligned} &(U_{11,n}P_1 + U_{12,n}Q_1)(U_{21,n}P_1 + U_{22,n}Q_1)^{-1} \\ &\quad = (U_{11,n}P_2 + U_{12,n}Q_2)(U_{21,n}P_2 + U_{22,n}Q_2)^{-1}. \end{aligned} \tag{6.65}$$

if and only if

$$\left\langle \begin{pmatrix} P_1 \\ Q_1 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} P_2 \\ Q_2 \end{pmatrix} \right\rangle. \tag{6.66}$$

Proof. Let $U_n := (U_{jk,n})_{j,k=0}^2$. From Lemma 6.7, (6.20), and Remark 6.11 we get that for each $z \in \mathbb{C} \setminus [a,b]$ the matrix $U_n(z)$ coincides with the nonsingular matrix $U_{1,n}(z)$.

(a) Let $\binom{P}{Q} \in \mathfrak{P}(-\tilde{J}_q, [a,b])$. Then there is a discrete subset $\mathcal{D}$ of $\mathbb{C} \setminus [a,b]$ such that the conditions (i), (ii), (iii), and (iv) of Definition 5.2 are fulfilled. Define $P_1 := U_{11,n}P + U_{12,n}Q$ and $Q_1 := U_{21,n}P + U_{22,n}Q$. From Lemma 6.12 we see that $\det P_1$ and $\det Q_1$ are complex-valued functions which are meromorphic in $\mathbb{C} \setminus [a,b]$ and which do not vanish identically. In particular, $\mathcal{D}_1 := \{w \in \mathbb{C} \setminus ([a,b] \cup \mathcal{D}) : \det Q_1(w) = 0\}$ is a discrete subset of $\mathbb{C} \setminus [a,b]$. Consequently, $\tilde{\mathcal{D}} := \mathcal{D} \cup \mathcal{D}_1$ is a discrete subset of $\mathbb{C} \setminus [a,b]$. Obviously, $S = P_1 Q_1^{-1}$. Let

$$\tilde{P} := PQ_1^{-1} \quad \text{and} \quad \tilde{Q} := QQ_1^{-1}.$$

The matrix-valued functions $S, \tilde{P}$, and $\tilde{Q}$ are holomorphic in $\mathbb{C} \setminus ([a,b] \cup \tilde{\mathcal{D}})$. We have

$$\begin{pmatrix} \tilde{P} \\ \tilde{Q} \end{pmatrix} = U_n^{-1} \begin{pmatrix} S \\ I \end{pmatrix} = U_{1,n}^{-1} \begin{pmatrix} S \\ I \end{pmatrix}.$$

Since $\binom{P}{Q}$ belongs to $\mathfrak{P}(-\tilde{J}_q, [a,b])$, we get

$$\begin{pmatrix} \tilde{P} \\ \tilde{Q} \end{pmatrix} \in \mathfrak{P}(-\tilde{J}_q, [a,b]). \tag{6.67}$$

In particular, $\tilde{P}$ and $\tilde{Q}$ are matrix-valued functions which are holomorphic in $\mathbb{C} \setminus ([a,b] \cup \tilde{\mathcal{D}})$, the condition

$$\operatorname{rank} \begin{pmatrix} \tilde{P}(z) \\ \tilde{Q}(z) \end{pmatrix} = q$$

is fulfilled for each $z \in \mathbb{C} \setminus ([a,b] \cup \tilde{\mathcal{D}})$, and the inequalities

$$\frac{1}{2 \operatorname{Im} z} \begin{pmatrix} \tilde{P}(z) \\ \tilde{Q}(z) \end{pmatrix}^* (-\tilde{J}_q) \begin{pmatrix} \tilde{P}(z) \\ \tilde{Q}(z) \end{pmatrix} \geq 0 \tag{6.68}$$

and

$$\frac{1}{2 \operatorname{Im} z} \begin{pmatrix} (z-a)\tilde{P}(z) \\ (b-z)\tilde{Q}(z) \end{pmatrix}^* (-\tilde{J}_q) \begin{pmatrix} (z-a)\tilde{P}(z) \\ (b-z)\tilde{Q}(z) \end{pmatrix} \geq 0 \tag{6.69}$$

hold for each $z \in \mathbb{C} \setminus (\mathbb{R} \cup \tilde{\mathcal{D}})$. Let $S_1 := S$ and let $S_2 : \mathbb{C} \setminus ([a,b] \cup \tilde{\mathcal{D}}) \to \mathbb{C}^{q \times q}$ be defined by (1.14). We consider an arbitrary $z \in \mathbb{C} \setminus (\mathbb{R} \cup \tilde{\mathcal{D}})$. From (6.68) we obtain then that (6.19) holds if $k = 1$. Now we are going to check that (6.19) is also true for $k = 2$. Lemma 6.7 yields that the matrix $U_{2,n}(z)$ is nonsingular.

From Proposition 6.10 we obtain then

$$[U_{2,n}(z)]^{-1} = \begin{pmatrix} (b-a)(z-a)I & 0 \\ 0 & \frac{b-z}{b-a}I \end{pmatrix} [U_{1,n}(z)]^{-1} \begin{pmatrix} \frac{1}{(z-a)(b-z)}I & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} I & s_0 z \\ 0 & I \end{pmatrix} \tag{6.70}$$

Using the identity

$$\begin{pmatrix} S_2(z) \\ I \end{pmatrix} = \begin{pmatrix} I & -s_0 z \\ 0 & I \end{pmatrix} \begin{pmatrix} (z-a)(b-z)I & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} S(z) \\ I \end{pmatrix},$$

and (6.70) we get

$$\begin{aligned} [U_{2,n}(z)]^{-1} \begin{pmatrix} S_2(z) \\ I \end{pmatrix} &= \begin{pmatrix} (b-a)(z-a)I & 0 \\ 0 & \frac{b-z}{b-a}I \end{pmatrix} [U_{1,n}(z)]^{-1} \begin{pmatrix} S(z) \\ I \end{pmatrix} \\ &= \begin{pmatrix} (b-a)I & 0 \\ 0 & \frac{1}{b-a}I \end{pmatrix} \begin{pmatrix} (z-a)\tilde{P}(z) \\ (b-z)\tilde{Q}(z) \end{pmatrix} \end{aligned} \tag{6.71}$$

and consequently

$$\begin{aligned} &\frac{1}{i(z-\overline{z})} \begin{pmatrix} S_2(z) \\ I \end{pmatrix}^* [U_{2,n}(z)]^{-*} \tilde{J}_q [U_{2,n}(z)]^{-1} \begin{pmatrix} S_2(z) \\ I \end{pmatrix} \\ &\quad = \frac{1}{i(z-\overline{z})} \begin{pmatrix} (z-a)\tilde{P}(z) \\ (b-z)\tilde{Q}(z) \end{pmatrix}^* \tilde{J}_q \begin{pmatrix} (z-a)\tilde{P}(z) \\ (b-z)\tilde{Q}(z) \end{pmatrix}. \end{aligned} \tag{6.72}$$

By virtue of (6.69) the right-hand side of (6.72) is nonnegative Hermitian. Thus (6.19) also holds in the case $k = 2$. Since (6.19) is valid from Lemma 6.4 and Lemma 6.7 we can conclude that for each $k \in \{1, 2\}$ the matrix $C_{k,n}^{[S]}(z)$ given by (6.1) is nonnegative Hermitian. This implies

$$\frac{S_k(z) - S_k^*(z)}{z - \overline{z}} \geq 0 \tag{6.73}$$

for each $k \in \{1, 2\}$. In particular, $\operatorname{Im} S(w) \geq 0$ for each $w \in \Pi_+ \setminus \mathcal{D}$. Thus one can easily see that S is holomorphic in Π_+ (compare, e.g., [DFK, Lemma 2.1.9]). Inequality (6.73) implies then $\operatorname{Im} S_2(w) \geq 0$ for each $w \in \Pi_+$. Application of Lemma 6.13 yields then that for each $x \in (-\infty, a) \setminus \mathcal{D}$ the matrix $S(x)$ is nonnegative Hermitian and that for each $x \in (b, +\infty) \setminus \mathcal{D}$ the matrix $-S(x)$ is nonnegative Hermitian. Since $\mathcal{D} \cap \mathbb{R}$ is a discrete subset of $\mathbb{R}$, the symmetry principle provides us that S is holomorphic in Π_- and that $S(z) = S^*(\overline{z})$ holds for all $z \in \Pi_-$. In order to check that S is also holomorphic at each point which belongs to $\mathbb{R} \setminus [a, b]$ one can argue in the same way as in the proof of [CDFK, Proposition 5.7]. Thus the matrix-valued function S is holomorphic in $\mathbb{C} \setminus [a, b]$. We obtain then that inequality (6.19) holds for each $k \in \{1, 2\}$ and each $z \in \mathbb{C} \setminus \mathbb{R}$. Application of Proposition 6.8 yields then that S belongs to $\mathcal{R}_q\left[[a, b], (s_j)_{j=0}^{2n}\right]$.

(b) Now we consider an arbitrary matrix-valued function S which belongs to $\mathcal{R}_q\left[[a, b], (s_j)_{j=0}^{2n}\right]$. Let

$$P := (I_q, 0) U_n^{-1} \begin{pmatrix} S \\ I \end{pmatrix} \quad \text{and} \quad Q := (0, I_q) U_n^{-1} \begin{pmatrix} S \\ I \end{pmatrix}. \tag{6.74}$$

From Lemma 6.7 and Remark 6.11 we see that the matrix-valued function U_n^{-1} is holomorphic in $\mathbb{C} \setminus [a,b]$. Hence P and Q are also holomorphic in $\mathbb{C} \setminus [a,b]$ and we obtain

$$\operatorname{rank}\begin{pmatrix} P(z) \\ Q(z) \end{pmatrix} = \operatorname{rank}\begin{pmatrix} S(z) \\ I \end{pmatrix} = q$$

for each $z \in \mathbb{C} \setminus [a,b]$. Let $S_1 := S$ and let $S_2 : \mathbb{C} \setminus [a,b] \to \mathbb{C}^{q\times q}$ be defined by (1.14). From Proposition 6.8 we obtain then that inequality (6.19) holds for each $k \in \{1,2\}$ and each $z \in \mathbb{C} \setminus \mathbb{R}$. Because of (6.74) for each $z \in \mathbb{C} \setminus [a,b]$ we have

$$\begin{pmatrix} P(z) \\ Q(z) \end{pmatrix} = [U_{1,n}(z)]^{-1} \begin{pmatrix} S(z) \\ I \end{pmatrix}. \tag{6.75}$$

From Proposition 6.10 we see that (6.70) holds for every choice of z in $\mathbb{C} \setminus [a,b]$. Using (6.75), (1.14), and (6.70), for each $z \in \mathbb{C} \setminus [a,b]$, we get

$$\begin{aligned}
\begin{pmatrix} (z-a)P(z) \\ (b-z)Q(z) \end{pmatrix} &= \begin{pmatrix} (z-a)I & 0 \\ 0 & (b-z)I \end{pmatrix} [U_{1,n}(z)]^{-1} \begin{pmatrix} S(z) \\ I \end{pmatrix} \\
&= \begin{pmatrix} (z-a)I & 0 \\ 0 & (b-z)I \end{pmatrix} [U_{1,n}(z)]^{-1} \begin{pmatrix} \frac{1}{(b-a)(b-z)}I & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} I & s_0 z \\ 0 & I \end{pmatrix} \begin{pmatrix} S_2(z) \\ I \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{b-a}I & 0 \\ 0 & (b-a)I \end{pmatrix} [U_{2,n}(z)]^{-1} \begin{pmatrix} S_2(z) \\ I \end{pmatrix}.
\end{aligned} \tag{6.76}$$

Thus the equations (6.75) and (6.76) and inequality (6.19) yield that (5.3) and (5.4) hold for every choice of z in $\mathbb{C} \setminus \mathbb{R}$. Consequently, the column pair $\binom{\tilde{P}}{\tilde{Q}}$ belongs to $\mathfrak{P}(\tilde{J}_q, [a,b])$. From (6.74) we see that

$$\begin{pmatrix} S \\ I \end{pmatrix} = \begin{pmatrix} U_{11,n}P + U_{12,n}Q \\ U_{21,n}P + U_{22,n}Q \end{pmatrix}$$

holds. This implies (6.64).

(c) Let $\binom{P_1}{Q_1}$ and $\binom{P_2}{Q_2}$ belong to $\mathfrak{P}(-\tilde{J}_q, [a,b])$. First suppose that (6.65) is valid. Lemma 6.12 yields

$$\begin{pmatrix} P_j \\ Q_j \end{pmatrix} = U_n^{-1} \begin{pmatrix} (U_{11,n}P_j + U_{12,n}Q_j)(U_{21,n}P_j + U_{22,n}Q_j)^{-1} \\ I \end{pmatrix} \cdot (U_{21,n}P_j + U_{22,n}Q_j)$$

for each $j \in \{1,2\}$. Because of (6.65) this implies

$$\begin{aligned}
\begin{pmatrix} P_1 \\ Q_1 \end{pmatrix} &= U_n^{-1} \begin{pmatrix} (U_{11,n}P_2 + U_{12,n}Q_2)(U_{21,n}P_2 + U_{22,n}Q_2)^{-1} \\ I \end{pmatrix} \\
&\quad \cdot (U_{21,n}P_1 + U_{22,n}Q_1) \\
&= \begin{pmatrix} P_2 \\ Q_2 \end{pmatrix} (U_{21,n}P_2 + U_{22,n}Q_2)^{-1}(U_{21,n}P_1 + U_{22,n}Q_1).
\end{aligned}$$

Consequently, (6.66) holds.

Conversely, it is readily checked that (6.66) is also sufficient for (6.65). □

Corollary 6.15. *If $(s_j)_{j=0}^{2n}$ is a sequence of complex $q \times q$ matrices such that the block Hankel matrices $\mathbf{H}_{1,n}$ and $\mathbf{H}_{2,n}$ are both positive Hermitian, then*

$$\mathcal{M}_{\geq}^{q}\left[[a,b],\mathfrak{B}\cap[a,b];(s_j)_{j=0}^{2n}\right]\neq\emptyset.$$

Proof. Use Remark 5.6 and Theorem 6.14. □

7. A necessary and sufficient condition for the existence of a solution of the moment problem

In this section, we turn our attention to the proof of Theorem 1.3 which contains a characterization of the case that the matricial moment problem on a finite interval considered in this paper has a solution. We suppose again that a and b are real numbers such that $a < b$ and that n is positive integer.

Remark 7.1. *Let $(s_j^{(1)})_{j=0}^{2n}$ and $(s_j^{(2)})_{j=0}^{2n}$ be sequences of complex $q \times q$ matrices, let α be a positive real number, and let $r_j := s_j^{(1)} + \alpha s_j^{(2)}$ for each integer j with $0 \leq j \leq 2n$. For $m \in \{1,2\}$, let*

$$\tilde{H}_{0,n}^{(m)} := (s_{j+k}^{(m)})_{j,k=0}^{n}, \quad \tilde{H}_{1,n-1}^{(m)} := (s_{j+k+1}^{(m)})_{j,k=0}^{n-1}, \quad \tilde{H}_{2,n-1}^{(m)} := (s_{j+k+2}^{(m)})_{j,k=0}^{n-1},$$

$$\mathbf{H}_{1,n}^{(m)} := \tilde{H}_{1,n}^{(m)}, \quad \text{and} \quad \mathbf{H}_{2,n}^{(m)} := -ab\tilde{H}_{0,n-1}^{(m)} + (a+b)\tilde{H}_{1,n-1}^{(m)} - \tilde{H}_{2,n-1}^{(m)}.$$

Suppose that the block Hankel matrices $\mathbf{H}_{1,n}^{(1)}$ and $\mathbf{H}_{2,n}^{(1)}$ are both nonnegative Hermitian and that the block Hankel matrices $\mathbf{H}_{1,n}^{(2)}$ and $\mathbf{H}_{2,n}^{(2)}$ are both positive Hermitian. Then the block Hankel matrices

$$\mathbf{Q}_{1,n} := (r_{j+k})_{j,k=0}^{n}, \quad \text{and} \quad \mathbf{Q}_{2,n} := (-abr_{j+k} + (a+b)r_{j+k+1} - r_{j+k+2})_{j,k=0}^{n-1} \quad (7.1)$$

are positive Hermitian as well.

Now we prove the necessary and sufficient condition for the existence of a solution of the moment problem in question.

Proof of Theorem 1.3. If $\mathcal{M}_{\geq}^{q}\left[[a,b],\mathfrak{B}\cap[a,b];(s_j)_{j,k=0}^{2n}\right]$ is nonempty, then Remark 3.4 yields that the matrices $\mathbf{H}_{1,n}$ and $\mathbf{H}_{2,n}$ are both nonnegative Hermitian. Conversely, now suppose that the matrices $\mathbf{H}_{1,n}$ and $\mathbf{H}_{2,n}$ are both nonnegative Hermitian. According to Lemma 6.1, let $(r_j)_{j=0}^{2n}$ be a sequence of complex $q \times q$ matrices such that the block Hankel matrices $\mathbf{Q}_{1,n}$ and $\mathbf{Q}_{2,n}$ defined by (7.1) are both positive Hermitian. For each real number ε with $0 < \varepsilon \leq 1$ and each integer j with $0 \leq j \leq 2n$, let $s_{j,\varepsilon} := s_j + \varepsilon r_j$. From Remark 7.1 we see that for each $\varepsilon \in (0,1]$ the matrices $(s_{j+k,\varepsilon})_{j,k=0}^{n}$ and $(-abs_{j+k,\varepsilon} + (a+b)s_{j+k+1,\varepsilon} - s_{j+k+2,\varepsilon})_{j,k=0}^{n-1}$ are both positive Hermitian. Corollary 6.15 provides us then that the set $\mathcal{M}_{\geq}^{q}\left[[a,b],\mathfrak{B}\cap[a,b];(s_{j,\varepsilon})_{j,k=0}^{2n}\right]$ is nonempty. Now let $(\varepsilon_m)_{m=1}^{\infty}$ be a sequence of real numbers belonging to the interval $(0,1]$ which satisfies

$$\lim_{m\to\infty}\varepsilon_m = 0.$$

For each positive integer m, we can choose then a nonnegative Hermitian $q \times q$ measure σ_m which belongs to $\mathcal{M}^q_{\geq}\left[[a,b], \mathfrak{B} \cap [a,b]; (s_{j,\varepsilon_m})_{j=0}^{2n}\right]$. Using the notation given in (3.2), we have $s_j^{[\sigma_m]} = s_{j,\varepsilon_m}$ for all positive integers m and all integers j which satisfy $0 \leq j \leq 2n$. Obviously, it follows

$$\sigma_m([a,b]) = s_0^{[\sigma_m]} = s_{0,\varepsilon_m} = s_0 + \varepsilon_m r_0 \leq s_0 + r_0 \tag{7.2}$$

for all positive integers m. In view of (7.2), application of the matricial version of the Helly-Prohorov theorem (see [FK, Satz 9]) provides us that there are a subsequence $(\sigma_{m_k})_{k=1}^{\infty}$ of the sequence $(\sigma_m)_{m=1}^{\infty}$ and a nonnegative Hermitian $q \times q$ measure $\sigma \in \mathcal{M}^q_{\geq}[[a,b], \mathfrak{B} \cap [a,b]]$ such that $(\sigma_{m_k})_{k=1}^{\infty}$ converges weakly to σ, i.e., such that

$$\lim_{k\to\infty} \int_{[a,b]} f d\sigma_{m_k} = \int_{[a,b]} f d\sigma$$

is satisfied for all continuous complex-valued functions defined on $[a,b]$. Therefore we can conclude then

$$s_j^{[\sigma]} = \lim_{k\to\infty} s_j^{[\sigma_{m_k}]} = \lim_{k\to\infty} (s_j + \varepsilon_{m_k} r_j) = s_j$$

for every integer j which satisfies $0 \leq j \leq 2n$. Hence σ belongs to $\mathcal{M}^q_{\geq}\left[[a,b], \mathfrak{B} \cap [a,b];\ (s_j)_{j=0}^{2n}\right]$. In particular, $\mathcal{M}^q_{\geq}\left[[a,b], \mathfrak{B} \cap [a,b];\ (s_j)_{j=0}^{2n}\right] \neq \emptyset$. □

Finally, let us give a remark concerning the scalar case $q = 1$. M.G. Krein [K, Theorem 4.2, p. 48] (see also [KN, Theorem 4.1, p. 110–111]) showed that the set $\mathcal{M}^1_{\geq}\left([a,b], \mathfrak{B} \cap [a,b]; (s_j)_{j=0}^{2n}\right)$ contains exactly one measure if and only if the block Hankel matrices $\mathbf{H}_{1,n}$ and $\mathbf{H}_{2,n}$ are both nonnegative Hermitian and at least one of them is not positive Hermitian.

Acknowledgement

The authors wish to express their gratitude to Helge Thiele for helpful discussions.

References

[AB] Alpay, D/Bolotnikov, V.: On a general moment problem and certain matrix equations, Linear Alg. Appl. 203–204 (1994), 3–43.

[BS] Bolotnikov, V.A.; Sakhnovich, L.A.: *On an operator approach to interpolation problems for Stieltjes functions*, Integral Equations and Operator Theory 35 (1999), 423–470.

[CDFK] Choque Rivero, A.E/Dyukarev, Y.M./Fritzsche, B./Kirstein, B.: A Truncated Matricial Moment Problem on a Finite Interval to appear in: Interpolation, Schur Functions and Moment Problems (Eds.: D. Alpay and I. Gohberg), Operator Theory: Advances and Applications, 165, Birkhäuser, Basel-Boston-Berlin 2006

[DFK] Dubovoj, V. K./Fritzsche, B./Kirstein, B.: Matricial Version of the Classical Schur Problem. *Teubner-Texte zur Mathematik Bd.* 129, Teubner, Stuttgart Leipzig 1992

[DK] Dyukarev, Yu.M.; Katsnelson, V.E.: *Multiplicative and additive classes of analytic matrix functions of Stieltjes type and associated interpolation problems* (Russian). Teor. Funkcii, Funkcional. Anal. i Prilozen., Part I: **36** (1981), 13–27; Part III: **41** (1984), 64–70. 36: 13–27, 1981. Engl. transl. in: Amer. Math. Soc. Transl., Series 2, 131 (1986), 55–70.

[Du] Dubovoj, V.K.: *Indefinite metric in the interpolation problem of Schur for analytic matrix functions* (Russian), Teor. Funkcii, Funkcional. Anal. i. Prilozen., Part I: **37** (1982), 14–26; Part II: **38** (1982), 32–39; Part III: **41** (1984), 55–64; Part IV: **42** (1984), 46–57; Part V: **45** (1986), 16–21; Part VI: **47** (1987), 112–119.

[Dy1] Dyukarev, Yu.M.: *Multiplicative and additive Stieltjes classes of analytic matrix-valued functions and interpolation problems connected with them.* II. (Russian), Teor. Funkcii, Funkcional. Anal. i Prilozen. **38** (1982), 40–48.

[FK] Fritzsche, B.; Kirstein, B.: *Schwache Konvergenz nichtnegativer hermitescher Borelmaße*, Wiss. Z. Karl-Marx-Univ. Leipzig, Math. Naturwiss. R. 37 (1988) 4, 375–398.

[FKK] Fritzsche, B.; Kirstein, B.; Krug, V.: *On several types of resolvent matrices of nondegenerate matricial Carathéodory problems* Linear Algebra Appl. 281 (1998), 137–170.

[G1] Golinskii, L.B.: *On the Nevanlinna-Pick problem in the generalized Schur class of analytic matrix functions* (Russian), in: Analysis in Indefinite Dimensional Spaces and Operator Theory (Ed.: V.A. Marčenko), Naukova Dumka, Kiev 1983, pp. 23–33.

[G2] Golinskii, L.B.: *On some generalization of the matricial Nevanlinna-Pick problem* (Russian), Izv. Akad. Nauk Armjan. SSR, Ser. Mat. **18** (1983), No. 3, 187–205.

[IS] Ivanchenko, T.S.; Sakhnovich, L.A.: *An operator approach to the Potapov scheme for the solution of interpolation problems*, in: Matrix and Operator Valued Functions – The Vladimir Petrovich Potapov Memorial Volume (Eds.: I. Gohberg, L.A. Sakhnovich), OT Series, Vol. 72, Birkhäuser, Basel-Boston-Berlin 1994, pp. 48–86.

[K] Krein, M.G.: *The ideas of P.L. Chebyshev and A.A. Markov in the theory of limit values of integrals and their further development* (Russian), Usp. Mat. Nauk 5 (1951), no. 4, 3–66, Engl. transl.: Amer. Math. Soc. Transl., Series 2, 12 (1959), 1–121.

[Ka1] Katsnelson, V.E.: *Continual analogues of the Hamburger-Nevanlinna theorem and fundamental matrix inequalities of classical problems* (Russian), Teor. Funkcii, Funkcional. Anal. i Prilozen., Part I: **36** (1981), 31–48; Part II: **37** (1982), 31–48; Part III: **39** (1983), 61–73; Part IV: **40** (1983), 79–90. Engl. transl. in: Amer. Math. Soc. Transl., Series 2, 136 (1987), 49–108.

[Ka2] Katsnelson, V.E.: *Methods of J-theory in continuous interpolation problems of analysis* (Russian), deposited in VINITI 1983. Engl. transl. by T. Ando, Sapporo 1985.

[Ka3] Katsnelson, V.E.: *On transformations of Potapov's fundamental matrix inequality*, in: Topics in Interpolation Theory (Eds.: H. Dym, B. Fritzsche, V.E. Katsnelson), OT Series, Birkhäuser, Basel-Boston-Berlin, 1997, pp. 253–281.

[KN] Krein, M.G.; Nudelman, A.A.: *The Markov Moment Problem and Extremal Problems,* Amer. Math. Soc. Transl., Vol. 50, Amer. Math. Soc., Providence, R.I. 1977.

[Ko1] Kovalishina, I.V.: *Analytic theory of a class of interpolation problems* (Russian), Izv. Akad. Nauk SSSR, Ser. Mat. **47** (1983), 455–497; Engl. transl. in: Math. USSR Izvestiya 22 (1984), 419–463.

[Ko2] Kovalishina, I.V.: *The multiple boundary interpolation problem for contractive matrix functions in the unit disc* (Russian), Teor. Funkcii, Funkcional. Anal. i Prilozen, **51** (1988), 38–55.

[Kt] Kats, I.S.: *On Hilbert spaces generated by Hermitian monotone matrix functions* (Russian), Zapiski Nauc.-issled. Inst. Mat. i Mekh. i Kharkov. Mat. Obsh. 22 (1950), 95–113.

[R1] Rosenberg, M.: The square integrability of matrix-valued functions with respect to a non-negative Hermitian measure. *Duke Math. J.* 31, (1964), 291–298.

[R2] Rosenberg, M.: Operators as spectral integrals of operator-valued functions from the study of multivariate stationary stochastic processes. *J. Multivariate Anal.* 4, (1974), 166–209.

[R3] Rosenberg, M.: Spectral integrals of operator-valued functions – II. From the study of stationary processes. *J. Multivariate Anal.* 6, (1976), 538–571.

[S1] Sakhnovich, L.A.: *Similarity of operators* (Russian), Sibirskii Matematicheskii Zhurnal 13 (1972), 868–883. English transl. in: Sibirian Mathematical Journal 13 (1972), 604–615.

[S2] Sakhnovich, L.A.: *Interpolation Theory and Its Applications,* Kluwer, Dordrecht 1997.

Abdon Eddy Choque Rivero
Instituto de Física y Matemáticas
Universidad Michoacana de San Nicolás de Hidalgo
Edificio C-3, Ciudad Universitaria, C.P. 58040
Morelia, Michoacana, Mexico
e-mail: abdon@ifm.umich.mx

Yuriy M. Dyukarev
Department of Physics
Kharkov State University
Svoboda Square 4
Kharkov 61077, Ukraine
e-mail: yury.m.dyukarev@univer.kharkov.ua

Bernd Fritzsche and Bernd Kirstein
Mathematisches Institut
Universität Leipzig
Augustusplatz 10/11
D-04109 Leipzig, Germany
e-mail: fritzsche@math.uni-leipzig.de
e-mail: kirstein@math.uni-leipzig.de

Operator Theory:
Advances and Applications, Vol. 176, 165–198

On the Irreducibility of a Class of Homogeneous Operators

Gadadhar Misra and Subrata Shyam Roy

Abstract. In this paper we construct a class of homogeneous Hilbert modules over the disc algebra $\mathcal{A}(\mathbb{D})$ as quotients of certain natural modules over the function algebra $\mathcal{A}(\mathbb{D}^2)$. These quotient modules are described using the jet construction for Hilbert modules. We show that the quotient modules obtained this way, belong to the class $\mathrm{B}_k(\mathbb{D})$ and that they are mutually inequivalent, irreducible and homogeneous.

Keywords. Homogeneous operators, Möbius group, Projective unitary representations, Cocycle, Imprimitivity, Hilbert modules, Jet construction, Reproducing kernels.

1. Introduction

Let $\mathcal{M}$ be a Hilbert space. All Hilbert spaces in this paper will be assumed to be complex and separable. Let $\mathcal{A}(\Omega)$ be the natural function algebra consisting of functions holomorphic in a neighborhood of the closure $\bar{\Omega}$ of some open, connected and bounded subset Ω of $\mathbb{C}^m$. The Hilbert space $\mathcal{M}$ is said to be a *Hilbert module* over $\mathcal{A}(\Omega)$ if $\mathcal{M}$ is a module over $\mathcal{A}(\Omega)$ and

$$\|f \cdot h\|_{\mathcal{M}} \leq C\|f\|_{\mathcal{A}(\Omega)}\|h\|_{\mathcal{M}} \text{ for } f \in \mathcal{A}(\Omega) \text{ and } h \in \mathcal{M},$$

for some positive constant C independent of f and h. It is said to be *contractive* if we also have $C \leq 1$.

Fix an inner product on the algebraic tensor product $\mathcal{A}(\Omega) \otimes \mathbb{C}^n$. Let the completion of $\mathcal{A}(\Omega) \otimes \mathbb{C}^n$ with respect to this inner product be the Hilbert space $\mathcal{M}$. Assume that the module action

$$\mathcal{A}(\Omega) \times \mathcal{A}(\Omega) \otimes \mathbb{C}^n \to \mathcal{A}(\Omega) \otimes \mathbb{C}^n$$

The research of the first author was supported in part by a grant from the DST-NSF Science and Technology Cooperation Programme. The second author was supported by the Indian Statistical Institute.

extends continuously to $\mathcal{A}(\Omega) \times \mathcal{M} \to \mathcal{M}$. With very little additional assumption on $\mathcal{M}$, we obtain a *quasi-free* Hilbert module (cf. [13]).

The simplest family of modules over $\mathcal{A}(\Omega)$ corresponds to evaluation at a point in the closure of Ω. For $\underline{z}$ in the closure of Ω, we make the one-dimensional Hilbert space $\mathbb{C}$ into the Hilbert module $\mathbb{C}_{\underline{z}}$, by setting $\varphi v = \varphi(z)v$ for $\varphi \in \mathcal{A}(\Omega)$ and $v \in \mathbb{C}$. Classical examples of contractive Hilbert modules are the Hardy and Bergman modules over the algebra $\mathcal{A}(\Omega)$.

Let G be a locally compact second countable group acting transitively on Ω. Let us say that the module $\mathcal{M}$ over the algebra $\mathcal{A}(\Omega)$ is homogeneous if $\varrho(f \circ \varphi)$ is unitarily equivalent to $\varrho(f)$ for all $\varphi \in G$. Here $\varrho : \mathcal{A}(\Omega) \to \mathcal{B}(\mathcal{M})$ is the homomorphism of the algebra $\mathcal{A}(\Omega)$ defined by $\varrho(f)h := f \cdot h$ for $f \in \mathcal{A}(\Omega)$ and $h \in \mathcal{M}$. It was shown in [19] that if the module $\mathcal{M}$ is irreducible and homogeneous then there exists a *projective unitary representation* $U : G \to \mathcal{U}(\mathcal{M})$ such that

$$U_\varphi^* \varrho(f) U_\varphi = \varrho(f \cdot \varphi),\ f \in \mathcal{A}(\Omega),\ \varphi \in G,$$

where $(f \cdot \varphi)(w) = f(\varphi \cdot w)$ for $w \in \Omega$.

A $*$-homomorphism ϱ of a C^*-algebra $\mathcal{C}$ and a unitary group representation U of G on the Hilbert space $\mathcal{M}$ satisfying the condition as above were first studied by Mackey and were called *Systems of Imprimitivity.* Mackey proved the Imprimitivity theorem which sets up a correspondence between induced representations of the group G and the Systems of Imprimitivity. The notion of homogeneity is obtained by compressing the systems of imprimitivities, in the sense of Mackey, to a subspace $\mathcal{N}$ of $\mathcal{M}$ and then restricting to a subalgebra of the C^*-algebra $\mathcal{C}$ (cf.[3]). However, it is not clear if the notion of homogeneity is in some correspondence with holomorphically induced representations, at least when the module $\mathcal{M}$ is assumed to be in $\mathrm{B}_k(\Omega)$.

An alternative description, in the particular case of the disc may be useful. The group of bi-holomorphic automorphisms Möb of the unit disc is $\{\varphi_{\theta,\alpha} : \theta \in [0, 2\pi) \text{ and } \alpha \in \mathbb{D}\}$, where

$$\varphi_{\theta,\alpha}(z) = e^{i\theta} \frac{z - \alpha}{1 - \bar{\alpha} z},\ z \in \mathbb{D}. \tag{1.1}$$

As a topological group (with the topology of locally uniform convergence) it is isomorphic to $\mathrm{PSU}(1,1)$ and to $\mathrm{PSL}(2,\mathbb{R})$.

An operator T from a Hilbert space into itself is said to be *homogeneous* if $\varphi(T)$ is unitarily equivalent to T for all φ in Möb which are analytic on the spectrum of T. The spectrum of a homogeneous operator T is either the unit circle $\mathbb{T}$ or the closed unit disc $\bar{\mathbb{D}}$, so that, actually, $\varphi(T)$ is unitarily equivalent to T for all φ in Möb. We say that a projective unitary representation U of Möb is *associated* with an operator T if

$$\varphi(T) = U_\varphi^* T U_\varphi$$

for all φ in Möb. We have already pointed out that if T is irreducible then it has an associated representation U. It is not hard to see that U is uniquely determined upto unitary equivalence.

Many examples (unitarily inequivalent) of homogeneous operators are known [6]. Since the direct sum (more generally direct integral) of two homogeneous operators is again homogeneous, a natural problem is the classification (up to unitary equivalence) of *atomic homogeneous operators*, that is, those homogeneous operators which can not be written as the direct sum of two homogeneous operators. In this generality, this problem remains unsolved. However, the irreducible homogeneous operators in the Cowen-Douglas class $\mathrm{B}_1(\mathbb{D})$ and $\mathrm{B}_2(\mathbb{D})$ have been classified (cf. [18] and [22]) and all the scalar shifts (not only the irreducible ones) which are homogeneous are known [7, List 4.1, p. 312]. Some recent results on classification of homogeneous bundles are in [8] and [15].

Clearly, irreducible homogeneous operators are atomic. Therefore, it is important to understand when a homogeneous operator is irreducible.

There are only two examples of atomic homogeneous operators known which are not irreducible. these are the multiplication operators – by the respective coordinate functions – on the Hilbert spaces $L^2(\mathbb{T})$ and $L^2(\mathbb{D})$. Both of these examples happen to be normal operators. We do not know if all atomic homogeneous operators possess an associated projective unitary representation. However, to every homogeneous operator in $\mathrm{B}_k(\mathbb{D})$, there exist an associated representation of the universal covering group of Möb [15, Theorem 4].

It turns out an irreducible homogeneous operator in $\mathrm{B}_2(\mathbb{D})$ is the compression of the tensor product of two homogeneous operators from $\mathrm{B}_1(\mathbb{D})$ (cf. [6]) to the ortho-compliment of a suitable invariant subspace. In the language of Hilbert modules, this is the statement that every homogeneous module in $\mathrm{B}_2(\mathbb{D})$ is obtained as quotient of the tensor product of two homogeneous modules in $\mathrm{B}_1(\mathbb{D})$ by the sub-module of functions vanishing to order 2 on $\triangle \subseteq \mathbb{D}^2$. However, beyond the case of rank 2, the situation is more complicated. The question of classifying homogeneous operators in the class $\mathrm{B}_k(\mathbb{D})$ amounts to classifying holomorphic and Hermitian vector bundles of rank k on the unit disc which are homogeneous. Classification problems such as this one are well known in the representation theory of locally compact second countable groups. However, in that context, there is no Hermitian structure present which makes the classification problem entirely algebraic. A complete classification of homogeneous operators in $\mathrm{B}_k(\mathbb{D})$ may still be possible using techniques from the theory of unitary representations of the Möbius group. Leaving aside, the classification problem of the homogeneous operators in $\mathrm{B}_k(\mathbb{D})$, we show that the "generalized Wilkins examples" (cf. [6]) are irreducible.

If one considers a bounded symmetric domain in $\mathbb{C}^m$, the classification question probably is even more complicated (cf. [4], [1]). Here part of the difficulty lies in the fact that no classification of the irreducible unitary representations of the group $\mathrm{Aut}(\Omega)$, the bi-holomorphic automorphism group of Ω, is known.

In the following section, we discuss reproducing kernels for a functional Hilbert space on a domain $\Omega \subseteq \mathbb{C}^m$ and the m-tuple of multiplication operators $\boldsymbol{M}$ of multiplication by coordinate functions. Although, our applications to the question of irreducibility is only for the multiplication operator $\boldsymbol{M}$ on a functional Hilbert space based on the unit disc $\mathbb{D}$, the more general discussion of this section is not any simpler in the one variable case.

In Subsection 2.2, we explain the realization of a m-tuple of operators $\boldsymbol{T}$ in the class $\mathrm{B}_k(\Omega)$ as the adjoint of a m-tuple of multiplication operators $\boldsymbol{M}$ on a Hilbert space of holomorphic functions, on the bounded connected open set $\Omega^* := \{w \in \mathbb{C}^m : \bar{w} \in \Omega\}$, possessing a reproducing kernel K. We point out, as in [11], that the normalized kernel $\widetilde{\widetilde{K}}$ obtained from the kernel K by requiring that $\widetilde{\widetilde{K}}(z,0) = 1$, for all $z \in \Omega$, determines the uniatry equivalence class of the m-tuple $\boldsymbol{T}$. We then obtain a criterion for the irreducibility of the m-tuple $\boldsymbol{T}$ in terms of the normalized kernel $\widetilde{\widetilde{K}}$. Roughly speaking, this says that the m-tuple of operators is irreducible if and only if the coefficients, in the power series expansion of $\widetilde{\widetilde{K}}$, are simultaneously irreducible. Following, [12] and [14], we describe the jet construction for Hilbert modules and discuss some examples.

In Section 3, we show that if $\mathcal{H}$ is a Hilbert space of holomorphic functions, on a bounded connected open set Ω and possesses a reproducing kernel K then it admits a natural multiplier representation of the automorphism group of Ω if K is *quasi-invariant*. We show that if K is quasi-invariant, then the corresponding multiplier representation intertwines $\boldsymbol{M}$ and $\varphi(\boldsymbol{M})$, that is, the m-tuple of multiplication operators $\boldsymbol{M}$ is *homogeneous*.

Our main results on irreducibility of certain class of homogeneous operators is in Section 4. The kernel $B^{(\alpha,\beta)}(z,w) = (1-z_1\bar{w}_1)^{-\alpha}(1-z_2\bar{w}_2)^{-\beta}$, $z = (z_1, z_2)$, $w = (w_1, w_2) \in \mathbb{D}^2$, determines a Hilbert module over the function algebra $\mathcal{A}(\mathbb{D}^2)$. We recall the computation of a matrix-valued kernel on the unit disc $\mathbb{D}$ using the jet construction for this Hilbert module which consists of holomorphic functions on the unit disc $\mathbb{D}$ taking values in $\mathbb{C}^n$. The multiplication operator on this Hilbert space is then shown to be irreducible by checking that all of the coefficients of the "normalized" matrix-valued kernel, obtained from the jet construction, cannot be simultaneously reducible.

In Section 5, we show that the kernel obtained from the jet construction is quasi-invariant and consequently, the corresponding multiplication operator is homogeneous. This proof involves the verification of a cocycle identity, which in turn, depends on a beautiful identity involving binomial coefficients.

Finally, in Section 6, we discuss some examples arising from the jet construction applied to a certain natural family of Hilbert modules over the algebra $\mathcal{A}(\mathbb{D}^3)$. A more systematic study of such examples is to be found in [20].

2. Reproducing kernels and the Cowen-Douglas class

2.1. Reproducing kernel

Let $\mathcal{L}(\mathbb{F})$ be the Banach space of all linear transformations on a Hilbert space $\mathbb{F}$ of dimension n for some $n \in \mathbb{N}$. Let $\Omega \subset \mathbb{C}^m$ be a bounded open connected set. A function $K : \Omega \times \Omega \to \mathcal{L}(\mathbb{F})$, satisfying

$$\sum_{i,j=1}^{p} \langle K(w^{(i)}, w^{(j)})\zeta_j, \zeta_i\rangle_{\mathbb{F}} \ \geq \ 0,\ w^{(1)}, \ldots, w^{(p)} \in \Omega,\ \zeta_1, \ldots, \zeta_p \in \mathbb{F},\ p > 0 \quad (2.2)$$

is said to be a *non negative definite* (*nnd*) *kernel* on Ω. Given such an nnd kernel K on Ω, it is easy to construct a Hilbert space $\mathcal{H}$ of functions on Ω taking values in $\mathbb{F}$ with the property

$$\langle f(w), \zeta\rangle_{\mathbb{F}} = \langle f, K(\cdot, w)\zeta\rangle, \text{ for } w \in \Omega,\ \zeta \in \mathbb{F}, \text{ and } f \in \mathcal{H}. \quad (2.3)$$

The Hilbert space $\mathcal{H}$ is simply the completion of the linear span of all vectors of the form $\mathcal{S} = \{K(\cdot, w)\zeta,\ w \in \Omega,\ \zeta \in \mathbb{F}\}$, where the inner product between two of the vectors from $\mathcal{S}$ is defined by

$$\langle K(\cdot, w)\zeta, K(\cdot, w')\eta\rangle = \langle K(w', w)\zeta, \eta\rangle, \text{ for } \zeta, \eta \in \mathbb{F}, \text{ and } w, w' \in \Omega, \quad (2.4)$$

which is then extended to the linear span $\mathcal{H}^\circ$ of the set $\mathcal{S}$. This ensures the reproducing property (2.3) of K on $\mathcal{H}^\circ$.

Remark 2.1. *We point out that although the kernel K is required to be merely* nnd*, the equation* (2.4) *defines a positive definite sesqui-linear form. To see this, simply note that $|\langle f(w), \zeta\rangle| = |\langle f, K(\cdot, w)\zeta\rangle|$ which is at most $\|f\|\langle K(w, w)\zeta, \zeta\rangle^{1/2}$ by the Cauchy-Schwarz inequality. It follows that if $\|f\|^2 = 0$ then $f = 0$.*

Conversely, let $\mathcal{H}$ be any Hilbert space of functions on Ω taking values in $\mathbb{F}$. Let $e_w : \mathcal{H} \to \mathbb{F}$ be the evaluation functional defined by $e_w(f) = f(w)$, $w \in \Omega$, $f \in \mathcal{H}$. If e_w is bounded for each $w \in \Omega$ then it admits a bounded adjoint $e_w^* : \mathbb{F} \to \mathcal{H}$ such that $\langle e_w f, \zeta\rangle = \langle f, e_w^* \zeta\rangle$ for all $f \in \mathcal{H}$ and $\zeta \in \mathbb{F}$. A function f in $\mathcal{H}$ is then orthogonal to $e_w^*(\mathcal{H})$ if and only if $f = 0$. Thus $f = \sum_{i=1}^{p} e_{w^{(i)}}^*(\zeta_i)$ with $w^{(1)}, \ldots, w^{(p)} \in \Omega,\ \zeta_1, \ldots, \zeta_p \in \mathbb{F},$ and $p > 0$, form a dense set in $\mathcal{H}$. Therefore we have

$$\|f\|^2 = \sum_{i,j=1}^{p} \langle e_{w^{(i)}} e_{w^{(j)}}^* \zeta_j, \zeta_i\rangle,$$

where $f = \sum_{i=1}^{n} e_{w^{(i)}}^*(\zeta_i),\ w^{(i)} \in \Omega,\ \zeta_i \in \mathcal{F}$. Since $\|f\|^2 \geq 0$, it follows that the kernel $K(z, w) = e_z e_w^*$ is non-negative definite as in (2.2). It is clear that $K(z, w)\zeta \in \mathcal{H}$ for each $w \in \Omega$ and $\zeta \in \mathbb{F}$, and that it has the reproducing property (2.3).

Remark 2.2. *If we assume that the evaluation functional e_w is surjective then the adjoint e_w^* is injective and it follows that $\langle K(w, w)\zeta, \zeta\rangle > 0$ for all non-zero vectors $\zeta \in \mathbb{F}$.*

There is a useful alternative description of the reproducing kernel K in terms of the orthonormal basis $\{e_k : k \geq 0\}$ of the Hilbert space $\mathcal{H}$. We think of the vector $e_k(w) \in \mathbb{F}$ as a column vector for a fixed $w \in \Omega$ and let $e_k(w)^*$ be the row vector $(\overline{e_k^1(w)}, \dots, \overline{e_k^n(w)})$. We see that

$$\begin{aligned} \langle K(z,w)\zeta, \eta\rangle &= \langle K(\cdot,w)\zeta, K(\cdot,z)\eta\rangle \\ &= \sum_{k=0}^{\infty} \langle K(\cdot,w)\zeta, e_k\rangle\langle e_k, K(\cdot,z)\eta\rangle \\ &= \sum_{k=0}^{\infty} \overline{\langle e_k(w), \zeta\rangle}\langle e_k(z), \eta\rangle \\ &= \sum_{k=0}^{\infty} \langle e_k(z)e_k(w)^*\zeta, \eta\rangle, \end{aligned}$$

for any pair of vectors $\zeta, \eta \in \mathbb{F}$. Therefore, we have the following very useful representation for the reproducing kernel K:

$$K(z,w) = \sum_{k=0}^{\infty} e_k(z)e_k(w)^*, \tag{2.5}$$

where $\{e_k : k \geq 0\}$ is any orthonormal basis in $\mathcal{H}$.

2.2. The Cowen-Douglas class

Let $\boldsymbol{T} = (T_1, \dots, T_m)$ be a d-tuple of commuting bounded linear operators on a separable complex Hilbert space $\mathcal{H}$. Define the operator $D_{\mathbf{T}} : \mathcal{H} \to \mathcal{H} \oplus \cdots \oplus \mathcal{H}$ by $D_{\mathbf{T}}(x) = (T_1x, \dots, T_mx)$, $x \in \mathcal{H}$. Let Ω be a bounded domain in $\mathbb{C}^m$. For $w = (w_1, \dots, w_m) \in \Omega$, let $\boldsymbol{T} - w$ denote the operator tuple $(T_1 - w_1, \dots, T_m - w_m)$. Let n be a positive integer. The m-tuple $\boldsymbol{T}$ is said to be in the Cowen-Douglas class $\mathrm{B}_n(\Omega)$ if

1. ran $D_{\mathbf{T}-w}$ is closed for all $w \in \Omega$
2. span $\{\ker D_{\mathbf{T}-w} : w \in \Omega\}$ is dense in $\mathcal{H}$
3. $\dim\ker D_{\mathbf{T}-w} = n$ for all $w \in \Omega$.

This class was introduced in [10]. The case of a single operator was investigated earlier in the paper [9]. In this paper, it is pointed out that an operator T in $\mathrm{B}_1(\Omega)$ is unitarily equivalent to the adjoint of the multiplication operator M on a reproducing kernel Hilbert space, where $(Mf)(z) = zf(z)$. It is not very hard to see that, more generally, a m-tuple $\boldsymbol{T}$ in $\mathrm{B}_n(\Omega)$ is unitarily equivalent to the adjoint of the m-tuple of multiplication operators $\boldsymbol{M} = (M_1, \dots, M_m)$ on a reproducing kernel Hilbert space [9] and [11, Remark 2.6 a) and b)]. Also, Curto and Salinas [11] show that if certain conditions are imposed on the reproducing kernel then the corresponding adjoint of the m-tuple of multiplication operators belongs to the class $\mathrm{B}_n(\Omega)$.

To a m-tuple $\boldsymbol{T}$ in $\mathrm{B}_n(\Omega)$, on the one hand, one may associate a holomorphic Hermitian vector bundle $E_{\mathbf{T}}$ on Ω (cf. [9]), while on the other hand, one may associate a normalized reproducing kernel K (cf. [11]) on a suitable sub-domain of $\Omega^* = \{w \in \mathbb{C}^m : \bar{w} \in \Omega\}$. It is possible to answer a number of questions regarding the m-tuple of operators $\boldsymbol{T}$ using either the vector bundle or the reproducing kernel. For instance, in the two papers [9] and [10], Cowen and Douglas show that the curvature of the bundle $E_{\mathbf{T}}$ along with a certain number of derivatives forms a complete set of unitary invariants for the operator $\boldsymbol{T}$ while Curto and Salinas [11] establish that the unitary equivalence class of the normalized kernel K is a complete unitary invariant for the corresponding m-tuple of multiplication operators. Also, in [9], it is shown that a single operator in $\mathrm{B}_n(\Omega)$ is reducible if and only if the associated holomorphic Hermitian vector bundle admits an orthogonal direct sum decomposition.

We recall the correspondence between a m-tuple of operators in the class $\mathrm{B}_n(\Omega)$ and the corresponding m-tuple of multiplication operators on a reproducing kernel Hilbert space on Ω.

Let $\boldsymbol{T}$ be a m-tuple of operators in $\mathrm{B}_n(\Omega)$. Pick n linearly independent vectors $\gamma_1(w), \dots, \gamma_n(w)$ in $\ker D_{\mathbf{T}-w}$, $w \in \Omega$. Define a map $\Gamma : \Omega \to \mathcal{L}(\mathbb{F}, \mathcal{H})$ by $\Gamma(w)\zeta = \sum_{i=0}^{n} \zeta_i \gamma_i(w)$, where $\zeta = (\zeta_1, \dots, \zeta_n) \in \mathbb{F}$, $\dim \mathbb{F} = n$. It is shown in [9, Proposition 1.11] and [11, Theorem 2.2] that it is possible to choose $\gamma_1(w), \dots, \gamma_n(w)$, w in some domain $\Omega_0 \subseteq \Omega$, such that Γ is holomorphic on Ω_0. Let $\mathcal{A}(\Omega, \mathbb{F})$ denote the linear space of all $\mathbb{F}$-valued holomorphic functions on Ω. Define $U_\Gamma : \mathcal{H} \to \mathcal{A}(\Omega_0^*, \mathbb{F})$ by

$$(U_\Gamma x)(w) = \Gamma(w)^* x, \; x \in \mathcal{H}, \; w \in \Omega_0. \tag{2.6}$$

Define a sesqui-linear form on $\mathcal{H}_\Gamma = \operatorname{ran} U_\Gamma$ by $\langle U_\Gamma f, U_\Gamma g \rangle_\Gamma = \langle f, g \rangle$, $f, g \in \mathcal{H}$. The map U_Γ is linear and injective. Hence $\mathcal{H}_\Gamma$ is a Hilbert space of $\mathbb{F}$-valued holomorphic functions on Ω_0^* with inner product $\langle \cdot, \cdot \rangle_\Gamma$ and U_Γ is unitary. Then it is easy to verify the following (cf. [11, Remarks 2.6]).

a) $K(z, w) = \Gamma(\bar{z})^* \Gamma(\bar{w})$, $z, w \in \Omega_0^*$ is the reproducing kernel for the Hilbert space $\mathcal{H}_\Gamma$.
b) $M_i^* U_\Gamma = U_\Gamma T_i$, where $(M_i f)(z) = z_i f(z)$, $z = (z_1, \dots, z_m) \in \Omega$.

An nnd kernel K for which $K(z, w_0) = I$ for all $z \in \Omega_0^*$ and some $w_0 \in \Omega$ is said to be normalized at w_0.

For $1 \le i \le m$, suppose that the operators $M_i : \mathcal{H} \to \mathcal{H}$ are bounded. Then it is easy to verify that for each fixed $w \in \Omega$, and $1 \le i \le m$,

$$M_i^* K(\cdot, w)\eta = \bar{w}_i K(\cdot, w)\eta \text{ for } \eta \in \mathbb{F}. \tag{2.7}$$

Differentiating (2.3), we also obtain the following extension of the reproducing property:

$$\langle (\partial_i^j f)(w), \eta \rangle = \langle f, \bar{\partial}_i^j K(\cdot, w)\eta \rangle \text{ for } 1 \le i \le m, \; j \ge 0, \; w \in \Omega, \; \eta \in \mathbb{F}, \; f \in \mathcal{H}. \tag{2.8}$$

Let $\boldsymbol{M} = (M_1, \dots, M_m)$ be the commuting m-tuple of multiplication operators and let $\boldsymbol{M}^*$ be the m-tuple $(M_1^*, \dots, M_m^*)$. It then follows from (2.7) that the eigenspace of the m-tuple $\boldsymbol{M}^*$ at $w \in \Omega^* \subseteq \mathbb{C}^m$ contains the n-dimensional subspace $\operatorname{r}an\, K(\cdot, \bar{w}) \subseteq \mathcal{H}$.

One may impose additional conditions on K to ensure that $\boldsymbol{M}$ is in $\mathrm{B}_n(\Omega^*)$. Assume that $K(w, w)$ is invertible for $w \in \Omega$. Fix $w_0 \in \Omega$ and note that $K(z, w_0)$

is invertible for z in some neighborhood $\Omega_0 \subseteq \Omega$ of w_0. Let K_{res} be the restriction of K to $\Omega_0 \times \Omega_0$. Define a kernel function K_0 on Ω_0 by

$$K_0(z,w) = \varphi(z)K(z,w)\varphi(w)^*, \ z,w \in \Omega_0, \tag{2.9}$$

where $\varphi(z) = K_{\mathrm{res}}(w_0,w_0)^{1/2}K_{\mathrm{res}}(z,w_0)^{-1}$. The kernel K_0 is said to be *normalized* at 0 and is characterized by the property $K_0(z,w_0) = I$ for all $z \in \Omega_0$. Let $\boldsymbol{M}_0$ denote the m-tuple of multiplication operators on the Hilbert space $\mathcal{H}$. It is not hard to establish the unitary equivalence of the two m-tuples $\boldsymbol{M}$ and $\boldsymbol{M}_0$ as in (cf. [11, Lemma 3.9 and Remark 3.8]). First, the restriction map res : $f \to f_{\mathrm{res}}$, which restricts a function in $\mathcal{H}$ to Ω_0 is a unitary map intertwining the m-tuple $\boldsymbol{M}$ on $\mathcal{H}$ with the m-tuple $\boldsymbol{M}$ on $\mathcal{H}_{\mathrm{res}}$ = ran res. The Hilbert space $\mathcal{H}_{\mathrm{res}}$ is a reproducing kernel Hilbert space with reproducing kernel K_{res}. Second, suppose that the m-tuples $\boldsymbol{M}$ defined on two different reproducing kernel Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ are in $\mathrm{B}_n(\Omega)$ and $X : \mathcal{H}_1 \to \mathcal{H}_2$ is a bounded operator intertwining these two operator tuples. Then X must map the joint kernel of one tuple in to the other, that is, $XK_1(\cdot,w)\boldsymbol{x} = K_2(\cdot,w)\Phi(w)\boldsymbol{x}$, $\boldsymbol{x} \in \mathbb{C}^n$, for some function $\Phi : \Omega \to \mathbb{C}^{n\times n}$. Assuming that the kernel functions K_1 and K_2 are holomorphic in the first and anti-holomorphic in the second variable, it follows, again as in [11, pp. 472], that Φ is anti-holomorphic. An easy calculation then shows that X^* is the multiplication operator $M_{\bar{\Phi}^{\mathrm{tr}}}$. If the two operator tuples are unitarily equivalent then there exists an unitary operator U intertwining them. Hence U^* must be of the form M_Ψ for some holomorphic function Ψ such that $\overline{\Psi(w)}^{\mathrm{tr}}$ maps the joint kernel of $(\boldsymbol{M} - w)^*$ isometrically onto the joint kernel of $(\boldsymbol{M} - w)^*$ for all $w \in \Omega$. The unitarity of U is equivalent to the relation $K_1(\cdot,w)\boldsymbol{x} = U^*K_2(\cdot,w)\overline{\Psi(w)}^{\mathrm{tr}}\boldsymbol{x}$ for all $w \in \Omega$ and $\boldsymbol{x} \in \mathbb{C}^n$. It then follows that

$$K_1(z,w) = \Psi(z)K_2(z,w)\overline{\Psi(w)}^{\mathrm{tr}}, \tag{2.10}$$

where $\Psi : \Omega_0 \subseteq \Omega \to \mathcal{GL}(\mathbb{F})$ is some holomorphic function. Here, $\mathcal{GL}(\mathbb{F})$ denotes the group of all invertible linear transformations on $\mathbb{F}$.

Conversely, if two kernels are related as above then the corresponding tuples of multiplication operators are unitarily equivalent since

$$M_i^* K(\cdot,w)\zeta = \bar{w}_i K(\cdot,w)\zeta, \ w \in \Omega, \ \zeta \in \mathbb{F},$$

where $(M_i f)(z) = z_i f(z)$, $f \in \mathcal{H}$ for $1 \le i \le m$.

Remark 2.3. *We observe that if there is a self adjoint operator X commuting with the m-tuple $\boldsymbol{M}$ on the Hilbert space $\mathcal{H}$ then we must have the relation*

$$\overline{\Phi(z)}^{\mathrm{tr}} K(z,w) = K(z,w)\Phi(w)$$

for some anti-holomorphic function $\Phi : \Omega \to \mathbb{C}^{n\times n}$. Hence if the kernel K is normalized then any projection P commuting with the m-tuple $\boldsymbol{M}$ is induced by a constant function Φ such that $\Phi(0)$ is an ordinary projection on $\mathbb{C}^n$.

In conclusion, what is said above shows that a m-tuple of operators in $\mathrm{B}_n(\Omega^*)$ admits a representation as the adjoint of a m-tuple of multiplication operators on a reproducing kernel Hilbert space of $\mathbb{F}$-valued holomorphic functions on Ω_0, where the reproducing kernel K may be assumed to be normalized. Conversely, the adjoint of the m-tuple of multiplication operators on the reproducing kernel Hilbert space associated with a normalized kernel K on Ω belongs to $\mathrm{B}_n(\Omega^*)$ if certain additional conditions are imposed on K (cf. [11]).

Our interest in the class $\mathrm{B}_n(\Omega)$ lies in the fact that the Cowen-Douglas theorem [9] provides a complete set of unitary invariants for operators which belong to this class. However, these invariants are somewhat intractable. Besides, often it is not easy to verify that a given operator is in the class $\mathrm{B}_n(\Omega)$. Although, we don't use the complete set of invariants that [9] provides, it is useful to ensure that the homogeneous operators that arise from the jet construction are in this class.

2.3. The jet construction

Let $\mathcal{M}$ be a Hilbert module over the algebra $\mathcal{A}(\Omega)$ for Ω a bounded domain in $\mathbb{C}^m$. Let $\mathcal{M}_k$ be the submodule of functions in $\mathcal{M}$ vanishing to order $k > 0$ on some analytic hyper-surface $\mathcal{Z}$ in Ω – the zero set of a holomorphic function φ in $\mathcal{A}(\Omega)$. A function f on Ω is said to vanish to order k on $\mathcal{Z}$ if it can be written $f = \varphi^k g$ for some holomorphic function g. The quotient module $\mathcal{Q} = \mathcal{M} \ominus \mathcal{M}_k$ has been characterized in [12]. This was done by a generalization of the approach in [2] to allow vector-valued kernel Hilbert modules. The basic result in [12] is that $\mathcal{Q}$ can be characterized as such a vector-valued kernel Hilbert space over the algebra $\mathcal{A}(\Omega)|_{\mathcal{Z}}$ of the restriction of functions in $\mathcal{A}(\Omega)$ to $\mathcal{Z}$ and multiplication by φ acts as a nilpotent operator of order k.

For a fixed integer $k > 0$, in this realization, $\mathcal{M}$ consists of $\mathbb{C}^k$-valued holomorphic functions, and there is an $\mathbb{C}^{k\times k}$-valued function $K(z,w)$ on $\Omega \times \Omega$ which is holomorphic in z and anti-holomorphic in w such that

(1) $K(\cdot,w)v$ is in $\mathcal{M}$ for w in Ω and v in $\mathbb{C}^k$;
(2) $\langle f, K(\cdot,w)v\rangle_{\mathcal{M}} = \langle f(w), v\rangle_{\mathbb{C}^k}$ for f in $\mathcal{M}$, w in Ω and v in $\mathbb{C}^k$; and
(3) $\mathcal{A}(\Omega)\mathcal{M} \subset \mathcal{M}$.

If we assume that $\mathcal{M}$ is in the class $\mathrm{B}_1(\Omega)$, then it is possible to describe the quotient module via a jet construction along the normal direction to the hypersurface $\mathcal{Z}$. The details are in [12]. In this approach, to every positive definite kernel $K : \Omega \times \Omega \to \mathbb{C}$, we associate a kernel $JK = \left(\!\left(\partial_1^i \bar{\partial}_1{}^j K\right)\!\right)_{i,j=0}^{k-1}$, where ∂_1 denotes differentiation along the normal direction to $\mathcal{Z}$. Then we may equip

$$J\mathcal{M} = \Big\{\mathbf{f} := \sum_{i=0}^{k-1} \partial_1^i f \otimes \varepsilon_i \in \mathcal{M} \otimes \mathbb{C}^k : f \in \mathcal{M}\Big\},$$

where $\varepsilon_0, \dots, \varepsilon_{k-1}$ are standard unit vectors in $\mathbb{C}^k$, with a Hilbert space structure via the kernel JK. The module action is defined by $\mathbf{f} \mapsto \mathbb{J}\mathbf{f}$ for $\mathbf{f} \in J\mathcal{M}$, where $\mathbb{J}$

is the array

$$\mathbb{J} = \begin{pmatrix} 1 & \dots & \dots & \dots & \dots & 0 \\ \partial_1 & 1 & & & & \vdots \\ \vdots & & \ddots & & & \vdots \\ \vdots & \binom{l}{j}\partial_1^{\ell-j} & & 1 & & \vdots \\ \vdots & & & & \ddots & 0 \\ \partial_1^{k-1} & \dots & \dots & \dots & \dots & 1 \end{pmatrix}$$

with $0 \leq \ell, j \leq k-1$. The module $J\mathcal{M}_{|\text{res } \mathcal{Z}}$ which is the restriction of $J\mathcal{M}$ to $\mathcal{Z}$ is then shown to be isomorphic to the quotient module $\mathcal{M} \ominus \mathcal{M}_k$.

We illustrate these results by means of an example. Let $\mathcal{M}^{(\alpha,\beta)}$ be the Hilbert module which corresponds to the reproducing kernel

$$B^{(\alpha,\beta)}(z,w) = \frac{1}{(1-z_1\bar{w}_1)^{\alpha}} \frac{1}{(1-z_2\bar{w}_2)^{\beta}},$$

$(z_1,z_2) \in \mathbb{D}^2$ and $(w_1,w_2) \in \mathbb{D}^2$. Let $\mathcal{M}_2^{(\alpha,\beta)}$ be the subspace of all functions in $\mathcal{M}^{(\alpha,\beta)}$ which vanish to order 2 on the diagonal $\{(z,z) : z \in \mathbb{D}\} \subseteq \mathbb{D} \times \mathbb{D}$. The quotient module $\mathcal{Q} := \mathcal{M}^{(\alpha,\beta)} \ominus \mathcal{M}_2^{(\alpha,\beta)}$ was described in [14] using an orthonormal basis for the quotient module $\mathcal{Q}$. This includes the calculation of the compression of the two operators, $M_1 : f \mapsto z_1 f$ and $M_2 : f \mapsto z_2 f$ for $f \in \mathcal{M}^{(\alpha,\beta)}$, on the quotient module $\mathcal{Q}$ (block weighted shift operators) with respect to this orthonormal basis. These are homogeneous operators in the class $\mathrm{B}_2(\mathbb{D})$ which were first discovered by Wilkins [22].

In [14], an orthonormal basis $\left\{e_p^{(1)}, e_p^{(2)}\right\}_{p=0}^{\infty}$ was constructed in the quotient module $\mathcal{M} \ominus \mathcal{M}_2^{(\alpha,\beta)}$. It was shown that the matrix

$$M_p^{(1)} = \begin{pmatrix} \frac{\binom{-(\alpha+\beta)}{p}^{1/2}}{\binom{-(\alpha+\beta)}{p+1}^{1/2}} & 0 \\ \left(\frac{\beta}{\alpha}\right)^{1/2} \frac{(\alpha+\beta+1)^{1/2}}{\left((\alpha+\beta+p)(\alpha+\beta+p+1)\right)^{1/2}} & \frac{\binom{-(\alpha+\beta+2)}{p-1}^{1/2}}{\binom{-(\alpha+\beta+2)}{p}^{1/2}} \end{pmatrix}$$

represents the operator M_1 which is multiplication by z_1 with respect to the orthonormal basis $\{e_p^{(1)}, e_p^{(2)}\}_{p=0}^{\infty}$. Similarly,

$$M_p^{(2)} = \begin{pmatrix} \frac{\binom{-(\alpha+\beta)}{p}^{1/2}}{\binom{-(\alpha+\beta)}{p+1}^{1/2}} & 0 \\ -\left(\frac{\alpha}{\beta}\right)^{1/2} \frac{(\alpha+\beta+1)^{1/2}}{\left((\alpha+\beta+p)(\alpha+\beta+p+1)\right)^{1/2}} & \frac{\binom{-(\alpha+\beta+2)}{p-1}^{1/2}}{\binom{-(\alpha+\beta+2)}{p}^{1/2}} \end{pmatrix}$$

represents the operator M_2 which is multiplication by z_2 with respect to the orthonormal basis $\{e_p^{(1)}, e_p^{(2)}\}_{p=0}^{\infty}$. Therefore, we see that $Q_1^{(p)} = \frac{1}{2}(M_1^{(p)} - M_2^{(p)})$ is a nilpotent matrix of index 2 while $Q_2^{(p)} = \frac{1}{2}(M_1^{(p)} + M_2^{(p)})$ is a diagonal matrix in case $\beta = \alpha$. These definitions naturally give a pair of operators Q_1 and Q_2 on the quotient module $\mathcal{Q}^{(\alpha,\beta)}$. Let f be a function in the bi-disc algebra $\mathcal{A}(\mathbb{D}^2)$ and

$$f(u_1,u_2) = f_0(u_1) + f_1(u_1)u_2 + f_2(u_1)u_2^2 + \cdots$$

be the Taylor expansion of the function f with respect to the coordinates $u_1 = \frac{z_1+z_2}{2}$ and $u_2 = \frac{z_1-z_2}{2}$. Now the module action for $f \in \mathcal{A}(\mathbb{D}^2)$ in the quotient module $\mathcal{Q}^{(\alpha,\beta)}$ is then given by

$$\begin{aligned} f\cdot h &= f(Q_1,Q_2)\cdot h \\ &= f_0(Q_1)\cdot h + f_1(Q_1)Q_2\cdot h \\ &\stackrel{\text{def}}{=} \begin{pmatrix} f_0 & 0 \\ f_1 & f_0 \end{pmatrix}\cdot\begin{pmatrix} h_1 \\ h_2 \end{pmatrix}, \end{aligned}$$

where $h = \binom{h_1}{h_2} \in \mathcal{Q}^{(\alpha,\beta)}$ is the unique decomposition obtained from realizing the quotient module as the direct sum $\mathcal{Q}^{(\alpha,\beta)} = \big(\mathcal{M}^{(\alpha,\beta)} \ominus \mathcal{M}_1^{(\alpha,\beta)}\big) \oplus \big(\mathcal{M}_1^{(\alpha,\beta)} \ominus \mathcal{M}_2^{(\alpha,\beta)}\big)$, where $\mathcal{M}_i^{(\alpha,\beta)}$, $i = 1,2$, are the submodules in $\mathcal{M}^{(\alpha,\beta)}$ consisting of all functions vanishing on $\mathcal{Z}$ to order 1 and 2 respectively.

We now calculate the curvature $\mathcal{K}^{(\alpha,\beta)}$ for the bundle $E^{(\alpha,\beta)}$ corresponding to the metric $B^{(\alpha,\beta)}(\mathbf{u},\mathbf{u})$, where $\mathbf{u} = (u_1,u_2) \in \mathbb{D}^2$. The curvature $\mathcal{K}^{(\alpha,\beta)}$ is easy to compute:

$$\mathcal{K}^{(\alpha,\beta)}(u_1,u_2) = (1-|u_1+u_2|^2)^{-2}\begin{pmatrix} \alpha & \alpha \\ \alpha & \alpha \end{pmatrix} + (1-|u_1-u_2|^2)^{-2}\begin{pmatrix} \beta & -\beta \\ -\beta & \beta \end{pmatrix}.$$

The restriction of the curvature to the hyper-surface $\{u_2 = 0\}$ is

$$\mathcal{K}^{(\alpha,\beta)}(u_1,u_2)_{|u_2=0} = (1-|u_1|^2)^{-2}\begin{pmatrix} \alpha+\beta & \alpha-\beta \\ \alpha-\beta & \alpha+\beta \end{pmatrix},$$

where $u_1 \in \mathbb{D}$. Thus we find that if $\alpha = \beta$, then the curvature is of the form $2\alpha(1-|u_1|^2)^{-2}I_2$.

We now describe the unitary map which is basic to the construction of the quotient module, namely,

$$h \mapsto \sum_{\ell=0}^{k-1} \partial_1^\ell h \otimes \varepsilon_\ell\Big|_{\text{res }\triangle}$$

for $h \in \mathcal{M}^{(\alpha,\beta)}$. For $k = 2$, it is enough to describe this map just for the orthonormal basis $\{e_p^{(1)}, e_p^{(2)} : p \geq 0\}$. A simple calculation shows that

$$\begin{aligned} e_p^{(1)}(z_1,z_2) &\mapsto \begin{pmatrix} \binom{-(\alpha+\beta)}{p}^{1/2} z_1^p \\ \beta\sqrt{\frac{p}{\alpha+\beta}}\binom{-(\alpha+\beta+1)}{p-1}^{1/2} z_1^{p-1} \end{pmatrix} \\ e_p^{(2)}(z_1,z_2) &\mapsto \begin{pmatrix} 0 \\ \sqrt{\frac{\alpha\beta}{\alpha+\beta}}\binom{-(\alpha+\beta+2)}{p-1}^{1/2} z_1^{p-1} \end{pmatrix}. \end{aligned} \tag{2.11}$$

This allows us to compute the 2×2 matrix-valued kernel function

$$K_{\mathcal{Q}}(\mathbf{z}, \mathbf{w}) = \sum_{p=0}^{\infty} e_p^{(1)}(\mathbf{z}) e_p^{(1)}(\mathbf{w})^* + \sum_{p=0}^{\infty} e_p^{(2)}(\mathbf{z}) e_p^{(2)}(\mathbf{w})^*, \; \mathbf{z}, \mathbf{w} \in \mathbb{D}^2$$

corresponding to the quotient module. Recall that $S(z, w) := (1 - z\bar{w})^{-1}$ is the Szegö kernel for the unit disc $\mathbb{D}$. We set $\mathbb{S}^r(z) := S(z, z)^r = (1 - |z|^2)^{-r}$, $r > 0$. A straight forward computation shows that

$$\begin{aligned}
& K_{\mathcal{Q}}(\mathbf{z}, \mathbf{z})_{|\text{res } \triangle} \\
&= \begin{pmatrix} \mathbb{S}(z)^{\alpha+\beta} & \beta z \mathbb{S}(z)^{\alpha+\beta+1} \\ \beta \bar{z} \mathbb{S}(z)^{\alpha+\beta+1} & \frac{\beta^2}{\alpha+\beta} \frac{d}{d|z|^2}\left(|z|^2 \mathbb{S}(z)^{\alpha+\beta+1}\right) + \frac{\beta\alpha}{\alpha+\beta} \mathbb{S}(z)^{\alpha+\beta+2} \end{pmatrix} \\
&= \left(\!\!\left(\mathbb{S}(z_1)^{\alpha} \partial^i \bar{\partial}^j \mathbb{S}(z_2)^{\beta}{}_{|\text{res } \triangle} \right)\!\!\right)_{i,j=0,1} \\
&= (JK)(\mathbf{z}, \mathbf{z})_{|\text{res } \mathcal{Z}}, \;\; \mathbf{z} \in \mathbb{D}^2,
\end{aligned}$$

where $\triangle = \{(z, z) \in \mathbb{D}^2 : z \in \mathbb{D}\}$. These calculations give an explicit illustration of one of the main theorems on quotient modules from [12, Theorem 3.4].

3. Multiplier representations

Let G be a locally compact second countable (lcsc) topological group acting transitively on the domain $\Omega \subseteq \mathbb{C}^m$. Let $\mathbb{C}^{n \times n}$ denote the set of $n \times n$ matrices over the complex field $\mathbb{C}$. We start with a cocycle J, that is, a holomorphic map $J_g : \Omega \to \mathbb{C}^{n \times n}$ satisfying the cocycle relation

$$J_{gh}(z) = J_h(z) J_g(h \cdot z), \text{ for all } g, h \in G, \; z \in \Omega. \tag{3.12}$$

Let $\mathrm{Hol}(\Omega, \mathbb{C}^n)$ be the linear space consisting of all holomorphic functions on Ω taking values in $\mathbb{C}^n$. We then obtain a natural (left) action U of the group G on $\mathrm{Hol}(\Omega, \mathbb{C}^n)$:

$$(U_{g^{-1}} f)(z) = J_g(z) f(g \cdot z), \; f \in \mathrm{Hol}(\Omega, \mathbb{C}^n), \; z \in \Omega. \tag{3.13}$$

Let e be the identity element of the group G. Note that the cocycle condition (3.12) implies, among other things, $J_e(z) = J_e(z)^2$ for all $z \in \Omega$.

Let $\mathbb{K} \subseteq G$ be the compact subgroup which is the stabilizer of 0. For h, k in $\mathbb{K}$, we have $J_{kh}(0) = J_h(0) J_k(0)$ so that $k \mapsto J_k(0)^{-1}$ is a representation of $\mathbb{K}$ on $\mathbb{C}^n$.

A positive definite kernel K on Ω defines an inner product on some linear subspace of $\mathrm{Hol}(\Omega, \mathbb{C}^n)$. The completion of this subspace is then a Hilbert space of holomorphic functions on Ω (cf. [2]). The natural action of the group G described above is seen to be unitary for an appropriate choice of such a kernel. Therefore, we first discuss these kernels in some detail.

Let $\mathcal{H}$ be a functional Hilbert space consisting of holomorphic functions on Ω possessing a reproducing kernel K. We will always assume that the m-tuple of multiplication operators $\boldsymbol{M} = (M_1, \ldots, M_m)$ on the Hilbert space $\mathcal{H}$ is bounded.

We also define the action of the group G on the space of multiplication operators – $g \cdot M_f = M_{f\circ g}$ for $f \in \mathcal{A}(\Omega)$ and $g \in G$. In particular, we have $g \cdot \boldsymbol{M} = \boldsymbol{M}_g$. We will say that the m-tuple $\boldsymbol{M}$ is *G-homogeneous* if the operator $g \cdot \boldsymbol{M}$ is unitarily equivalent to $\boldsymbol{M}$ for all $g \in G$. $g \mapsto U_{g^{-1}}$ defined in (3.13) leaves $\mathcal{H}$ invariant. The following theorem says that the reproducing kernel of such a Hilbert space must be *quasi invariant* under the G action.

A version of the following Theorem appears in [16] for the unit disc. However, the proof here, which is taken from [16], is for a more general domain Ω in $\mathbb{C}^m$.

Theorem 3.1. *Suppose that $\mathcal{H}$ is a Hilbert space which consists of holomorphic functions on Ω and possesses a reproducing kernel K on which the m-tuple $\boldsymbol{M}$ is irreducible and bounded. Then the following are equivalent.*

1. *The m-tuple $\boldsymbol{M}$ is G-homogeneous.*
2. *The reproducing kernel K of the Hilbert space $\mathcal{H}$ transforms, for some cocycle $J_g : \Omega \to \mathbb{C}^{n\times n}$, according to the rule*
$$K(z,w) = J_g(z)K(g\cdot z, g\cdot w)J_g(w)^*,\ z,w\in\Omega.$$
3. *The operator $U_{g^{-1}} : f \mapsto M_{J_g} f\circ g$ for $f \in \mathcal{H}$ is unitary.*

Proof. Assuming that K is quasi-invariant, that is, K satisfies the transformation rule, we see that the linear transformation U defined in (3.13) is unitary. To prove this, note that

$$\begin{aligned}\langle U_{g^{-1}}K(z,w)\boldsymbol{x}, U_{g^{-1}}K(z,w')\boldsymbol{y}\rangle &= \langle J_g(z)K(g\cdot z,w)\boldsymbol{x}, J_g(z)K(g\cdot z,w')\boldsymbol{y}\rangle\\ &= \langle K(z,\tilde w)J_g(\tilde w)^{*-1}\boldsymbol{x}, K(z,\tilde w')J_g(\tilde w')^{*-1}\boldsymbol{y}\rangle\\ &= \langle K(\tilde w',\tilde w)J_g(\tilde w)^{*-1}\boldsymbol{x}, J_g(\tilde w')^{*-1}\boldsymbol{y}\rangle\\ &= \langle J_g(\tilde w')^{-1}K(\tilde w',\tilde w)J_g(\tilde w)^{*-1}\boldsymbol{x}, \boldsymbol{y}\rangle\\ &= \langle K(g\cdot\tilde w', g\cdot\tilde w)\boldsymbol{x},\boldsymbol{y}\rangle,\end{aligned}$$

where $\tilde w = g^{-1}\cdot w$ and $\tilde w' = g^{-1}\cdot w'$. Hence

$$\langle K(g\cdot\tilde w', g\cdot\tilde w)\boldsymbol{x},\boldsymbol{y}\rangle = \langle K(w',w)\boldsymbol{x},\boldsymbol{y}\rangle.$$

It follows that the map $U_{g^{-1}}$ is isometric. On the other hand, if U of (3.13) is unitary then the reproducing kernel K of the Hilbert space $\mathcal{H}$ satisfies

$$K(z,w) = J_g(z)K(g\cdot z, g\cdot w)J_g(w)^*. \tag{3.14}$$

This follows from the fact that the reproducing kernel has the expansion (2.5) for some orthonormal basis $\{e_\ell : \ell \geq 0\}$ in $\mathcal{H}$. The uniqueness of the reproducing kernel implies that the expansion is independent of the choice of the orthonormal basis. Consequently, we also have $K(z,w) = \sum_{\ell=0}(U_{g^{-1}}e_\ell)(z)(U_{g^{-1}}e_\ell)(w)^*$ which verifies the equation (3.14). Thus we have shown that U is unitary if and only if the reproducing kernel K transforms according to (3.14).

We now show that the m-tuple $\boldsymbol{M}$ is homogeneous if and only if $f \mapsto M_{J_g} f\circ g$ is unitary. The eigenvector at w for $g.\boldsymbol{M}$ is clearly $K(\cdot, g^{-1}\cdot w)$. It is not hard, using the unitary operator U_Γ in (2.6), to see that that $g^{-1}\cdot\boldsymbol{M}$ is unitarily equivalent to

$\boldsymbol{M}$ on a Hilbert space $\mathcal{H}_g$ whose reproducing kernel is $K_g(z,w) = K(g\cdot z, g\cdot w)$ and the unitary U_Γ is given by $f \mapsto f \circ g$ for $f \in \mathcal{H}$. However, the homogeneity of the m-tuple $\boldsymbol{M}$ is equivalent to the existence of a unitary operator intertwining the m-tuple of multiplication on the two Hilbert spaces $\mathcal{H}$ and $\mathcal{H}_g$. As we have pointed out in Section 2.2, this unitary operator is induced by a multiplication operator M_{J_g}, where J_g is a holomorphic function (depends on g) such that $K_g(z,w) = J_g(z)K(z,w)\overline{J_g(w)}^{\mathrm{tr}}$. The composition of these two unitaries is $f \mapsto M_{J_g} f \circ g$ and is therefore a unitary. □

The discussion below and the corollary following it is implicit in [16]. Let g_z be an element of G which maps 0 to z, that is $g_z \cdot 0 = z$. We could then try to define possible kernel functions $K : \Omega \times \Omega \to \mathbb{C}^{n\times n}$ satisfying the transformation rule (3.14) via the requirement

$$K(g_z \cdot 0, g_z \cdot 0) = (J_{g_z}(0))^{-1} K(0,0)(J_{g_z}(0)^*)^{-1}, \tag{3.15}$$

choosing any positive operator $K(0,0)$ on $\mathbb{C}^n$ which commutes with $J_k(0)$ for all $k \in \mathbb{K}$. Then the equation (3.15) determines the function K unambiguously as long as $J_k(0)$ is unitary for $k \in \mathbb{K}$. Pick $g \in G$ such that $g \cdot 0 = z$. Then $g = g_z k$ for some $k \in \mathbb{K}$. Hence

$$\begin{aligned} K(g_z k \cdot 0, g_z k \cdot 0) &= (J_{g_z k}(0))^{-1} K(0,0)(J_{g_z k}(0)^*)^{-1} \\ &= \big(J_k(0) J_{g_z}(k\cdot 0)\big)^{-1} K(0,0)\big(J_{g_z}(k\cdot 0)^* J_k(0)^*\big)^{-1} \\ &= (J_{g_z}(0))^{-1}(J_k(0))^{-1} K(0,0)(J_k(0)^*)^{-1}(J_{g_z}(0)^*)^{-1} \\ &= (J_{g_z}(0))^{-1} K(0,0)(J_{g_z}(0)^*)^{-1} \\ &= K(g_z \cdot 0, g_z \cdot 0). \end{aligned}$$

Given the definition (3.15), where the choice of $K(0,0) = A$ involves as many parameters as the number of irreducible representations of the form $k \mapsto J_k(0)^{-1}$ of the compact group $\mathbb{K}$, one can polarize (3.15) to get $K(z,w)$. In this approach, one has to find a way of determining if K is non-negative definite, or for that matter, if $K(\cdot, w)$ is holomorphic on all of Ω for each fixed but arbitrary $w \in \Omega$. However, it is evident from the definition (3.15) that

$$\begin{aligned} K(h\cdot z, h\cdot z) &= J_h(g_z\cdot 0)^{-1} J_{g_z}(0)^{-1} A J_{g_z}(0)^{*-1}(J_h(g_z\cdot 0)^*)^{-1} \\ &= J_h(z)^{-1} K(z,z) J_h(z)^{*-1} \end{aligned}$$

for all $h \in G$. Polarizing this equality, we obtain

$$K(h\cdot z, h\cdot w) = J_h(z)^{-1} K(z,w) J_h(w)^{*-1}$$

which is the identity (3.14). It is also clear that the linear span of the set $\{K(\cdot, w)\zeta : \ w \in \Omega,\ \zeta \in \mathbb{C}^n\}$ is stable under the action (3.13) of G:

$$g \mapsto J_g(z) K(g\cdot z, w)\zeta = K(z, g^{-1}\cdot w) J_g(g^{-1}w)^{*-1}\zeta,$$

where $J_g(g^{-1}w)^{*-1}\zeta$ is a fixed element of $\mathbb{C}^n$.

Corollary 3.2. *If $J : G \times \Omega \to \mathbb{C}^{n\times n}$ is a cocycle and g_z is an element of G which maps 0 to z then the kernel $K : \Omega \times \Omega \to \mathbb{C}^{n\times n}$ defined by the requirement*

$$K(g_z \cdot 0, g_z \cdot 0) = (J_{g_z}(0))^{-1} K(0,0) (J_{g_z}(0)^*)^{-1}$$

is quasi-invariant, that is, it transforms according to (3.14).

4. Irreducibility

In Section 2.2, we have already pointed out that any Hilbert space $\mathcal{H}$ of scalar-valued holomorphic functions on $\Omega \subset \mathbb{C}^m$ with a reproducing kernel K determines a line bundle $\mathcal{E}$ on $\Omega^* := \{\bar{w} : w \in \Omega\}$. The fibre of E at $\bar{w} \in \Omega^*$ is spanned by $K(.,w)$. We can now construct a rank $(n+1)$ vector bundle $J^{(n+1)}\mathcal{E}$ over Ω^*. A holomorphic frame for this bundle is $\{\bar{\partial}_2^l K(.,w) : 0 \le l \le k, w \in \Omega\}$, and as usual, this frame determines a metric for the bundle which we denote by $J^{(n+1)}K$, where

$$J^{(n+1)}K(w,w) = \left(\!\left(\langle \bar{\partial}_2^j K(.,w), \bar{\partial}_2^i K(.,w)\rangle\right)\!\right)_{i,j=0}^n = \left(\!\left(\bar{\partial}_2^j \partial_2^i K(w,w)\right)\!\right)_{i,j=0}^n, w \in \Omega.$$

Recall that the kernel function on $\mathbb{D}^2$, $B^{(\alpha,\beta)} : \mathbb{D}^2 \times \mathbb{D}^2 \longrightarrow \mathbb{C}$ is defined by

$$B^{(\alpha,\beta)}(z,w) = (1 - z_1\bar{w_1})^{-\alpha}(1 - z_2\bar{w_2})^{-\beta},$$

for $z = (z_1, z_2) \in \mathbb{D}^2$ and $w = (w_1, w_2) \in \mathbb{D}^2$, $\alpha, \beta > 0$. Take $\Omega = \mathbb{D}^2, K = B^{(\alpha,\beta)}$. Notice that the Hilbert space $\mathcal{M}^{(\alpha,\beta)}$ corresponding to the kernel function $B^{(\alpha,\beta)}$ is the tensor product of the two Hilbert spaces $\mathcal{M}^{(\alpha)}$ and $\mathcal{M}^{(\beta)}$. These are determined by the two kernel functions $\mathbb{S}^\alpha(z,w) = (1 - z\bar{w})^{-\alpha}$ and $\mathbb{S}^\beta(z,w) = (1 - z\bar{w})^{-\beta}$, $z, w \in \mathbb{D}$, respectively.

It follows from [12] that $h_{n+1}(z) = J^{(n+1)}B^{(\alpha,\beta)}(z,z)_{|\text{res } \triangle}$ is a metric for the Hermitian anti-holomorphic vector bundle $J^{(n+1)}\mathcal{E}_{|\text{res } \triangle}$ over $\triangle = \{(z,z) : z \in \mathbb{D}\} \subseteq \mathbb{D}^2$. However, $J^{(n+1)}\mathcal{E}_{|\text{res } \triangle}$ is a Hermitian holomorphic vector bundle over $\triangle^* = \{(\bar{z},\bar{z}) : z \in \mathbb{D}\}$, that is, $\bar{z}$ is the holomorphic variable in this description. Thus $\partial f = 0$ if and only if f is holomorphic on $\triangle^*$. To restore the usual meaning of ∂ and $\bar{\partial}$, we interchange the roles of z and $\bar{z}$ in the metric which amounts to replacing h_{n+1} by its transpose.

As shown in [12], this Hermitian anti-holomorphic vector bundle $J^{(n+1)}\mathcal{E}_{|\text{res } \triangle}$ defined over the diagonal subset $\triangle$ of the bidisc $\mathbb{D}^2$ gives rise to a reproducing kernel Hilbert space $J^{(n+1)}\mathcal{H}$. The reproducing kernel for this Hilbert space is $J^{(n+1)}B^{(\alpha,\beta)}(z,w)$ which is obtained by polarizing $J^{(n+1)}B^{(\alpha,\beta)}(z,z) = h_{n+1}(z)^t$.

Lemma 4.1. *Let α, β be two positive real numbers and $n \ge 1$ be an integer. Let $\mathcal{M}_n$ be the ortho-complement of the subspace of $\mathcal{M}^{(\alpha)} \otimes \mathcal{M}^{(\beta)}$ (viewed as a Hilbert space of analytic functions on the bi-disc $\mathbb{D} \times \mathbb{D}$) consisting of all the functions vanishing to order k on the diagonally embedded unit disc $\triangle := \{(z,z) : z \in \mathbb{D}\}$. The compressions to $\mathcal{M}_n$ of $M^{(\alpha)} \otimes I$ and $I \otimes M^{(\beta)}$ are homogeneous operators with a common associated representation.*

Proof. For each real number $\alpha > 0$, let $\mathcal{M}^{(\alpha)}$ be the Hilbert space completion of the inner product space spanned by $\{f_k : k \in \mathbb{Z}^+\}$ where the f_k's are mutually orthogonal vectors with norms given by

$$\|f_k\|^2 = \frac{\Gamma(1+k)}{\Gamma(\alpha+k)},\ k \in \mathbb{Z}^+.$$

(Up to scaling of the norm, this Hilbert space may be identified, via non-tangential boundary values, with the Hilbert space of analytic functions on $\mathbb{D}$ with reproducing kernel $(z,w) \mapsto (1 - z\bar{w})^{-\alpha}$.) The representation D^+_α lives on $\mathcal{M}^{(\alpha)}$, and is given (at least on the linear span of the f_k's) by the formula

$$D^+_\alpha(\varphi^{-1})f = (\varphi')^{\alpha/2} f \circ \varphi,\ \varphi \in \text{Möb}.$$

Clearly the subspace $\mathcal{M}_n$ is invariant under the Discrete series representation $\pi := D^+_\alpha \otimes D^+_\beta$ associated with both the operators $M^{(\alpha)} \otimes I$ and $I \otimes M^{(\beta)}$. It is also co-invariant under these two operators. An application of Proposition 2.4 in [5] completes the proof of the lemma. □

The subspace $\mathcal{M}_n$ consists of those functions $f \in \mathcal{M}$ which vanish on $\triangle$ along with their first n derivatives with respect to z_2. As it turns out, the compressions to $\mathcal{M} \ominus \mathcal{M}_n$ of $M^{(\alpha)} \otimes I$ is the multiplication operator on the Hilbert space $J^{(n+1)}\mathcal{H}_{|\text{res }\triangle}$ which we denote $M^{(\alpha,\beta)}$. An application of [12, Proposition 3.6] shows that the adjoint M^* of the multiplication operator M is in $\mathrm{B}_{n+1}(\mathbb{D})$.

Theorem 4.2. *The multiplication operator $M := M^{(\alpha,\beta)}$ is irreducible.*

The proof of this theorem will be facilitated by a series of lemmas which are proved in the sequel. We set, for now, $K(z,w) = J^{(n+1)}B^{(\alpha,\beta)}(z,w)$. Let $\widetilde{K}(z,w) = K(0,0)^{-1/2}K(z,w)K(0,0)^{-1/2}$, so that $\widetilde{K}(0,0) = I$. Also, let $\widetilde{\widetilde{K}}(z,w) = \widetilde{K}(z,0)^{-1}\widetilde{K}(z,w)\widetilde{K}(0,w)^{-1}$. This ensures that $\widetilde{\widetilde{K}}(z,0) = I$ for $z \in \mathbb{D}$, that is, $\widetilde{\widetilde{K}}$ is a normalized kernel. Each of the kernels K, $\widetilde{K}$ and $\widetilde{\widetilde{K}}$ admit a power series expansion, say, $K(z,w) = \sum_{m,\,p\,\geq\,0} a_{mp}\, z^m \bar{w}^p$, $\widetilde{K}(z,w) = \sum_{m,\,p\,\geq\,0} \widetilde{a}_{mp}\, z^m \bar{w}^p$, and $\widetilde{\widetilde{K}}(z,w) = \sum_{m,\,p\,\geq\,0} \widetilde{\widetilde{a}}_{mp}\, z^m \bar{w}^p$ for $z,\ w \in \mathbb{D}$, respectively. Here the coefficients a_{mp} and $\widetilde{a}_{mp}$ and $\widetilde{\widetilde{a}}_{mp}$ are $(n+1) \times (n+1)$ matrices for $m,\ p \geq 0$. In particular, $\widetilde{a}_{mp} = K(0,0)^{-1/2}a_{mp}\,K(0,0)^{-1/2} = a_{00}^{-1/2}a_{mp}\,a_{00}^{-1/2}$ for $m, p \geq 0$. Also, let us write $K(z,w)^{-1} = \sum_{m,\,p\,\geq\,0} b_{mp}\, z^m \bar{w}^p$ and $\widetilde{K}(z,w)^{-1} = \sum_{m,\,p\,\geq\,0} \widetilde{b}_{mp}\, z^m \bar{w}^p$, $z, w \in \mathbb{D}$. Again, the coefficients b_{mp} and $\widetilde{b}_{mp}$ are $(n+1) \times (n+1)$ matrices for $m,\ p \geq 0$. However, $\widetilde{\widetilde{a}}_{00} = I$ and $\widetilde{\widetilde{a}}_{m0} = \widetilde{\widetilde{a}}_{0p} = 0$ for $m,\ p \geq 1$.

The following lemma is from [11, Theorem 3.7, Remark 3.8 and Lemma 3.9]. The proof was discussed in Section 2.2.

Lemma 4.3. *The multiplication operators on Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ with reproducing kernels $K_1(z,w)$ and $K_2(z,w)$ respectively, are unitarily equivalent if and only if $K_2(z,w) = \Psi(z)K_1(z,w)\overline{\Psi(w)}^t$, where Ψ is an invertible matrix-valued holomorphic function.*

The proof of the lemma below appears in [16, Lemma 5.2] and is discussed in Section 2.2, see Remark 2.3.

Lemma 4.4. *The multiplication operator M on the Hilbert space $\mathcal{H}$ with reproducing kernel K is irreducible if and only if there is no non-trivial projection P on $\mathbb{C}^{n+1}$ commuting with all the coefficients in the power series expansion of the normalized kernel $\widetilde{\widetilde{K}}(z,w)$.*

We will prove irreducibility of M by showing that only operators on $\mathbb{C}^{n+1}$ which commutes with all the coefficients of $\widetilde{\widetilde{K}}(z,w)$ are scalars. It turns out that the coefficients of $z^k\bar{w}$ for $2 \le k \le n+1$, that is, the coefficients $\widetilde{\widetilde{a}}_{k1}$ for $2 \le k \le n+1$ are sufficient to reach the desired conclusion.

Lemma 4.5. *The coefficient of $z^k\bar{w}$ is*

$$\widetilde{\widetilde{a}}_{k1} = \sum_{s=1}^{k} \widetilde{b}_{s0}\widetilde{a}_{k-s,1} + \widetilde{a}_{k1}$$

for $1 \le k \le n+1$.

Proof. Let us denote the coefficient of $z^k\bar{w}^l$ in the power series expansion of $\widetilde{\widetilde{K}}(z,w)$ is $\widetilde{\widetilde{a}}_{kl}$ for $k,l \ge 0$. We see that

$$\begin{aligned}
\widetilde{\widetilde{a}}_{kl} &= \sum_{s=0}^{k}\sum_{t=0}^{l} \widetilde{b}_{s0}\widetilde{a}_{k-s,l-t}\widetilde{b}_{0t} \\
&= \sum_{s=1}^{k}\sum_{t=1}^{l} \widetilde{a}_{s0}\widetilde{a}_{k-s,l-t}\widetilde{b}_{0t} + \sum_{s=1}^{k} \widetilde{b}_{s0}\widetilde{a}_{k-s,l} + \sum_{t=1}^{l} \widetilde{a}_{k,l-t}\widetilde{b}_{0t} + \widetilde{a}_{kl}
\end{aligned}$$

as $\widetilde{a}_{00} = \widetilde{b}_{00} = I$. Also,

$$\begin{aligned}
\widetilde{\widetilde{a}}_{k1} &= \sum_{s=1}^{k} \widetilde{b}_{s0}\widetilde{a}_{k-s,0}\widetilde{b}_{01} + \sum_{s=1}^{k} \widetilde{b}_{s0}\widetilde{a}_{k-s,1} + \widetilde{a}_{k0}\widetilde{b}_{01} + \widetilde{a}_{k1} \\
&= \Big(\sum_{s=0}^{k} \widetilde{b}_{s0}\widetilde{a}_{k-s,0}\Big)\widetilde{b}_{01} + \sum_{s=1}^{k} \widetilde{b}_{s0}\widetilde{a}_{k-s,1} + \widetilde{a}_{k1} \\
&= \sum_{s=1}^{k} \widetilde{b}_{s0}\widetilde{a}_{k-s,1} + \widetilde{a}_{k1}
\end{aligned}$$

as $\widetilde{b}_{00} = I$ and coefficient of z^k in $\widetilde{K}(z,w)^{-1}\widetilde{K}(z,w) = \sum_{s=0}^{k} \widetilde{b}_{s0}\widetilde{a}_{k-s,0} = 0$ for $k \ge 1$. □

Now we compute some of the coefficients of $K(z,w)$ which are useful in computing $\widetilde{\widetilde{a}}_{k1}$. In what follows, we will compute only the non-zero entries of the matrices involved, that is, *all those entries which are not specified are assumed to be zero.*

Lemma 4.6. *$(a_{00})_{kk} = k!(\beta)_k$ for $0 \le k \le n$, $(a_{m0})_{r,r+m} = \frac{(m+r)!}{m!}(\beta)_{m+r}$ and $(a_{m+1,1})_{r,r+m} = \frac{(m+r)!}{m!}(\beta)_{m+r}\big(\alpha + (1+\frac{r}{m+1})(\beta+m+r)\big)$ for $0 \le r \le n-m, 0 \le m \le n$, where $(x)_0 = 1, (x)_d = x(x+1)\dots(x+d-1)$, for any positive integer d, is the Pochhammer symbol.*

Proof. The coefficient of $z^p\bar{w}^q$ in $J^{(n+1)}B^{(\alpha,\beta)}(z,w)$ is the same as the coefficient of $z^p\bar{z}^q$ in $J^{(n+1)}B^{(\alpha,\beta)}(z,z)$. So, $(a_{00})_{kk} =$ constant term in $\bar{\partial}_2^k\partial_2^k\big(\mathbb{S}(z_1)^\alpha\mathbb{S}(z_2)^\beta\big)|_\triangle$. Now,

$$\begin{aligned}
&\bar{\partial}_2^k\partial_2^k\big(\mathbb{S}(z_1)^\alpha\mathbb{S}(z_2)^\beta\big)|_\triangle\\
&\quad= \bar{\partial}_2^k\partial_2^k\big(\mathbb{S}(z_1)^\alpha\mathbb{S}(z_2)^\beta\big)|_\triangle\\
&\quad= \mathbb{S}(z_1)^\alpha(\beta)_k\bar{\partial}_2^k\big(\mathbb{S}(z_2)^{\beta+k}\bar{z}_2^k\big)|_\triangle\\
&\quad= \big(\mathbb{S}(z_1)^\alpha(\beta)_k\sum_{l=0}^{k}\binom{k}{l}\bar{\partial}_2^{k-l}(\mathbb{S}(z_2)^{\beta+k})\bar{\partial}_2^l(\bar{z_2}^k)\big)|_\triangle\\
&\quad= \big(\mathbb{S}(z_1)^\alpha(\beta)_k\sum_{l=0}^{k}\binom{k}{l}(\beta+k)_{k-l}\mathbb{S}(z_2)^{\beta+k+(k-l)}{z_2}^{k-l}l!\binom{k}{l}\bar{z_2}^{k-l}\big)|_\triangle,
\end{aligned}$$

that is, $\big(a_{00}\big)_{kk} = k!(\beta)_k$ for $0 \le k \le n$.

We see that the coefficient of z^m in $\bar{\partial}_2^{m+r}\partial_2^r\big(\mathbb{S}(z_1)^\alpha\mathbb{S}(z_2)^\beta\big)|_\triangle$ is $(a_{m0})_{r,r+m}$. Thus

$$\begin{aligned}
&\bar{\partial}_2^{m+r}\partial_2^r\big(\mathbb{S}(z_1)^\alpha\mathbb{S}(z_2)^\beta\big)|_\triangle\\
&= \mathbb{S}(z_1)^\alpha(\beta)_r\bar{\partial}_2^{m+r}\big(\mathbb{S}(z_2)^{\beta+r}\bar{z}_2^r\big)|_\triangle\\
&= \big(\mathbb{S}(z_1)^\alpha(\beta)_r\sum_{l=0}^{m+r}\tbinom{m+r}{l}\bar{\partial}_2^{m+r-l}(\mathbb{S}(z_2)^{\beta+r})\bar{\partial}_2^l(\bar{z_2}^r)\big)|_\triangle\\
&= \big(\mathbb{S}(z_1)^\alpha(\beta)_r\sum_{l=0}^{m+r}\tbinom{m+r}{l}(\beta+r)_{m+r-l}\mathbb{S}(z_2)^{\beta+2r+m-l)}{z_2}^{m+r-l}l!\binom{r}{l}\bar{z_2}^{r-l}\big)|_\triangle.
\end{aligned}$$

Therefore, the term containing z^m occurs only when $l = r$ in the sum above, that is, $(a_{m0})_{r,r+m} = (\beta)_r\binom{m+r}{r}(\beta+r)_m r! = \frac{(m+r)!}{m!}(\beta)_{m+r}$, for $0 \le r \le n-m, 0 \le m \le n$.

The coefficient of $z^{m+1}\bar{z}$ in $\bar{\partial}_2^{m+r}\partial_2^r\big(\mathbb{S}(z_1)^\alpha\mathbb{S}(z_2)^\beta\big)|_\triangle$ is $(a_{m+1,1})_{r,r+m}$. For any real analytic function f on $\mathbb{D}$, for now, let $\big(f(z,\bar{z})\big)_{(p,q)}$ denote the coefficient of $z^p\bar{z}^q$ in $f(z,\bar{z})$. We have

$$\begin{aligned}
\big(a_{m+1,1}\big)_{r,r+m} &= \big(\bar{\partial}_2^{m+r}\partial_2^r\big(\mathbb{S}(z_1)^\alpha\mathbb{S}(z_2)^\beta\big)|_\triangle\big)_{(m+1,1)}\\
&= \Big((\beta)_r\sum_{l=0}^{m+r}\tbinom{m+r}{l}(\beta+r)_{m+r-l}\mathbb{S}(z)^{\alpha+\beta+r+(m+r-l)}z^{m+r-l}l!\binom{r}{l}\bar{z}^{r-l}\Big)_{(m+1,1)}.
\end{aligned}$$

The terms containing $z^{m+1}\bar{z}$ occurs in the sum above, only when $l = r$ and $l = r-1$, that is,

$$\begin{aligned}
\big(a_{m+1,1}\big)_{r,r+m} &= \Bigg((\beta)_r r!\Big(\tbinom{m+r}{r}(\beta+r)_m \mathbb{S}(z)^{\alpha+\beta+m+r} z^m \\
&\qquad + \tbinom{m+r}{r-1}(\beta+r)_{m+1}\mathbb{S}(z)^{\alpha+\beta+m+r+1} z^{m+1}\bar{z}\Big)\Bigg)_{(m+1,1)} \\
&= \Bigg((\beta)_r r!\Bigg(\frac{(m+r)!}{r!m!}(\beta+r)_m(1+(\alpha+\beta+m+r)|z|^2)z^m \\
&\qquad + \frac{(m+r)!r}{r!(m+1)!}(\beta+r)_{m+1}\mathbb{S}(z)^{\alpha+\beta+m+r+1} z^{m+1}\bar{z}\Bigg)\Bigg)_{(m+1,1)} \\
&= \frac{(m+r)!}{m!}(\beta)_{m+r}\big((\alpha+\beta+m+r) + \frac{r}{m+1}(\beta+m+r)\big) \\
&= \frac{(m+r)!}{m!}(\beta)_{m+r}\big(\alpha + (1+\frac{r}{m+1})(\beta+m+r)\big),
\end{aligned}$$

for $0 \le r \le n-m, 0 \le m \le n$, where we have followed the convention: $\binom{p}{q} = 0$ for a negative integer q. □

Lemma 4.7. *Let c_{k0} denote $a_{00}^{1/2}\widetilde{b}_{k0}a_{00}^{1/2}$. For $0 \le r \le n-k$, $0 \le k \le n$, $(c_{k0})_{r,r+k} = \frac{(-1)^k(r+k)!}{k!}(\beta)_{r+k}$.*

Proof. We have $\widetilde{K}(z,w)^{-1} = a_{00}^{1/2}K(z,w)^{-1}a_{00}^{1/2} = \sum_{mn\ge 0}\big(a_{00}^{1/2}\widetilde{b}_{mn}a_{00}^{1/2}\big)z^m\bar{w}^n$. Hence $\widetilde{b}_{mn} = a_{00}^{1/2}b_{mn}a_{00}^{1/2}$ for $m,n \ge 0$. By invertibility of a_{00}, we see that $\widetilde{b}_{k0}$ and c_{k0} uniquely determine each other for $k \ge 0$. Since $(\widetilde{b}_{k0})_{k\ge 0}$ are uniquely determined as the coefficients of power series expansion of $\widetilde{K}(z,w)^{-1}$, it is enough to prove that

$$\sum_{l=0}^{m}\widetilde{a}_{m-l,0}\widetilde{b}_{l0} = 0 \quad \text{for} \quad 1 \le m \le n.$$

Equivalently, we must show that

$$\sum_{l=0}^{m}(a_{00}^{-1/2}a_{m-l,0}a_{00}^{-1/2})(a_{00}^{-1/2}c_{l0}a_{00}^{-1/2}) = 0$$

which amounts to showing

$$a_{00}^{-1/2}\left(\sum_{l=0}^{m}a_{m-l,0}a_{00}^{-1}c_{l0}\right)a_{00}^{-1/2} = 0 \quad \text{for} \quad 1 \le m \le n.$$

It follows from Lemma 4.6 that

$$(a_{m-l,0})_{r,r+(m-l)} = \frac{(m-l+r)!}{(m-l)!}(\beta)_{m-l+r} \quad \text{and} \quad (a_{00})_{rr} = r!(\beta)_r.$$

Therefore

$$\begin{aligned}(a_{m-l,0}a_{00}^{-1})_{r,r+(m-l)} &= (a_{m-l,0})_{r,r+(m-l)}(a_{00}^{-1})_{r+(m-l),r+(m-l)}\\ &= \frac{(m-l+r)!}{(m-l)!}(\beta)_{m-l+r}\big((m-l+r)!(\beta)_{m-l+r}\big)^{-1}\\ &= \frac{1}{(m-l)!}.\end{aligned}$$

We also have

$$\begin{aligned}(a_{m-l,0}a_{00}^{-1}c_{l0})_{r,r+m} &= (a_{m-l,0}a_{00}^{-1})_{r,r+(m-l)}(c_{l0})_{r+(m-l),r+(m-l)+l}\\ &= \frac{1}{(m-l)!}\frac{(-1)^l(r+m)!}{l!}(\beta)_{r+m}\\ &= \frac{(-1)^l(r+m)!}{(m-l)!l!}(\beta)_{r+m}\end{aligned}$$

for $0\le l\le m, 0\le r\le n-m, 1\le m\le n$. Now observe that

$$\begin{aligned}\left(\sum_{l=0}^{m}a_{m-l,0}a_{00}^{-1}c_{l0}\right)_{r,r+m} &= (r+m)!(\beta)_{m+r}\sum_{l=0}^{m}\frac{(-1)^l}{(m-l)!l!}\\ &= \frac{(r+m)!}{m!}(\beta)_{m+r}\sum_{l=0}^{m}(-1)^l\binom{m}{l}=0,\end{aligned}$$

which completes the proof of this lemma. □

Lemma 4.8. $(\widetilde{\widetilde{a}}_{k1})_{n-k+1,n}$ *is a non-zero real number, for* $2\le k\le n+1, n\ge 1$. *All other entries of* $\widetilde{\widetilde{a}}_{k1}$ *are zero.*

Proof. From Lemma 4.5 and Lemma 4.7, we know that

$$\begin{aligned}\widetilde{\widetilde{a}}_{k1} &= \sum_{s=1}^{k}\widetilde{b}_{s0}\widetilde{a}_{k-s,1}+\widetilde{a}_{k1}\\ &= \sum_{s=1}^{k}(a_{00}^{-1/2}c_{s0}a_{00}^{-1/2})(a_{00}^{-1/2}a_{k-s,1}a_{00}^{-1/2})+a_{00}^{-1/2}a_{k1}a_{00}^{-1/2}.\end{aligned}$$

Consequently, $a_{00}^{1/2}\widetilde{\widetilde{a}}_{k1}a_{00}^{1/2}=\sum_{s=1}^{k}c_{s0}a_{00}^{-1}a_{k-s,1}+a_{k1}$ for $1\le k\le n+1$. By Lemma 4.6 and Lemma 4.7, we have

$$\begin{aligned}(c_{s0}a_{00}^{-1})_{r,r+s} &= (c_{s0})_{r,r+s}(a_{00}^{-1})_{r+s,r+s}\\ &= \frac{(-1)^s(r+s)!}{s!}(\beta)_{r+s}\big((r+s)!(\beta)_{r+s}\big)^{-1}=\frac{(-1)^s}{s!},\end{aligned}$$

for $0 \le r \le n-s, 0 \le s \le k, 1 \le k \le n+1$.

$$\begin{aligned}&(a_{k-s,1})_{r,r+(k-s-1)}\\ &\quad= \frac{(k+r-s-1)!}{(k-s-1)!}(\beta)_{r+k-s-1}\big(\alpha + (1+\frac{r}{k-s})(\beta+r+k-s-1)\big),\end{aligned}$$

for $k-s-1 \ge 0, 2 \le k \le n+1$. Now,

$$\begin{aligned}&(c_{s0}a_{00}^{-1}a_{k-s,1})_{r+s,r+s+(k-s-1)}\\ &\quad= (c_{s0}a_{00}^{-1})_{r,r+s}(a_{k-s,1})_{r+s,r+s+(k-s-1)}\\ &\quad= \frac{(-1)^s}{s!}\frac{(r+k-1)!}{(k-s-1)!}(\beta)_{r+k-1}\big(\alpha+(1+\frac{r+s}{k-s})(\beta+r+k-1)\big),\end{aligned}$$

for $1 \le s \le k-1, 0 \le r \le n-k+1, 1 \le k \le n+1$. Hence

$$\begin{aligned}&(c_{s0}a_{00}^{-1}a_{k-s,1})_{r+s,r+k-1}\\ &\quad= \frac{(-1)^s}{s!}\frac{(r+k-1)!}{(k-s-1)!}(\beta)_{r+k-1}\big(\alpha+\frac{k+r}{k-s}(\beta+r+k-1)\big).\end{aligned}$$

Since $\overline{K(z,w)}^t = K(w,z)$, it follows that $a_{mn} = \overline{a_{nm}}^t$ for $m,n \ge 0$. Thus, by Lemma 4.6, $(a_{01})_{r+1,r} = (r+1)!(\beta)_{r+1}$, for $0 \le r \le n-1$, $(c_{k0}a_{00}^{-1})_{r,r+k} = \frac{(-1)^k}{k!}$, for $0 \le r \le n-k, 1 \le k \le n+1$ and

$$(c_{k0}a_{00}^{-1}a_{01})_{r,r+k-1} = (c_{k0}a_{00}^{-1})_{r,r+k}(a_{01})_{r+k,r+k-1} = \frac{(-1)^k}{k!}(r+k)!(\beta)_{r+k},$$

$0 \le r \le n-k, 1 \le k \le n+1$. Now, for $0 \le r \le n-k, 2 \le k \le n+1$. Since $c_{00} = a_{00}$, we clearly have

$$\begin{aligned}(a_{00}^{1/2}\widetilde{\widetilde{a}}_{k1}a_{00}^{1/2})_{r,r+k-1} &= \Big(\sum_{s=1}^{k} c_{s0}a_{00}^{-1}a_{k-s,1} + a_{k1}\Big)_{r,r+k-1}\\ &= \Big(\sum_{s=0}^{k-1} c_{s0}a_{00}^{-1}a_{k-s,1} + c_{k0}a_{00}^{-1}a_{01}\Big)_{r,r+k-1}\\ &= \sum_{s=0}^{k-1}\tfrac{(-1)^s(k+r-1)!}{s!(k-s-1)!}(\beta)_{r+k-1}\big(\alpha+\tfrac{k+r}{k-s}(\beta+r+k-1)\big) + \frac{(-1)^k(r+k)!}{k!}(\beta)_{r+k}\\ &= \alpha(\beta)_{r+k-1}\tfrac{(k+r-1)!}{(k-1)!}\sum_{s=0}^{k-1}(-1)^s\tbinom{k-1}{s} + (\beta)_{k+r}\big(\sum_{s=0}^{k-1}\tfrac{(-1)^s(k+r)!}{s!(k-s)!}+\tfrac{(-1)^k(k+r)!}{k!}\big)\\ &= \tfrac{(k+r)!}{k!}(\mathbb{S}(z_2)^{\beta+k+(k-l)}\beta)_{k+r}\sum_{s=0}^{k}(-1)^s\binom{k}{s}.\end{aligned}$$

Therefore $(a_{00}^{1/2}\widetilde{\widetilde{a}}_{k1}a_{00}^{1/2})_{r,r+k-1} = 0$. Now, $c_{00} = a_{00}$ and $(c_{k0}a_{00}^{-1}a_{01})_{n-k+1,n} = 0$ for $2 \le k \le n+1$. Hence

$$\begin{aligned}
(a_{00}^{1/2}\widetilde{\widetilde{a}}_{k1}a_{00}^{1/2})_{n-k+1,n} &= \Big(\sum_{s=1}^{k} c_{s0}a_{00}^{-1}a_{k-s,1} + a_{k1}\Big)_{n-k+1,n}\\
&= \Big(\sum_{s=0}^{k-1} c_{s0}a_{00}^{-1}a_{k-s,1}\Big)_{n-k+1,n}\\
&= \sum_{s=0}^{k-1} \tfrac{(-1)^s(k+(n-k+1)-1)!}{s!(k-s-1)!}(\beta)_n\Big(\alpha + \tfrac{k+(n-k+1)}{k-s}(\beta+n)\Big)\\
&= n!(\beta)_n\Big(\alpha\sum_{s=0}^{k-1}\tfrac{(-1)^s}{s!(k-1-s)!} + (n+1)(\beta+n)\sum_{s=0}^{k-1}\tfrac{(-1)^s}{s!(k-s)!}\Big)\\
&= n!(\beta)_n\Big(\tfrac{\alpha}{(k-1)!}\sum_{s=0}^{k-1}(-1)^s\tbinom{k-1}{s} + \tfrac{(n+1)(\beta+n)}{k!}\sum_{s=0}^{k}(-1)^s\tbinom{k}{s} - \tfrac{(-1)^k(n+1)(\beta+n)}{k!}\Big)\\
&= 0 + 0 - n!(\beta)_n\tfrac{(-1)^k(n+1)(\beta+n)}{k!}\\
&= \tfrac{(-1)^{k+1}(n+1)!(\beta)_{n+1}}{k!},\ \text{ for } 2 \le k \le n+1.
\end{aligned}$$

Since a_{00} is a diagonal matrix with positive diagonal entries, $\widetilde{\widetilde{a}}_{k1}$ has the form as stated in the lemma, for $2 \le k \le n+1, n \ge 1$. □

Here is a simple lemma from matrix theory which will be useful for us in the sequel.

Lemma 4.9. *Let $\{A_k\}_{k=0}^{n-1}$ are $(n+1)\times(n+1)$ matrices such that $(A_k)_{kn} = \lambda_k \ne 0$ for $0 \le k \le n-1$, $n \ge 1$. If $AA_k = A_kA$ for some $(n+1)\times(n+1)$ matrix $A = \big(\!\!\big(a_{ij}\big)\!\!\big)_{i,j=0}^{n}$ for $0 \le k \le n-1$, then A is upper triangular with equal diagonal entries.*

Proof. $(AA_k)_{in} = a_{ik}(A_k)_{kn} = a_{ik}\lambda_k$ and $(A_kA)_{kj} = (A_k)_{kn}a_{nj} = \lambda_k a_{nj}$ for $0 \le i,j \le n, 0 \le k \le n-1$. Putting $i = k$ and $j = n$, we have $(AA_k)_{kn} = a_{kk}\lambda_k$ and $(A_kA)_{kn} = \lambda_k a_{nn}$. By hypothesis we have $a_{kk}\lambda_k = \lambda_k a_{nn}$. As $\lambda_k \ne 0$, this implies that $a_{kk} = a_{nn}$ for $0 \le k \le n-1$, which is same as saying that A has equal diagonal entries. Now observe that $(A_kA)_{ij} = 0$ if $i \ne k$ for $0 \le j \le n$, which implies that $(A_kA)_{in} = 0$ if $i \ne k$. By hypothesis this is same as $(AA_k)_{in} = a_{ik}\lambda_k = 0$ if $i \ne k$. This implies $a_{ik} = 0$ if $i \ne k, 0 \le i \le n, 0 \le k \le n-1$, which is a stronger statement than saying A is upper triangular. □

Lemma 4.10. *If an $(n+1)\times(n+1)$ matrix A commutes with $\widetilde{\widetilde{a}}_{k1}$ and $\widetilde{\widetilde{a}}_{1k}$ for $2 \le k \le n+1, n \ge 1$, then A is a scalar.*

Proof. It follows from Lemma 4.8 and Lemma 4.9 that if A commutes with $\widetilde{\widetilde{a}}_{k1}$ for $2 \le k \le n+1$, then A is upper triangular with equal diagonal entries. As the entries of $\widetilde{\widetilde{a}}_{k1}$ are real, $(\widetilde{\widetilde{a}}_{1k}) = (\widetilde{\widetilde{a}}_{k1})^t$. If A commutes with $\widetilde{\widetilde{a}}_{1k}$ for $2 \le k \le n+1$, then by a similar proof as in Lemma 4.9, it follows that A is lower triangular with equal diagonal entries. So A is both upper triangular and lower triangular with equal diagonal entries, hence A is a scalar. $\square$

This sequence of lemmas put together constitutes a proof of Theorem 4.2.

For homogeneous operators in the class $\mathrm{B}_2(\mathbb{D})$, we have a proof of reducibility that avoids the normalization of the kernel. This proof makes use of the fact that if such an operator is reducible then each of the direct summands must belong to the class $\mathrm{B}_1(\mathbb{D})$. We give a precise formulation of this phenomenon along with a proof below. Let K be a positive definite kernel on $\mathbb{D}^2$ and $\mathcal{H}$ be the corresponding Hilbert space. Assume that the pair (M_1, M_2) on $\mathcal{H}$ is in $\mathrm{B}_1(\mathbb{D}^2)$. The operator M^* is the adjoint of the multiplication operator on Hilbert space $J^{(2)}\mathcal{H}_{|\mathrm{res}\ \triangle}$ which consists of $\mathbb{C}^2$-valued holomorphic function on $\mathbb{D}$ and possesses the reproducing kernel $J^{(2)}K(z,w)$. The operator M^* is in $\mathrm{B}_2(\mathbb{D})$ (cf. [12, Proposition 3.6]).

Proposition 4.11. *The operator M^* on Hilbert space $J^{(2)}\mathcal{H}_{|\mathrm{res}\ \triangle}$ is irreducible.*

Proof. If possible, let M^* be reducible, that is, $M^* = T_1 \oplus T_2$ for some $T_1, T_2 \in \mathrm{B}_1(\mathbb{D})$ by [9, Proposition 1.18], which is same as saying by [9, Proposition 1.18] that the associated bundle E_{M^*} is reducible. A metric on the associated bundle E_{M^*} is given by $h(z) = J^{(2)}K(z,z)^t$. So, there exists a holomorphic change of frame $\psi : \mathbb{D} \longrightarrow GL(2,\mathbb{C})$ such that $\overline{\psi(z)}^t h(z)\psi(z) = \begin{pmatrix} h_1(z) & 0 \\ 0 & h_2(z) \end{pmatrix}$ for $z \in \mathbb{D}$, where h_1 and h_2 are metrics on the associated line bundles E_{T_1} and E_{T_2} respectively. So $\psi(z)^{-1}\mathcal{K}_h(z)\psi(z) = \begin{pmatrix} \mathcal{K}_{h_1}(z) & 0 \\ 0 & \mathcal{K}_{h_2}(z) \end{pmatrix}$, where $\mathcal{K}_h(z) = \bar\partial(h^{-1}\partial h)(z)$ is the curvature of the bundle E_{M^*} with respect to the metric h and $\mathcal{K}_{h_i}(z)$ are the curvatures of the bundles E_{T_i} for $i = 1, 2$ as in [9, pp. 211]. A direct computation shows that $\mathcal{K}_h(z) = \begin{pmatrix} \alpha & -2\beta(\beta+1)(1-|z|^2)^{-1}\bar z \\ 0 & \alpha+2\beta+2 \end{pmatrix}(1-|z|^2)^{-2}$. Thus the matrix $\psi(z)$ diagonalizes $\mathcal{K}_h(z)$ for $z \in \mathbb{D}$. It follows that $\psi(z)$ is determined, that is, the columns of $\psi(z)$ are eigenvectors of $\mathcal{K}_h(z)$ for $z \in \mathbb{D}$. These are uniquely determined up to multiplication by non-vanishing scalar-valued functions f_1 and f_2 on $\mathbb{D}$. Now one set of eigenvectors of $\mathcal{K}_h(z)$ is given by $\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -\beta\bar z \\ 1-|z|^2 \end{pmatrix}\right\}$ and it is clear that there does not exist any non-vanishing scalar-valued function f_2 on $\mathbb{D}$ such that $f_2(z)\begin{pmatrix} -\beta\bar z \\ 1-|z|^2 \end{pmatrix}$ is an eigenvector for $\mathcal{K}_h(z)$ whose entries are holomorphic functions on $\mathbb{D}$. Hence there does not exist any holomorphic change of frame $\psi : \mathbb{D} \longrightarrow GL(2,\mathbb{C})$ such that $\overline{\psi}^t h\psi = \begin{pmatrix} h_1 & 0 \\ 0 & h_2 \end{pmatrix}$ on $\mathbb{D}$. Hence M^* is irreducible. $\square$

Theorem 4.12. *The operators $T = M^{(\alpha,\beta)}$ and $\widetilde{T} := M^{(\widetilde\alpha,\widetilde\beta)}$ are unitarily equivalent if and only if $\alpha = \widetilde\alpha$ and $\beta = \widetilde\beta$.*

One of the implications is trivial. To prove the other implication, recall that [12, Proposition 3.6] $T, \widetilde{T} \in \mathrm{B}_{n+1}(\mathbb{D})$. It follows from [9] that if $T, \widetilde{T} \in \mathrm{B}_{n+1}(\mathbb{D})$ are unitarily equivalent then the curvatures $\mathcal{K}_T, \mathcal{K}_{\widetilde{T}}$ of the associated bundles E_T and $E_{\widetilde{T}}$ respectively, are unitarily equivalent as matrix-valued real-analytic functions on $\mathbb{D}$. In particular, this implies that $\mathcal{K}_T(0)$ and $\mathcal{K}_{\widetilde{T}}(0)$ are unitarily equivalent. Therefore, we compute $\mathcal{K}_T(0)$ and $\mathcal{K}_{\widetilde{T}}(0)$. Let $\mathcal{K}_T(h)$ denote the curvature of the bundle E_T with respect to the metric $h(z) := \widetilde{\widetilde{K}}(z,z)^t$.

Lemma 4.13. *The curvature $\mathcal{K}_T(h)(0)$ at 0 of the bundle E_T equals the coefficient of $z\bar{z}$ in the normalized kernel $\widetilde{\widetilde{K}}$, that is, $\mathcal{K}_T(h)(0) = \widetilde{\widetilde{a}}_{11}^{\,t}$.*

Proof. The curvature of the bundle E_T with respect to the metric $h(z) := \widetilde{\widetilde{K}}(z,z)^t$ is $\mathcal{K}_T(h) = \bar{\partial}(h^{-1}\partial h)$. If $h(z) = \sum_{m,n\geq 0} h_{mn} z^m \bar{z}^n$, then $h_{mn} = \widetilde{\widetilde{a}}_{mn}^{\,t}$ for $m, n \geq 0$. So, $h_{00} = I$ and $h_{m0} = h_{0n} = 0$ for $m, n \geq 1$. Hence $\mathcal{K}_T(h)(0) = \bar{\partial} h^{-1}(0)\partial h(0) + h^{-1}(0)\bar{\partial}\partial h(0) = (\bar{\partial} h^{-1}(0))h_{10} + h_{00}^{-1}h_{11} = h_{11} = \widetilde{\widetilde{a}}_{11}^{\,t}$. □

Lemma 4.14. $\big(\mathcal{K}_T(0)\big)_{ii} = \alpha$, *for $i = 0, \ldots, n-1$ and* $\big(\mathcal{K}_T(0)\big)_{nn} = \alpha + (n+1)(\beta+n)$ *for $n \geq 1$.*

Proof. From Lemma 4.13 and Lemma 4.5 we know that $\mathcal{K}_T(0) = \widetilde{\widetilde{a}}_{11}^{\,t} = \big(\widetilde{a}_{11} + \widetilde{b}_{10}\widetilde{a}_{01}\big)^t$. Thus $\mathcal{K}_T(0)$ is the transpose of $a_{00}^{-1/2}(a_{11} + c_{10}a_{00}^{-1}a_{01})a_{00}^{-1/2}$ by Lemma 4.7. Now, by Lemma 4.6 and Lemma 4.7, $\big(c_{10}\big)_{r,r+1} = -(r+1)!(\beta)_{r+1}$ for $0 \leq r \leq n-1$, $\big(a_{00}\big)_{rr} = r!(\beta)_r$, $\big(a_{11}\big)_{rr} = r!(\beta)_r\big(\alpha + (r+1)(\beta+r)\big)$ for $0 \leq r \leq n$ and $\big(a_{01}\big)_{r+1,r} = (r+1)!(\beta)_{r+1}$ for $0 \leq r \leq n-1$. Therefore, $\big(c_{10}a_{00}^{-1}a_{01}\big)_{rr} = -(r+1)!(\beta)_{r+1}$ for $0 \leq r \leq n-1$. Also, $\big(a_{11} + c_{10}a_{00}^{-1}a_{01}\big)_{rr} = \alpha r!(\beta)_{r+1}$ for $0 \leq r \leq n-1$, and $\big(a_{11} + c_{10}a_{00}^{-1}a_{01}\big)_{nn} = n!(\beta)_n(\alpha + (n+1)(\beta+n))$. Finally, $\mathcal{K}_T(h)(0) = \widetilde{\widetilde{a}}_{11}^{\,t} = \widetilde{\widetilde{a}}_{11}$, as $\widetilde{\widetilde{a}}_{11}$ is a diagonal matrix with real entries. In fact, $\big(\mathcal{K}_T(0)\big)_{ii} = \alpha$, for $i = 0, \ldots, n-1$ and $\big(\mathcal{K}_T(0)\big)_{nn} = \alpha + (n+1)(\beta+n)$. □

We now see that M and $\widetilde{M}$ are unitarily equivalent implies that $\alpha = \widetilde{\alpha}$ and $\alpha + (n+1)(\beta+n) = \widetilde{\alpha} + (n+1)(\widetilde{\beta}+n)$, that is, $\alpha = \widetilde{\alpha}$ and $\beta = \widetilde{\beta}$.

5. Homogeneity of the operator $M^{(\alpha,\beta)}$

Theorem 5.1. *The multiplication operator $M := M^{(\alpha,\beta)}$ on $J^{(n+1)}\mathcal{H}$ is homogeneous.*

This theorem is a particular case of the Lemma 4.1. A proof first appeared in [6, Theorem 5.2.]. We give an alternative proof of this theorem by showing that that the kernel is quasi-invariant, that is,

$$K(z,w) = J_{\varphi^{-1}}(z)K\big(\varphi^{-1}(z), \varphi^{-1}(w)\big)\overline{J_{\varphi^{-1}}(w)}^{\,t}$$

for some cocycle

$$J : \text{Möb} \times \mathbb{D} \longrightarrow \mathbb{C}^{(n+1)\times(n+1)},\ \varphi \in \text{Möb},\ z,\ w \in \mathbb{D}.$$

First we prove that $K(z,z) = J_{\varphi^{-1}}(z)K\big(\varphi^{-1}(z),\varphi^{-1}(z)\big)\overline{J_{\varphi^{-1}}(z)}^t$ and then polarize to obtain the final result. We begin with a series of lemmas.

Lemma 5.2. *Suppose that* $J : \text{Möb} \times \mathbb{D} \longrightarrow \mathbb{C}^{(n+1)\times(n+1)}$ *is a cocycle. Then the following are equivalent*

1. $K(z,z) = J_{\varphi^{-1}}(z)K\big(\varphi^{-1}(z),\varphi^{-1}(z)\big)\overline{J_{\varphi^{-1}}(z)}^t$ *for all* $\varphi \in$ Möb *and* $z \in \mathbb{D}$;
2. $K(0,0) = J_{\varphi^{-1}}(0)K\big(\varphi^{-1}(0),\varphi^{-1}(0)\big)\overline{J_{\varphi^{-1}}(0)}^t$ *for all* $\varphi \in$ Möb.

Proof. One of the implications is trivial. To prove the other implication, note that

$$\begin{aligned} J_{\varphi_1^{-1}}(0)K\big(\varphi_1^{-1}(0),\varphi_1^{-1}(0)\big)\overline{J_{\varphi_1^{-1}}(0)}^t &= K(0,0) \\ &= J_{\varphi_2^{-1}}(0)K\big(\varphi_2^{-1}(0),\varphi_2^{-1}(0)\big)\overline{J_{\varphi_2^{-1}}(0)}^t \end{aligned}$$

for any $\varphi_1, \varphi_2 \in$Möb and $z \in \mathbb{D}$. Now pick $\psi \in$Möb such that $\psi^{-1}(0) = z$ and taking $\varphi_1 = \psi, \varphi_2 = \psi\varphi$ in the previous identity we see that

$$\begin{aligned} &J_{\psi^{-1}}(0)K\big(\psi^{-1}(0),\psi^{-1}(0)\big)\overline{J_{\psi^{-1}}(0)}^t \\ &= J_{\varphi^{-1}\psi^{-1}}(0)K\big(\varphi^{-1}\psi^{-1}(0),\varphi^{-1}\psi^{-1}(0)\big)\overline{J_{\varphi^{-1}\psi^{-1}}(0)}^t \\ &= J_{\psi^{-1}}(0)J_{\varphi^{-1}}(\psi^{-1}(0))K\big(\varphi^{-1}(z),\varphi^{-1}(z)\big)\overline{J_{\varphi^{-1}}(\psi^{-1}(0))}^t\overline{J_{\psi^{-1}}(0)}^t \end{aligned}$$

for $\varphi \in \text{Möb}, z \in \mathbb{D}$. Since $J_{\psi^{-1}}(0)$ is invertible, it follows from the equality of first and third quantities that

$$K\big(\psi^{-1}(0),\psi^{-1}(0)\big) = J_{\varphi^{-1}}(\psi^{-1}(0))K\big(\varphi^{-1}\psi^{-1}(0),\varphi^{-1}\psi^{-1}(0)\big)\overline{J_{\varphi^{-1}}(\psi^{-1}(0))}^t.$$

This is the same as $K(z,z) = J_{\varphi^{-1}}(z)K\big(\varphi^{-1}(z),\varphi^{-1}(z)\big)\overline{J_{\varphi^{-1}}(z)}^t$ by the choice of ψ. The proof of this lemma is therefore complete. □

Let $\mathcal{J}_{\varphi^{-1}}(z) = (J_{\varphi^{-1}}(z)^t)^{-1}, \varphi \in$Möb, $z \in \mathbb{D}$, where X^t denotes the transpose of the matrix X. Clearly, $(J_{\varphi^{-1}}(z)^t)^{-1}$ satisfies the cocycle property if and only if $\mathcal{J}_{\varphi^{-1}}(z)$ does and they uniquely determine each other. It is easy to see that the condition

$$K(0,0) = J_{\varphi^{-1}}(0)K\big(\varphi^{-1}(0),\varphi^{-1}(0)\big)\overline{J_{\varphi^{-1}}(0)}^t$$

is equivalent to

$$h\big(\varphi^{-1}(0)\big) = \overline{\mathcal{J}_{\varphi^{-1}}(0)}^t h(0)\mathcal{J}_{\varphi^{-1}}(0), \tag{5.16}$$

where $h(z)$ is the transpose of $K(z,z)$ as before. It will be useful to define the two functions

(i) $c : \text{Möb}\times\mathbb{D} \longrightarrow \mathbb{C},\ c(\varphi^{-1},z) = (\varphi^{-1})'(z)$;
(ii) $p : \text{Möb}\times\mathbb{D} \longrightarrow \mathbb{C},\ p(\varphi^{-1},z) = \frac{\overline{ta}}{1+\overline{ta}z}$

for $\varphi_{t,a} \in$ Möb, $t \in \mathbb{T}$, $a \in \mathbb{D}$. We point out that the function c is the well-known cocycle for the group Möb.

Lemma 5.3. *With notation as above, we have*

(a) $\varphi_{t,a}^{-1} = \varphi_{\bar{t},-ta}$
(b) $\varphi_{s,b}\varphi_{t,a} = \varphi_{\frac{s(t+\bar{a}b)}{1+ta\bar{b}},\frac{a+\bar{t}b}{1+\overline{ta}b}}$
(c) $c(\varphi^{-1},\psi^{-1}(z))c(\psi^{-1}(z)) = c(\varphi^{-1}\psi^{-1},z)$ *for* $\varphi,\psi \in$Möb, $z\in\mathbb{D}$
(d) $p(\varphi^{-1},\psi^{-1}(z))c(\psi^{-1},z) + p(\psi^{-1},z) = p(\varphi^{-1}\psi^{-1},z)$ for $\varphi,\psi \in$Möb, $z\in\mathbb{D}$.

Proof. The proof of (a) is a mere verification. We note that

$$\varphi_{s,b}(\varphi_{t,a}(z)) = s\frac{t\frac{z-a}{1-\bar{a}z}-b}{1-\bar{b}t\frac{z-a}{1-\bar{a}z}} = s\frac{tz-ta-b+\bar{a}bz}{1-\bar{a}z-t\bar{b}z+ta\bar{b}} = \frac{s(t+\bar{a}b)}{1+ta\bar{b}}\frac{z-\frac{ta+b}{t+\bar{a}b}}{1-\frac{\bar{a}+t\bar{b}}{1+ta\bar{b}}z},$$

which is (b). The chain rule gives (c). To prove (d), we first note that for $\varphi=\varphi_{t,a}$ and $\psi=\varphi_{s,b}$, if $\psi^{-1}\varphi^{-1}=\varphi_{t',a'}$ for some $(t',a')\in\mathbb{T}\times\mathbb{D}$ then

$$\overline{t'a'} = \frac{\bar{s}(\bar{t}+a\bar{b})}{1+\overline{ta}b}\frac{\bar{a}+t\bar{b}}{1+ta\bar{b}} = \frac{\bar{s}(\bar{b}+\overline{ta})}{1+\overline{ta}b}.$$

It is now easy to verify that

$$\begin{aligned}
p(\varphi^{-1},\psi^{-1}(z))c(\psi^{-1},z)+p(\psi^{-1},z) &= \frac{\overline{ta}}{1+\overline{ta}\psi_{s,b}^{-1}(z)}\frac{\bar{s}(1-|b|^2)}{(1+\overline{sb}z)^2}+\frac{\overline{sb}}{1+\overline{sb}z}\\
&= \frac{\bar{s}(\bar{b}+\overline{ta})}{1+\overline{ta}b+\bar{s}(\bar{b}+\overline{ta})z}\\
&= \frac{\bar{s}\frac{\bar{b}+\overline{ta}}{1+\overline{ta}b}}{1+\frac{\bar{s}(\bar{b}+\overline{ta})}{1+\overline{ta}b}z}\\
&= p(\varphi^{-1}\psi^{-1},z).
\end{aligned}$$

□

Let

$$\big(\mathcal{J}_{\varphi^{-1}}(z)\big)_{ij} = c(\varphi^{-1},z)^{-\frac{\alpha+\beta}{2}-n}\frac{(\beta)_j}{(\beta)_i}\binom{j}{i}c(\varphi^{-1},z)^{n-j}p(\varphi^{-1},z)^{j-i} \tag{5.17}$$

for $0\le i\le j\le n$.

Lemma 5.4. $J_{\varphi^{-1}}(z)$ *defines a cocycle for the group* Möb.

Proof. To say that $J_{\varphi^{-1}}(z)$ satisfies the cocycle property is the same as saying $\mathcal{J}_{\varphi^{-1}}(z)$ satisfies the cocycle property, which is what we will verify. Thus we want to show that $\big(\mathcal{J}_{\psi^{-1}}(z)\mathcal{J}_{\varphi^{-1}}(\psi^{-1}(z))\big)_{ij} = \big(\mathcal{J}_{\varphi^{-1}\psi^{-1}}(z)\big)_{ij}$ for $0\le i,j\le n$. We note that $\mathcal{J}_{\varphi^{-1}}(z)$ is upper triangular, as the product of two upper triangular matrices is again upper triangular, it suffices to prove this equality for $0\le i\le j\le n$.

Clearly, we have

$$
\begin{aligned}
\big(\mathcal{J}_{\psi^{-1}}(z)\mathcal{J}_{\varphi^{-1}}(\psi^{-1}(z))\big)_{ij} &= \sum_{k=i}^{j}\big(\mathcal{J}_{\psi^{-1}}(z)\big)_{ik}\big(\mathcal{J}_{\varphi^{-1}}(\psi^{-1}(z))\big)_{kj}\\
&= c(\psi^{-1},z)^{-\frac{\alpha+\beta}{2}-n}c(\varphi^{-1},\psi^{-1}(z))^{-\frac{\alpha+\beta}{2}-n}\sum_{k=i}^{j}\Big(\frac{(\beta)_k}{(\beta)_i}\binom{k}{i}c(\psi^{-1},z)^{n-k}\\
&\quad p(\psi^{-1},z)^{k-i}\frac{(\beta)_j}{(\beta)_k}\binom{j}{k}c(\varphi^{-1},\psi^{-1}(z))^{n-j}p(\varphi^{-1},\psi^{-1}(z))^{j-k}\Big)\\
&= c(\varphi^{-1}\psi^{-1},z)^{-\frac{\alpha+\beta}{2}-n}\frac{(\beta)_j}{(\beta)_i}c(\psi^{-1},z)^{n-j}c(\varphi^{-1},\psi^{-1}(z))^{n-j}\\
&\qquad\sum_{k=i}^{j}\frac{j!}{i!(k-i)!(j-k)!}c(\psi^{-1},z)^{j-k}p(\varphi^{-1},\psi^{-1}(z))^{j-k}p(\psi^{-1},z)^{k-i}\\
&= c(\varphi^{-1}\psi^{-1},z)^{-\frac{\alpha+\beta}{2}-n}\frac{(\beta)_j}{(\beta)_i}\binom{j}{i}c(\varphi^{-1}\psi^{-1},z)^{n-j}\\
&\qquad\sum_{k=i}^{j}\binom{j-i}{k-i}c(\psi^{-1},z)^{j-k}p(\varphi^{-1},\psi^{-1}(z))^{j-k}p(\psi^{-1},z)^{k-i}\\
&= c(\varphi^{-1}\psi^{-1},z)^{-\frac{\alpha+\beta}{2}-n}\frac{(\beta)_j}{(\beta)_i}\binom{j}{i}c(\varphi^{-1}\psi^{-1},z)^{n-j}\\
&\qquad\sum_{k=0}^{j-i}\binom{j-i}{k}c(\psi^{-1},z)^{(j-i)-k}p(\varphi^{-1},\psi^{-1}(z))^{(j-i)-k}p(\psi^{-1},z)^{k}\\
&= c(\varphi^{-1}\psi^{-1},z)^{-\frac{\alpha+\beta}{2}-n}\frac{(\beta)_j}{(\beta)_i}\binom{j}{i}c(\varphi^{-1}\psi^{-1},z)^{n-j}\\
&\qquad\qquad\Big(c(\psi^{-1},z)p(\varphi^{-1},\psi^{-1}(z))+p(\psi^{-1},z)\Big)^{j-i}\\
&= c(\varphi^{-1}\psi^{-1},z)^{-\frac{\alpha+\beta}{2}-n}\frac{(\beta)_j}{(\beta)_i}\binom{j}{i}c(\varphi^{-1}\psi^{-1},z)^{n-j}p(\varphi^{-1}\psi^{-1},z)^{j-i}\\
&= \big(\mathcal{J}_{\varphi^{-1}\psi^{-1}}(z)\big)_{ij},
\end{aligned}
$$

for $0\le i\le j\le n$. The penultimate equality follows from Lemma 5.3. □

We need the following beautiful identity to prove (5.16). We provide two proofs, the first one is due to C. Verughese and the second is due to B. Bagchi.

Lemma 5.5. *For nonnegative integers $j\ge i$ and $0\le k\le i$, we have*

$$\sum_{l=0}^{i-k}(-1)^l(l+k)!\binom{i}{l+k}\binom{j}{l+k}\binom{l+k}{l}(a+j)_{i-l-k} = k!\binom{i}{k}\binom{j}{k}(a+k)_{i-k},$$

for all $a\in\mathbb{C}$.

Proof. Here is the first proof due to C. Verughese: For any integer $i \geq 1$ and $a \in \mathbb{C} \setminus \mathbb{Z}$, we have

$$\begin{aligned}
&\sum_{l=0}^{i-k}(-1)^l(l+k)!\binom{i}{l+k}\binom{j}{l+k}\binom{l+k}{l}(a+j)_{i-l-k}\\
&= \frac{i!j!}{k!\Gamma(a+j)}\sum_{l=0}^{i-k}\frac{(-1)^l}{l!(i-k-l)!}\frac{\Gamma(a+j+i-l-k)}{\Gamma(j-l-k+1)}\\
&= \frac{i!j!}{k!(i-k)!\Gamma(a+j)\Gamma(1-a-i)}\sum_{l=0}^{i-k}(-1)^l\tbinom{i-k}{l}B(a+j+i-k-l,1-a-i)\\
&= \frac{i!j!}{k!(i-k)!\Gamma(a+j)\Gamma(1-a-i)}\sum_{l=0}^{i-k}(-1)^l\tbinom{i-k}{l}\int_0^1 t^{a+j+i-k-l-1}(1-t)^{-a-i}dt\\
&= \frac{i!j!}{k!(i-k)!\Gamma(a+j)\Gamma(1-a-i)}\int_0^1\sum_{l=0}^{i-k}(-1)^l\tbinom{i-k}{l}t^{a+j+i-k-l-1}(1-t)^{-a-i}dt\\
&= \frac{i!j!}{k!(i-k)!\Gamma(a+j)\Gamma(1-a-i)}\int_0^1 (1-t)^{-a-i}t^{a+j-1}\big(\sum_{l=0}^{i-k}(-1)^l\tbinom{i-k}{l}t^{i-k-l}\big)dt\\
&= \frac{i!j!}{k!(i-k)!\Gamma(a+j)\Gamma(1-a-i)}\int_0^1 (1-t)^{-a-i}t^{a+j-1}(t-1)^{i-k}dt\\
&= \frac{(-1)^{i-k}i!j!}{k!(i-k)!\Gamma(a+j)\Gamma(1-a-i)}B(a+j,1-a-k)\\
&= \frac{(-1)^{i-k}i!j!}{k!(i-k)!\Gamma(a+j)\Gamma(1-a-i)}\frac{\Gamma(a+j)\Gamma(1-a-k)}{\Gamma(1+j-k)}\\
&= \frac{(-1)^{i-k}i!j!}{k!(i-k)!\Gamma(1-a-i)}\frac{\Gamma(1-a-k)}{(j-k)!}\\
&= (-1)^{i-k}k!\binom{i}{k}\binom{j}{k}\frac{\Gamma(1-a-k)}{\Gamma(1-a-i)}\\
&= k!\binom{i}{k}\binom{j}{k}\frac{\Gamma(1-a)}{(-1)^i\Gamma(1-a-i)}\frac{(-1)^k\Gamma(1-a-k)}{\Gamma(1-a)}\\
&= k!\binom{i}{k}\binom{j}{k}\frac{\Gamma(1-a)}{\cos i\pi\Gamma(1-a-i)}\frac{\cos k\pi\Gamma(1-a-k)}{\Gamma(1-a)}\\
&= k!\binom{i}{k}\binom{j}{k}\frac{\Gamma(a+i)}{\Gamma(a+k)}.
\end{aligned}$$

Since we have an equality involving a polynomial of degree $i-k$ for all a in $\mathbb{C}\setminus\mathbb{Z}$, it follows that the equality holds for all $a \in \mathbb{C}$.

Here is another proof due to B. Bagchi: Since

$$\tbinom{-x}{n} = \tfrac{-x(-x-1)\cdots(-x-n+1)}{n!} = (-1)^n\tbinom{x+n-1}{n}$$

and

$$(x)_n = x(x+1)\cdots(x+n-1) = n!\tbinom{x+n-1}{n},$$

it follows that

$$\begin{aligned}
&\sum_{l=0}^{i-k}(-1)^l(l+k)!\binom{i}{l+k}\binom{j}{l+k}\binom{l+k}{l}(a+j)_{i-l-k}\\
&= \frac{i!j!}{k!}\sum_{l=0}^{i-k}\frac{(-1)^l}{l!(i-k-l)!(j-k-l)!}(i-k-l)!\binom{a+j+i-k-l-1}{i-k-l}\\
&= \frac{i!j!}{k!(j-k)!}\sum_{l=0}^{i-k}\frac{(-1)^l(j-k)!}{l!(j-k-l)!}(-1)^{i-k-l}\binom{-a-j}{i-k-l}\\
&= i!\binom{j}{k}(-1)^{i-k}\sum_{l=0}^{i-k}\binom{j-k}{l}\binom{-a-j}{i-k-l}\\
&= i!\binom{j}{k}(-1)^{i-k}\binom{-a-k}{i-k}\\
&= i!\binom{j}{k}(-1)^{i-k}(-1)^{i-k}\binom{a+i-1}{i-k}\\
&= k!\binom{i}{k}\binom{j}{k}(a+k)_{i-k},
\end{aligned}$$

where the equality after the last summation symbol follows from Vandermonde's identity. □

Lemma 5.6. *For* $\varphi \in$Möb *and* $\mathcal{J}_{\varphi^{-1}}(z)$ *as in* (5.17),

$$h\big(\varphi^{-1}(0)\big) = \overline{\mathcal{J}_{\varphi^{-1}}(0)}^t h(0)\mathcal{J}_{\varphi^{-1}}(0).$$

Proof. It is enough to show that

$$h\big(\varphi^{-1}(0)\big)_{ij} = \big(\overline{\mathcal{J}_{\varphi^{-1}}(0)}^t h(0)\mathcal{J}_{\varphi^{-1}}(0)\big)_{ij}, \text{ for } 0 \le i \le j \le n.$$

Let $\varphi = \varphi_{t,z}$, $t \in \mathbb{T}$, and $z \in \mathbb{D}$. Since $\big(h\big(\varphi^{-1}(0)\big)\big)_{ij} = \big(h(z)\big)_{ij}$, it follows that

$$\begin{aligned}
\big(h\big(\varphi^{-1}(0)\big)\big)_{ij} &= \bar\partial_2^i\partial_2^j\big(\mathbb{S}(z_1)^\alpha\mathbb{S}(z_2)^\beta\big)|_\triangle\\
&= (\beta)_j\mathbb{S}(z_1)^\alpha\bar\partial_2^i\big(\mathbb{S}(z_2)^{\beta+j}\bar z_2^j\big)|_\triangle\\
&= (\beta)_j\mathbb{S}(z_1)^\alpha\sum_{r=0}^{i}\binom{i}{r}\bar\partial_2^{(i-r)}\big(\mathbb{S}(z_2)^{\beta+j}\big)\bar\partial_2^r(\bar z_2^j)|_\triangle\\
&= (\beta)_j\mathbb{S}(z_1)^\alpha\sum_{r=0}^{i}\binom{i}{r}(\beta+j)_{i-r}\mathbb{S}(z_2)^{\beta+j+(i-r)}z_2^{i-r}r!\binom{j}{r}\bar z_2^{j-r}|_\triangle\\
&= (\beta)_j\mathbb{S}(z)^{\alpha+\beta+i+j}\bar z^{j-i}\sum_{r=0}^{i}r!\binom{i}{r}\binom{j}{r}(\beta+j)_{i-r}\mathbb{S}(z)^{-r}|z|^{2(i-r)},
\end{aligned}$$

for $i \le j$. Clearly,

$$\big(\mathcal{J}_{\varphi^{-1}}(0)\big)_{ij} = c(\varphi^{-1},0)^{-\frac{\alpha+\beta}{2}-n}\frac{(\beta)_j}{(\beta)_i}\binom{j}{i}c(\varphi^{-1},0)^{n-j}p(\varphi^{-1},0)^{j-i}$$

and

$$h(0)_{ii} = i!(\beta)_i,$$

$0 \le i \le j \le n$. We have

$$\begin{aligned}
\big(\overline{\mathcal{J}_{\varphi^{-1}}(0)}^t h(0)\mathcal{J}_{\varphi^{-1}}(0)\big)_{ij} &= \sum_{k=0}^{j}\big(\overline{\mathcal{J}_{\varphi^{-1}}(0)}^t h(0)\big)_{ik}\big(\mathcal{J}_{\varphi^{-1}}(0)\big)_{kj}\\
&= \sum_{k=0}^{i}\sum_{k=0}^{j}\big(\overline{\mathcal{J}_{\varphi^{-1}}(0)}^t\big)_{ik}\big(h(0)\big)_{kk}\big(\mathcal{J}_{\varphi^{-1}}(0)\big)_{kj}\\
&= \sum_{k=0}^{\min(i,j)}\big(\overline{\mathcal{J}_{\varphi^{-1}}(0)}^t\big)_{ik}\big(h(0)\big)_{kk}\big(\mathcal{J}_{\varphi^{-1}}(0)\big)_{kj}.
\end{aligned}$$

Now, for $0 \le i \le j \le n$,

$$\begin{aligned}
&\sum_{k=0}^{\min(i,j)}\big(\overline{\mathcal{J}_{\varphi^{-1}}(0)}^t\big)_{ik}\big(h(0)\big)_{kk}\big(\mathcal{J}_{\varphi^{-1}}(0)\big)_{kj}\\
&= |c(\varphi^{-1},0)|^{-\alpha-\beta-2n}\sum_{k=0}^{i}\Big(\frac{(\beta)_i}{(\beta)_k}\binom{i}{k}\overline{c(\varphi^{-1},0)}^{n-i}\,\overline{p(\varphi^{-1},0)}^{i-k}k!(\beta)_k\frac{(\beta)_j}{(\beta)_k}\\
&\qquad\qquad\binom{j}{k}c(\varphi^{-1},0)^{n-j}p(\varphi^{-1},0)^{j-k}\Big)\\
&= \mathbb{S}(z)^{\alpha+\beta+2n}\sum_{k=0}^{i}\frac{k!(\beta)_i(\beta)_j}{(\beta)_k}\binom{i}{k}\binom{j}{k}\\
&\qquad\qquad\big(t\mathbb{S}(z)\big)^{-n+i}(tz)^{i-k}\big(\bar{t}\mathbb{S}(z)\big)^{-n+j}(\overline{tz})^{j-k}\\
&= (\beta)_j\mathbb{S}(z)^{\alpha+\beta+i+j}\sum_{k=0}^{i}k!\binom{i}{k}\binom{j}{k}\frac{(\beta)_i}{(\beta)_k}z^{i-k}\bar{z}^{j-k}\\
&= (\beta)_j\mathbb{S}(z)^{\alpha+\beta+i+j}\bar{z}^{j-i}\sum_{k=0}^{i}k!\binom{i}{k}\binom{j}{k}\frac{(\beta)_i}{(\beta)_k}|z|^{2(i-k)}.
\end{aligned}$$

Clearly, to prove the desired equality we have to show that

$$\begin{aligned}
&\sum_{r=0}^{i}r!\binom{i}{r}\binom{j}{r}(\beta+j)_{i-r}\mathbb{S}(z)^{-r}|z|^{2(i-r)}\\
&\qquad = \sum_{k=0}^{i}k!\binom{i}{k}\binom{j}{k}\frac{(\beta)_i}{(\beta)_k}|z|^{2(i-k)} \qquad (5.18)
\end{aligned}$$

for $0 \leq i \leq j \leq n$.

But

$$\begin{aligned}
&\sum_{r=0}^{i} r! \binom{i}{r}\binom{j}{r} (\beta + j)_{i-r} (1 - |z|^2)^r |z|^{2(i-r)} \\
&= \sum_{r=o}^{i} r! \binom{i}{r}\binom{j}{r} (\beta + j)_{i-r} \sum_{l=0}^{r} (-1)^l \binom{r}{l} |z|^{2l} |z|^{2(i-r)} \\
&= \sum_{l=0}^{i} \sum_{r=l}^{i} (-1)^l r! \binom{i}{r}\binom{j}{r}\binom{r}{l} (\beta + j)_{i-r} |z|^{2(i-(r-l))} \\
&= \sum_{l=0}^{i} \sum_{r=0}^{i-l} (-1)^l (r+l)! \binom{i}{r+l}\binom{j}{r+l}\binom{r+l}{l} (\beta + j)_{i-r-l} |z|^{2(i-r)}.
\end{aligned}$$

For $0 \leq k \leq i - l$, the coefficient of $|z|^{2(i-k)}$ in the left-hand side of (5.18) is

$$\sum_{l=0}^{i} (-1)^l (k+l)! \binom{i}{k+l}\binom{j}{k+l}\binom{k+l}{l} (\beta + j)_{i-k-l},$$

which is the same as

$$\sum_{l=0}^{i-k} (-1)^l (k+l)! \binom{i}{k+l}\binom{j}{k+l}\binom{k+l}{l} (\beta + j)_{i-k-l},$$

for $0 \leq l \leq i - k \leq i$. So, to complete the proof we have to show that

$$\sum_{l=0}^{i-k} (-1)^l (k+l)! \binom{i}{k+l}\binom{j}{k+l}\binom{k+l}{l} (\beta + j)_{i-k-l} = k! \binom{i}{k}\binom{j}{k} \frac{(\beta)_i}{(\beta)_k},$$

for $0 \leq k \leq i, i \leq j$. But this follows from Lemma 5.5. □

6. The case of the tri-disc $\mathbb{D}^3$

We discuss the jet construction for $\mathbb{D}^3$. Let $K : \mathbb{D}^3 \times \mathbb{D}^3 \longrightarrow \mathbb{C}$ be a reproducing kernel. Following the jet construction of [12], we define

$$J^{(1,1)}K(z,w) = \begin{pmatrix} K(z,w) & \partial_2 K(z,w) & \partial_3 K(z,w) \\ \bar\partial_2 K(z,w) & \partial_2\bar\partial_2 K(z,w) & \bar\partial_2\partial_3 K(z,w) \\ \bar\partial_3 K(z,w) & \partial_2\bar\partial_3 K(z,w) & \bar\partial_3\partial_3 K(z,w) \end{pmatrix}, \; z, \, w \in \mathbb{D}^3.$$

As before, to retain the usual meaning of ∂ and $\bar\partial$ we replace $J^{(1,1)}K(z,w)$ by its transpose. For simplicity of notation, we let $G(z,w) := J^{(1,1)}K(z,w)^t$. In this notation, choosing the kernel function K on $\mathbb{D}^3$ to be

$$K(z,w) = (1 - z_1\bar w_1)^{-\alpha}(1 - z_2\bar w_2)^{-\beta}(1 - z_3\bar w_3)^{-\gamma},$$

we have

$$G(z,w)=\begin{pmatrix}(1-z\bar w)^2 & \beta z(1-z\bar w) & \gamma z(1-z\bar w)\\ \beta\bar w(1-z\bar w) & \beta(1+\beta z\bar w) & \beta\gamma z\bar w\\ \gamma\bar w(1-z\bar w) & \beta\gamma z\bar w & \gamma(1+\gamma z\bar w)\end{pmatrix}(1-z\bar w)^{-\alpha-\beta-\gamma-2},$$

for $z,w\in\mathbb{D},\ \alpha,\beta,\gamma>0$.

Theorem 6.1. *The adjoint of the multiplication operator M^* on the Hilbert space of $\mathbb{C}^3$-valued holomorphic functions on $\mathbb{D}$ with reproducing kernel G is in $\mathrm{B}_3(\mathbb{D})$. It is homogeneous and reducible. Moreover, M^* is unitarily equivalent to $M_1^*\oplus M_2^*$ for a pair of irreducible homogeneous operators M_1^* and M_2^* from $\mathrm{B}_2(\mathbb{D})$ and $\mathrm{B}_1(\mathbb{D})$ respectively.*

Proof. Although homogeneity of M^* follows from [6, Theorem 5.2.], we give an independent proof using the ideas we have developed in this note. Let

$$\tilde{\tilde G}(z,w)=G(0,0)^{1/2}G(z,0)^{-1}G(z,w)G(0,w)^{-1}G(0,0)^{1/2}.$$

Evidently, $\tilde{\tilde G}(z,0)=I$, that is, $\tilde{\tilde G}$ is a normalized kernel. The form of $\tilde{\tilde G}(z,w)$ for $z,w\in\mathbb{D}$ is $(1-z\bar w)^{-\alpha-\beta-\gamma-2}$ times

$$\begin{pmatrix}\begin{matrix}(1-z\bar w)^2-(\beta+\gamma)(1-z\bar w)z\bar w\\ +(\beta+\gamma)(1+\beta+\gamma)z^2\bar w^2\end{matrix} & -\sqrt{\beta}(1+\beta+\gamma)z^2\bar w & -\sqrt{\gamma}(1+\beta+\gamma)z^2\bar w\\ -\sqrt{\beta}(1+\beta+\gamma)z\bar w^2 & 1+\beta z\bar w & \sqrt{\beta\gamma}z\bar w\\ -\sqrt{\gamma}(1+\beta+\gamma)z\bar w^2 & \sqrt{\beta\gamma}z\bar w & 1+\gamma z\bar w\end{pmatrix}.$$

Let $U=\frac{1}{\sqrt{\beta+\gamma}}\begin{pmatrix}1&0&0\\0&\sqrt{\beta}&\sqrt{\gamma}\\0&-\sqrt{\gamma}&\sqrt{\beta}\end{pmatrix}$ which is unitary on $\mathbb{C}^3$. By a direct computation, we see that the equivalent normalized kernel $U\tilde{\tilde G}(z,w)\overline{U}^t$ is equal to the direct sum $G_1(z,w)\oplus G_2(z,w)$, where $G_2(z,w)=(1-z\bar w)^{-\alpha-\beta-\gamma-2}$ and

$$G_1(z,w)=\begin{pmatrix}\begin{matrix}(1-z\bar w)^2-(\beta+\gamma)(1-z\bar w)z\bar w\\ +(\beta+\gamma)(1+\beta+\gamma)z^2\bar w^2\end{matrix} & -\sqrt{\beta+\gamma}(1+\beta+\gamma)z^2\bar w\\ -\sqrt{\beta+\gamma}(1+\beta+\gamma)z\bar w^2 & 1+(\beta+\gamma)z\bar w\end{pmatrix}(1-z\bar w)^{-\alpha-\beta-\gamma-2}.$$

It follows that M^* is unitarily equivalent to a reducible operator by an application of Lemma 4.3, that is, M^* is reducible. If we replace β by $\beta+\gamma$ in Theorem 4.2 take $n=1$, then

$$K(z,w)=\begin{pmatrix}(1-z\bar w)^2 & (\beta+\gamma)z(1-z\bar w)\\ (\beta+\gamma)\bar w(1-z\bar w) & (\beta+\gamma)(1+(\beta+\gamma)z\bar w)\end{pmatrix}(1-z\bar w)^{-\alpha-\beta-\gamma-2},$$

for $z,w\in\mathbb{D}$. We observe that

$$G_1(z,w)=K(0,0)^{1/2}K(z,0)^{-1}K(z,w)K(0,w)^{-1}K(0,0)^{1/2}$$

and $G_1(z,0)=I$, as is to be expected. The multiplication operator corresponding to G_1, which we denote by M_1, is unitarily equivalent to $M^{(\alpha,\beta+\gamma)}$ by Lemma 4.3. Hence it is in $\mathrm{B}_2(\mathbb{D})$ by [12, Proposition 3.6]. Since both homogeneity and irreducibility are invariant under unitary equivalence it follows, by an easy application of Lemma 4.3, Theorem 4.2 and Theorem 5.1 that M_1^* is a irreducible homogeneous operator in $\mathrm{B}_2(\mathbb{D})$. Irreducibility of M_1^* also follows from Proposition 4.11. Let M_2 be the multiplication operator on the Hilbert space of scalar-valued holomorphic functions with reproducing kernel G_2. Again, M_2^* is in $\mathrm{B}_1(\mathbb{D})$. The operator M_2 is irreducible by [9, Corollary 1.19]. Homogeneity of M_2^* was first established in [17], see also [22]. An alternative proof is obtained when we observe that $\Gamma : \text{Möb} \times \mathbb{D} \longrightarrow \mathbb{C}$, where $\Gamma_{\varphi^{-1}}(z) = \big((\varphi^{-1})'(z)\big)^{\frac{\alpha+\beta+\gamma}{2}+1}$ is a cocycle such that $G_2(z,w) = \Gamma_{\varphi^{-1}}(z) G_2\big(\varphi^{-1}(z), \varphi^{-1}(w)\big)\overline{\Gamma_{\varphi^{-1}}(z)}$ for $z, w \in \mathbb{D}, \varphi \in$Möb. Now we conclude that M^* is homogeneous as it is unitarily equivalent to the direct sum of two homogeneous operators. Also M^* is in $\mathrm{B}_3(\mathbb{D})$ being the direct sum of two operators from the Cowen-Douglas class. □

References

[1] J. Arazy and G. Zhang, *Homogeneous multiplication operators on bounded symmetric domains*, J. Func. Anal. **202** (2003), 44–66.

[2] N. Aronszajn, *Theory of reproducing kernels*, Trans. Amer.Math. Soc. **68** (1950), 337–404.

[3] B. Bagchi and G. Misra, *Homogeneous tuples of operators and systems of imprimitivity*, Contemporary Mathematics, **185** (1995), 67–76.

[4] ———, *Homogeneous tuples of multiplication operators on twisted Bergman space*, J. Funct. Anal. **136** (1996), 171–213.

[5] ———, *Constant characteristic functions and Homogeneous operators*, J. Operator Theory, **37** (1997), 51–65.

[6] ———, *Homogeneous operators and the projective representations of the Möbius group: a survey*, Proc. Indian Acad. Sci. Math. Sci **111** (2001), no. 4, 415–437.

[7] ———, *The homogeneous shifts*, J. Funct. Anal. **204** (2003), 293–319.

[8] I. Biswas and G. Misra, $\widetilde{\mathrm{SL}(2,\mathbb{R})}$-*homogeneous vector bundles*, preprint, 2006.

[9] M.J. Cowen and R.G. Douglas, *Complex geometry and operator theory*, Acta Math. **141** (1978), 187 – 261.

[10] ———, *On operators possessing an open set of eigenvalues*, Memorial Conf. for Féjer-Riesz, Colloq. Math. Soc. J. Bolyai, 1980, pp. 323–341.

[11] R.E. Curto and N. Salinas, *Generalized Bergman kernels and the Cowen-Douglas theory*, Amer. J. Math. **106** (1984), 447–488.

[12] R.G. Douglas, G. Misra and C. Varughese, *On quotient modules- the case of arbitrary multiplicity*, J. Funct. Anal. **174**(2000), 364–398.

[13] R.G. Douglas and G. Misra, *On quasi-free Hilbert modules*, New York J. Math., **11** (2005), 547–561.

[14] *Equivalence of quotient Hilbert modules – II*, To appear, Trans. Amer. Math. Soc.

[15] A. Korányi and G. Misra, *New constructions of homogeneous operators*, C. R. Acad. Sci. Paris, ser. I **342** (2006), 933–936.

[16] A. Korányi and G. Misra, *Homogeneous operators on Hilbert spaces of holomorphic functions – I*, preprint, 2006.

[17] G. Misra, *Curvature and the backward shift operators*, Proc. Amer. Math. Soc. **91** (1984), no. 1, 105–107.

[18] G. Misra, *Curvature and discrete series representations of* $SL_2(\mathbb{R})$, J. Int. Eqn. Oper. Th. **9** (1986), 452–459.

[19] G. Misra and N.S.N. Sastry, *Homogeneous tuples of operators and holomorphic discrete series representation of some classical groups*, J. Operator Theory **24** (1990), 23–32.

[20] G. Misra and S.S. Roy, *Curvature calculations for homogeneous operators*, preprint, 2006.

[21] L. Peng and G. Zhang, *Tensor products of holomorphic representations and bilinear differential operators*, J. Funct. Anal. **210** (2004), 171–192.

[22] D.R. Wilkins, *Homogeneous vector bundles and Cowen-Douglas operators*, Internat. J. Math. **4** (1993), no. 3, 503–520.

Gadadhar Misra and Subrata Shyam Roy
Indian Statistical Institute,
R. V. College Post Office,
Bangalore – 560 059, India
e-mail: `gm@isibang.ac.in`
e-mail: `ssroy@isibang.ac.in`

Operator Theory:
Advances and Applications, Vol. 176, 199–254

Canonical Forms for Symmetric and Skewsymmetric Quaternionic Matrix Pencils

Leiba Rodman

Abstract. Canonical forms are given for pairs of quaternionic matrices, or equivalently matrix pencils, with various symmetry properties, under strict equivalence and symmetry respecting congruence. Symmetry properties are induced by involutory antiautomorphisms of the quaternions which are different from the quaternionic conjugation. Some applications are developed, in particular, canonical forms for quaternionic matrices that are symmetric or skewsymmetric with respect to symmetric or skewsymmetric quaternion-valued inner products. Another application concerns joint numerical cones of pairs of skewsymmetric quaternionic matrices.

Mathematics Subject Classification (2000). 15A33.

Keywords. Quaternionic matrices, involutory antiautomorphisms, symmetric matrix pairs, skewsymmetric matrix pairs, canonical forms.

1. Introduction

The theory of bilinear and quadratic forms over fields is central in many areas of mathematics and its applications. A key part of this theory is understanding of transformations of coordinates leading to simplifications of the original form, and eventually to canonical representations. Many applications require simultaneous consideration of two bilinear or quadratic forms, often with various symmetry properties. For the real and complex fields the theory of reduction of a pair of forms with symmetries is classical, see for example [46], [34], [35] and numerous references there.

In this paper, we focus on the skew field of real quaternions. Recently, there has been a renewed interest in all aspects of quaternionic mathematics, motivated in part by diverse applications, in particular in physics, (we reference here only

Research supported in part by NSF grant DMS-0456625.

two books [1], [28] among many on this subject) control systems [5], [39], and numerical analysis [19], [12]. Recent works on quaternionic linear algebra and matrix theory include [32], [47], [50], [18], [2], [51]. There is an extensive recent literature on hyperholomorphic functions – quaternionic analogues of analytic functions in complex analysis – and related topics; we mention here only [15], [3], [4], [27].

In contrast to the classical theory and abundant literature on bilinear, sesquilinear, and quadratic forms and pairs of forms over the reals and the complexes, for quaternionic forms the literature is not nearly as extensive. Nevertheless, it goes back at least six decades; [30], [37], [10] are early works on this subject. A list (far from complete) of more recent works includes [16], [17], [32], [43], [44], [45]; we mention also the books [8] and [11]; the latter contains a detailed treatment of the subject.

In this paper we derive canonical forms for pairs of quaternionic matrices with various symmetry properties, under strict equivalence and symmetry respecting congruences, and develop some applications, in particular, canonical forms for quaternionic matrices that are symmetric or skewsymmetric with respect to symmetric or skewsymmetric quaternionic inner products. The symmetry is induced by an involutory antiautomorphism of the quaternions, and in this paper only so-called nonstandard involutory antiautomorphisms are considered, i.e., those that are different from the antiautomorphism of the quaternionic conjugation. Another application involves numerical cones of pairs of skewsymmetric quaternionic matrices.

Our treatment is based on matrix theory techniques, in the spirit of the classical treatise [20], [21], and later developed in [46], [34], [35], where the corresponding problems were addressed for real and complex matrices. The present paper is largely expository, and many of the results presented here are not new.

We now turn to a more precise description of the problems studied here. Denote by H the standard skew field of real quaternions. Consider pencils of quaternionic matrices $A + tB$, where A and B are $m \times n$ matrices with entries in H (notation: $A, B \in \mathsf{H}^{m\times n}$), and t is an independent real variable. Canonical form of the pencil $A + tB$ under *strict equivalence*:

$$A + tB \qquad \mapsto \qquad P(A + tB)Q,$$

where $P \in \mathsf{H}^{m\times m}$ and $Q \in \mathsf{H}^{n\times n}$ are invertible matrices, is known as the Kronecker form. Equivalently, it is the canonical form of ordered pairs of matrices (A, B) under the group action $(A, B) \mapsto (PAQ, PBQ)$, with invertible quaternionic matrices P and Q.

If ϕ is an involutory antiautomorphism of H (see Section 2 for more details), and $A \in \mathsf{H}^{m\times n}$, we denote by A_ϕ the $n \times m$ quaternionic matrix obtained by applying ϕ entrywise to the *transposed* matrix A^T. A matrix $A \in \mathsf{H}^{n\times n}$ is called *ϕ-symmetric*, resp. *ϕ-skewsymmetric*, if $A_\phi = A$, resp. $A_\phi = -A$. For ϕ the *standard antiautomorphism*, i.e., $\phi(x) = \overline{x}$, $x \in \mathsf{H}$, where by $\overline{x}$ we denote the quaternionic conjugation, we have that A is ϕ-symmetric if and only if A is *hermitian*: $A = A^*$, and A is ϕ-skewsymmetric if and only if A is skew-hermitian: $A = -A^*$. Two square size quaternionic matrices A and B are said to be *ϕ-congruent* if $A = S_\phi BS$ for some invertible quaternionic matrix S.

We are concerned in this paper with the following main problems, for nonstandard involutory antiautomorphisms ϕ:

Problem 1.1. *Study the canonical forms, their properties and applications for pairs of ϕ-symmetric $n \times n$ quaternionic matrices under strict equivalence and under the action of simultaneous ϕ-congruence defined as follows:*

$$(A, B), \quad A, B \ \phi\text{-symmetric} \qquad \mapsto \qquad (S_\phi AS, S_\phi BS), \quad S \text{ invertible.}$$

Problem 1.2. *Study the canonical forms under strict equivalence and under simultaneous ϕ-congruence, their properties and applications, for pairs of ϕ-skewsymmetric quaternionic matrices.*

We also consider Problem 1.1 for pairs of quaternionic matrices that are symmetric simultaneously with respect to several nonstandard involutory antiautomorphisms.

Analogous problems for pairs of matrices (A, B), where A is ϕ-symmetric and B is ϕ-skewsymmetric, as well as for symmetric/skewsymmetric pairs of quaternionic matrices with respect to the standard antiautomorphism of the quaternionic conjugation, will be studied separately [41], [42].

Problems 1.1 and 1.2 in the context of quaternionic conjugation have been studied in [44], where the canonical forms were obtained, and see also [32].

Besides the introduction, the paper consists of seven sections. Sections 2, 3, and 4 contain an exposition of some basic facts of the skew field H and of quaternionic linear algebra. This material serves as a background for the rest of the paper. Many results there, when convenient and instructive, are given with proofs, especially in cases when proofs are not widely available. Section 5 is devoted to matrix polynomials of real variable with quaternionic coefficients. The main theme here is the analogue of the Smith form for matrix polynomials with coefficients in a field. In contrast with the case of (commutative) field, generally there is no simple uniqueness criteria. Based on this material, the Kronecker form for quaternionic matrix pencils is presented in Section 6. The canonical forms for pairs of ϕ-symmetric quaternionic matrices are presented, with full proofs, in Section 7. Here ϕ is a fixed nonstandard involutory antiautomorphism. It turns out that the forms are the same under strict equivalence and under ϕ-congruence. As an application, canonical forms of quaternionic matrices that are symmetric with respect to a nondegenerate ϕ-symmetric inner products are derived. Finally, in Section 8 we give, again with full proofs, canonical forms for pairs of ϕ-skewsymmetric quaternionic matrices. In contrast with ϕ-symmetric pairs, here strictly equivalent pairs need not be ϕ-congruent. In fact, every strictly equivalent class of pairs of ϕ-skewsymmetric matrices consists of finitely many classes of ϕ-congruence, which are distinguished from each other by a certain *sign characteristic.* This phenomenon is well known in other contexts of pairs of matrices with various symmetry properties, see for example [24] for complex or real matrices that are selfadjoint or unitary with respect to a nondegenerate indefinite inner product on C^n or on R^n.

We denote by R, C and H the real and complex fields and the skew field of real quaternions, respectively. The standard imaginary units in H will be denoted $\mathsf{i}, \mathsf{j}, \mathsf{k}$; thus, $\mathsf{i}^2 = \mathsf{j}^2 = \mathsf{k}^2 = -1$ and $\mathsf{jk} = -\mathsf{kj} = \mathsf{i}$, $\mathsf{ij} = -\mathsf{ji} = \mathsf{k}$, $\mathsf{ki} = -\mathsf{ik} = \mathsf{j}$. In the sequel, C will be identified with $\mathsf{R} + \mathsf{i}\mathsf{R}$. For a quaternion $x = a_0 + a_1\mathsf{i} + a_2\mathsf{j} + a_3\mathsf{k}$, $a_0, a_1, a_2, a_3 \in \mathsf{R}$, we let $\mathfrak{R}(x) := a_0$ and $\mathfrak{V}(x) := a_1\mathsf{i} + a_2\mathsf{j} + a_3\mathsf{k}$ be the *real* and the *vector* parts of x, respectively. The conjugate quaternion $a_0 - a_1\mathsf{i} - a_2\mathsf{j} - a_3\mathsf{k}$ is denoted by $\overline{x}$, and $|x| = \sqrt{a_0^2 + a_1^2 + a_2^2 + a_3^2}$ stands for the norm of x.

The set $\mathsf{H}^{n\times 1}$ of n-component columns with quaternionic entries will be considered as a right quaternionic vector space. We denote by $\operatorname{Span}\{x_1, \dots, x_k\}$ the quaternionic vector subspace of $\mathsf{H}^{n\times 1}$ spanned by the vectors $x_1, \dots, x_k \in \mathsf{H}^{n\times 1}$, and by $\operatorname{Span}_{\mathsf{R}}\{x_1, \dots, x_k\}$ the *real* vector subspace spanned by the same vectors. The standard Euclidean inner product on $\mathsf{H}^{n\times 1}$ is defined by

$$(x, y) := y^* x \in \mathsf{H}, \qquad x, y \in \mathsf{H}^{n\times 1},$$

where $y^* := \overline{y^T} \in \mathsf{H}^{1\times n}$. We have

$$(x\alpha, y\beta) = \overline{\beta}(x, y)\alpha, \quad x, y \in \mathsf{H}^{n\times 1}, \quad \alpha, \beta \in \mathsf{H}.$$

Similarly, the set $\mathsf{H}^{1\times n}$ of n-component rows with quaternionic entries will be considered as a left quaternionic vector space.

The block diagonal matrix with the diagonal blocks $X_1, \dots, X_p$ (in this order) will be denoted $\operatorname{diag}(X_1, \dots, X_p)$ or $X_1 \oplus \cdots \oplus X_p$. The identity matrix of size u is denoted I_u, and the $u \times v$ zero matrix is denoted $0_{u\times v}$, often abbreviated 0_u if $u = v$.

2. Algebra of quaternions

In this section we recall basic facts about the quaternionic algebra, some not so well known. For the readers' convenience, we present also some results that are not strictly needed in the sequel; often proofs are provided. Besides the general references on quaternionic linear algebra given in the introduction, we refer also to [33].

An ordered triple of quaternions (q_1, q_2, q_3) is said to be a *units triple* if

$$\begin{aligned} q_1^2 = q_2^2 = q_3^2 = -1, &\qquad q_1q_2 = -q_2q_1 = q_3, \\ q_2q_3 = -q_3q_2 = q_1, &\qquad q_3q_1 = -q_1q_3 = q_2. \end{aligned} \tag{2.1}$$

Proposition 2.1. *An ordered triple (q_1, q_2, q_3), $q_j \in \mathsf{H}$, is a units triple if and only if there exists a 3×3 real orthogonal matrix $P = [p_{\alpha,\beta}]_{\alpha,\beta=1}^3$ with determinant 1 such that*

$$q_\alpha = p_{1,\alpha}\mathsf{i} + p_{2,\alpha}\mathsf{j} + p_{3,\alpha}\mathsf{k}, \quad \alpha = 1, 2, 3. \tag{2.2}$$

Proof. A straightforward computation verifies that $x \in \mathsf{H}$ satisfies $x^2 = -1$ if and only if

$$x = a_1\mathsf{i} + a_2\mathsf{j} + a_3\mathsf{k},$$

where $a_1, a_2, a_3 \in \mathsf{R}$ and $a_1^2 + a_2^2 + a_3^2 = 1$. Thus, we may assume that q_α are given by (2.2) with the vectors $p_\alpha := (p_{1,\alpha}, p_{2,\alpha}, p_{3,\alpha})^T \in \mathsf{R}^{3\times 1}$ having Euclidean norm 1, for $\alpha = 1, 2, 3$. Next, observe that

$$q_u q_v = -p_u^T p_v + \begin{bmatrix} \mathsf{i} & \mathsf{j} & \mathsf{k} \end{bmatrix} (p_u \times p_v), \quad u, v \in \{1, 2, 3\},$$

where in the right-hand side $\times$ denotes the cross product of vectors in $\mathsf{R}^{3\times 1}$. The result of Proposition 2.1 now follows easily. ☐

In particular, for every units triple (q_1, q_2, q_3) the quaternions $1, q_1, q_2, q_3$ form a basis of the real vector space H.

Quaternions $x, y \in \mathsf{H}$ are said to be *similar* if

$$x = \alpha^{-1} y \alpha \tag{2.3}$$

for some $\alpha \in \mathsf{H} \setminus \{0\}$.

Proposition 2.2. *Two quaternions*

$$x = a_0 + a_1\mathsf{i} + a_2\mathsf{j} + a_3\mathsf{k} \quad \textit{and} \quad y = b_0 + b_1\mathsf{i} + b_2\mathsf{j} + b_3\mathsf{k}$$

are similar if and only if

$$\mathfrak{R}(x) = \mathfrak{R}(y) \quad \text{and} \quad |\mathfrak{V}(x)| = |\mathfrak{V}(y)|. \tag{2.4}$$

In the sequel we will use the term "similar" for quaternions x and y that satisfy (2.4); of course, by Proposition 2.2 this is equivalent to the similarity relation defined by (2.3).

Next, we consider endomorphisms and antiendomorphisms of quaternions.

A map $\phi : \mathsf{H} \longrightarrow \mathsf{H}$ is called an *endomorphism*, resp., an *antiendomorphism*, if $\phi(xy) = \phi(x)\phi(y)$, resp., $\phi(xy) = \phi(y)\phi(x)$ for all $x, y \in \mathsf{H}$, and $\phi(x + y) = \phi(x) + \phi(y)$ for all $x, y \in \mathsf{H}$.

Proposition 2.3. *Let ϕ be an endomorphism or an antiendomorphism of H. Assume that ϕ does not map H into zero. Then ϕ is one-to-one and onto H, thus ϕ is in fact an automorphism or an antiautomorphism. Moreover, ϕ is real linear, and representing ϕ as a 4×4 real matrix with respect to the basis $\{1, \mathsf{i}, \mathsf{j}, \mathsf{k}\}$, we have:*

(a) *ϕ is an automorphism if and only if*

$$\phi = \begin{bmatrix} 1 & 0 \\ 0 & T \end{bmatrix}, \tag{2.5}$$

where T is a 3×3 real orthogonal matrix of determinant 1.

(b) *ϕ is an antiautomorphism if and only if ϕ has the form (2.5) where T is a 3×3 real orthogonal matrix of determinant -1.*

(c) *ϕ is an involutory antiautomorphism if and only if*

$$\phi = \begin{bmatrix} 1 & 0 \\ 0 & T \end{bmatrix},$$

where either $T = -I_3$ or T is a 3×3 real orthogonal symmetric matrix with eigenvalues $1,\ 1,\ -1$.

Proof. Clearly, ϕ is one-to-one (because the kernel of ϕ is a two-sided ideal of H, and H has no nontrivial two sided ideals), and $\phi(x) = x$ for every real rational x.

We will use an observation which can be easily checked by straightforward algebra. Namely, *if $x \in \mathsf{H}$ is nonreal, then the commutant of x, namely, the set of all $y \in \mathsf{H}$ such that $xy = yx$ coincides with the set of all quaternions of the form $a + bx$, where $a, b \in \mathsf{R}$.*

Next, we prove that ϕ maps reals into reals. Arguing by contradiction, assume that $\phi(x)$ is nonreal for some real x. Since $xy = yx$ for every $y \in \mathsf{H}$, we have that $\phi(\mathsf{H})$ is contained in the commutant of $\phi(x)$, i.e., by the above observation

$$\phi(\mathsf{H}) \subseteq \mathsf{R} + \mathsf{R}\phi(x).$$

However, the set $\mathsf{R} + \mathsf{R}\phi(x)$ contains only two square roots of -1, namely,

$$\pm \frac{\mathfrak{V}(\phi(x))}{|\mathfrak{V}(\phi(x))|}.$$

On the other hand, H contains a continuum of square roots of -1. Since ϕ maps square roots of -1 onto square roots of -1, ϕ cannot be one-to-one, a contradiction. Thus, ϕ maps reals into reals, and the restriction of ϕ to R is a nonzero endomorphism of the field of real numbers. Now R has no nontrivial endomorphisms (indeed, any nonzero endomorphism of R fixes every rational number, and since only nonnegative real numbers have real square roots, any such endomorphism is also order preserving, and these properties easily imply that any such endomorphism must be the identity). Therefore, we must have $\phi(x) = x$ for all $x \in \mathsf{R}$. Now clearly ϕ is a real linear map. Representing ϕ as a 4×4 matrix with respect to the basis $\{1, \mathsf{i}, \mathsf{j}, \mathsf{k}\}$, we obtain the result of part (a) from Proposition 2.1. For part (b), note that if ϕ is an antiautomorphism, then a composition of ϕ with any fixed antiautomorphism is an automorphism. Taking a composition of ϕ with the antiautomorphism of standard conjugation $\mathsf{i} \longrightarrow -\mathsf{i}$, $\mathsf{j} \longrightarrow -\mathsf{j}$, $\mathsf{k} \longrightarrow -\mathsf{k}$, we see by part (a) that the composition has the form (2.5) with T a real orthogonal matrix of determinant 1. Since the standard conjugation has the form (2.5) with $T = -I$, we obtain the result of part (b). Finally, clearly ϕ is involutory if and only if the matrix T of (2.5) has eigenvalues ± 1, and (c) follows at once from (a) and (b). $\square$

If the former case of (c) holds, then ϕ is the standard conjugation, and we say that ϕ is *standard.* If the latter case of (c) holds, then we have a family of involutory antiautomorphisms parameterized by one-dimensional real subspaces (representing eigenvectors of T corresponding to the eigenvalue -1) in R^3, and we say that these involutory antiautomorphisms are *nonstandard.*

The following fact will be useful in the sequel. The abbreviation "iaa" stands throughout the paper for "involutory antiautomorphism".

Lemma 2.4. *Let ϕ_1 and ϕ_2 be two distinct nonstandard iaa's. Then for any $\alpha \in \mathsf{H}$, the equality $\phi_1(\alpha) = \phi_2(\alpha)$ holds if and only if $\phi_1(\alpha) = \phi_2(\alpha) = \alpha$.*

Proof. The part "if" is obvious. To prove the "only if" part, let T_1 and T_2 be the 3×3 real orthogonal symmetric matrices with eigenvalues $1, 1, -1$ such that

$$\phi_j = \begin{bmatrix} 1 & 0 \\ 0 & T_j \end{bmatrix}, \quad j = 1, 2,$$

as in Proposition 2.3(c). Then the 'only if' part amounts to the following: If $T_1 \neq T_2$ and $T_1 x = T_2 x$ for some $x \in \mathsf{R}^{3\times 1}$, then $T_1 x = T_2 x = x$. Considering the orthogonal complement of a common eigenvector of T_1 and T_2 corresponding to the eigenvalue 1, the proof reduces to the statement that if $\widehat{T_1} \neq \widehat{T_2}$ are 2×2 real orthogonal symmetric matrices with determinants -1, then $\det(\widehat{T_1} - \widehat{T_2}) \neq 0$. This can be verified by elementary matrix manipulations. □

Proposition 2.3 allows us to prove easily, using elementary linear algebra, the following well-known fact:

Proposition 2.5. *Every automorphism of H is inner, i.e., if $\phi : \mathsf{H} \to \mathsf{H}$ is an automorphism, then there exists $\alpha \in \mathsf{H} \setminus \{0\}$ such that*

$$\phi(x) = \alpha^{-1} x \alpha, \qquad \forall\, x \in \mathsf{H}. \tag{2.6}$$

Proof. An elementary (but tedious) calculation shows that for

$$\alpha = a + b\mathsf{i} + c\mathsf{j} + d\mathsf{k} \in \mathsf{H} \setminus \{0\},$$

the 4×4 matrix representing the R-linear transformation $x \mapsto \alpha^{-1} x \alpha$ in the standard basis $\{1, \mathsf{i}, \mathsf{j}, \mathsf{k}\}$ is equal to

$$\begin{bmatrix} 1 & 0 \\ 0 & S \end{bmatrix},$$

where

$$S := \frac{1}{|\alpha|^2} \left(\begin{bmatrix} b \\ c \\ d \end{bmatrix} \begin{bmatrix} b & c & d \end{bmatrix} + \begin{bmatrix} a & d & -c \\ -d & a & b \\ c & -b & a \end{bmatrix}^2 \right). \tag{2.7}$$

In view of Proposition 2.3, the matrix S is orthogonal with determinant 1, because the transformation $x \mapsto \alpha^{-1} x \alpha$ is an automorphism. Conversely, every 3×3 real orthogonal matrix T with determinant 1 has the form of the right-hand side of (2.7): Indeed, if $T = I$, choose $a \neq 0$ and $b = c = d = 0$. If $T \neq I$, choose $(b, c, d)^T$ as a unit length eigenvector of T corresponding to the eigenvalue 1. Observe that $(b, c, d)^T$ is also an eigenvector of the matrix in the right-hand side of (2.7) corresponding to the eigenvalue 1. Since the trace of the right-hand side of (2.7) is equal to

$$1 + \frac{2}{a^2 + b^2 + c^2 + d^2}(a^2 - b^2 - c^2 - d^2),$$

it remains to choose a so that

$$\frac{a^2 - b^2 - c^2 - d^2}{a^2 + b^2 + c^2 + d^2}$$

coincides with the real part of those eigenvalues of T which are different from 1. The choice of $a \in \mathsf{R}$ is always possible because the real part of the eigenvalues of T other than 1 is between -1 and 1 and is not equal to 1. $\square$

If ϕ is a nonstandard iaa, we denote by $\operatorname{Inv}(\phi)$ the set of quaternions left invariant by ϕ:

$$\operatorname{Inv}(\phi) := \{x \in \mathsf{H} \,:\, \phi(x) = x\}.$$

Clearly, $\operatorname{Inv}(\phi)$ is a 3-dimensional real vector space spanned by 1 and a pair of quaternions q_1 and q_2 satisfying the equalities $q_1^2 = q_2^2 = -1$ and $q_1 q_2 = -q_2 q_1$.

3. Quaternionic linear algebra

In this section we present basic and mostly well-known facts on quaternionic matrices. General references for this and related material are the books [11], [47], as well as a recent review paper [50]; and see also the references in [50].

To start with, we consider extensions of iaa's to quaternionic matrices in a natural way. Let ϕ be an iaa of H. If $X \in \mathsf{H}^{m\times n}$, then $\phi(X) \in \mathsf{H}^{m\times n}$ stands for the matrix obtained from X by applying ϕ to X entrywise. For $A \in \mathsf{H}^{m\times n}$, we let

$$A_\phi = \phi(A^T) \in \mathsf{H}^{n\times m},$$

where A^T is the transpose of A. In particular, for ϕ the quaternionic conjugation A_ϕ is just the conjugate transpose A^* of A. The following algebra is easily observed:

Proposition 3.1.

(a) $(\alpha A + \beta B)_\phi = A_\phi \phi(\alpha) + B_\phi \phi(\beta)$, *for* $\alpha, \beta \in \mathsf{H}$, $A, B \in \mathsf{H}^{m\times n}$.
(b) $(A\alpha + B\beta)_\phi = \phi(\alpha) A_\phi + \phi(\beta) B_\phi$, *for* $\alpha, \beta \in \mathsf{H}$, $A, B \in \mathsf{H}^{m\times n}$.
(c) $(AB)_\phi = B_\phi A_\phi$, *for* $A \in \mathsf{H}^{m\times n}$, $B \in \mathsf{H}^{n\times p}$.
(d) $(A_\phi)_\phi = A$, *where* $A \in \mathsf{H}^{m\times n}$.
(e) *If* $A \in \mathsf{H}^{n\times n}$ *is invertible, then* $(A_\phi)^{-1} = (A^{-1})_\phi$.

Next, we present useful embeddings of quaternionic matrices into the algebras of real and complex matrices.

With every units triple $q = (q_1, q_2, q_3)$ we associate an embedding of $\mathsf{H}^{m\times n}$ into $\mathsf{R}^{4m\times 4n}$, as follows. Write $x \in \mathsf{H}$ as a linear combination

$$x = a_0 + a_1 q_1 + a_2 q_2 + a_3 q_3, \quad a_0, a_1, a_2, a_3 \in \mathsf{R}.$$

Then we define

$$\mho_{q,\mathsf{R}}(x) = \begin{bmatrix} a_0 & -a_1 & a_3 & -a_2 \\ a_1 & a_0 & -a_2 & -a_3 \\ -a_3 & a_2 & a_0 & -a_1 \\ a_2 & a_3 & a_1 & a_0 \end{bmatrix},$$

and for

$$X = [x_{\alpha,\beta}]_{\alpha=1,\beta=1}^{m,n} \in \mathsf{H}^{m\times n}$$

define

$$\mho_{q,\mathsf{R}}(X) = [\mho_{q,\mathsf{R}}(x_{\alpha,\beta})]_{\alpha=1,\beta=1}^{m,n} \in \mathsf{R}^{4m\times 4n}.$$

The algebraic properties of the map $\mho_{q,\mathsf{R}}$ are well known:

Proposition 3.2. *The map $\mho_{q,\mathsf{R}}$ is a one-to-one *-homomorphism of real algebras:*

$$\mho_{q,\mathsf{R}}(aX+bY) = a\mho_{q,\mathsf{R}}(X) + b\mho_{q,\mathsf{R}}(Y), \quad \forall \;\; X,Y \in \mathsf{H}^{m\times n}, \quad a,b \in \mathsf{R};$$
$$\mho_{q,\mathsf{R}}(XY) = \mho_{q,\mathsf{R}}(X)\mho_{q,\mathsf{R}}(Y), \quad \forall \;\; X \in \mathsf{H}^{m\times n}, \;\; Y \in \mathsf{H}^{n\times p};$$
$$\mho_{q,\mathsf{R}}(X^*) = (\mho_{q,\mathsf{R}}(X))^T, \quad \forall \;\; X \in \mathsf{H}^{m\times n}.$$

The proof is straightforward.

Next, let $q=(q_1,q_2,q_3)$ be a units triple. Write $X \in \mathsf{H}^{m\times n}$ in the form

$$X = X_{11} + q_1X_{12} + (X_{21}+q_1X_{22})q_2,$$

where $X_{11}, X_{12}, X_{21}, X_{22} \in \mathsf{R}^{m\times n}$. Then we define

$$\mho_{q,\mathsf{C}}(X) = \begin{bmatrix} X_{11}+\mathrm{i}X_{12} & X_{21}+\mathrm{i}X_{22} \\ -X_{21}+\mathrm{i}X_{22} & X_{11}-\mathrm{i}X_{12} \end{bmatrix} \in \mathbb{C}^{2m\times 2n}. \tag{3.1}$$

For the well-known general properties of the map $\mho_{q,\mathsf{C}}$ see [3], [50], for example. Again, it is a *-homomorphism:

Proposition 3.3. *The map $\mho_{q,\mathsf{C}}$ is a one-to-one *-homomorphism of real algebras:*

$$\mho_{q,\mathsf{C}}(aX+bY) = a\mho_{q,\mathsf{C}}(X) + b\mho_{q,\mathsf{C}}(Y), \quad \forall \;\; X,Y \in \mathsf{H}^{m\times n}, \quad a,b \in \mathsf{R};$$
$$\mho_{q,\mathsf{C}}(XY) = \mho_{q,\mathsf{C}}(X)\mho_{q,\mathsf{C}}(Y), \quad \forall \;\; X \in \mathsf{H}^{m\times n}, \;\; Y \in \mathsf{H}^{n\times p};$$
$$\mho_{q,\mathsf{C}}(X^*) = (\mho_{q,\mathsf{C}}(X))^*, \quad \forall \;\; X \in \mathsf{H}^{m\times n}.$$

We note the following connection between the action of a nonstandard iaa and the complex representation $\mho_{q,\mathsf{C}}$:

Proposition 3.4. *Let ϕ be a nonstandard iaa, and let $q=(q_1,q_2,q_3)$ be a units triple such that*

$$\mathrm{Inv}\,(\phi) = \mathrm{Span}_{\mathsf{R}}\{1, q_1, q_3\}.$$

Then for the map

$$\mho_{q,\mathsf{C}}(X) \,:\, \mathsf{H}^{m\times n} \;\longrightarrow\; \mathsf{C}^{2m\times 2n}$$

given by (3.1) *we have*

$$\mho_{q,\mathsf{C}}(X_\phi) = (\mho_{q,\mathsf{C}}(X))^T. \tag{3.2}$$

The proof is by a straightforward verification.

Proposition 3.5. *Let $A \in \mathsf{H}^{m\times n}$. Then the columns of A are linearly independent* (*as elements of a right quaternionic vector space*) *if and only A is left invertible, i.e., $BA = I$ for some $B \in \mathsf{H}^{n\times m}$.*

Analogously, the rows of A are linearly independent (*as elements of a left quaternionic vector space*) *if and only if A is right invertible.*

A proof follows from a rank decomposition: If k is the *rank*, i.e., the rank of the column space of A as a right quaternionic space, which is the same as the rank of the row space of A as a left quaternionic space, then

$$A = P \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix} Q$$

for some invertible matrices P and Q. The existence of a rank decomposition is proved by standard linear algebra arguments.

Next, we consider eigenvalues. A quaternion α is said to be an *eigenvalue* of $A \in \mathsf{H}^{n\times n}$ if $Ax = x\alpha$ for some nonzero $x \in \mathsf{H}^{n\times 1}$. We denote by $\sigma(A) \subseteq \mathsf{H}$ the *spectrum*, i.e., the set of eigenvalues of A.

Proposition 3.6. *Let $A \in \mathsf{H}^{n\times n}$. Then:*

(a) *The spectrum of A is nonempty.*
(b) *If $\alpha \in \sigma(A)$, and $\beta \in \mathsf{H}$ is similar to α, then also $\beta \in \sigma(A)$.*

A topological proof of part (a) is given in [6], and a proof based on the map $\mho_{q,\mathsf{C}}$ goes back to [37].

Proof. Since $\mathsf{H}^{n\times n}$ is a finite-dimensional real algebra, there exists a monic scalar polynomial

$$p(t) = a_0 + a_1 t + \cdots + a_{m-1}t^{m-1} + t^m \tag{3.3}$$

with real coefficients $a_0, \dots, a_{m-1}$ such that

$$p(A) := a_0 I + a_1 A + \cdots + a_{m-1}A^{m-1} + A^m = 0. \tag{3.4}$$

Suppose that (3.3) is such polynomial of minimal degree. Consider two cases: (a) The polynomial (3.3) has a real root a. Then $A - aI$ is not invertible, for otherwise the polynomial $\widetilde{p}(t) := p(t)/(t-a)$ would be a polynomial with real coefficients of degree smaller than that of $p(t)$ such that $\widetilde{p}(A) = 0$, a contradiction with the choice of $p(t)$. By Proposition 3.5 the columns of $A - aI$ are linearly dependent, and we obtain that $Ax = xa$ for some nonzero $x \in \mathsf{H}^{n\times 1}$. (b) The polynomial (3.3) has a nonreal root $a + b\mathsf{i}$. Then arguing analogously to the case (a), we obtain

$$(A^2 - 2aA + (a^2 + b^2)I)x = 0$$

for some nonzero $x \in \mathsf{H}^{n\times 1}$. Then either

$$Ax = x(a + b\mathsf{i}),$$

or

$$A(-x(a+b\mathsf{i}) + Ax) = (-x(a+b\mathsf{i}) + Ax)(a - b\mathsf{i}), \quad -x(a+b\mathsf{i}) + Ax \neq 0.$$

Thus, in all cases there exists an eigenvalue of A, and (a) is proved.

Part (b) is obvious: If $Ax = x\alpha$ and $\beta = \gamma^{-1}\alpha\gamma$ for some $\gamma \in \mathsf{H} \setminus \{0\}$, then $A(x\gamma) = (x\gamma)\beta$. □

A monic real polynomial $p(t)$ of minimal degree such that $p(A) = 0$ is called the *minimal polynomial* of $A \in \mathsf{H}^{n\times n}$. It is easy to see that the minimal polynomial of A is unique, and it generates the ideal of all real polynomials $\widetilde{p}(t)$ such that $\widetilde{p}(A) = 0$. (Here A is fixed.)

Consider next the Jordan form. Introduce the Jordan blocks

$$J_m(\lambda) = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \vdots & \vdots & & \lambda & 1 \\ 0 & 0 & \cdots & 0 & \lambda \end{bmatrix} \in \mathsf{H}^{m\times m},$$

where $\lambda \in \mathsf{H}$. We will also use the real Jordan blocks

$$J_{2m}(a \pm \mathrm{i}b) = \begin{bmatrix} a & b & 1 & 0 & \cdots & 0 & 0 \\ -b & a & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & a & b & \cdots & 0 & 0 \\ 0 & 0 & -b & a & \cdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & & 0 & 1 \\ 0 & 0 & 0 & 0 & & a & b \\ 0 & 0 & 0 & 0 & & -b & a \end{bmatrix} \in \mathsf{R}^{2m\times 2m},$$

where $a \in \mathsf{R}$ and $b \in \mathsf{R} \setminus \{0\}$.

The Jordan form of a quaternionic matrix is given next:

Proposition 3.7. *Let $A \in \mathsf{H}^{n\times n}$. Then there exists an invertible $S \in \mathsf{H}^{n\times n}$ such that*

$$S^{-1}AS = J_{m_1}(\alpha_1) \oplus \cdots \oplus J_{m_r}(\alpha_r), \quad \alpha_1, \ldots, \alpha_r \in \mathsf{H}. \tag{3.5}$$

Moreover, the right-hand side of (3.5) *is uniquely determined by A, up to permutation of diagonal blocks, and up to replacement of each α_j with any similar quaternion β_j.*

Jordan form for quaternions and other material on linear algebra over the quaternions may be found in many sources, for example, [50], and earlier papers [49],[48]. Similarity reduction of matrices over more general skew fields was studied in [14] and [31].

The elements $\alpha_1, \ldots, \alpha_r$ in (3.5) are obviously eigenvalues of A. It is easy to see that every eigenvalue of $J_m(\alpha)$ must be similar to α. Thus:

$$\sigma(A) = \left\{\beta_1^{-1}\alpha_1\beta_1, \beta_2^{-1}\alpha_2\beta_2, \ldots, \beta_r^{-1}\alpha_r\beta_r \,:\, \beta_1, \ldots, \beta_r \in \mathsf{H} \setminus \{0\}\right\}.$$

Proof. For proof of existence of Jordan form (3.5), we refer the reader to transparent proofs given in [48], [6]. For the proof of uniqueness, first of all observe that if $\alpha = \gamma^{-1}\beta\gamma$ for some $\gamma \in \mathsf{H} \setminus \{0\}$, then $J_m(\alpha)$ and $J_m(\beta)$ are similar:

$$J_m(\alpha) \cdot \gamma^{-1}I = \gamma^{-1}I \cdot J_m(\beta).$$

Hence without loss of generality we may assume that $\alpha_j \in \mathsf{C}$, $j = 1, 2, \ldots, r$. We apply now the map $\mho_{q,\mathsf{C}}$ to both sides of (3.5), with $q = (\mathsf{i}, \mathsf{j}, \mathsf{k})$. The result is

$$\begin{aligned}&(\mho_{q,\mathsf{C}}(S))^{-1}\mho_{q,\mathsf{C}}(A)\mho_{q,\mathsf{C}}(S)\\ &= \begin{bmatrix} J_{m_1}(\alpha_1) & 0 \\ 0 & J_{m_1}(\overline{\alpha_1}) \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} J_{m_r}(\alpha_r) & 0 \\ 0 & J_{m_r}(\overline{\alpha_r}) \end{bmatrix}.\end{aligned}$$

Now use the familiar uniqueness property of the Jordan form of complex matrices to obtain the required uniqueness properties of the Jordan form (3.5). □

As a by-product of the proof, we have the first part of the next theorem. By the *complex Jordan form*, resp. *real Jordan form*, of a complex, resp. real, matrix X we mean the familiar Jordan form under similarity $X \mapsto S^{-1}XS$, where S is an invertible complex, resp. real, matrix.

Theorem 3.8.

(a) *If* $$J_{m_1}(\alpha_1) \oplus \cdots \oplus J_{m_r}(\alpha_r), \quad \alpha_1 = a_1 + \mathsf{i}b_1, \ldots, \alpha_r = a_r + \mathsf{i}b_r \in \mathsf{C} \tag{3.6}$$ *is a Jordan form of $A \in \mathsf{H}^{n\times n}$, then*
$$\begin{bmatrix} J_{m_1}(\alpha_1) & 0 \\ 0 & J_{m_1}(\overline{\alpha_1}) \end{bmatrix} \oplus \cdots \oplus \begin{bmatrix} J_{m_r}(\alpha_r) & 0 \\ 0 & J_{m_r}(\overline{\alpha_r}) \end{bmatrix}$$
is the complex Jordan form of $\mho_{q,\mathsf{C}}(A)$, for every units triple q.

(b) *Let* (3.6) *be a Jordan form of $A \in \mathsf{H}^{n\times n}$, where $\alpha_1, \ldots, \alpha_s$ are real and $\alpha_{s+1}, \ldots, \alpha_r$ are nonreal. Then for every units triple q the real Jordan form of $\mho_{q,\mathsf{R}}(A)$ is*
$$\begin{aligned}&(J_{m_1}(\alpha_1) \oplus J_{m_1}(\alpha_1) \oplus J_{m_1}(\alpha_1) \oplus J_{m_1}(\alpha_1)) \oplus \cdots\\ &\cdots \oplus (J_{m_s}(\alpha_s) \oplus J_{m_s}(\alpha_s) \oplus J_{m_s}(\alpha_s) \oplus J_{m_s}(\alpha_s)) \oplus\\ &\begin{bmatrix} J_{2m_{s+1}}(a_{s+1} \pm \mathsf{i}b_{s+1}) & 0 \\ 0 & J_{2m_{s+1}}(a_{s+1} \pm \mathsf{i}b_{s+1}) \end{bmatrix} \oplus \cdots\\ &\cdots \oplus \begin{bmatrix} J_{2m_r}(a_r \pm \mathsf{i}b_r) & 0 \\ 0 & J_{2m_r}(a_r \pm \mathsf{i}b_r) \end{bmatrix}.\end{aligned}$$

The second part of the theorem is obtained easily for a particular units triple $(\mathsf{i}, \mathsf{j}, \mathsf{k})$, and for a general units triple (q_1, q_2, q_3) use the fact the map $1 \mapsto 1$, $q_1 \mapsto \mathsf{i}$, $q_2 \mapsto \mathsf{j}$, $q_2 \mapsto \mathsf{k}$ generates an automorphism of H which must be inner by Proposition 2.5. It follows that for any two units triples q and q', the matrices $\mho_{q,\mathsf{C}}(A)$ and $\mho_{q',\mathsf{C}}(A)$ are similar, and the matrices $\mho_{q,\mathsf{R}}(A)$ and $\mho_{q',\mathsf{R}}(A)$ are similar.

The Jordan form allows us to identify the minimal polynomial of a quaternionic matrix:

Proposition 3.9. *Let $A \in \mathsf{H}^{n\times n}$. Assume that $\beta_1, \ldots, \beta_u$ are the distinct real eigenvalues of A (if any), and $\beta_{u+1}, \ldots, \beta_{u+v} \in \mathsf{H}$ is a maximal set of nonreal pairwise nonsimilar eigenvalues of A (if any). Furthermore, in a Jordan form of A,* *let m_j be the largest size of Jordan blocks corresponding to the eigenvalue β_j, $j = 1, 2, \ldots, u + v$. Then the minimal polynomial of A is given by*

$$(t - \beta_1)^{m_1} \cdots (t - \beta_u)^{m_u} \prod_{j=1}^{v} (t^2 - 2\Re(\beta_{u+j})t + |\beta_{u+j}|^2)^{m_{u+j}}.$$

The proof is immediate from the Jordan form of A.

Corollary 3.10. *The matrices $A \in \mathsf{H}^{n\times n}$, A^*, and A_ϕ for every nonstandard iaa, all have the same spectrum.*

Proof. If

$$S^{-1}AS = J_{m_1}(\alpha_1) \oplus \cdots \oplus J_{m_r}(\alpha_r)$$

is a Jordan form of A, then

$$S^* A^* (S^*)^{-1} = J_{m_1}(\alpha_1)^* \oplus \cdots \oplus J_{m_r}(\alpha_r)^*.$$

Observe that $J_{m_j}(\alpha_j)^*$ is similar to $J_{m_j}(\alpha_j)$ over H to obtain the statement concerning A^*. For A_ϕ, combine Proposition 3.4 and Theorem 3.8(a), and use the fact that a complex matrix and its transposed have the same spectrum. □

Proposition 3.11. *Let $A \in \mathsf{H}^{m\times m}$, $B \in \mathsf{H}^{n\times n}$. Then the equation*

$$AX = XB, \qquad X \in \mathsf{H}^{m\times n}$$

has only the trivial solution $X = 0$ if and only if

$$\sigma(A) \cap \sigma(B) = \emptyset. \tag{3.7}$$

Proof. If (3.7) holds, then by Theorem 3.8 we have

$$\sigma(\mho_{q,\mathsf{R}}(A)) \cap \sigma(\mho_{q,\mathsf{R}}(B)) = \emptyset.$$

Thus, by the familiar result for the real field (see [20] or [36], for example) the equation

$$\mho_{q,\mathsf{R}}(A)\mho_{q,\mathsf{R}}(X) = \mho_{q,\mathsf{R}}(X)\mho_{q,\mathsf{R}}(B)$$

has only the trivial solution. Conversely, if $\alpha \in \sigma(A) \cap \sigma(B)$, then by Corollary 3.10 there exist nonzero $x, y \in \mathsf{H}^{n\times n}$ such that

$$Ax = x\alpha, \quad B^* y = y\alpha^*.$$

Then $y^*B = \alpha y^*$, and $X := xy^* \neq 0$ satisfies the equality $AX = XB$. □

4. Canonical forms of symmetric matrices

Before going on to pairs of symmetric/skewsymmetric matrices, we review in this section the canonical forms for just one symmetric/skewsymmetric matrix. Such canonical form is well known for Hermitian matrices, but it is not as well known for matrices that are symmetric with respect to a nonstandard iaa.

Let ϕ be an iaa of H. A matrix $A \in \mathsf{H}^{n\times n}$ is said to be *symmetric with respect to ϕ*, or *ϕ-symmetric*, if $A_\phi = A$. If A is a standard iaa, then ϕ-symmetric matrix means Hermitian. Clearly, if A is ϕ-symmetric, then so is $S_\phi AS$ for any invertible matrix $S \in \mathsf{H}^{n\times n}$. Canonical forms with respect to this action are given next.

Theorem 4.1.

(a) *Assume that the iaa ϕ is standard. Then for every ϕ-symmetric $A \in \mathsf{H}^{n\times n}$ there exists an invertible S such that*

$$S_\phi AS = \begin{bmatrix} I_r & 0 & 0 \\ 0 & -I_s & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

for some nonnegative integers r and s such that $r+s \le n$. Moreover, r and s are uniquely determined by A.

(b) *Assume that the iaa ϕ is nonstandard. Then for every ϕ-symmetric $A \in \mathsf{H}^{n\times n}$ there exists an invertible S such that*

$$S_\phi AS = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}, \tag{4.1}$$

for a nonnegative integer $r \le n$. Moreover, r is uniquely determined by A.

Proof. Part (a) is well known (see, e.g., [11, Section 12.5],[50], or [47]), and part (b) is contained in [44, Theorem 1].

For completeness we provide a proof of (b). First of all, the uniqueness of r is obvious: r coincides with the rank of A. To show the existence of invertible S such that (4.1) holds, we need a lemma.

Lemma 4.2. *Let ϕ be a nonstandard iaa. The for every $x \in \mathsf{H}$ such that $x = \phi(x)$ there exists $y \in \mathsf{H}$ such that $x = \phi(y)y$.*

Proof of the lemma. It follows from Proposition 2.3 that the R-linear map ϕ is involutory automorphically similar to ϕ_0, where ϕ_0 is the nonstandard iaa defined by $\phi(\mathsf{i}) = \mathsf{i}$, $\phi(\mathsf{j}) = \mathsf{j}$, $\phi(\mathsf{k}) = -\mathsf{k}$. In other words, we have $\phi = \psi^{-1}\phi_0\psi$, where ψ is an involutory automorphism of H. Therefore, it suffices to prove the lemma for ϕ_0. Writing $x = \phi_0(x)$ in the form

$$x = u + v\mathsf{i} + w\mathsf{j}, \quad u, v, w \in \mathsf{R},$$

we have to solve the equation

$$u + v\mathsf{i} + w\mathsf{j} = (a + b\mathsf{i} + c\mathsf{j} + d\mathsf{k})(a + b\mathsf{i} + c\mathsf{j} - d\mathsf{k}) \tag{4.2}$$

for $a, b, c, d \in \mathsf{R}$. The right-hand side of this equation is equal to

$$a^2 - b^2 - c^2 + d^2 + (2ab - 2dc)\mathsf{i} + (2ac + 2bd)\mathsf{j}.$$

Thus, equation (4.2) has a solution if and only if for every vector $(u, v, w) \in \mathsf{R}^{1\times 3}$ there exist $a, b, c, d \in \mathsf{R}$ such that

$$u = a^2 - b^2 - c^2 + d^2, \quad v = 2ab - 2dc, \quad w = 2ac + 2bd. \tag{4.3}$$

Assume that $v^2 + w^2 > 0$ (if $v = w = 0$, let $a = c = 0$ and one of b and d equal to zero to satisfy (4.3)). Solve the second and third equations in (4.3) for b and c:

$$b = \frac{av + dw}{2a^2 + 2d^2}, \quad c = \frac{aw - vd}{2a^2 + 2d^2}.$$

Substitute these expression for b and c in the first equation in (4.3). After simplification, the first equation in (4.3) is reduced to

$$-4q^2 + 4qu + v^2 + w^2 = 0, \quad \text{where} \quad q := a^2 + d^2. \tag{4.4}$$

Obviously, the equation in (4.4) has a positive solution q. Thus, the system (4.3) has a solution a, b, c, d for every triple of real numbers u, v, w. □

We now return to the proof of part (b), and proceed by induction on n. The case $n = 1$ is taken care of by Lemma 4.2. Assume part (b) is already proved for $(n-1) \times (n-1)$ matrices. Given a ϕ-symmetric matrix $A = [a_{i,j}]_{i,j=1}^n \in \mathsf{H}^{n\times n}$, we first suppose that A has a nonzero diagonal entry. By row and column permutation, we suppose that $a_{1,1} \neq 0$. Obviously, $\phi(a_{1,1}) = a_{1,1}$, and by Lemma 4.2 there exists $y \in \mathsf{H}$ such that the matrix

$$B = [b_{i,j}]_{i,j=1}^n := \operatorname{diag}(\phi(y), 1, 1, \ldots, 1) \quad A \quad \operatorname{diag}(y, 1, 1, \ldots, 1)$$

has $b_{1,1} = 1$. Transforming $B = B_\phi$ further, let

$$C = S_\phi B S,$$

where

$$S = \begin{bmatrix} 1 & -b_{1,2} & -b_{1,3} & \cdots & -b_{1,n} \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

We see that C is a direct sum: $C = 1 \oplus C'$. Now the induction hypothesis completes the proof in the case A has a nonzero diagonal entry.

Suppose all diagonal entries of A are zero. Then we apply a transformation $A \mapsto S_\phi A S$, where $S \in \mathsf{H}^{n\times n}$ is upper triangular with 1's on the diagonal and a sole nonzero off diagonal entry, to make a nonzero diagonal entry in $S_\phi A S$, thereby reducing the proof to the case already considered (we exclude the trivial situation when A is the zero matrix). Indeed, assuming $n = 2$ for simplicity of notation, we have

$$\begin{aligned} S_\phi A S &= \begin{bmatrix} 1 & 0 \\ \phi(q) & 1 \end{bmatrix} \begin{bmatrix} 0 & a_{1,2} \\ \phi(a_{1,2}) & 0 \end{bmatrix} \begin{bmatrix} 1 & q \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & a_{1,2} \\ \phi(a_{1,2}) & \phi(a_{1,2})q + \phi(q)a_{1,2} \end{bmatrix}. \end{aligned}$$

It remains to choose $q \in \mathsf{H}$ so that

$$\phi(a_{1,2})q + \phi(q)a_{1,2} \neq 0. \tag{4.5}$$

Since $a_{1,2} \neq 0$ (because we have excluded the trivial case in which A has no nonzero entries), such q is easily found; for example, if $\phi(a_{1,2}) \neq a_{1,2}$, take $q = -1$. □

One application of Theorem 4.1(b) has to do with maximal ϕ-neutral subspaces. Let ϕ be a nonstandard, and let $A \in \mathsf{H}^{n\times n}$ be a ϕ-symmetric matrix. A (quaternionic) subspace $\mathcal{M} \subseteq \mathsf{H}^{n\times 1}$ is said to be A-*neutral* if $x_\phi Ax = 0$ for every $x \in \mathcal{M}$.

Theorem 4.3. *Let $A \in \mathsf{H}^{n\times n}$ be ϕ-symmetric, and let* $\operatorname{rank}(A)$ *be the rank of A, i.e., the dimension of the column space of A as a right quaternionic vector space. Then the maximal dimension of an A-neutral subspace is equal to $n - \frac{1}{2}(\operatorname{rank}(A))$ if $\operatorname{rank}(A)$ is even, and to $n - \frac{1}{2}(\operatorname{rank}(A)+1)$ if $\operatorname{rank}(A)$ is odd.*

Proof. By Theorem 4.1(b) we may assume that $A = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$, where $r = \operatorname{rank}(A)$. Now the proof is easily reduced to the case when $\operatorname{rank}(A) = n$, i.e., $A = I_n$. The existence of an A-neutral subspace of dimension $\lfloor n/2 \rfloor$ follows by consideration of Span $\begin{bmatrix} 1 \\ \alpha \end{bmatrix}$ as an I_2-neutral subspace, where $\alpha \in \mathsf{H}$ is such that $\alpha^2 = -1$ and $\phi(\alpha) = \alpha$. Let $x_1, \dots, x_p$ be a basis in an I_n-neutral subspace $\mathcal{M}$, and let $y_j = \overline{[\phi(x_j)]}$, $j = 1, 2, \dots, p$, where the vector $[\phi(x_j)] \in \mathsf{H}^{n\times 1}$ is obtained from x_j by componentwise application of ϕ. Then $y_1, \dots, y_p$ are linearly independent and orthogonal (in the sense of the Euclidean inner product) to $\mathcal{M}$. Therefore we must have $2p \le n$. $\square$

The result of Theorem 4.1 can be extended to matrices that are symmetric simultaneously with respect to several iaa's.

Let Φ be a set of iaa's. A matrix $A \in \mathsf{H}^{n\times n}$ is said to be Φ-*symmetric* if $A_\phi = A$ for all $\phi \in \Phi$. Let

$$\operatorname{Inv}(\Phi) := \bigcap\nolimits_{\phi\in\Phi} \operatorname{Inv}(\phi) = \{\alpha \in \mathsf{H} \,:\, \phi(\alpha) = \alpha \quad \forall \;\; \phi \in \Phi\}.$$

Clearly, $\operatorname{Inv}(\Phi) \supseteq \mathsf{R}$.

Theorem 4.4.

(a) *Assume that* $\operatorname{Inv}(\Phi) = \mathsf{R}$. *Then for every Φ-symmetric $A \in \mathsf{H}^{n\times n}$ there exists an invertible S such that*

$$S_\phi = S \quad \forall \;\; \phi \in \Phi \tag{4.6}$$

$$\textit{and} \qquad S_\phi A S = \begin{bmatrix} I_r & 0 & 0 \\ 0 & -I_s & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

for some nonnegative integers r and s such that $r + s \le n$. Moreover, r and s are uniquely determined by A.

(b) *Assume that* $\operatorname{Inv}(\Phi) \ne \mathsf{R}$. *Then for every Φ-symmetric $A \in \mathsf{H}^{n\times n}$ there exists an invertible S such that* (4.6) *holds and*

$$S_\phi A S = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}, \tag{4.7}$$

for a nonnegative integer $r \le n$. Moreover, r is uniquely determined by A.

Proof. If Φ is a one-element set, this result is Theorem 4.1. Otherwise, by Proposition 2.3, in the case (b), $\mathrm{Inv}\,(\Phi) = \mathrm{Span}_{\mathsf{R}}\,\{1, q\}$ for some $q \in \mathsf{H}$ such that $q^2 = -1$. Without loss of generality we assume that $q = \mathsf{i}$; thus, $\mathrm{Inv}\,(\Phi) = \mathsf{C}$. Moreover, using Lemma 2.4 for $\phi_1, \phi_2 \in \Phi$, $\phi_1 \neq \phi_2$, we have

$$X \in \mathsf{H}^{n\times n}, \quad X_{\phi_1} = X_{\phi_2} \quad \Longleftrightarrow \quad X \in \mathsf{C}^{n\times n}.$$

Now the result in the case (b) follows from the fact that any matrix $A \in \mathsf{C}^{n\times n}$ can be reduced to a form $\begin{bmatrix} I_{\mathrm{rank}\,(A)} & 0 \\ 0 & 0 \end{bmatrix}$ using the complex congruence transformations $A \longrightarrow S^T AS$, $A, S \in \mathsf{C}^{n\times n}$, S invertible (this can be proved using the well-known Lagrange's method of reduction of bilinear forms, see [20], for example).

In the case (a), we argue similarly. If Φ is a one-element set (in which case it must be the standard iaa), we are done by Theorem 4.1. Next, consider the case when Φ consists of the standard iaa and a non-empty set Φ_0 of nonstandard iaa's. If $\mathrm{Inv}\,(\Phi_0) = \mathsf{R}$, the proof essentially reduces to the congruence of real matrices, and if $\mathrm{Inv}\,(\Phi_0) = \mathrm{Span}_{\mathsf{R}}\,\{1, q\}$ for some $q \in \mathsf{H}$ such that $q^2 = -1$, then the proof essentially reduces to the $*$-congruence of complex matrices $A \Longrightarrow S^*AS$, $A, S \in \mathsf{C}^{n\times n}$, S invertible; cf. the preceding paragraph. It remains to consider the case when the standard iaa does not belong to Φ. Using Lemma 2.4, we again reduce the proof to the case of real symmetric matrix A and invertible real matrix S. $\square$

5. Matrix polynomials with quaternionic coefficients

Let $\mathsf{H}(t)$ be the noncommutative ring of polynomials with quaternionic coefficients, with the real independent variable t. Note that t commutes with the quaternions. Therefore, for every fixed $t_0 \in \mathsf{R}$, the evaluation map

$$f \;\mapsto\; f(t_0), \quad f(t) \in \mathsf{H}(t),$$

is well defined as a unital homomorphism of real algebras $\mathsf{H}(t) \longrightarrow \mathsf{H}$.

Let $p(t), q(t) \in \mathsf{H}(t)$. A polynomial $q(t)$ is called a *divisor* of $p(t)$ if $p(t) = q(t)s(t)$ and $p(t) = r(t)q(t)$ for some $s(t), r(t) \in \mathsf{H}(t)$. A polynomial q is said to be a *total divisor* of p if $\alpha q(t)\alpha^{-1}$ is a divisor of $p(t)$ for every $\alpha \in \mathsf{H} \setminus \{0\}$, or equivalently, if $q(t)$ is a divisor of $\beta p(t)\beta^{-1}$ for all $\beta \in \mathsf{H} \setminus \{0\}$.

Let $\mathsf{H}(t)^{m\times n}$ be the set of all $m \times n$ matrices with entries in $\mathsf{H}(t)$, which will be called *matrix polynomials* with the standard operations of addition, right and left multiplication by quaternions, and matrix multiplication: If $A(t) \in \mathsf{H}(t)^{m\times n}$ and $B(t) \in \mathsf{H}(t)^{n\times p}$, then $A(t)B(t) \in \mathsf{H}(t)^{m\times p}$.

In this section we recall the Smith form for matrix polynomials with quaternionic entries. The form is found in [31, Chapter 3], in a more general context of non-commutative principal ideal domains, and in [39] (in a different set up of Laurent polynomials). For matrix polynomials over a (commutative) field, and more generally over commutative principal ideal domains, the Smith form is widely known, see, for example, the books [20], [23], [22], [36]. Our exposition follows the general outline for the standard proof of the Smith form for matrix polynomials over fields.

A matrix polynomial $A(t) \in \mathsf{H}(t)^{n\times n}$ is said to be *elementary* if it can be represented as a product (in any order) of diagonal $n\times n$ polynomials with constant nonzero quaternions on the diagonal, and of $n \times n$ polynomials with 1's on the diagonal and a sole nonzero off diagonal entry. Constant permutation matrices are elementary, as it follows from the equality (given in [31])

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

A matrix polynomial $A(t) \in \mathsf{H}(t)^{m\times n}$ is said to be *unimodular* if

$$A(t)B(t) = B(t)A(t) \equiv I \tag{5.1}$$

for some matrix polynomial $B(t) \in \mathsf{H}(t)^{n\times m}$. (In this case, clearly $m = n$.) Obviously, every elementary matrix polynomial is unimodular. We shall see later that the converse also holds.

Theorem 5.1. *Let $A(t) \in \mathsf{H}(t)^{m\times n}$. Then there exist elementary matrix polynomials $D(t) \in \mathsf{H}(x)^{m\times m}$, $E(t) \in \mathsf{H}(t)^{n\times n}$, and monic (i.e., with leading coefficient equal to 1) scalar polynomials $a_1(t), a_2(t), \ldots, a_r(t) \in \mathsf{H}(t)$, $0 \le r \le \min\{m, n\}$, such that*

$$D(t)A(t)E(t) = \operatorname{diag}(a_1(t), a_2(t), \ldots, a_r(t), 0, \ldots, 0), \tag{5.2}$$

where $a_j(t)$ is a total divisor of $a_{j+1}(t)$, for $j = 1, 2, \ldots, r-1$.

The right-hand side of (5.2) will be called the *Smith form* of $A(t)$. Theorem 5.1 in the more general context of matrix polynomials with coefficients in a division ring was proved in [38]; see also [31].

Proof. **Step 1**. We prove the existence of the diagonal form (5.2) with monic polynomials $a_1(t), \ldots, a_r(t)$, but not necessarily with the property that $a_j(t)$ is a total divisor of $a_{j+1}(t)$, for $j = 1, 2, \ldots, r-1$.

If $A(t) \equiv 0$ we are done. Otherwise, let α be the minimal degree of nonzero entries of all matrix polynomials of the form $D_1(t)A(t)E_1(t)$, where $D_1(t)$ and $E_1(t)$ are elementary matrix polynomials. Thus,

$$D_2(t)A(t)E_2(t) = B(t)$$

for some elementary matrix polynomials $D_2(t)$ and $E_2(t)$, where

$$B(t) = [b_{p,q}(t)]_{p=1,q=1}^{m,n}$$

is such that the degree of $b_{1,1}(t)$ is equal to α. Consider $b_{1,q}(t)$ with $q = 2, \ldots, n$. Division with remainder

$$b_{1,q}(t) = b_{1,1}(t)c_{1,q}(t) + d_{1,q}(t), \quad c_{1,q}(t), d_{1,q}(t) \in \mathsf{H}(t),$$

where the degree of $d_{1,q}(t)$ is smaller than α, shows that we must have $d_{1,q}(t) \equiv 0$, for otherwise the product $B(t)F(t)$, where the elementary matrix polynomial $F(t)$ has 1 on the main diagonal, $-c_{1,q}(t)$ in the $(1, q)$th position, and zero everywhere else, will have a nonzero entry of degree less than α, a contradiction with the definition of α. Thus,

$$b_{1,q}(t) = b_{1,1}(t)c_{1,q}(t), \quad q = 2, 3, \ldots, n,$$

for some $c_{1,q}(t) \in \mathsf{H}(t)$. Now it is easy to see that there exists an elementary matrix polynomial $E_3(t)$ of the form $E_3(t) = I + \widehat{E}(t)$, where $\widehat{E}(t)$ may have nonzero entries only in the positions $(1,2), \ldots, (1,n)$, such that $B(t)E_3(t)$ has zero entries in the positions $(1,2), \ldots, (1,n)$. Similarly, we prove that there exists an elementary matrix polynomial $D_3(t)$ such that $D_3(t)B(t)E_3(t)$ has zero entries in all positions $(1,q)$, $q \neq 1$ and $(p,1)$, $p \neq 1$. Now induction on $m+n$ completes the proof of existence of (5.2). The base of induction, namely the cases when $m = 1$ or $n = 1$, can be dealt with analogously.

Step 2. Now clearly r is an invariant of the diagonal form (5.2) with monic polynomials $a_j(t)$. Indeed, r coincides with the maximal rank of the matrix $A(t)$, over all real t. In verifying this fact, use the property that each $a_j(t)$, being a monic polynomial of real variable with quaternionic coefficients, may have only a finite number of real roots.

Step 3. Consider a diagonal form (5.2) with monic $a_1(t), \ldots, a_r(t)$ that have the following additional property: The degrees $\delta_1, \ldots, \delta_r$ of $a_1(t), \ldots, a_r(t)$, respectively, are such that for any other diagonal form

$$\widetilde{D}(t)A(t)\widetilde{E}(t) = \operatorname{diag}(b_1(t), b_2(t), \ldots, b_r(t), 0, \ldots, 0),$$

where $\widetilde{D}(t)$ and $\widetilde{E}(t)$ are elementary matrix polynomials, and $b_1(t), b_2(t), \ldots, b_r(t)$ are monic scalar polynomials, the degrees $\delta'_1, \ldots, \delta'_r$ of $b_1(t), \ldots, b_r(t)$, respectively, satisfy the inequalities:

$$\delta_1 \leq \delta'_1; \quad \text{if } \delta_1 = \delta'_1, \text{ then } \delta_2 \leq \delta'_2; \quad \text{if } \delta_j = \delta'_j \text{ for } j = 1, 2, \text{ then } \delta_3 \leq \delta'_3; \quad \ldots;$$

$$\text{if } \delta_j = \delta'_j \text{ for } j = 1, 2, \ldots, r-1, \text{ then } \delta_r \leq \delta'_r.$$

Obviously, a diagonal form $\operatorname{diag}(a_1(t), a_2(t), \ldots, a_r(t), 0, \ldots, 0)$ of $A(t)$ with these properties does exist. We claim that then a_j is a total divisor of a_{j+1}, for $j = 1, 2, \ldots, r-1$. Suppose not, and assume that a_j is not a total divisor of a_{j+1}, for some j. Then there exists $\alpha \in \mathsf{H} \setminus \{0\}$ such that

$$\alpha a_{j+1}(t)\alpha^{-1} = d(t)a_j(t) + s(t),$$

where $d(t), s(t) \in \mathsf{H}(t)$, $s(t) \neq 0$, is such that the degree of s is smaller than δ_j, the degree of a_j. (If it happens that

$$\alpha a_{j+1}(t)\alpha^{-1} = a_j(t)d(t) + s(t),$$

the subsequent argument is completely analogous.) We have

$$\begin{bmatrix} 1 & 0 \\ d & \alpha \end{bmatrix} \begin{bmatrix} a_j & 0 \\ 0 & a_{j+1} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\alpha^{-1} & \alpha^{-1} \end{bmatrix} = \begin{bmatrix} a_j & 0 \\ -s & \alpha a_{j+1}\alpha^{-1} \end{bmatrix}.$$

Since the degree of s is smaller than δ_j, Step 1 of the proof shows that for some elementary 2×2 matrix polynomials $D'(t)$ and $E'(t)$, we have

$$D'(t) \begin{bmatrix} a_j(t) & 0 \\ 0 & a_{j+1}(t) \end{bmatrix} E'(t) = \begin{bmatrix} a'_j(t) & 0 \\ 0 & a'_{j+1}(t) \end{bmatrix},$$

where $a'_j(t)$ and $a'_{j+1}(t)$ are monic polynomials and the degree of $a'_j(t)$ is smaller than δ_j. Using $\begin{bmatrix} a'_j(t) & 0 \\ 0 & a'_{j+1}(t) \end{bmatrix}$ in place of $\begin{bmatrix} a_j(t) & 0 \\ 0 & a_{j+1}(t) \end{bmatrix}$ in

$$\operatorname{diag}(a_1(t), a_2(t), \ldots, a_r(t), 0, \ldots, 0),$$

we obtain a contradiction with the choice of $a_1(t), \ldots, a_r(t)$. □

Obviously, if $\operatorname{diag}(a_1(t), a_2(t), \ldots, a_r(t), 0, \ldots, 0)$ is a Smith form of $A(t)$, then so is $\operatorname{diag}(\alpha_1^{-1} a_1(t)\alpha_1, \alpha_2^{-1} a_2(t)\alpha_2, \ldots, \alpha_r^{-1} a_r(t)\alpha_r, 0, \ldots, 0)$, for any choice of nonzero quaternions $\alpha_1, \ldots, \alpha_r$. However, the converse is generally false: A Smith form of $A(t)$ need not be unique up to a replacement of each $a_j(t)$ with $\alpha_j^{-1} a_j(t)\alpha_j$, $\alpha_j \in \mathsf{H} \setminus \{0\}$. See [26], [25] for a characterization of the nonuniqueness of the Smith form for matrix polynomials with coefficients in a division ring.

Corollary 5.2. *A matrix polynomial $A(t)$ is elementary if and only if $A(t)$ is unimodular.*

Proof. The "only if" part was already observed. Assume that $A(t)$ is such that (5.1) holds for some matrix polynomial $B(t)$. Clearly, $A(t) \in \mathsf{H}(t)^{n \times n}$ for some n. Without loss of generality we may assume that $A(t)$ is in the form (5.2). Now equation (5.1) implies that $a_j(t)b_j(t) \equiv 1$ for some $b_j(t) \in \mathsf{H}(t)$, hence $a_j(t)$ is a nonzero constant and we are done. □

The *rank* $r(A(t))$ of a matrix polynomial $A(t) \in \mathsf{H}(t)^{m \times n}$ is defined as the maximal rank of quaternionic matrices $A(t_0)$, over all $t_0 \in \mathsf{R}$. It is easy to see that the rank of $A(t)$ coincides with integer r of Theorem 5.1. Also, if the rank of $A(t_0)$ is smaller than a certain integer q for infinitely many real values of t_0, then $r(A(t)) < q$. This follows from the fact that a non identically zero quaternionic polynomial $s(t) \in \mathsf{H}(t)$ can have only finitely many real zeros (this can be easily seen by first assuming without loss of generality that $s(t)$ is monic, and then by considering the real part of the polynomial equation $s(t) = 0$). Note that in contrast a non identically zero quaternionic polynomial may have a continuum of non real quaternionic zeros, for example, $t^2 + 1$.

Corollary 5.3. *Let $A(t) \in \mathsf{H}(t)^{m \times n}$.*

(a) *Assume that $r(A(t)) < n$. (This condition is automatically satisfied if $m < n$.) Then there exists $Y(t) \in \mathsf{H}^{n \times (n - r(A(t)))}$ such that*

$$A(t)Y(t) \equiv 0$$

and $r(Y(t)) = n - r(A(t))$.

(b) *Assume that* $r(A(t)) < m$. (*This condition is automatically satisfied if* $m > n$.) *Then there exists* $Y(t) \in \mathsf{H}^{(m-r(A(t)))\times m}$ *such that*

$$Y(t)A(t) \equiv 0$$

and $r(Y(t)) = m - r(A(t))$.

Proof. Part (a). Without loss of generality we may assume that $A(t)$ is in the form of the right-hand side of (5.2). Now the hypothesis of the corollary implies that the right-most column of $A(t)$ is zero, and we may take $y(t) = (0, 0, \ldots, 0, 1)^T$. The proof of (b) is analogous. □

6. The Kronecker form

In this section we present the Kronecker form of a pair of quaternionic matrices, with a complete proof modeled after the standard proof for complex (or real) matrices, see for example [21] and [46], also the book [23]. Although an alternative proof can be given using the map $\mho_{q,\mathsf{C}}$ (see [48], where this approach was used to derive the Jordan form for quaternionic matrices), we prefer the direct proof, as it can be applicable to matrices over more general division rings.

6.1. Statement of the Kronecker form

Consider two matrix polynomials of degree at most one, often called *matrix pencils*: $A_1 + tB_1$ and $A_2 + tB_2$, where $A_1, B_1, A_2, B_2 \in \mathsf{H}^{m\times n}$. The polynomials $A_j + tB_j$, $j = 1, 2$, are called *strictly equivalent* if

$$A_1 = PA_2Q, \quad B_1 = PB_2Q$$

for some invertible matrices $P \in \mathsf{H}^{m\times m}$ and $Q \in \mathsf{H}^{n\times n}$. We develop here canonical form of matrix pencils under strict equivalence.

The Kronecker form of matrix polynomials $A + tB$ is formulated in Theorem 6.1 below. In a less explicit form it is found in [43] and [16].

Introduce special matrix polynomials:

$$L_\varepsilon(t) = [0_{\varepsilon\times 1} \;\; I_\varepsilon] + t[I_\varepsilon \;\; 0_{\varepsilon\times 1}] \in \mathsf{H}^{\varepsilon\times(\varepsilon+1)}.$$

Here ε is a positive integer. Also, the following standard real symmetric matrices will be used:

$$F_m = \begin{bmatrix} 0 & \cdots & \cdots & 0 & 1 \\ \vdots & & & 1 & 0 \\ \vdots & & & & \vdots \\ 0 & 1 & & & \vdots \\ 1 & 0 & \cdots & \cdots & 0 \end{bmatrix} = F_m^{-1}, \tag{6.1}$$

$$G_m = \begin{bmatrix} 0 & \cdots & \cdots & 1 & 0 \\ \vdots & & & 0 & 0 \\ \vdots & & & & \vdots \\ 1 & 0 & & & \vdots \\ 0 & 0 & \cdots & \cdots & 0 \end{bmatrix} = \begin{bmatrix} F_{m-1} & 0 \\ 0 & 0 \end{bmatrix}. \tag{6.2}$$

Note the equalities

$$\widetilde{G}_m := \begin{bmatrix} 0 & 0 \\ 0 & F_{m-1} \end{bmatrix} = F_m G_m F_m. \tag{6.3}$$

Theorem 6.1. *Every pencil $A + tB \in \mathsf{H}(t)^{m\times n}$ is strictly equivalent to a matrix pencil with the block diagonal form:*

$$\begin{aligned} 0_{u\times v} \oplus L_{\varepsilon_1\times(\varepsilon_1+1)} \oplus \cdots \oplus L_{\varepsilon_p\times(\varepsilon_p+1)} \oplus L^T_{\eta_1\times(\eta_1+1)} \oplus L^T_{\eta_q\times(\eta_q+1)} \\ \oplus (I_{k_1} + tJ_{k_1}(0)) \oplus \cdots \oplus (I_{k_r} + tJ_{k_r}(0)) \\ \oplus (tI_{\ell_1} + J_{\ell_1}(\alpha_1)) \oplus \cdots \oplus (tI_{\ell_s} + J_{\ell_s}(\alpha_s)), \end{aligned} \tag{6.4}$$

where $\varepsilon_1 \le \cdots \le \varepsilon_p$; $\eta_1 \le \cdots \le \eta_q$; $k_1 \le \cdots \le k_r$; $\ell_1 \le \cdots \le \ell_s$ are positive integers, and $\alpha_1, \ldots, \alpha_s \in \mathsf{H}$.

Moreover, the integers u, v, and ε_i, η_j, k_w are uniquely determined by the pair A, B, and the part

$$(tI_{\ell_1} + J_{\ell_1}(\alpha_1)) \oplus \cdots \oplus (tI_{\ell_s} + J_{\ell_s}(\alpha_s))$$

is uniquely determined by A and B up to a permutation of the diagonal blocks and up to replacing each α_j with any quaternion similar to α_j.

The following terminology is used in connection with the Kronecker form (6.4) of the matrix pencil $A + tB$. The integers $\varepsilon_1 \le \cdots \le \varepsilon_p$ and $\eta_1 \le \cdots \le \eta_q$ are called the *left indices* and the *right indices*, respectively, of $A + tB$. The integers $k_1 \le \cdots \le k_r$ are called the *indices at infinity* of $A + tB$. The quaternions $\alpha_1, \ldots, \alpha_s$ are called the *eigenvalues* of $A + tB$; they are uniquely determined up to permutation and similarity. Note that if B is invertible, then the eigenvalues of $A + tB$ are exactly the eigenvalues of $B^{-1}A$ (or the eigenvalues of AB^{-1}). For a fixed eigenvalue α of $A + tB$, let $i_1 < \cdots < i_w$ be all the subscripts in (6.4) such that $\alpha_{i_1}, \ldots, \alpha_{i_w}$ are similar to α; then the integers $\ell_{i_1}, \ldots, \ell_{i_w}$ are called the *indices* of the eigenvalue α of $A + tB$. Note that there may be several indices at infinity that are equal to a fixed positive integer; the same remark applies to the indices of a fixed eigenvalue of $A + tB$, to the left indices of $A + tB$, and to the right indices of $A + tB$.

6.2. Proof of Theorem 6.1.

We start with preliminary results.

Theorem 6.2. *Let* $A + tB \in \mathsf{H}(t)^{m\times n}$ *be a matrix pencil such that the rank of* $A + tB$ *is smaller than* n. *Assume that*

$$Ax = Bx = 0, \quad x \in \mathsf{H}^{n\times 1} \quad \Longrightarrow \quad x = 0. \tag{6.5}$$

Then $A + tB$ *is strictly equivalent to a direct sum*

$$L_\varepsilon(t) \oplus (A' + tB') \tag{6.6}$$

for some $\varepsilon > 0$.

Proof. With some changes, the proof follows the proof of the reduction theorem in [21, Chapter XII], [20, Chapter 2].

By Corollary 5.3, there exists a nonzero polynomial $y(t) \in \mathsf{H}(t)^{n\times 1}$ such that

$$(A + tB)y(t) \equiv 0. \tag{6.7}$$

Let $\varepsilon - 1 \geq 1$ be the smallest degree of a nonzero $y(t)$ for which (6.7) holds. (We cannot have $\varepsilon - 1 = 0$ in view of hypothesis (6.5).) Write

$$y(t) = \sum_{j=0}^{\varepsilon-1} t^j(-1)^j y_j, \quad y_0, \ldots, y_{\varepsilon-1} \in \mathsf{H}^{n\times 1}, \quad y_{\varepsilon-1} \neq 0.$$

Then equation (6.7) reads

$$\begin{bmatrix} A & 0_{m\times n} & \cdots & 0_{m\times n} \\ B & A & \cdots & 0_{m\times n} \\ 0_{m\times n} & B & \cdots & \vdots \\ \vdots & \vdots & \ddots & A \\ 0_{m\times n} & 0_{m\times n} & \cdots & B \end{bmatrix} \begin{bmatrix} y_0 \\ -y_1 \\ \vdots \\ (-1)^{\varepsilon-1} y_{\varepsilon-1} \end{bmatrix} = 0. \tag{6.8}$$

Denote by $Z_{(\varepsilon+1)m\times\varepsilon n}(A + tB)$ the $(\varepsilon + 1)m \times \varepsilon n$ matrix in the left-hand side of (6.8). In view of (6.8) we have

$$\operatorname{rank} Z_{(\varepsilon+1)m\times\varepsilon n}(A + tB) < \varepsilon n, \tag{6.9}$$

and since $\varepsilon - 1$ is the smallest possible degree of a nonzero vector $y(x)$ for which (6.7) holds, we have

$$\operatorname{rank} Z_{(q+1)m\times qn}(A + tB) = qn, \qquad q = 1, 2, \ldots, \varepsilon - 1. \tag{6.10}$$

Next, we prove that the vectors

$$Ay_1, \ldots, Ay_{\varepsilon-1} \in \mathsf{H}^{m\times 1} \tag{6.11}$$

are linearly independent. Assume the contrary, and let

$$Ay_h = \sum_{j=1}^{h-1} Ay_j \alpha_{h-j}, \qquad \alpha_1, \ldots, \alpha_{h-1} \in \mathsf{H},$$

for some h, $2 \leq h \leq \varepsilon - 1$. (If Ay_1 were equal to 0, then in view of (6.8) we would have $Ay_0 = By_0 = 0$, hence $y_0 = 0$ by (6.5), and so $y(t)/t$ would be a polynomial of degree smaller than $\varepsilon - 1$ still satisfying (6.7), a contradiction with the choice of $\varepsilon - 1$.) Equation (6.8) gives

$$By_{h-1} = \sum_{j=1}^{h-1} By_{j-1}\alpha_{h-j},$$

in other words,

$$B\widetilde{y}_{h-1} = 0, \qquad \widetilde{y}_{h-1} := y_{h-1} - \sum_{j=0}^{h-2} y_j \alpha_{h-1-j}.$$

Introducing the vectors

$$\widetilde{y}_{h-2} := y_{h-2} - \sum_{j=0}^{h-3} y_j \alpha_{h-2-j}, \quad \widetilde{y}_{h-3} := y_{h-3} - \sum_{j=0}^{h-4} y_j \alpha_{h-3-j}, \quad \ldots,$$

$$\widetilde{y}_1 := y_1 - y_0\alpha_1, \quad \widetilde{y}_0 := y_0,$$

we obtain using (6.8) the equalities

$$A\widetilde{y}_{h-1} = B\widetilde{y}_{h-2}, \quad \ldots, \quad A\widetilde{y}_1 = B\widetilde{y}_0, \quad A\widetilde{y}_0 = 0.$$

Thus,

$$\widetilde{y}(t) = \sum_{j=0}^{h-1} t^j (-1)^j \widetilde{y}_j$$

is a nonzero (because $\widetilde{y}_0 \neq 0$) polynomial of degree smaller than $h \leq \varepsilon - 1$ satisfying (6.7), a contradiction with the choice of $\varepsilon - 1$.

Now it is easy to see that the vectors $y_0, \ldots, y_{\varepsilon-1}$ are linearly independent. Indeed, if

$$y_0\alpha_0 + \cdots + y_{\varepsilon-1}\alpha_{\varepsilon-1} = 0, \qquad \alpha_j \in \mathsf{H},$$

then applying A and using the linear independence of (6.11), we obtain $\alpha_1 = \cdots = \alpha_{\varepsilon-1} = 0$, and since $y_0 \neq 0$, the equality $\alpha_0 = 0$ follows as well.

Let now $Q \in \mathsf{H}^{n\times n}$ be a invertible matrix whose first columns ε are $y_0, \ldots, y_{\varepsilon-1}$ (in that order), and let $P \in \mathsf{H}^{m\times m}$ be an invertible matrix whose first columns $\varepsilon - 1$ are $Ay_1, \ldots, Ay_{\varepsilon-1}$. (The existence of such P and Q follows easily from properties of bases of quaternionic subspaces; see, for example, [47, Proposition 1.6].) Define the matrices A' and B' by the equality

$$(A + tB)Q = P(A' + tB').$$

Using (6.8), it is easily seen that

$$A' + tB' = \begin{bmatrix} L_{\varepsilon-1}(t) & D + tF \\ 0_{(m-(\varepsilon-1))\times\varepsilon} & A'' + tB'' \end{bmatrix}, \tag{6.12}$$

for some matrices

$$A'', B'' \in \mathsf{H}^{(m-(\varepsilon-1))\times(n-\varepsilon)} \quad \text{and} \quad D, F \in \mathsf{H}^{(\varepsilon-1)\times(n-\varepsilon)}.$$

We obviously have

$$\begin{aligned} & Z_{\varepsilon m\times(\varepsilon-1)n}(A+tB) \\ = P' & \begin{bmatrix} Z_{\varepsilon(\varepsilon-1)\times(\varepsilon-1)\varepsilon}(L_{\varepsilon-1}(t)) & * \\ 0 & Z_{\varepsilon(m-(\varepsilon-1))\times(\varepsilon-1)(n-\varepsilon)}(A''+tB'') \end{bmatrix} Q' \end{aligned} \tag{6.13}$$

for some invertible matrices P' and Q'. One easily checks that the matrix

$$Z_{\varepsilon(\varepsilon-1)\times(\varepsilon-1)\varepsilon}(L_{\varepsilon-1}(t))$$

is invertible. Since by (6.10)

$$\operatorname{rank} Z_{\varepsilon m\times(\varepsilon-1)n}(A+tB) = (\varepsilon-1)n,$$

in view of (6.13) we have

$$\operatorname{rank} Z_{\varepsilon(m-(\varepsilon-1))\times(\varepsilon-1)(n-\varepsilon)}(A''+tB'') = (\varepsilon-1)(n-\varepsilon). \tag{6.14}$$

Finally, we shall prove that by applying a suitable strict equivalence, one can reduce the matrix polynomial in the right-hand side of (6.12) to the form (6.6). More precisely, we shall show that for some quaternionic matrices X and Y of suitable sizes the equality

$$\begin{aligned} L_{\varepsilon-1}(t) \oplus (A''+tB'') &= \begin{bmatrix} I_{\varepsilon-1} & Y \\ 0 & I_{m-(\varepsilon-1))} \end{bmatrix} \begin{bmatrix} L_{\varepsilon-1}(t) & D+tF \\ 0_{(m-(\varepsilon-1))\times\varepsilon} & A''+tB'' \end{bmatrix} \\ &= \begin{bmatrix} I_\varepsilon & -X \\ 0 & I_{n-\varepsilon} \end{bmatrix} \end{aligned} \tag{6.15}$$

holds. Equality (6.15) can be rewritten in the form

$$L_{\varepsilon-1}(t)X = D + tF + Y(A''+tB''). \tag{6.16}$$

Introduce notation for the entries of D, F, X, for the rows of Y, and for the columns of A'' and B'':

$$D = [d_{i,k}], \quad F = [f_{i,k}], \quad X = [x_{j,k}],$$

where

$$i = 1, 2, \ldots, \varepsilon-1; \quad k = 1, 2, \ldots, n-\varepsilon; \quad j = 1, 2, \ldots, \varepsilon;$$

and

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{\varepsilon-1} \end{bmatrix}, \quad A'' = \begin{bmatrix} a_1 & a_2 & \cdots & a_{n-\varepsilon} \end{bmatrix}, \quad B'' = \begin{bmatrix} b_1 & b_2 & \cdots & b_{n-\varepsilon} \end{bmatrix}.$$

Then, equating the entries in the kth column of both sides of (6.16), we see that (6.16) is equivalent to the following system of scalar equations, with unknowns $x_{j,k}$ and y_j:

$$\begin{array}{rcll} x_{2,k} + tx_{1,k} & = & d_{1,k} + tf_{1,k} + y_1 a_k + ty_1 b_k, & k = 1, 2, \ldots, n - \varepsilon, \\ x_{3,k} + tx_{2,k} & = & d_{2,k} + tf_{2,k} + y_2 a_k + ty_2 b_k, & k = 1, 2, \ldots, n - \varepsilon, \\ \vdots & \vdots & \vdots & \\ x_{\varepsilon,k} + tx_{\varepsilon-1,k} & = & d_{\varepsilon-1,k} + tf_{\varepsilon-1,k} + y_{\varepsilon-1} a_k + ty_{\varepsilon-1} b_k, & k = 1, 2, \ldots, n - \varepsilon. \end{array} \tag{6.17}$$

We first solve the system of equations

$$\begin{array}{rcll} y_1 a_k - y_2 b_k & = & f_{2,k} - d_{1,k}, & k = 1, 2, \ldots, n - \varepsilon, \\ y_2 a_k - y_3 b_k & = & f_{3,k} - d_{2,k}, & k = 1, 2, \ldots, n - \varepsilon, \\ \vdots & \vdots & \vdots & \\ y_{\varepsilon-2} a_k - y_{\varepsilon-1} b_k & = & f_{\varepsilon-1,k} - d_{\varepsilon-2,k}, & k = 1, 2, \ldots, n - \varepsilon. \end{array} \tag{6.18}$$

Indeed, (6.18) can be rewritten in the form

$$\begin{bmatrix} y_1 & -y_2 & \cdots & (-1)^{\varepsilon-1} y_{\varepsilon-2} & (-1)^{\varepsilon} y_{\varepsilon-1} \end{bmatrix} Z_{(\varepsilon-1)(m-\varepsilon+1)\times(\varepsilon-2)(n-\varepsilon)}(A'' + tB'') =$$
$$\left[[f_{2,k} - d_{1,k}]_{k=1}^{n-\varepsilon} \ \ [f_{3,k} - d_{2,k}]_{k=1}^{n-\varepsilon} \ \cdots \ [f_{\varepsilon-2,k} - d_{\varepsilon-3,k}]_{k=1}^{n-\varepsilon} \ \ [f_{\varepsilon-1,k} - d_{\varepsilon-2,k}]_{k=1}^{n-\varepsilon} \right].$$

This equation can be always solved for $y_1, \ldots, y_{\varepsilon-1}$, because by (6.14) and Proposition 3.5 the matrix $Z_{(\varepsilon-1)(m-\varepsilon+1)\times(\varepsilon-2)(n-\varepsilon)}(A'' + tB'')$ is left invertible. Once (6.18) is satisfied, a solution of (6.17) is easily obtained by setting

$$x_{1,k} = f_{1,k} + y_1 b_k, \quad x_{2,k} = f_{2,k} + y_2 b_k, \ \ldots, \ x_{\varepsilon-1,k} = f_{\varepsilon-1,k} + y_{\varepsilon-1} b_k.$$

This concludes the proof of Theorem 6.2. □

The dual statement of Theorem 6.2 reads as follows:

Theorem 6.3. *Let $A + tB \in \mathsf{H}(x)^{m\times n}$ be a matrix pencil such that the rank of $A + tB$ is smaller than m. Assume that*

$$xA = xB = 0, \quad x \in \mathsf{H}^{1\times m} \quad \Longrightarrow \quad x = 0. \tag{6.19}$$

Then $A + tB$ is strictly equivalent to a direct sum

$$(L_\varepsilon(t))^T \oplus (A' + tB') \tag{6.20}$$

for some $\varepsilon > 0$.

The proof is reduced to Theorem 6.2 upon considering the matrix pencil $A^* + tB^*$.

We are now ready to prove Theorem 6.1. First, we prove the existence of the form (6.4). The proof proceeds by induction on the sizes m and n. Using Theorems 6.2 and 6.3, we easily reduce the proof to the case when $m = n$ and $\operatorname{rank}(A+t_0B) = n$ for some real t_0. In other words, the matrix $A+t_0B$ is invertible. Replacing A by $A + t_0B$ and making a change of variable $t' = t - t_0$, we reduce the proof to the case when A is invertible. Now, an application of the Jordan form of $A^{-1}B$ easily completes the proof of existence of the form (6.4).

For the uniqueness, use the map $\mho_{q,\mathsf{C}}$, with $q = (\mathsf{i}, \mathsf{j}, \mathsf{k})$. Assume that (6.4) (where the α_j are taken to be complex numbers) and

$$\begin{aligned} 0_{u'\times v'} \oplus L_{\varepsilon'_1\times(\varepsilon'_1+1)} \oplus \cdots \oplus L_{\varepsilon'_{p'}\times(\varepsilon'_{p'}+1)} \oplus L^T_{\eta'_1\times(\eta'_1+1)} \oplus L^T_{\eta'_{q'}\times(\eta'_{q'}+1)} \\ \oplus(I_{k'_1} + tJ_{k'_1}(0)) \oplus \cdots \oplus (I_{k'_{r'}} + tJ_{k'_{r'}}(0)) \\ \oplus(tI_{\ell'_1} + J_{\ell'_1}(\alpha'_1)) \oplus \cdots \oplus (tI_{\ell'_s} + J_{\ell'_{s'}}(\alpha'_{s'})), \end{aligned} \tag{6.21}$$

are two block diagonal forms which are strictly equivalent to $A + tB$. Here

$$\varepsilon'_1 \le \cdots \le \varepsilon'_{p'};\ \eta'_1 \le \cdots \le \eta'_{q'};\ k'_1 \le \cdots \le k'_{r'};\quad \ell'_1 \le \cdots \le \ell'_{s'}$$

are positive integers, and $\alpha'_1, \ldots, \alpha'_{s'} \in \mathsf{C}$. Denoting by $A_0 + tB_0$ the form (6.4), and by $A'_0 + tB'_0$ the form (6.21), we have

$$A_0 + tB_0 = P(A'_0 + tB'_0)Q$$

for some invertible quaternionic matrices P and Q. Applying $\mho_{\{\mathsf{i},\mathsf{j},\mathsf{k}\},\mathsf{C}}$, we obtain

$$\begin{aligned} &\mho_{\{\mathsf{i},\mathsf{j},\mathsf{k}\},\mathsf{C}}(A_0) + t\mho_{\{\mathsf{i},\mathsf{j},\mathsf{k}\},\mathsf{C}}(B_0) \\ &\quad = \mho_{\{\mathsf{i},\mathsf{j},\mathsf{k}\},\mathsf{C}}(P)(\mho_{\{\mathsf{i},\mathsf{j},\mathsf{k}\},\mathsf{C}}(A'_0) + t\mho_{\{\mathsf{i},\mathsf{j},\mathsf{k}\},\mathsf{C}}(B'_0))\mho_{\{\mathsf{i},\mathsf{j},\mathsf{k}\},\mathsf{C}}(Q), \end{aligned}$$

where $\mho_{\{\mathsf{i},\mathsf{j},\mathsf{k}\},\mathsf{C}}(P)$ and $\mho_{\{\mathsf{i},\mathsf{j},\mathsf{k}\},\mathsf{C}}(Q)$ are invertible complex matrices. On the other hand, the definition of $\mho_{\{\mathsf{i},\mathsf{j},\mathsf{k}\},\mathsf{C}}$ gives

$$\mho_{\{\mathsf{i},\mathsf{j},\mathsf{k}\},\mathsf{C}}(X) = \begin{bmatrix} X & 0 \\ 0 & \overline{X} \end{bmatrix}, \quad X \in \{A_0, A'_0, B_0, B'_0\}.$$

Now the uniqueness of the Kronecker canonical form over C for the complex matrix pencil $\mho_{\{\mathsf{i},\mathsf{j},\mathsf{k}\},\mathsf{C}}(A_0)+t\mho_{\{\mathsf{i},\mathsf{j},\mathsf{k}\},\mathsf{C}}(B_0)$ (see, for example, [21, Chapter II] or [23]) yields the desired uniqueness statement in Theorem 6.1. $\square$

7. Canonical forms for symmetric matrix pencils

We fix a nonstandard iaa ϕ throughout Subsection 7.1.

Consider a matrix pencil $A + tB$, $A, B \in \mathsf{H}^{n\times n}$. The matrix pencil $A + tB$ is said to be *ϕ-symmetric* if $A_\phi = A$ and $B_\phi = B$. Two matrix pencils $A + tB$ and $A' + tB'$ are said to be *ϕ-congruent* if

$$A' + tB' = S_\phi(A + tB)S$$

for some invertible $S \in \mathsf{H}^{n\times n}$. Clearly, the ϕ-congruence of matrix pencils is an equivalence relation, and ϕ-congruent matrix pencils are strictly equivalent. However, in general strictly equivalent matrix pencils need not be ϕ-congruent. Moreover, any pencil which is ϕ-congruent to a ϕ-symmetric matrix pencil, is itself ϕ-symmetric.

7.1. Canonical form

In this subsection, we give the canonical forms for ϕ-symmetric matrix pencils, under strict equivalence and ϕ-congruence. It turns out that these forms are the same:

Theorem 7.1.

(a) *Every ϕ-symmetric matrix pencil $A + tB$ is strictly equivalent to a ϕ-symmetric matrix pencil of the form*

$$\begin{aligned} & 0_u \quad \oplus \\ \oplus \quad & \left(t \begin{bmatrix} 0 & 0 & F_{\varepsilon_1} \\ 0 & 0 & 0 \\ F_{\varepsilon_1} & 0 & 0 \end{bmatrix} + G_{2\varepsilon_1+1} \right) \oplus \cdots \oplus \left(t \begin{bmatrix} 0 & 0 & F_{\varepsilon_p} \\ 0 & 0 & 0 \\ F_{\varepsilon_p} & 0 & 0 \end{bmatrix} + G_{2\varepsilon_p+1} \right) \\ \oplus \quad & (F_{k_1} + tG_{k_1}) \oplus \cdots \oplus (F_{k_r} + tG_{k_r}) \\ \oplus \quad & ((t+\alpha_1) F_{\ell_1} + G_{\ell_1}) \oplus \cdots \oplus \left((t+\alpha_q) F_{\ell_q} + G_{\ell_q}\right). \end{aligned} \tag{7.1}$$

Here, $\varepsilon_1 \le \cdots \le \varepsilon_p$ and $k_1 \le \cdots \le k_r$ are positive integers, and $\alpha_j \in \mathrm{Inv}(\phi)$. The form (7.1) is uniquely determined by A and B up to permutation of the diagonal blocks, and up to replacement of each α_j with a quaternion $\beta_j \in$ $\mathrm{Inv}(\phi)$ such that

$$\mathfrak{R}(\alpha_j) = \mathfrak{R}(\beta_j) \quad \text{and} \quad |\mathfrak{V}(\alpha_j)| = |\mathfrak{V}(\beta_j)|. \tag{7.2}$$

(b) *Every ϕ-symmetric matrix pencil $A + tB$ is ϕ-congruent to a ϕ-symmetric matrix pencil of the form* (7.1)*, with the same uniqueness conditions as in the part* (a)*.*

(c) *ϕ-symmetric matrix pencils are ϕ-congruent if and only if they are strictly equivalent.*

Note that in (7.1) one can replace each G_v by $\widetilde{G}_v$ (cf. (6.3)).

Proof. Part (c) obviously follows from parts (a) and (b).

We prove (a). Let (6.4) be the Kronecker form of $A + tB$, where we assume $\alpha_j \in \mathrm{Inv}(\phi)$. Note that the three matrix pencils

$$\begin{gathered} I_{k_i} + tJ_{k_i}(0), \quad (I_{k_i})_\phi + t(J_{k_i}(0))_\phi = I_{k_i} + tJ_{k_i}(0)^T, \\ F_{k_i} + tG_{k_i} = (F_{k_i})_\phi + t(G_{k_i})_\phi \end{gathered} \tag{7.3}$$

are strictly equivalent. Indeed, the strict equivalence of the first two matrix pencils in (7.3) follows at once from the similarity (over the reals) of $J_{k_i}(0)$ and $J_{k_i}(0)^T$, and the strict equivalence of the last two matrix pencils in (7.3) follows from the equality

$$F_k J_k(\alpha)^T = \alpha F_k + G_k, \qquad \alpha \in \mathsf{H}. \tag{7.4}$$

Also, one checks analogously that the three matrix pencils

$$\begin{gathered} tI_{\ell_s} + J_{\ell_s}(\alpha_s), \quad t(I_{\ell_s})_\phi + (J_{\ell_s}(\alpha_s))_\phi = tI_{\ell_s} + J_{\ell_s}(\alpha_s)^T, \\ (t+\alpha_s)F_{\ell_s} + G_{\ell_s}, \end{gathered} \tag{7.5}$$

where $\alpha_s \in \operatorname{Inv}(\phi)$, are strictly equivalent. Since

$$A_\phi = A, \quad B_\phi = B, \quad \text{and} \quad (L_{\varepsilon,\varepsilon+1})_\phi = L^T_{\varepsilon,\varepsilon+1},$$

the uniqueness statement in Theorem 6.1 implies that $p = q$ and $\varepsilon_i = \eta_i$ for $i = 1, 2, \ldots, p$. Now note the equality

$$(I_\varepsilon \oplus F_{\varepsilon+1})(L_{\varepsilon,\varepsilon+1} \oplus L^T_{\varepsilon,\varepsilon+1})(I_{\varepsilon+1} \oplus F_\varepsilon)F_{2\varepsilon+1} = t \begin{bmatrix} 0 & 0 & F_\varepsilon \\ 0 & 0 & 0 \\ F_\varepsilon & 0 & 0 \end{bmatrix} + G_{2\varepsilon+1}. \quad (7.6)$$

Using equalities (7.3), (7.5), and (7.6), existence of the form (7.1) follows. The uniqueness of (7.1) (up to permutation of the diagonal blocks and up to replacements subject to (7.2)) follows from the corresponding uniqueness statement of Theorem 6.1.

We now prove (b). The following lemma (attributed to Hua [29], see [46], [34] for a proof) will be useful:

Lemma 7.2. *Let $p(\lambda)$ be a real scalar polynomial having one of the forms $p(\lambda) = \lambda - a$ with $a > 0$, or $p(\lambda) = (\lambda - a)(\lambda - \overline{a})$ with $a \in \mathsf{C} \setminus \mathsf{R}$. Then for every positive integer m there exists a scalar polynomial $f_m(\lambda)$ with real coefficients such that $\lambda \equiv (f_m(\lambda))^2 (\operatorname{mod}(p(\lambda))^m)$.*

Let (7.1) be the canonical form of $A + tB$ under strict equivalence, as in (a). We follow *mutatis mutandis* the line of argument as presented in the proof of Theorem 6.1 in [34] (which in turn is based on earlier sources, primarily [46]). We have

$$A + tB = P(A_0 + tB_0)Q \quad (7.7)$$

for some invertible $P, Q \in \mathsf{H}^{n\times n}$, where $A_0 + tB_0$ is in the form (7.1). Replacing $A + tB$ by the matrix pencil $Q_\phi^{-1}(A + tB)Q^{-1}$, we may (and do) assume that $Q = I$. Then the equalities $A_\phi = A$, $B_\phi = B$ and $(A_0)_\phi = A_0$, $(B_0)_\phi = B_0$ imply

$$PA_0 = A_0P_\phi, \quad PB_0 = B_0P_\phi.$$

Therefore, for every scalar polynomial with real coefficients $f(t)$ we also have

$$\begin{aligned} f(P)A_0 &= A_0(f(P))_\phi, \\ f(P)B_0 &= B_0(f(P))_\phi. \end{aligned} \quad (7.8)$$

We now consider several separate cases.

Case 1. *P has only real eigenvalues, and there is only one distinct real eigenvalue of P.*

Assume first that the eigenvalue γ of P is positive. From the Jordan form of P it easily follows that for the polynomial with real coefficients $(t - \gamma^{-1})^n$, we have $(P^{-1} - \gamma^{-1})^n = 0$. By Lemma 7.2 there exists a polynomial $f_n(t)$ with real coefficients such that $t = (f_n(t))^2 \ (\operatorname{mod}(t - \gamma^{-1})^n)$. Then $P^{-1} = (f_n(P^{-1}))^2$, and since P^{-1} itself is a polynomial of P with real coefficients, $P^{-1} = (f(P))^2$ for some polynomial $f(t)$ with real coefficients. Let $R = f(P)$, and use (7.8) to obtain

$$\begin{aligned} R(A+tB)R_\phi &= RP(A_0+tB_0)R_\phi = RP(A_0+tB_0)(f(P))_\phi \\ &= f(P)Pf(P)(A_0+tB_0) = A_0+tB_0, \end{aligned} \tag{7.9}$$

and the existence of the form (7.1) under ϕ-congruence follows.

Now assume that the eigenvalue of P is negative. Then, arguing as in the preceding paragraph, it is found that $(-P)^{-1} = (f(-P))^2$ for some polynomial $f(t)$ with real coefficients. As in (7.9) it follows that

$$R(A+tB)R_\phi = -(A_0+tB_0),$$

where $R = f(-P)$. Thus, to complete the proof of existence of the form (7.1) in Case 1, it remains to show that each of the blocks

$$Z_1 := t\begin{bmatrix} 0 & 0 & F_\varepsilon \\ 0 & 0 & 0 \\ F_\varepsilon & 0 & 0 \end{bmatrix} + G_{2\varepsilon+1} \tag{7.10}$$

$$Z_2 := F_{k_1} + tG_{k_1}, \tag{7.11}$$

and

$$Z_3 := (t+\alpha)F_\ell + G_\ell, \quad \alpha \in \mathrm{Inv}\,(\phi) \tag{7.12}$$

is ϕ-congruent to its negative. This follows from the equalities

$$\begin{bmatrix} I_\varepsilon & 0 \\ 0 & -I_{\varepsilon+1} \end{bmatrix} Z_1 \begin{bmatrix} I_\varepsilon & 0 \\ 0 & -I_{\varepsilon+1} \end{bmatrix} = -Z_1, \tag{7.13}$$

and

$$\widetilde{q}Z_2\widetilde{q} = -Z_2, \quad qZ_3q = -Z_3,$$

where $\widetilde{q}, q \in \mathrm{Inv}\,(\phi)$ are such that $\widetilde{q}^2 = -1$, and $q^2 = -1$ if α is real, and $q = \mathfrak{V}\,(\alpha)/|\mathfrak{V}\,(\alpha)|$ if α is nonreal.

Case 2. *All eigenvalues of P have equal real parts, and their vector parts are non zero and have equal Euclidean lengths.*

Let $\beta \in \mathsf{H}$ be an eigenvalue (necessarily nonreal) of P. Denote $\alpha = \beta^{-1}$. In view of the Jordan form of P, for some positive integer w (which may be taken the largest size of Jordan blocks in the Jordan form of P) we have $g(P^{-1}) = 0$, where $g(t) = ((t-\beta)(t-\overline{\beta}))^w$ is a polynomial with real coefficients. Using Lemma 7.2, we find a polynomial $g_w(t)$ with real coefficients such that $t = (g_w(t))^2 (\mathrm{mod}\, g(t))$. Then $P^{-1} = (g_w(P^{-1}))^2$. Since P is itself a matrix root of a polynomial with real coefficients, namely, $\widehat{g}(P) = 0$, where

$$\widehat{g}(t) = ((t-\alpha)(t-(\overline{\alpha})))^w,$$

it follows that P^{-1} is also a polynomial of P with real coefficients. Now argue as in Case 1 (when the eigenvalue of P is positive).

Case 3. *All other possibilities* (*not covered in Cases* 1 *and* 2).

In this case, using the Jordan form, P can be written in the form

$$P = S \text{ diag} (P_1, P_2, \dots, P_r)\ S^{-1} \tag{7.14}$$

where S is an invertible matrix and $P_1, \dots, P_r$ are matrices of sizes $n_1 \times n_1, \dots, n_r \times n_r$, respectively, such that any pair of eigenvalues λ of P_j and μ of P_i $(i \neq j)$ satisfies the following condition:

$$\text{either } \mathfrak{R}(\lambda) \neq \mathfrak{R}(\mu) \text{ or } |\mathfrak{V}(\lambda)| \neq |\mathfrak{V}(\mu)|, \text{ or both.} \tag{7.15}$$

We also have $r \geq 2$ (the situations when $r = 1$ are covered in cases 1 and 2). Substituting (7.14) in the equality $A + tB = P(A_0 + tB_0)$, we obtain

$$\left(P_1^{-1} \oplus \cdots \oplus P_r^{-1}\right) (\widetilde{A} + t\widetilde{B}) = \widetilde{A}_0 + t\widetilde{B}_0, \tag{7.16}$$

where

$\widetilde{A} = S^{-1}A(S_\phi)^{-1}$, $\widetilde{B} = S^{-1}B(S_\phi)^{-1}$, $\widetilde{A}_0 = S^{-1}A_0(S_\phi)^{-1}$, $\widetilde{B}_0 = S^{-1}B_0(S_\phi)^{-1}$.

Partition the matrix $\widetilde{A}$:

$$\widetilde{A} = [M_{ij}]_{i,j=1}^r ,$$

where M_{ij} is of the size $n_i \times n_j$. Since $\widetilde{A}$ and $\widetilde{A}_0$ are ϕ-symmetric, (7.16) implies

$$P_i^{-1} M_{ij} = M_{ij} \left((P_j)_\phi\right)^{-1} .$$

An elementary calculation shows that if (7.15) holds for two nonzero quaternions λ and μ, then the same property holds for the inverses λ^{-1} and μ^{-1}. Note also that in view of Corollary 3.10, $\nu \in \sigma(((P_j)_\phi)^{-1})$ if and only if $\nu \in \sigma(P_j^{-1})$. Thus, in view of these remarks, we have

$$\sigma(P_i^{-1}) \cap \sigma(((P_j)_\phi)^{-1}) = \emptyset \quad \text{for} \quad i \neq j.$$

Now by Proposition 3.11 we have $M_{ij} = 0$ $(i \neq j)$. In other words,

$$\widetilde{A} = M_{11} \oplus \cdots \oplus M_{rr}.$$

Similarly, $\widetilde{B} = N_{11} \oplus \cdots \oplus N_{rr}$, where N_{ii} is of size $n_i \times n_i$.

Now induction is used on the size of the matrix pencil $A + tB$ to complete the proof of part (b). The base of induction, when A and B are scalar quaternions, is easily verified (use Lemma 4.2). □

7.2. Symmetry with respect to several antiautomorphisms

The canonical forms of the preceding subsection may be extended to the situations when the quaternionic matrix pencil is symmetric with respect to several nonstandard iaa's. As we will see, this set up encompasses as a particular case complex symmetric matrix pencils as well.

Let Φ be a (nonempty) set of nonstandard iaa's. A matrix pencil $A + tB$ is said to be Φ-*symmetric* if $A_\phi = A$ and $B_\phi = B$ for all $\phi \in \Phi$. Two matrix pencils $A + tB$ and $A' + tB'$ are said to be Φ-*congruent* if

$$A' + tB' = S_\phi(A + tB)S$$

for some invertible $S \in \mathsf{H}^{n\times n}$ such that $S_\phi = S_{\phi'}$ for all $\phi, \phi' \in \Phi$. Clearly, the Φ-congruence of matrix pencils is an equivalence relation, and Φ-congruent matrix pencils are strictly equivalent. Any pencil which is Φ-congruent to a Φ-symmetric pencil is itself Φ-symmetric.

Denote

$$\mathrm{Inv}\,(\Phi) := \bigcap_{\phi\in\Phi} \mathrm{Inv}\,(\phi) = \{\alpha \in \mathsf{H} \, : \, \phi(\alpha) = \alpha \quad \forall \quad \phi \in \Phi\}.$$

Clearly, $\mathrm{Inv}\,(\Phi) \supseteq \mathsf{R}$.

Theorem 7.3. *Assume the set of nonstandard iaa's Φ is such that*

$$\mathrm{Inv}\,(\Phi) \neq \mathsf{R}. \tag{7.17}$$

Then:

(a) *Every Φ-symmetric matrix pencil $A+tB$ is strictly equivalent to a Φ-symmetric matrix pencil of the form* (7.17)*, where $\alpha_j \in \mathrm{Inv}\,(\Phi)$ for $j = 1, 2, \ldots, q$. Moreover, the form* (7.17) *is unique up to permutation of the diagonal blocks, and up to replacement of each α_j with a similar quaternion $\beta_j \in \mathrm{Inv}\,(\Phi)$.*

(b) *Every Φ-symmetric matrix pencil $A + tB$ is Φ-congruent to a Φ-symmetric matrix pencil of the form* (7.17)*, with the same uniqueness conditions as in the part* (a)*.*

(c) *Φ-symmetric matrix pencils are Φ-congruent if and only if they are strictly equivalent.*

Proof. If Φ consists only of one element, this is Theorem 7.1.

Assume that $\Phi = \{\phi_1, \phi_2\}$, where ϕ_1 and ϕ_2 are distinct nonstandard iaa's. It is easy to see from Proposition 2.3 that

$$\mathrm{Inv}\,(\Phi) = \mathrm{Span}_{\mathsf{R}}\{1, q\}$$

for some $q \in \mathsf{H}$ such that $q^2 = -1$. Without loss of generality we assume that $q = \mathsf{i}$; thus, $\mathrm{Inv}\,(\Phi) = \mathsf{C}$. In view of Lemma 2.4 we have

$$T \in \mathsf{H}^{n\times n}, \quad T_{\phi_1} = T_{\phi_2} \quad \Longleftrightarrow \quad T \in \mathsf{C}^{n\times n}.$$

Now clearly a matrix pencil $A + tB$ is Φ-symmetric if and only if the matrices A and B are complex and symmetric. The result now follows from the well-known canonical form of pairs of complex symmetric matrices, see, for example, [46, Theorem 1]. Same argument completes the proof in the general case when $\mathrm{Inv}\,(\Phi)$ is a two-dimensional real vector space.

The case when $\mathrm{Inv}\,(\Phi)$ is one-dimensional, i.e., $\mathrm{Inv}\,(\Phi) = \mathsf{R}$, is excluded by the hypothesis. $\square$

7.3. Symmetric inner products

Let Φ be a set of iaa's with the property (7.17). Thus, Φ contains only nonstandard iaa's. Let

$$\mathsf{V}_\Phi := \{\alpha \in \mathsf{H} \,:\, \alpha_\phi = \alpha_\psi \quad \forall \;\; \phi, \psi \in \Phi\},$$

and

$$\mathsf{V}_\Phi^{n\times 1} := \{x \in \mathsf{H}^{n\times 1} \,:\, x_\phi = x_\psi \quad \forall \;\; \phi, \psi \in \Phi\}.$$

For example, if Φ is a one element set, then $\mathsf{V}_\Phi = \mathsf{H}$ and $\mathsf{V}_\Phi^{n\times 1} = \mathsf{H}^{n\times 1}$. Clearly, $\mathsf{V}_\Phi^{n\times 1}$ is a real vector subspace of $\mathsf{H}^{n\times 1}$, and

$$\mathsf{V}_\Phi^{n\times 1} \supseteq (\mathrm{Inv}\,(\Phi))^{n\times 1} \supseteq \mathsf{R}^{n\times 1}.$$

In view of (7.17), we actually have $\mathsf{V}_\Phi^{n\times 1} \neq \mathsf{R}^{n\times 1}$.

Let $H \in \mathsf{H}^{n\times n}$ be an invertible matrix such that

$$H_\phi = H \quad \text{for every} \quad \phi \in \Phi. \tag{7.18}$$

The matrix H defines a *symmetric inner product* (or H-symmetric inner product, if H is to be emphasized) on $\mathsf{V}_\Phi^{n\times 1}$ by the formula

$$[x, y]_H := y_\phi H x, \quad x, y \in \mathsf{V}_\Phi^{n\times 1},$$

where ϕ is any (arbitrary) element of Φ. (The symmetric inner product depends also on Φ, of course; we suppress this dependence in our notation.)

The following properties of the H-symmetric inner product are immediate:

Proposition 7.4.

(1) $[x\alpha, y\beta]_H = \phi(\beta)[x, y]_H\alpha$, *for all* $x, y \in \mathsf{V}_\Phi^{n\times 1}$ *and all* $\alpha, \beta \in \mathsf{V}_\Phi$.
(2) $[x_1 + x_2, y]_H = [x_1, y]_H + [x_2, y]_H$, *for all* $x_1, x_2, y \in \mathsf{V}_\Phi^{n\times 1}$.
(3) $[x, y_1 + y_2]_H = [x, y_1]_H + [x, y_2]_H$, *for all* $x, y_1, y_2 \in \mathsf{V}_\Phi^{n\times 1}$.
(4) *If* $[x, y]_H = 0$ *for all* $y \in \mathsf{V}_\Phi^{n\times 1}$, *or* $[y, x]_H = 0$ *for all* $y \in \mathsf{V}_\Phi^{n\times 1}$, *then* $x = 0$.
(5) $[x, y]_H = \phi([y, x]_H)$ *for all* $x, y \in \mathsf{V}_\Phi^{n\times 1}$ *and all* $\phi \in \Phi$.

Proof. We indicate the verification of (4) only. If Φ is one element set, we have $\mathsf{V}_\Phi = \mathsf{H}$, and (4) is obvious. If Φ has more than one element, then $\mathsf{V}_\Phi = \mathrm{Span}\,\{1, q\}$ for some $q \in \mathsf{H}$ with $q^2 = -1$, and, upon identifying V_Φ with C, the H-symmetric inner product reduces to a nondegenerate bilinear inner product in $\mathsf{C}^{n\times 1}$, in which case (4) is obvious again. $\square$

A matrix $A \in \mathsf{H}^{n\times n}$ is said to be *H-symmetric* if

$$[Ax, y]_H = [x, Ay]_H \quad \forall \quad x, y \in \mathsf{V}_\Phi^{n\times 1},$$

or equivalently if the equality $HA = A_\phi H$ holds for every $\phi \in \Phi$.

It is easy to see that A is H-symmetric if and only if $S^{-1}AS$ is $S_\phi HS$-symmetric, for any invertible matrix $S \in \mathsf{H}^{n\times n}$ such that $S_\phi = S_\psi$ for all $\phi, \psi \in \Phi$. Canonical form under this action is given in the next theorem.

Theorem 7.5. *Let $H \in \mathsf{H}^{n\times n}$ be an invertible matrix such that* (7.18) *holds, and let A be H-symmetric. Then there exists an invertible matrix S such that $S_\phi = S_\psi$ for all $\phi, \psi \in \Phi$ and the matrices $S^{-1}AS$ and $S_\phi HS$ (for some or every $\phi \in \Phi$) have the form*

$$S_\phi HS = F_{\ell_1} \oplus \cdots \oplus F_{\ell_q}, \quad S^{-1}AS = J_{\ell_1}(\alpha_1) \oplus \cdots \oplus J_{\ell_q}(\alpha_q), \tag{7.19}$$

where $\alpha_1, \ldots, \alpha_q \in \mathrm{Inv}\,(\Phi)$. Moreover, the form (7.19) *is unique up to permutation of the diagonal blocks, and up to replacement of each α_j with a similar quaternion $\beta_j \in \mathrm{Inv}\,(\Phi)$.*

For the proof apply Theorem 7.3 (with G_ℓ replaced by $\widetilde{G}_\ell$) to the Φ-symmetric pencil $HA + tH$, and take advantage of the invertibility of H.

8. Canonical forms for skewsymmetric matrix pencils

We fix a nonstandard iaa ϕ throughout this section. A matrix pencil $A + tB$ is called *ϕ-skewsymmetric* if $A_\phi = -A$ and $B_\phi = -B$.

In this section we consider ϕ-skewsymmetric matrix pencils, their canonical forms, and various applications.

8.1. Canonical form

We start with the canonical form:

Theorem 8.1. *Fix $\beta \in \mathsf{H}$ such that $\phi(\beta) = -\beta$, $|\mathfrak{V}(\beta)| = 1$.*

(a) *Every ϕ-skewsymmetric matrix pencil $A + tB$, i.e., such that both $A \in \mathsf{H}^{n\times n}$ and $B \in \mathsf{H}^{n\times n}$ are ϕ-skewsymmetric, is strictly equivalent to a ϕ-skewsymmetric matrix pencil of the form*

$$\begin{aligned}
0_u \quad & \oplus \quad \left(t \begin{bmatrix} 0 & 0 & F_{\varepsilon_1} \\ 0 & 0 & 0 \\ -F_{\varepsilon_1} & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & F_{\varepsilon_1} & 0 \\ -F_{\varepsilon_1} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right) \oplus \cdots \\
& \oplus \quad \left(t \begin{bmatrix} 0 & 0 & F_{\varepsilon_p} \\ 0 & 0 & 0 \\ -F_{\varepsilon_p} & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & F_{\varepsilon_p} & 0 \\ -F_{\varepsilon_p} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right) \\
& \oplus \quad (\beta F_{k_1} + t\beta G_{k_1}) \oplus \cdots \oplus (\beta F_{k_r} + t\beta G_{k_r}) \\
& \oplus \quad ((t + \gamma_1)\beta F_{m_1} + \beta G_{m_1}) \oplus \cdots \oplus \big((t + \gamma_p)\beta F_{m_p} + \beta G_{m_p}\big) \\
& \oplus \quad \left((t + \alpha_1) \begin{bmatrix} 0 & F_{\ell_1} \\ -F_{\ell_1} & 0 \end{bmatrix} + \begin{bmatrix} 0 & G_{\ell_1} \\ -G_{\ell_1} & 0 \end{bmatrix} \right) \oplus \cdots \\
& \oplus \quad \left((t + \alpha_q) \begin{bmatrix} 0 & F_{\ell_q} \\ -F_{\ell_q} & 0 \end{bmatrix} + \begin{bmatrix} 0 & G_{\ell_q} \\ -G_{\ell_q} & 0 \end{bmatrix} \right).
\end{aligned} \tag{8.1}$$

Here, $\varepsilon_1 \le \cdots \le \varepsilon_p$ and $k_1 \le \cdots \le k_r$ are positive integers, $\alpha_1, \ldots, \alpha_q \in \mathrm{Inv}\,(\phi)\backslash\mathsf{R}$, and $\gamma_1, \ldots, \gamma_p$ are real.

The form (8.1) *is uniquely determined by A and B up to permutation of the diagonal blocks, and up to replacement of each α_j with a quaternion $\beta_j \in \mathrm{Inv}\,(\phi)$ such that*

$$\mathfrak{R}(\alpha_j) = \mathfrak{R}(\beta_j) \quad \text{and} \quad |\mathfrak{V}(\alpha_j)| = |\mathfrak{V}(\beta_j)|. \tag{8.2}$$

(b) *Every ϕ-skewsymmetric matrix pencil $A + tB$ is ϕ-congruent to a ϕ-skewsymmetric matrix pencil of the form*

$$\begin{aligned}
0_u \quad &\oplus \quad \left(t \begin{bmatrix} 0 & 0 & F_{\varepsilon_1} \\ 0 & 0 & 0 \\ -F_{\varepsilon_1} & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & F_{\varepsilon_1} & 0 \\ -F_{\varepsilon_1} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right) \oplus \cdots \\
&\oplus \quad \left(t \begin{bmatrix} 0 & 0 & F_{\varepsilon_p} \\ 0 & 0 & 0 \\ -F_{\varepsilon_p} & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & F_{\varepsilon_p} & 0 \\ -F_{\varepsilon_p} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right) \\
&\oplus \quad \delta_1 \left(\beta F_{k_1} + t\beta G_{k_1} \right) \oplus \cdots \oplus \delta_r \left(\beta F_{k_r} + t\beta G_{k_r} \right) \\
&\oplus \quad \eta_1 \left((t + \gamma_1)\beta F_{m_1} + \beta G_{m_1} \right) \oplus \cdots \oplus \eta_p \left((t + \gamma_p)\beta F_{m_p} + \beta G_{m_p} \right) \\
&\oplus \quad \left((t + \alpha_1) \begin{bmatrix} 0 & F_{\ell_1} \\ -F_{\ell_1} & 0 \end{bmatrix} + \begin{bmatrix} 0 & G_{\ell_1} \\ -G_{\ell_1} & 0 \end{bmatrix} \right) \oplus \cdots \\
&\oplus \quad \left((t + \alpha_q) \begin{bmatrix} 0 & F_{\ell_q} \\ -F_{\ell_q} & 0 \end{bmatrix} + \begin{bmatrix} 0 & G_{\ell_q} \\ -G_{\ell_q} & 0 \end{bmatrix} \right).
\end{aligned} \tag{8.3}$$

Here ε_i, k_j, α_m and γ_s are as in the part (a), *and $\delta_1, \ldots, \delta_r$ and $\eta_1, \ldots, \eta_p$ are signs ± 1.*

The form (8.3) *under ϕ-congruence is uniquely determined by A and B up to permutation of the diagonal blocks, and up to replacement of each α_j with a similar quaternion $\beta_j \in \mathrm{Inv}\,(\phi)$.*

In view of Proposition 2.3(c), up to multiplication by -1, there exists exactly one quaternion β such that $\phi(\beta) = -\beta$ and $|\mathfrak{V}(\beta)| = 1$. Note also that $\beta^2 = -1$.

In connection with the form (8.3), note the following ϕ-congruence:

$$\begin{aligned}
&\begin{bmatrix} \beta I_\varepsilon & 0 \\ 0 & I_{\varepsilon+1} \end{bmatrix} \left(t \begin{bmatrix} 0 & 0 & F_\varepsilon \\ 0 & 0 & 0 \\ -F_\varepsilon & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & F_\varepsilon & 0 \\ -F_\varepsilon & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right) \begin{bmatrix} -\beta I_\varepsilon & 0 \\ 0 & I_{\varepsilon+1} \end{bmatrix} \\
&\qquad = \beta \left(t \begin{bmatrix} 0 & 0 & F_\varepsilon \\ 0 & 0 & 0 \\ F_\varepsilon & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & F_\varepsilon & 0 \\ F_\varepsilon & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right).
\end{aligned} \tag{8.4}$$

The signs δ_i and η_j of (8.3) form the *sign characteristic* of the ϕ-skew-symmetric pencil $A + tB$. In view of the equalities

$$(\beta F_{k_i} + t\beta G_{k_i})(-\beta F_{k_i}) = I + tJ_{k_i}(0) \tag{8.5}$$

and

$$(t\beta F_{m_j} + \gamma_j \beta F_{m_j} + \beta G_{m_j})(-\beta F_{m_j}) = tI + J_{m_j}(\gamma_j), \quad \gamma_j \in \mathsf{R}, \tag{8.6}$$

the sign characteristic assigns a sign ± 1 to every index corresponding to a real eigenvalue of $A + tB$, as well as to every index at infinity. The signs are uniquely determined by $A+tB$, up to permutations of signs that correspond to equal indices of the same real eigenvalue, and permutations of signs that correspond to the equal indices at infinity.

Most of the rest of this section will be devoted to the proof of this theorem.

In the next lemma, the properties of the indices in the Kronecker form of ϕ-skewsymmetric pencils are given.

Lemma 8.2. *Let $A + tB \in \mathsf{H}^{n\times n}$, where A and B are ϕ-skewsymmetric. Then:*

(1) *the left indices of the pencil $A + tB$ coincide with its right indices;*
(2) *for every nonreal eigenvalue α of $A+tB$, the indices of $A+tB$ that correspond to α appear in pairs. In other words, for every nonreal eigenvalue α and every positive integer k, the number of blocks $tI_{\ell_j} + J_{\ell_j}(\alpha_j)$ in the Kronecker form of $A + tB$ for which $\ell_j = k$ and α_j is similar to α, is even.*

Proof. Part (1) follows from the uniqueness of the Kronecker form of $A + tB$; indeed, the left indices of $A + tB$ coincide with the right indices of $-(A_\phi + tB_\phi) = A + tB$. For the proof of Part (2), apply the map $\mho_{q,\mathsf{C}}$ for a suitable q, to transform A and B into a pair of complex skewsymmetric matrices $\mho_{q,\mathsf{C}}(A)$ and $\mho_{q,\mathsf{C}}(B)$ (see Proposition 3.4). If $A_0 + tB_0$ is the Kronecker form of $A + tB$, then $\mho_{q,\mathsf{C}}(A_0) + t\mho_{q,\mathsf{C}}(B_0)$ is the Kronecker form of $\mho_{q,\mathsf{C}}(A) + t\mho_{q,\mathsf{C}}(B)$. By the known result for Kronecker forms of pairs of skewsymmetric complex matrices [46, Section 4], in the form $\mho_{q,\mathsf{C}}(A) + t\mho_{q,\mathsf{C}}(B)$ the indices that correspond to every fixed nonzero eigenvalue appear in pairs. Using Theorem 3.8, we see that this property for $\mho_{q,\mathsf{C}}(A) + t\mho_{q,\mathsf{C}}(B)$ is equivalent to the statement in part (2). $\square$

In contrast with ϕ-symmetric matrix pencils, in the case of skewsymmetric pencils there is no statement for Φ-skewsymmetric pencils that parallels Theorem 8.1, where Φ is a set of at least two iaa's such that $\mathrm{Inv}\,(\Phi) \neq \mathsf{R}$. Indeed, for any such Φ there is no nonzero $\beta \in \mathsf{H}$ with the property that $\phi(\beta) = -\beta$ for every $\phi \in \Phi$. A closer examination (cf. the proof of Theorem 7.3) shows that in this case we recover the well-known canonical form of a pair of complex skewsymmetric matrices, as in [46], for example.

Also in contrast with ϕ-symmetric matrix pencils, strictly equivalent ϕ-skewsymmetric matrix pencils need not be ϕ-congruent. In fact:

Theorem 8.3. *Let $A + tB$ be a ϕ-skewsymmetric $n \times n$ quaternionic matrix pencil. Then the strict equivalence class of $A + tB$ consists of not more than 2^n ϕ-congruence classes.*

The upper bound 2^n is achieved for

$$A + tB = \mathrm{diag}\,((t + \gamma_1)\beta, (t + \gamma_2)\beta, \ldots, (t + \gamma_n)\beta)\,,$$

where $\gamma_1, \ldots, \gamma_n$ are distinct real numbers.

Proof. Assume that $A+tB$ is strictly equivalent to the form (8.1), and assume that in the form (8.1) there are exactly s distinct blocks of the types $\beta F_{k_i} + t\beta G_{k_i}$ and $(t + \gamma_i)\beta F_{m_i} + \beta G_{m_i}$. Denote these s distinct blocks by $K_1, \ldots, K_s$, and further assume that the block K_j appears in (8.1) exactly v_j times, for $j = 1, 2, \ldots, s$. Define the integer

$$q := \prod_{j=1}^{s} (v_j + 1).$$

Theorem 8.1 shows that the strict equivalence class of $A + tB$ contains exactly q ϕ-congruence classes. So the maximal number of ϕ-congruence classes in a strict equivalence class of a ϕ-skewsymmetric $n \times n$ quaternionic matrix pencil is equal to

$$\max\{\prod_{j=1}^{s} (v_j + 1)\},$$

where the maximum is taken over all tuples of positive integers $(v_1, \ldots, v_s)$ such that $v_1 + \cdots + v_s \leq n$. It is easy to see that the maximum is achieved for $s = n$ and $v_1 = \cdots = v_n = 1$ (use the elementary inequality $v + 1 < (v - u + 1)(u + 1)$ for every pair of positive integers u, v such that $1 \leq u \leq v - 1$). $\square$

A criterion for ϕ-skewsymmetric matrix pencils with the property that their strict equivalence to other ϕ-skewsymmetric matrix pencils implies ϕ-congruence is given next.

Theorem 8.4. *Let $A + tB$ be a ϕ-skewsymmetric matrix pencil. Assume that the following property holds:*

$$\operatorname{rank}(A + tB) = \operatorname{rank} A = \operatorname{rank} B \quad \forall \ t \in \mathsf{R}. \tag{8.7}$$

Then a ϕ-skewsymmetric matrix pencil $A' + tB'$ is ϕ-congruent to $A + tB$ if and only if $A' + tB'$ is strictly equivalent to $A + tB$.

Conversely, if a ϕ-skewsymmetric matrix pencil $A + tB$ has the property that every ϕ-skewsymmetric matrix pencil that is strictly equivalent to $A+tB$ is actually ϕ-congruent to $A + tB$, then (8.7) *holds.*

The proof of this theorem follows easily from Theorem 8.1 by inspection of (8.3). Indeed, property (8.7) is equivalent to the absence of blocks $\delta_j(\beta F_{k_j} + t\beta G_{k_j})$ and $\eta_j((t + \gamma_j)\beta F_{m_j} + \beta G_{m_j})$ in (8.3).

As another corollary of Theorem 8.1, we obtain a canonical form for ϕ-skewsymmetric matrices under ϕ-congruence:

Theorem 8.5. *Let β be as in Theorem* 8.1. *Then for every ϕ-skewsymmetric matrix $A \in \mathsf{H}^{n\times n}$ there exists an invertible $S \in \mathsf{H}^{n\times n}$ such that*

$$S_\phi A S = \begin{bmatrix} \beta I_p & 0 & 0 \\ 0 & -\beta I_q & 0 \\ 0 & 0 & 0_{n-p-q} \end{bmatrix}.$$

Moreover, the integers p and q are uniquely determined by A.

For the proof, apply Theorem 8.1 with A as given and $B = 0$.

The triple $(p, q, n-p-q)$ of Theorem 8.5 will be called the β-*signature* of the ϕ-skewsymmetric matrix $A \in \mathsf{H}^{n\times n}$. Since β is determined by ϕ up to multiplication by -1, the β-signature of A depends on β. In fact, the $-\beta$-signature of A is the negative of the β-signature of A.

To illustrate Theorem 8.5, let $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Then

$$S_\phi AS = \begin{bmatrix} \beta & 0 \\ 0 & -\beta \end{bmatrix}, \quad \text{where} \quad S = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ \beta & -\beta \end{bmatrix}.$$

8.2. Proof of Theorem 8.1

The rather long proof of Theorem 8.1 will be done in several steps.

We start with a lemma.

Lemma 8.6. (a) *The pencil $\beta F_p + t\beta G_p$ is not ϕ-congruent to its negative $-\beta F_p - t\beta G_p$.*

(b) *The pencil $(t+\gamma)\beta F_m + \beta G_m$, where $\gamma \in \mathsf{R}$, is not ϕ-congruent to its negative.*

Proof. It will be convenient to assume (without essential loss of generality) that $\phi(\mathsf{i}) = \mathsf{i}$, $\phi(\mathsf{j}) = \mathsf{j}$, $\phi(\mathsf{k}) = -\mathsf{k}$, and $\beta = \mathsf{k}$.

We dispose first of the easy case $p = 1$. Arguing by contradiction, assume that k is ϕ-congruent to $-\mathsf{k}$: $\phi(s)\mathsf{k}s = -\mathsf{k}$ for some $s \in \mathsf{H}$. Writing $s = a + b\mathsf{i} + c\mathsf{j} + d\mathsf{k}$, $a, b, c, d \in \mathsf{R}$, we have

$$(a + b\mathsf{i} + c\mathsf{j} - d\mathsf{k})\mathsf{k}(a + b\mathsf{i} + c\mathsf{j} + d\mathsf{k}) = -\mathsf{k},$$

which leads to a contradiction upon equating the coefficients of k in both sides of this equation: $a^2 + b^2 + c^2 + d^2 = -1$.

From now on we assume $p \geq 2$. Suppose the contrary:

$$S_\phi(\mathsf{k}F_p + t\mathsf{k}G_p)S = -\mathsf{k}F_p - t\mathsf{k}G_p,$$

for some (necessarily invertible) $S \in \mathsf{H}^{p\times p}$. Thus

$$S_\phi \mathsf{k}F_p S = -\mathsf{k}F_p, \quad S_\phi \mathsf{k}G_p S = -\mathsf{k}G_p. \tag{8.8}$$

Taking inverses in the first equation in (8.8) we obtain

$$S^{-1}\mathsf{k}F_p S_\phi^{-1} = -\mathsf{k}F_p. \tag{8.9}$$

We prove by induction on m that

$$S_\phi \mathsf{k}(G_pF_p)^m G_p S = -\mathsf{k}(G_pF_p)^m G_p. \tag{8.10}$$

The base of induction, when $m = 0$, is given in (8.8). Assuming (8.10) is proved for $m - 1$, we have

$$\begin{aligned} S_\phi \mathsf{k}(G_pF_p)^m G_p S &= -S_\phi \mathsf{k}(G_pF_p)^{m-1} G_p \mathsf{k}F_p \mathsf{k}G_p S \\ &= -S_\phi \mathsf{k}(G_pF_p)^{m-1} G_p S S^{-1} \mathsf{k}F_p S_\phi^{-1} S_\phi \mathsf{k}G_p S \\ &= -(-\mathsf{k}(G_pF_p)^{m-1} G_p)(-\mathsf{k}F_p)(-\mathsf{k}G_p), \end{aligned}$$

where to obtain the last equation we have used the induction hypothesis, equality (8.9), and the second equality in (8.8). Thus,

$$S_\phi \mathsf{k}(G_pF_p)^m G_p S = -\mathsf{k}(G_pF_p)^m G_p,$$

as required. In particular,

$$S_\phi \mathsf{k} H S = -\mathsf{k}H, \quad H := (G_pF_p)^{p-2}G_p, \tag{8.11}$$

and note that H has 1 in the top left corner and zeros everywhere else. In particular, H is positive semidefinite and nonzero. Write $S = A + \mathsf{i}B + \mathsf{j}C + \mathsf{k}D$, where $A, B, C, D \in \mathsf{R}^{p\times p}$. Then

$$S_\phi = A^T + \mathsf{i}B^T + \mathsf{j}C^T - \mathsf{k}D^T,$$

and equation (8.11) takes the form

$$(A^T + \mathsf{i}B^T + \mathsf{j}C^T - \mathsf{k}D^T)\mathsf{k}H(A + \mathsf{i}B + \mathsf{j}C + \mathsf{k}D) = -\mathsf{k}H. \tag{8.12}$$

Equating the coefficient of k in the left and in the right-hand sides of (8.12), we obtain

$$A^THA + B^THB + C^THC + D^THD = -H.$$

This equality is contradictory because the left-hand side is positive semidefinite in view of the inertia theorem, whereas the right-hand side is negative semidefinite and nonzero. This proves part (a).

For the part (b) we argue similarly; note that equalities (8.8) and (8.9) follow from the assumption to the contrary in this case as well. □

Next, we prove the part (a) of Theorem 8.1. Let $A+tB$ be a ϕ-skewsymmetric matrix pencil. By Lemma 8.2, the Kronecker form of $A+tB$ consists of blocks of the following types (besides the zero block 0_u):

(1) $L_{\varepsilon\times(\varepsilon+1)} \oplus L^T_{\varepsilon\times(\varepsilon+1)}$;
(2) $I + tJ_k(0)$;
(3) $tI_m + J_m(\gamma)$, $\gamma \in \mathsf{R}$;
(4) $(tI_\ell + J_\ell(\alpha)) \oplus (tI_\ell + J_\ell(\alpha))$, $\alpha \in \mathrm{Inv}\,(\phi) \setminus \mathsf{R}$.

Now each of the blocks (1)–(4) is strictly equivalent to one of the blocks in (8.1). Indeed, we have the following equalities:

$$\begin{bmatrix} I_{\varepsilon+1} & 0 \\ 0 & -F_\varepsilon \end{bmatrix} \left(t \begin{bmatrix} 0 & 0 & F_\varepsilon \\ 0 & 0_1 & 0 \\ -F_\varepsilon & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & F_\varepsilon & 0 \\ -F_\varepsilon & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right) \begin{bmatrix} F_\varepsilon & 0 \\ 0 & I_{\varepsilon+1} \end{bmatrix}$$

$$= \begin{bmatrix} 0 & L_{\varepsilon\times(\varepsilon+1)} \\ L^T_{\varepsilon\times(\varepsilon+1)} & 0 \end{bmatrix};$$

$$(tI_\ell + J_\ell(\alpha)) \oplus (tI_\ell + J_\ell(\alpha)) \begin{bmatrix} 0 & F_\ell \\ -F_\ell & 0 \end{bmatrix} = (t+\alpha) \begin{bmatrix} 0 & F_\ell \\ -F_\ell & 0 \end{bmatrix} + \begin{bmatrix} 0 & G_\ell \\ -G_\ell & 0 \end{bmatrix},$$

for $\alpha \in \mathrm{Inv}\,(\phi) \setminus \mathsf{R}$, as well as (8.5), (8.6). This proves the existence of the form (8.1) for $A + tB$; the uniqueness of (8.1), as indicated in the statement of (a), follows from the uniqueness of the Kronecker form of $A + tB$ (see Theorem 6.1).

For the proof of (b), we mimic the proof of Theorem 7.1. Let $A + tB$ be a ϕ-skewsymmetric matrix pencil. Using (a), there exist invertible matrices P and Q such that

$$(A + tB)\, Q = P\,(A_0 + tB_0)\,, \tag{8.13}$$

where $A_0 + tB_0$ is given by (8.3). We show first that $A + tB$ and $A_0 + tB_0$ are ϕ-congruent, for some choice of the signs $\delta_1, \ldots, \delta_r$ and $\eta_1, \ldots, \eta_p$.

Without loss of generality assume $Q = I$. Then we have

$$PA_0 = A = -A_\phi = -(A_0)_\phi P_\phi = A_0 P_\phi, \quad PB_0 = B_0 P_\phi, \tag{8.14}$$

and therefore

$$F(P)A_0 = A_0(F(P))_\phi, \quad F(P)B_0 = B_0(F(P))_\phi \tag{8.15}$$

for every polynomial $F(t)$ with real coefficients.

Assume first that P has only real eigenvalues and P has only one distinct real eigenvalue. Then the proof proceeds as in Case 1 (under the hypothesis that the eigenvalue of P is positive) of the proof of Theorem 7.1. If P has only real eigenvalues, exactly one distinct real eigenvalue, and the eigenvalue is negative, then the proof is reduced to the first case as in the proof of Theorem 7.1; however, by Lemma 8.6 the block $\beta F_k + t\beta G_k$ is not ϕ-congruent to its negative, and the block $(t + \gamma)\beta F_m + \beta G_m$ (where γ is real) is not ϕ-congruent to its negative. This accounts for the possibility of the signs δ_i and η_j. On the other hand, the blocks

$$K_{2\varepsilon+1} := t \begin{bmatrix} 0 & 0 & F_\varepsilon \\ 0 & 0 & 0 \\ -F_\varepsilon & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & F_\varepsilon & 0 \\ -F_\varepsilon & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and

$$(t + \alpha) \begin{bmatrix} 0 & F_\ell \\ -F_\ell & 0 \end{bmatrix} + \begin{bmatrix} 0 & G_\ell \\ -G_\ell & 0 \end{bmatrix}, \quad \alpha \in \mathrm{Inv}\,(\phi) \setminus \mathsf{R},$$

are ϕ-congruent to their negatives. Indeed,

$$\begin{bmatrix} I_\ell & 0 \\ 0 & -I_\ell \end{bmatrix} \begin{bmatrix} 0 & (t+\alpha)F_\ell + G_\ell \\ -\left((t+\alpha)F_\ell + G_\ell\right) & 0 \end{bmatrix} \begin{bmatrix} I_\ell & 0 \\ 0 & -I_\ell \end{bmatrix}$$
$$= \begin{bmatrix} 0 & -\left((t+\alpha)F_\ell + G_\ell\right) \\ (t+\alpha)F_\ell + G_\ell & 0 \end{bmatrix}$$

and

$$\begin{bmatrix} I_\varepsilon & 0 \\ 0 & -I_{\varepsilon+1} \end{bmatrix} K_{2\varepsilon+1} \begin{bmatrix} I_\varepsilon & 0 \\ 0 & -I_{\varepsilon+1} \end{bmatrix} = -K_{2\varepsilon+1}.$$

If all eigenvalues of P have the equal real parts and their vector parts have the same nonzero Euclidean lengths, then the proof is reduced to the first case as in the proof of Theorem 7.1. Finally, in the remaining cases use the quaternionic Jordan form of P, and the following decomposition:

$$P = S\,(\mathrm{diag}\,(P_1, P_2, \ldots, P_r))\, S^{-1},$$

where S is invertible, and $P_1, \ldots, P_r$ are quaternionic matrices of sizes $n_1 \times n_1, \ldots, n_r \times n_r$, respectively, such that

$$\lambda \in \sigma(P_i) \quad \Longrightarrow \quad \mu \notin \sigma(P_j) \quad \text{for } j \neq i, \quad \text{for } \mu \text{ similar to } \lambda \tag{8.16}$$

(i.e., any similar pair of eigenvalues of P is confined to just one block, P_j). Then argue as in Case 3 of the proof of Theorem 7.1. This concludes the proof of existence of the form (8.3).

It will be convenient to prove uniqueness of (8.3) in a separate subsection.

8.3. Proof of Theorem 8.1(b): uniqueness

We prove here the uniqueness (up to allowed permutations and replacements) of the form (8.3). The proof will be modeled after the proof of uniqueness of the canonical form for pairs of complex Hermitian matrices (see, for example, [34, Section 8]).

A preliminary result will be needed:

Lemma 8.7. *If*

$$S_\phi \begin{bmatrix} \beta I_p & 0 \\ 0 & -\beta I_q \end{bmatrix} S = \begin{bmatrix} \beta I_{p'} & 0 \\ 0 & -\beta I_{q'} \end{bmatrix}, \quad p+q = p'+q' = n,$$

for some $S \in \mathsf{H}^{n\times n}$, then $p = p'$ and $q = q'$.

Proof. We may assume $\phi(\mathsf{i}) = \mathsf{i}$, $\phi(\mathsf{j}) = \mathsf{j}$, $\phi(\mathsf{k}) = -\mathsf{k}$, and $\beta = \mathsf{k}$. Write $S = A + B\mathsf{i} + C\mathsf{j} + D\mathsf{k}$, where $A, B, C, D \in \mathsf{R}^{n\times n}$. Then

$$(A^T + B^T\mathsf{i} + C^T\mathsf{j} - D^T\mathsf{k}) \begin{bmatrix} \mathsf{k}I_p & 0 \\ 0 & -\mathsf{k}I_q \end{bmatrix} (A + B\mathsf{i} + C\mathsf{j} + D\mathsf{k}) = \begin{bmatrix} \mathsf{k}I_{p'} & 0 \\ 0 & -\mathsf{k}I_{q'} \end{bmatrix}. \tag{8.17}$$

Separating in this equality the real matrix coefficients of each of 1, i, j, and k, the equality (8.17) is easily seen to be equivalent to the following:

$$\begin{bmatrix} A^T & B^T & C^T & D^T \\ D^T & -C^T & B^T & -A^T \\ -B^T & A^T & D^T & -C^T \\ C^T & D^T & -A^T & -B^T \end{bmatrix} \operatorname{diag}(I_p, -I_q, I_p, -I_q, I_p, -I_q, I_p, -I_q)$$

$$\cdot \begin{bmatrix} A & D & -B & C \\ B & -C & A & D \\ C & B & D & -A \\ D & -A & -C & -B \end{bmatrix} = \operatorname{diag}(I_{p'}, -I_{q'}, I_{p'}, -I_{q'}, I_{p'}, -I_{q'}, I_{p'}, -I_{q'}).$$

Now the inertia theorem for real symmetric matrices implies $4p = 4p'$, as required. $\square$

We introduce the following terminology. Let $A_1 + tB_1$ be a ϕ-skewsymmetric pencil of the form (8.3). Then we say that

$$\delta_1 (\beta F_{k_1} + t\beta G_{k_1}) \oplus \cdots \oplus \delta_r (\beta F_{k_r} + t\beta G_{k_r})$$

is the β-*part* of A_1+tB_1, and for every real γ,

$$\oplus_{\{j:\gamma_j=\gamma\}}\left(\eta_j\left((t+\gamma_j)\beta F_{m_j}+\beta G_{m_j}\right)\right)$$

is the γ-*part* of A_1+tB_1. We will write A_1+tB_1 in the following form (perhaps, after permutation of blocks):

$$A_1+tB_1=0_u\oplus(A_{1,\flat}+tB_{1,\flat})\oplus(A_{1,\sharp}+tB_{1,\sharp})\oplus(A_{1,\beta}+tB_{1,\beta})\oplus\oplus_{j=1}^{m}(A_{1,\gamma_j}+tB_{1,\gamma_j}), \tag{8.18}$$

where $A_{1,\beta}+tB_{1,\beta}$ is the β-part of A_1+tB_1, the block $A_{1,\gamma_j}+tB_{1,\gamma_j}$ is the γ_j-part of A_1+tB_1 (the numbers $\gamma_1,\ldots,\gamma_m$ are real, distinct, and arranged in the increasing order: $\gamma_1<\cdots<\gamma_m$), the part $A_{1,\flat}+tB_{1,\flat}$ consists of a direct sum of blocks

$$t\begin{bmatrix} 0 & 0 & F_{\varepsilon_i} \\ 0 & 0 & 0 \\ -F_{\varepsilon_i} & 0 & 0 \end{bmatrix}+\begin{bmatrix} 0 & F_{\varepsilon_i} & 0 \\ -F_{\varepsilon_i} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad i=1,2,\ldots,p,$$

and the part $A_{1,\sharp}+tB_{1,\sharp}$ consists of the blocks

$$(t+\alpha_i)\begin{bmatrix} 0 & F_{\ell_i} \\ -F_{\ell_i} & 0 \end{bmatrix}+\begin{bmatrix} 0 & G_{\ell_i} \\ -G_{\ell_i} & 0 \end{bmatrix}, \quad i=1,2,\ldots,q, \quad \alpha_i\in \mathrm{Inv}\,(\phi)\setminus \mathsf{R}.$$

We say that the form (8.18), with the indicated properties, is a *standard form.*

Lemma 8.8. *Let two ϕ-skewsymmetric matrix pencils A_1+tB_1 and A_2+tB_2 be given in the standard forms* (8.18) *and*

$$A_2+tB_2=0_{u'}\oplus(A_{2,\flat}+tB_{2,\flat})\oplus(A_{2,\sharp}+tB_{2,\sharp})\oplus(A_{2,\beta}+tB_{2,\beta})\oplus\oplus_{j=1}^{m'}(A_{2,\gamma_j'}+tB_{2,\gamma_j'}).$$

Assume that A_1+tB_1 and A_2+tB_2 are ϕ-congruent. Then

$$u=u',\quad m=m',\quad \gamma_j=\gamma_j' \text{ for } j=1,2,\ldots,m, \tag{8.19}$$

the ϕ-skewsymmetric matrix pencils $A_{1,\flat}+tB_{1,\flat}$ and $A_{2,\flat}+tB_{2,\flat}$ are ϕ-congruent, the ϕ-skewsymmetric matrix pencils $A_{1,\sharp}+tB_{1,\sharp}$ and $A_{2,\sharp}+tB_{2,\sharp}$ are ϕ-congruent, the β-parts $A_{1,\beta}+tB_{1,\beta}$ and $A_{2,\beta}+tB_{2,\beta}$ are ϕ-congruent, and for each γ_j, the γ_j-parts $A_{1,\gamma_j}+tB_{1,\gamma_j}$ and $A_{2,\gamma_j}+tB_{2,\gamma_j}$ are ϕ-congruent.

Proof. Equalities (8.19) follow from the uniqueness of the Kronecker form of A_1+tB_1 (which is the same as the Kronecker form of A_2+tB_2). For the same reason, $A_{1,\flat}+tB_{1,\flat}$ is strictly equivalent to $A_{2,\flat}+tB_{2,\flat}$, and therefore, permuting if necessary the blocks, we may (and will) assume that

$$A_{1,\flat}+tB_{1,\flat}=A_{2,\flat}+tB_{2,\flat}.$$

Analogously, we may assume that

$$A_{1,\sharp}+tB_{1,\sharp}=A_{2,\sharp}+tB_{2,\sharp}.$$

Also, the β-parts of A_1+tB_1 and A_2+tB_2 are strictly equivalent, as well as the γ_j-parts of A_1+tB_1 and A_2+tB_2, for every fixed γ_j.

For the uniformity of notation, we let

$$M_0 + tN_0 := 0_u,$$

$$M_i + tN_i := t\begin{bmatrix} 0 & 0 & F_{\varepsilon_i} \\ 0 & 0 & 0 \\ -F_{\varepsilon_i} & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & F_{\varepsilon_i} & 0 \\ -F_{\varepsilon_i} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad i = 1, 2, \ldots, p, \quad (8.20)$$

$$\begin{aligned} M_{p+i} + tN_{p+i} \quad := \quad & \bigoplus_{s=1}^{w_i} \left((t+\alpha_i) \begin{bmatrix} 0 & F_{\ell_{i,s}} \\ -F_{\ell_{i,s}} & 0 \end{bmatrix} + \begin{bmatrix} 0 & G_{\ell_{i,s}} \\ -G_{\ell_{i,s}} & 0 \end{bmatrix} \right), \\ & i = 1, 2, \ldots, q, \end{aligned} \quad (8.21)$$

where $\alpha_1, \ldots, \alpha_q \in \mathrm{Inv}\,(\phi) \setminus \mathsf{R}$ are mutually nonsimilar,

$$M_{p+q+1} + tN_{p+q+1} := A_{1,\beta} + tB_{1,\beta}, \quad M'_{p+q+1} + tN'_{p+q+1} := A_{2,\beta} + tB_{2,\beta},$$

$$M_{p+q+1+j} + tN_{p+q+1+j} := A_{1,\gamma_j} + tB_{1,\gamma_j}, \quad j = 1, 2, \ldots, m;$$

$$M'_{p+q+1+j} + tN'_{p+q+1+j} := A_{2,\gamma_j} + tB_{2,\gamma_j}, \quad j = 1, 2, \ldots, m.$$

Let $n_j \times n_j$ be the size of $M_j + tN_j$, for $j = 1, 2, \ldots, p+q+1+m$ (note that $n_j \times n_j$ is also the size of $M'_j + tN'_j$, for $j = p+q+1, p+q+2, \ldots, p+q+1+m$). Write

$$\begin{aligned} T\left(\oplus_{j=0}^{p+q+1+m} (M_j + tN_j)\right) \quad &= \quad \left(\oplus_{j=0}^{p+q} (M_j + tN_j) \oplus \oplus_{j=p+q+1}^{p+q+1+m} \left(M'_j + tN'_j\right)\right) S, \\ S \quad &= \quad T_\phi^{-1}, \end{aligned} \quad (8.22)$$

for some invertible T, and partition T and S conformably with (8.22):

$$T = (T^{(ij)})_{i,j=0}^{p+q+1+m}, \quad S = (S^{(ij)})_{i,j=0}^{p+q+1+m},$$

where $T^{(ij)}$ and $S^{(ij)}$ are $n_i \times n_j$. We then have from (8.22):

$$T^{(ij)} M_j = M_i S^{(ij)}; \quad T^{(ij)} N_j = N_i S^{(ij)} \quad (i, j = 0, \ldots, p+q+1+m), \quad (8.23)$$

where in the right-hand sides M_i and N_i are replaced by M'_i and N'_i, respectively, for $i = p+q+1, p+q+2, \ldots, p+q+m+1$. In particular,

$$T^{(0j)} M_j = 0, \quad T^{(0j)} N_j = 0 \quad (j = 0, \ldots, p+q+1+m),$$

which immediately implies the equalities $T^{(0j)} = 0$ for $j = p+1, \ldots, p+q+1+m$. Since

$$\mathrm{Range}\, M_j + \mathrm{Range}\, N_j = \mathsf{H}^{n_j}, \qquad (j = 1, \ldots, p),$$

we also obtain $T^{(0j)} = 0$ for $j = 1, \ldots, p$. Therefore, in view of the equality $S = T_\phi^{-1}$, also $S^{(i0)} = 0$ for $i = 1, \ldots, p+q+1+m$. Now clearly the matrices

$$\widetilde{T} := \left[T^{(ij)}\right]_{i,j=1}^{p+q+1+m} \qquad \text{and} \quad \widetilde{S} := \left[S^{(ij)}\right]_{i,j=1}^{p+q+1+m} = (\widetilde{T}_\phi)^{-1}$$

are invertible, and we have

$$\widetilde{T}\left(\oplus_{j=1}^{p+q+1+m}(M_j+tN_j)\right)=\left(\oplus_{j=1}^{p+q}(M_j+tN_j)\oplus\oplus_{j=p+q+1}^{p+q+1+m}\left(M_j'+tN_j'\right)\right)\widetilde{S}. \tag{8.24}$$

Next, consider the equalities (8.23) for $i,j=p+1,\dots,p+q+1+m$ and $i\neq j$ (these hypotheses on i and j will be assumed throughout the present paragraph). If $i,j\neq p+q+1$, then N_i and N_j are both invertible, and using (8.23) we obtain

$$T^{(ij)}M_j=M_iS^{(ij)}=M_iN_i^{-1}T^{(ij)}N_j, \tag{8.25}$$

hence

$$T^{(ij)}M_jN_j^{-1}=M_iN_i^{-1}T^{(ij)}$$

(if $i>p+q+1$, then we use M_i' and N_i' in place of M_i and N_i, respectively, in the right-hand side of (8.25)). A computation shows that

$$M_jN_j^{-1}=\oplus_{s=1}^{w_{j-p}}\left(J_{\ell_{j-p,s}}(\alpha_{j-p})\oplus J_{\ell_{j-p,s}}(\alpha_{j-p})\right),\quad \text{for}\quad j=p+1,\dots,p+q,$$

and $M_jN_j^{-1}=M_j'(N_j')^{-1}$ is a Jordan matrix with eigenvalue $\gamma_{j-p-q-1}$ for $j=p+q+2,\dots,p+q+1+m$. By Proposition 3.11, we obtain

$$T^{(ij)}=0\quad \text{for } i,j\in\{p+1,\dots,p+q,p+q+2,\dots,p+q+1+m\},\ i\neq j.$$

If $i\neq p+q+1$ but $j=p+q+1$, then N_i and M_j are invertible, and (8.23) leads to

$$\begin{aligned}T^{(ij)}&=M_iS^{(ij)}M_j^{-1}=M_iN_i^{-1}T^{(ij)}N_jM_j^{-1}\\&=(M_iN_i^{-1})^2T^{(ij)}(N_jM_j^{-1})^2\\&=\cdots=(M_iN_i^{-1})^wT^{(ij)}(N_jM_j^{-1})^w=0\end{aligned} \tag{8.26}$$

for sufficiently large positive integer w, because $N_jM_j^{-1}$ is easily seen to be nilpotent. (If $i>p+q+1$, replace M_i and N_i by M_i' and N_i', respectively, in (8.26).) Finally, if $i=p+q+1$ and $j\neq p+q+1$, then analogously (8.23) gives

$$\begin{aligned}T^{(ij)}&=N_i'S^{(ij)}N_j^{-1}=N_i'(M_i')^{-1}T^{(ij)}M_jN_j^{-1}\\&=(N_i'(M_i')^{-1})^2T^{(ij)}(M_jN_j^{-1})^2\\&=\cdots=(N_i'(M_i')^{-1})^wT^{(ij)}(N_jM_j^{-1})^w,\end{aligned} \tag{8.27}$$

and since $N_i'(M_i')^{-1}$ is nilpotent, the equality $T^{(ij)}=0$ follows. Thus $T^{(ij)}=0$ for $i,j=p+1,\dots,p+q+1+m$ and $i\neq j$. Analogously, we obtain the equalities $S^{(ij)}=0$ for $i,j=p+1,\dots,p+q+1+m$ and $i\neq j$.

Now consider $S^{(ij)}$ for $i=1,\dots,p$ and $j=p+1,\dots,p+q+1+m$ (these hypotheses on i and j will be assumed throughout the current paragraph). Assume first $j=p+q+1$. Then (8.23) gives

$$T^{(i,p+q+1)}M_{p+q+1}=M_iS^{(i,p+q+1)},\quad T^{(i,p+q+1)}N_{p+q+1}=N_iS^{(i,p+q+1)},$$

and therefore

$$M_iS^{(i,p+q+1)}M_{p+q+1}^{-1}N_{p+q+1}=N_iS^{(i,p+q+1)}. \tag{8.28}$$

Using the form (8.20) of M_i and N_i, and equating the bottom rows in the left- and right-hand sides of (8.28), we find that the first row of $S^{(i,p+q+1)}$ is zero. Then consideration of the next to the bottom row of (8.28) implies that the second row of $S^{(i,p+q+1)}$ is zero. Continuing in this fashion it is found that the top ε_i rows of $S^{(i,p+q+1)}$ consist of zeros. Applying a similar argument to the equality

$$T^{(p+q+1,i)}N_i = N'_{p+q+1}((M'_{p+q+1})^{-1}T^{(p+q+1,i)}M_i),$$

it is found that the first ε_i columns of $T^{(p+q+1,i)}$ consist of zeros. Now assume $j \neq p+q+1$. Then (8.23) gives

$$T^{(ij)}M_j = M_iS^{(ij)}; \quad T^{(ij)}N_j = N_iS^{(ij)},$$

and consequently

$$N_iS^{(ij)}N_j^{-1}M_j = M_iS^{(ij)}. \tag{8.29}$$

In view of the form (8.20) of M_i and N_i the (ε_i+1)th row in the left-hand side of (8.29) is zero. But (ε_i+1)th row in the right-hand side is the negative of the ε_ith row of $S^{(ij)}$. So the ε_ith row of $S^{(ij)}$ is equal to zero. By considering successively the (ε_i+2)th, (ε_i+3)th, etc. row on both sides of (8.29), it is found that the first ε_i rows of $S^{(ij)}$ are zeros. Next, use the equalities

$$T^{(ji)}M_i = M_jS^{(ji)}, \quad T^{(ji)}N_i = N_jS^{(ji)},$$

with M_j, N_j replaced by M'_j, N'_j, respectively, if $j > p+q+1$, to obtain

$$T^{(ji)}M_i = M_jS^{(ji)} = M_j(N_j^{-1}T^{(ji)}N_i).$$

Arguing as above, it follows that the first ε_i columns of $T^{(ji)}$ consist of zeros.

It has been shown that matrix $\widetilde{T}$ has the following form:

$$\widetilde{T} = \left[\begin{array}{c|c|c} T^{(11)} & \cdots & T^{(1p)} \\ \vdots & \vdots & \vdots \\ T^{(p1)} & \cdots & T^{(pp)} \\ \hline 0_{n_{p+1}\times\epsilon_1} \quad * & \cdots & 0_{n_{p+1}\times\epsilon_p} \quad * \\ 0_{n_{p+2}\times\epsilon_1} \quad * & \cdots & 0_{n_{p+2}\times\epsilon_p} \quad * \\ \vdots & \vdots & \vdots \\ 0_{n_{p+q+1+m}\times\epsilon_1} \quad * & \cdots & 0_{n_{p+q+1+m}\times\epsilon_p} \quad * \end{array}\right]$$

(p leftmost block columns), and

$$\widetilde{T} = \begin{bmatrix} T^{(1,p+1)} & T^{(1,p+2)} & \cdots & T^{(1,p+q+1+m)} \\ \vdots & \vdots & \vdots & \vdots \\ T^{(p,p+1)} & T^{(p,p+2)} & \cdots & T^{(p,p+q+1+m)} \\ T^{(p+1,p+1)} & 0 & \cdots & 0 \\ 0 & T^{(p+2,p+2)} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & T^{(p+q+1+m,p+q+1+m)} \end{bmatrix}$$

($q+1+m$ rightmost block columns). Let $W = T^{-1}$, and partition conformably with (8.18):

$$W = \left[W^{(ij)}\right]_{i,j=0}^{p+q+1+m}, \quad \widetilde{W} := \widetilde{T}^{-1} = \left[W^{(ij)}\right]_{i,j=1}^{p+q+1+m}.$$

In view of (8.24) we have

$$\widetilde{W}\left(\oplus_{j=1}^{p+q}(M_j + tN_j) \oplus \oplus_{j=p+q+1}^{p+q+1+m}\left(M'_j + tN'_j\right)\right) = \left(\oplus_{j=1}^{p+q+1+m}(M_j + tN_j)\right)\widetilde{S}^{-1}. \tag{8.30}$$

A similar argument shows that for $i = 1,\dots,p$; $j = p+1,\dots,p+q+1+m$, the first ε_i columns of $W^{(ji)}$ are zeros. Also, (8.30) implies $W^{(ij)} = 0$ for $i,j = p+1,\dots,p+q+1+m$ and $i \neq j$, which is proved exactly in the same way as $T^{(ij)} = 0$ ($i,j = p+1,\dots,p+q+1+m$, $i \neq j$) was proved using (8.24).

Denoting by $\widetilde{W}^{(ik)}$ the matrix formed by the $(\varepsilon_k + 1)$ right most columns of $W^{(ik)}$, and by $\widetilde{S}^{(kj)}$ the matrix formed by the $(\varepsilon_k + 1)$ bottom rows of $S^{(kj)}$ (here $i,j = p+1,\dots,p+q+1+m$; $k = 1,\dots,p$) we have

$$W^{(ik)}(M_k + tN_k)S^{(kj)} = \begin{bmatrix} 0 & \widetilde{W}^{(ik)} \end{bmatrix}\begin{bmatrix} 0_{\varepsilon_k} & * \\ * & 0_{\varepsilon_k+1} \end{bmatrix}\begin{bmatrix} 0 \\ \widetilde{S}^{(kj)} \end{bmatrix} = 0, \tag{8.31}$$

where the form (8.20) of M_k and N_k was used.

Hence, using the properties of $W^{(ij)}$ and $S^{(ij)}$ verified above, the following equalities are obtained, where $i,j \in \{p+1,\dots,p+q+1+m\}$:

$$\sum_{k=1}^{p+q} W^{(ik)}(M_k + tN_k)S^{(kj)} + \sum_{k=p+q+1}^{p+q+1+m} W^{(ik)}\left(M'_k + tN'_k\right)S^{(kj)} = 0$$

if $i \neq j$, and

$$\begin{aligned}\sum_{k=1}^{p+q} W^{(ik)}(M_k + tN_k)S^{(kj)} \quad &+ \sum_{k=p+q+1}^{p+q+1+m} W^{(ik)}\left(M'_k + tN'_k\right)S^{(kj)} \\ &= \quad W^{(ii)}\left(M_i + tN_i\right)S^{(ii)}\end{aligned} \tag{8.32}$$

if $i = j$. (For $i = p+q+1,\dots,p+q+1+m$, replace $W^{(ii)}(M_i + tN_i)S^{(ii)}$ with $W^{(ii)}\left(M'_i + tN'_i\right)S^{(ii)}$ in (8.32).) We now obtain, by a computation starting with (8.30), where (8.32) is used:

$$\begin{aligned} &\operatorname{diag}\left[M_i + tN_i\right]_{i=p+q+1}^{p+q+1+m} \\ = &\left[\sum_{k=1}^{p+q} W^{(ik)}(M_k + tN_k)S^{(kj)} + \sum_{k=p+q+1}^{p+q+1+m} W^{(ik)}\left(M'_k + tN'_k\right)S^{(kj)}\right]_{i,j=p+q+1}^{p+q+1+m} \\ = &\operatorname{diag}\left[W^{(ii)}\right]_{i=p+q+1}^{p+q+1+m}\operatorname{diag}\left[M'_i + tN'_i\right]_{i=p+q+1}^{p+q+1+m}\operatorname{diag}\left[S^{(ii)}\right]_{i=p+q+1}^{p+q+1+m}. \end{aligned} \tag{8.33}$$

Note that the matrix $\operatorname{diag}\left[M_i + tN_i\right]_{i=p+q+1}^{p+q+1+m}$ is clearly invertible for some real value of t, hence we see from (8.33) that the matrices $W^{(ii)}$ and $S^{(ii)}$ are invertible

($i = p+q+1, \dots, p+q+1+m$). Since also $\widetilde{S} = \widetilde{W_\phi}$, the equality (8.33) shows that for every $i = p+q+1, \dots, p+q+1+m$, the ϕ-skewsymmetric matrix pencils $M_i + tN_i$ and $M_i' + tN_i'$ are ϕ-congruent, as required. ☐

We now return to the proof of uniqueness in Theorem 8.1(b). Let $A + tB$ be a ϕ-skewsymmetric matrix pencil which is congruent to two forms (8.3). The uniqueness part of Theorem 8.1(a) guarantees that apart from permutations of blocks, these two forms can possibly differ only in the signs δ_j and η_k. Lemma 8.8 allows us to reduce the proof to cases when either only blocks of the form

$$\delta_1 \left(\beta F_{k_1} + t\beta G_{k_1}\right) \oplus \cdots \oplus \delta_r \left(\beta F_{k_r} + t\beta G_{k_r}\right) \tag{8.34}$$

are present, or only blocks of the form

$$\eta_1 \left((t+\gamma)\beta F_{m_1} + \beta G_{m_1}\right) \oplus \cdots \oplus \eta_p \left((t+\gamma)\beta F_{m_p} + \beta G_{m_p}\right), \quad \gamma \in \mathsf{R} \tag{8.35}$$

are present. We now consider each of these two cases separately.

We start with the form (8.35). So it is assumed that a ϕ-skewsymmetric matrix pencil $A + tB$ is ϕ-congruent to (8.35), as well as to (8.35) with possibly different signs $\widetilde{\eta}_j$. We have to prove that in fact the form

$$\widetilde{\eta}_1 \left((t+\gamma)\beta F_{m_1} + \beta G_{m_1}\right) \oplus \cdots \oplus \widetilde{\eta}_p \left((t+\gamma)\beta F_{m_p} + \beta G_{m_p}\right), \tag{8.36}$$

is obtained from (8.35) after a permutation of blocks. Write

$$T \left(\oplus_{j=1}^p \eta_j \left((t+\gamma)\beta F_{m_j} + \beta G_{m_j}\right)\right) = \left(\oplus_{j=1}^p \widetilde{\eta}_j \left((t+\gamma)\beta F_{m_j} + \beta G_{m_j}\right)\right) S, \tag{8.37}$$

where T is an invertible quaternionic matrix and $S = (T_\phi)^{-1}$, and partition

$$T = [T_{ij}]_{i,j=1}^p, \quad S = [S_{ij}]_{i,j=1}^p,$$

where T_{ij} and S_{ij} are $m_i \times m_j$. Then

$$\begin{aligned} T_{ij}\left(\eta_j \beta F_{m_j}\right) &= (\widetilde{\eta}_i \beta F_{m_i}) S_{ij}, \\ T_{ij}\left(\eta_j \left(\gamma\beta F_{m_j} + \beta G_{m_j}\right)\right) &= \left(\widetilde{\eta}_i \left(\gamma\beta F_{m_i} + \beta G_{m_i}\right)\right) S_{ij}, \end{aligned}$$

for $i, j = 1, 2, \dots, p$, and therefore

$$T_{ij}(\eta_j(\gamma\beta F_{m_j} + \beta G_{m_j})) = (\widetilde{\eta}_i(\gamma\beta F_{m_i} + \beta G_{m_i}))\widetilde{\eta}_i \beta^{-1} F_{m_i} T_{ij}(\eta_j \beta F_{m_j}). \tag{8.38}$$

Equality (8.38) implies

$$T_{ij} V_j = V_i T_{ij}, \tag{8.39}$$

where

$$V_i = \left(\gamma F_{m_i} + G_{m_i}\right) F_{m_i}^{-1} = J_{m_i}(\gamma),$$

the $m_i \times m_i$ Jordan block with eigenvalue γ. An inspection of the equality (8.39) shows that T_{ij} has the form

$$T_{ij} = \begin{bmatrix} 0 & \widetilde{T}_{ij} \end{bmatrix} \quad (\text{if } m_i \le m_j) \tag{8.40}$$

or

$$T_{ij} = \begin{bmatrix} \widetilde{T}_{ij} \\ 0 \end{bmatrix} \quad (\text{if } m_i \ge m_j), \tag{8.41}$$

where $\widetilde{T}_{ij}$ is an upper triangular Toeplitz matrix of size $\min(m_i, m_j) \times \min(m_i, m_j)$. Permuting blocks in (8.35) if necessary, it can be assumed that the sizes m_j are arranged in nondecreasing order. Let $u < v$ be indices such that

$$m_i < m_{u+1} = m_{u+2} = \cdots = m_v < m_j \quad \text{for all } i \leq u \text{ and for all } j > v.$$

Now for $u < i \leq v$, $u < k \leq v$, in view of (8.37) we obtain the following equality, where δ_{ik} is the Kronecker symbol, i.e., $\delta_{ik} = 1$ if $i = k$ and $\delta_{ik} = 0$ if $i \neq k$:

$$\delta_{ik}\widetilde{\eta}_i\beta F_{m_i} = \sum_{j=1}^{p} T_{ij}\left(\eta_j\beta F_{m_j}\right)(T_{kj})_\phi \tag{8.42}$$

$$= \sum_{j=1}^{u} T_{ij}(\eta_j\beta F_{m_j})(T_{kj})_\phi + \sum_{j=u+1}^{v} T_{ij}(\eta_j\beta F_{m_j})(T_{kj})_\phi + \sum_{j=v+1}^{p} T_{ij}(\eta_j\beta F_{m_j})(T_{kj})_\phi.$$

In view of (8.41) the lower left corner in the first sum in the right-hand side of (8.42) is zero. Using (8.40), it is easily verified that the lower left corner in the third sum is also zero. The lower left corner in the second sum in the right-hand side of (8.42) is equal to

$$\sum_{j=u+1}^{v} t_{ij}\eta_j\beta(t_{kj})_\phi,$$

where t_{ij} is the entry on the main diagonal of T_{ij}. Thus,

$$\delta_{ik}\widetilde{\eta}_i\beta = \sum_{j=u+1}^{v} t_{ij}\eta_j\beta(t_{kj})_\phi.$$

It follows that

$$\widetilde{\eta}_{u+1}\beta \oplus \cdots \oplus \widetilde{\eta}_v\beta = [t_{ik}]_{i,k=u+1}^{v}\left(\eta_{u+1}\beta \oplus \cdots \oplus \eta_v\beta)\right)\left([t_{ik}]_{i,k=u+1}^{v}\right)_\phi. \tag{8.43}$$

Now Lemma 8.7 guarantees that the two systems $\{\eta_{u+1}, \ldots, \eta_v\}$ and $\{\widetilde{\eta}_{u+1}, \ldots, \widetilde{\eta}_v\}$ have the same number of $+1$'s (and also the same number of -1's). Thus, within each set of blocks of equal size m_j, the number of η_j's which are equal to $+1$ (resp. to -1) coincides with the number of $\widetilde{\eta}_j$'s which are equal to $+1$ (resp. to -1). This shows that (8.36) is indeed obtained from (8.35) after a permutation of blocks.

Finally, assume that $A + tB$ is ϕ-congruent to (8.34), and also ϕ-congruent to

$$\widetilde{\delta}_1\left(\beta F_{k_1} + t\beta G_{k_1}\right) \oplus \cdots \oplus \widetilde{\delta}_r\left(\beta F_{k_r} + t\beta G_{k_r}\right)$$

with possibly different signs $\widetilde{\delta}_j$, $j = 1, 2, \ldots, r$. Arguing as in the preceding case, we obtain the equalities

$$T_{ij}(\delta_j\beta F_{k_j}) = (\widetilde{\delta}_i\beta F_{k_i})S_{ij},$$
$$T_{ij}\left(\delta_j\beta G_{k_j}\right) = \left(\widetilde{\delta}_i\beta G_{k_i}\right)S_{ij}.$$

The proof that the form

$$\oplus_{j=1}^{r}\widetilde{\delta}_j\left(F_{k_j}+G_{k_j}\right)$$

is obtained from (8.34) after a permutation of blocks, proceeds from now on in the same way (letting $\gamma=0$) as the proof that (8.35) and (8.36) are the same up to a permutation of blocks.

The proof of the uniqueness part of Theorem 8.1(b) is complete.

8.4. Skewsymmetric inner products

As in the preceding subsections, we fix a nonstandard iaa ϕ. Let $H\in\mathsf{H}^{n\times n}$ be an invertible matrix such that

$$H_\phi=-H. \tag{8.44}$$

The matrix H defines a *skewsymmetric inner product* (or H-skewsymmetric inner product, if H is to be emphasized) on $\mathsf{H}^{n\times 1}$ by the formula

$$[x,y]_H:=y_\phi Hx,\quad x,y\in\mathsf{H}^{n\times 1}.$$

(The skewsymmetric inner product depends also on ϕ, but we suppress this dependence in our notation.)

A matrix $A\in\mathsf{H}^{n\times n}$ is said to be *H-symmetric* if

$$[Ax,y]_H=[x,Ay]_H\quad\forall\quad x,y\in\mathsf{H}^{n\times 1},$$

or equivalently if the equality $HA=A_\phi H$ holds. In turn, the equality $HA=A_\phi H$ is equivalent to the matrix HA being ϕ-skewsymmetric: $(HA)_\phi=-HA$.

It is easy to see that A is H-symmetric if and only if $S^{-1}AS$ is $S_\phi HS$-symmetric, for any invertible matrix $S\in\mathsf{H}^{n\times n}$. Canonical form under this action on pairs of quaternionic matrices (H,A), where H is ϕ-skewsymmetric and A is H-symmetric, is given next.

Theorem 8.9. *Fix $\beta\in\mathsf{H}$ such that $\phi(\beta)=-\beta$ and $|\mathfrak{V}(\beta)|=1$. Let $H=-H_\phi\in\mathsf{H}^{n\times n}$ be an invertible matrix, and let A be H-symmetric. Then there exists an invertible matrix S such that the matrices $S^{-1}AS$ and $S_\phi HS$ have the form*

$$\begin{aligned}
S_\phi HS &= \eta_1\beta F_{m_1}\oplus\cdots\oplus\eta_p\beta F_{m_p}\oplus\begin{bmatrix}0 & F_{\ell_1}\\ -F_{\ell_1} & 0\end{bmatrix}\oplus\cdots\oplus\begin{bmatrix}0 & F_{\ell_q}\\ -F_{\ell_q} & 0\end{bmatrix},\\
S^{-1}AS &= J_{m_1}(\gamma_1)^T\oplus\cdots\oplus J_{m_p}(\gamma_p)^T\oplus\begin{bmatrix}J_{\ell_1}(\alpha_1)^T & 0\\ 0 & J_{\ell_1}(\alpha_1)^T\end{bmatrix}\oplus\\
&\cdots\oplus\begin{bmatrix}J_{\ell_q}(\alpha_q)^T & 0\\ 0 & J_{\ell_q}(\alpha_q)^T\end{bmatrix},
\end{aligned}\tag{8.45}$$

where $\eta_1,\ldots,\eta_p$ are signs ±1, the quaternions $\alpha_1,\ldots,\alpha_q\in\mathrm{Inv}\,(\phi)\setminus\mathsf{R}$, and $\gamma_1,\ldots,\gamma_p$ are real.

Moreover, the form (8.45) *is unique up to permutations of the diagonal blocks, and up to replacements of each α_j by a similar quaternion $\beta_j\in\mathrm{Inv}\,(\phi)$.*

In formula (8.45), any of the $J_{m_j}(\gamma_j)^T$'s can be replaced by $J_{m_j}(\gamma_j)$, and any of the $J_{\ell_j}(\alpha_j)^T$'s can be replaced by $J_{\ell_j}(\beta_j)$, where $\beta_j \in \mathrm{Inv}\,(\phi)$ is any quaternion similar to α_j. Use the equalities

$$F_{m_j} J_{m_j}(\gamma_j)^T F_{m_j} = J_{m_j}(\gamma_j), \quad j = 1, 2, \ldots, p,$$

and similar equalities for the blocks $J_{\ell_j}(\alpha_j)^T$ to verify this.

Proof. We apply Theorem 8.1 to the ϕ-skewsymmetric pencil $HA + tH$. Using the invertibility of H, we obtain $S_\phi HS$ as in (8.45), and $S^{-1}AS$ in the form

$$\begin{aligned} S^{-1}AS &= F_{m_1}(\gamma_1 F_{m_1} + G_{m_1}) \oplus \cdots \oplus F_{m_p}\left(\gamma_p F_{m_p} + G_{m_p}\right) \\ &\oplus \begin{bmatrix} 0 & F_{\ell_1} \\ -F_{\ell_1} & 0 \end{bmatrix}^{-1} \left(\alpha_1 \begin{bmatrix} 0 & F_{\ell_1} \\ -F_{\ell_1} & 0 \end{bmatrix} + \begin{bmatrix} 0 & G_{\ell_1} \\ -G_{\ell_1} & 0 \end{bmatrix}\right) \oplus \cdots \\ &\oplus \begin{bmatrix} 0 & F_{\ell_q} \\ -F_{\ell_q} & 0 \end{bmatrix}^{-1} \left(\alpha_q \begin{bmatrix} 0 & F_{\ell_q} \\ -F_{\ell_q} & 0 \end{bmatrix} + \begin{bmatrix} 0 & G_{\ell_q} \\ -G_{\ell_q} & 0 \end{bmatrix}\right), \end{aligned} \tag{8.46}$$

where $\eta_j \in \{1, -1\}$, the quaternions $\alpha_j, \ldots, \alpha_q \in \mathrm{Inv}\,(\phi) \setminus \mathsf{R}$, and $\gamma_1, \ldots, \gamma_p$ are real. It remains to observe that (8.46) simplifies to (8.45). □

8.5. Numerical ranges and numerical cones

Let ϕ be a nonstandard iaa. For a pair of ϕ-skewsymmetric $n \times n$ quaternionic matrices (A, B), we define the ϕ-*numerical range*

$$W_\phi(A, B) := \{(x_\phi Ax, x_\phi Bx) \,:\, x \in \mathsf{H}^{n\times 1}, \;\; \|x\| = 1\} \;\subseteq\; \mathsf{H}^2$$

and the ϕ-*numerical cone*

$$C_\phi(A, B) := \{(x_\phi Ax, x_\phi Bx) \,:\, x \in \mathsf{H}^{n\times 1}\} \;\subseteq\; \mathsf{H}^2.$$

Since $\phi(x_\phi Ax) = -x_\phi Ax$ and $\phi(x_\phi Bx) = -x_\phi Bx$, we clearly have that

$$W_\phi(A, B) \subseteq C_\phi(A, B) \subseteq \{(y_1\beta, y_2\beta) \,:\, y_1, y_2 \in \mathsf{R}\}.$$

Here β as before is a fixed quaternion such that $\phi(\beta) = -\beta$ and $\|\beta\| = 1$.

Proposition 8.10. *The ϕ-numerical range $W_\phi(A, B)$ and cone $C_\phi(A, B)$ are convex.*

Proof. We need to prove the convexity of $W_\phi(A, B)$, then the convexity of $C_\phi(A, B)$ will follow easily. Write

$$x = x_1 + x_2\mathsf{i} + x_3\mathsf{j} + x_4\mathsf{k} \in \mathsf{H}^{n\times 1}, \quad x_1, x_2, x_3, x_4 \in \mathsf{R}^{n\times 1},$$

$$A = A_1 + A_2\mathsf{i} + A_3\mathsf{j} + A_4\mathsf{k} \in \mathsf{H}^{n\times n}, \quad A_1, A_2, A_3, A_4 \in \mathsf{R}^{n\times n}.$$

Then we have $x_\phi Ax = a\beta$, where the real number a is represented as a bilinear form of

$$\widetilde{x} := \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \in \mathsf{R}^{4n\times 1}.$$

In other words, there exists a real symmetric $4n \times 4n$ matrix $\widetilde{A}$ such that

$$a = \widetilde{x}^T \widetilde{A} \widetilde{x}, \quad \widetilde{x} \in \mathsf{R}^{4n\times 1}.$$

Note also that $\|x\| = \|\widetilde{x}\|$. Now clearly

$$W_\phi(A,B) = W^r(\widetilde{A},\widetilde{B})\beta,$$

where

$$W^r(X,Y) := \{(x^T X x, x^T Y x) \,:\, x \in \mathsf{R}^{4n\times 1}, \quad \|x\| = 1\}$$

is the *real* numerical range of the pair (X,Y) of two real symmetric $4n\times 4n$ matrices X and Y. By the well-known result on convexity of real numerical ranges of pairs of real symmetric matrices (see [9]), the result follows. ☐

In studies of numerical ranges and their generalizations, characterization of geometric shapes in terms of the algebraic properties of the matrices involved is often of interest (see, for example, [7]). We present a result that characterizes the situations when the ϕ-numerical cone is contained in a half-plane bounded by a line passing through the origin. We identify here $\mathsf{R}^2\beta$ with R^2.

Theorem 8.11. *The following statements are equivalent for a pair of ϕ-skewsymmetric $n \times n$ matrices (A,B):*

(1) *$C_\phi(A,B)$ is contained in a half-plane;*
(2) *the pencil $A + tB$ is ϕ-congruent to a pencil of the form $\beta A' + t\beta B'$, where A' and B' are real symmetric matrices such that some linear combination $(\sin\mu)A' + (\cos\mu)B'$, $0 \le \mu < 2\pi$ is positive semidefinite;*

Proof. Observe that implication (2) $\Longrightarrow$ (1) is evident.

We prove (1) $\Longrightarrow$ (2). Since $C_\phi(A,B) = C_\phi(S_\phi AS, S_\phi BS)$ for any invertible quaternionic matrix S, we may (and do) assume without loss of generality that the pair (A,B) is given in the canonical form of Theorem 8.1(b).

It will be convenient to consider particular blocks first.

Claim. *Let*

$$A_0 = \begin{bmatrix} 0 & \alpha F_\ell + G_\ell \\ -\alpha F_\ell - G_\ell & 0 \end{bmatrix}, \quad B_0 = \begin{bmatrix} 0 & F_\ell \\ -F_\ell & 0 \end{bmatrix},$$

where $\alpha \in \mathrm{Inv}\,(\phi) \setminus \mathsf{R}$. Then

$$C_\phi(A_0,B_0) = \mathsf{R}^2\beta. \tag{8.47}$$

To verify (8.47), first observe that adding a real nonzero multiple of B_0 to A_0 does not alter the property (8.47). Thus, we may assume that the real part of

α is zero. Next, replacing if necessary ϕ by a similar nonstandard iaa, we may also assume that $\beta = \mathsf{k}$ and $\alpha = a\mathsf{i} + b\mathsf{j}$, for some real a and b not both zero. It will be proved that

$$(\mathsf{R}\mathsf{k}, 0) \subseteq \left\{ \left([\phi(y) \ \ 0 \ \ldots \ 0 \ \ \phi(z)]A_0 \begin{bmatrix} y \\ 0 \\ \vdots \\ 0 \\ z \end{bmatrix}, [\phi(y) \ \ 0 \ \ldots \ 0 \ \ \phi(z)]B_0 \begin{bmatrix} y \\ 0 \\ \vdots \\ 0 \\ z \end{bmatrix} \right) \right\}, \tag{8.48}$$

where in the right-hand side $y, z \in \mathsf{H}$ are arbitrary, and

$$(0, \mathsf{R}\mathsf{k}) \subseteq \left\{ \left([\phi(y) \ \ 0 \ \ldots \ 0 \ \ \phi(z)]A_0 \begin{bmatrix} y \\ 0 \\ \vdots \\ 0 \\ z \end{bmatrix}, [\phi(y) \ \ 0 \ \ldots \ 0 \ \ \phi(z)]B_0 \begin{bmatrix} y \\ 0 \\ \vdots \\ 0 \\ z \end{bmatrix} \right) \right\}, \tag{8.49}$$

where again $y, z \in \mathsf{H}$ in the right-hand side are arbitrary. In view of the convexity of $C_\phi(A_0, B_0)$ this will suffice to prove (8.47). Write

$$\begin{bmatrix} y \\ z \end{bmatrix} = x_1 + x_2\mathsf{i} + x_3\mathsf{j} + x_4\mathsf{k}, \quad x_1, x_2, x_3, x_4 \in \mathsf{R}^{2\times 1}.$$

Then

$$[\phi(y)\, 0 \ \ldots \ 0\, \phi(z)]B_0 \begin{bmatrix} y \\ 0 \\ \vdots \\ 0 \\ z \end{bmatrix} = (x_1^T + x_2^T\mathsf{i} + x_3^T\mathsf{j} - x_4^T\mathsf{k}) \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} (x_1 + x_2\mathsf{i} + x_3\mathsf{j} + x_4\mathsf{k}),$$

which in turn is equal to

$$\mathsf{k} \begin{bmatrix} x_1^T & x_2^T & x_3^T & x_4^T \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}. \tag{8.50}$$

Analogously,

$$[\phi(y)\ \ 0\ \dots\ 0\ \ \phi(z)]A_0 \begin{bmatrix} y \\ 0 \\ \vdots \\ 0 \\ z \end{bmatrix}$$

$$= (x_1^T + x_2^T \mathsf{i} + x_3^T \mathsf{j} - x_4^T \mathsf{k}) \begin{bmatrix} 0 & a\mathsf{i} + b\mathsf{j} \\ -a\mathsf{i} - b\mathsf{j} & 0 \end{bmatrix} (x_1 + x_2\mathsf{i} + x_3\mathsf{j} + x_4\mathsf{k}),$$

which is equal to

$$\mathsf{k} \begin{bmatrix} x_1^T & x_2^T & x_3^T & x_4^T \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & -b & 0 & a & 0 & 0 \\ 0 & 0 & b & 0 & -a & 0 & 0 & 0 \\ 0 & b & 0 & 0 & 0 & 0 & 0 & -a \\ -b & 0 & 0 & 0 & 0 & 0 & a & 0 \\ 0 & -a & 0 & 0 & 0 & 0 & 0 & -b \\ a & 0 & 0 & 0 & 0 & 0 & b & 0 \\ 0 & 0 & 0 & a & 0 & b & 0 & 0 \\ 0 & 0 & -a & 0 & -b & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}. \tag{8.51}$$

Assuming $x_3 = x_4 = 0$ (if $b \neq 0$) or $x_2 = x_4 = 0$ (if $a \neq 0$) in (8.50) and (8.51) shows the inclusion (8.48), and assuming $x_1 = x_4 = 0$ shows (8.49).

Assume now that statement (1) holds. In view of the Claim, and taking advantage of formula (8.4), we may further assume that $A = \widetilde{A}\beta$ and $B = \widetilde{B}\beta$, where $\widetilde{A}$ and $\widetilde{B}$ are real symmetric $n \times n$ matrices. Now clearly statement (2) follows. $\square$

Other criteria for Theorem 8.11(2) to hold, in terms of the pair of real symmetric matrices (A', B'), are given in [34] and [13].

Acknowledgment

The author thanks J. Baez and R. M. Guralnick for helpful consultations.

References

[1] S.L. Adler. *Quaternionic Quantum Mechanics and Quantum Fields*. Oxford University Press, New York, 1995.

[2] D. Alpay, A.C.M. Ran, and L. Rodman. Basic classes of matrices with respect to quaternionic indefinite inner product spaces. *Linear Algebra and its Applications*, **416** (2006), 242–269.

[3] D. Alpay and M. Shapiro. Reproducing kernel quaternionic Pontryagin spaces. *Integral Equations and Operator Theory*, **50** (2004), 431–476.

[4] D. Alpay, M. Shapiro, and D. Volok. Rational hyperholomorphic functions in R^4. *J. Funct. Anal.*, **221** (2005), 122–149.

[5] P. L.F. Arena, G. Muscato, and M.G. Xibilia. *Neural Networks in Multidimensional Domains*. Lecture Notes in Control and Information Sciences, Vol. 234, Springer, London, 1998.

[6] A. Baker. Right eigenvalues for quaternionic matrices: a topological approach. *Linear Algebra Appl.*, **286** (1999), 303–309.

[7] V. Bolotnikov, C.-K. Li, P.R. Meade, C. Mehl, and L. Rodman. Shells of matrices in indefinite inner product spaces. *Electron. J. Linear Algebra* , **9** (2002), 67–92.

[8] N. Bourbaki. *Éléments de mathématique. Livre II: Algèbre. Chapitre 9: Formes sesquilinéaires et formes quadratiques*, Hermann, Paris, 1959. (French.)

[9] L. Brickman. On the field of values of a matrix. *Proc. Amer. Math. Soc.* **12** (1961), 61–66.

[10] J.L. Brenner. Matrices of quaternions. *Pacific J. of Math.*, **1** (1951), 329–335.

[11] E. Brieskorn. *Lineare Algebra und analytische Geometrie*. Volumes I and II. Friedr. Vieweg and Sohn, Braunschweig, 1985.

[12] A. Bunse-Gerstner, R. Byers, and V. Mehrmann. A chart of numerical methods for structured eigenvalue problems, *SIAM J. Matrix Anal. Appl.*, **13** (1992), 419–453.

[13] W.S. Cheung, C.-K. Li, and L. Rodman. Operators with numerical range in a closed half-plane. To appear in *Taiwanese Journal of Mathematics*.

[14] P.M. Cohn. Similarity reduction of matrices over a skew field. *Math. Z.*, **132** (1973), 151–163.

[15] F. Colombo, I. Sabadini, F. Sommen, and D.C. Struppa. *Analysis of Dirac Systems and Computational Algebra*, Birkhäuser, Boston, 2004.

[16] D.Z. Djokovic. Classification of pairs consisting of a linear and a semilinear map. *Linear and Multilinear Algebra* 20 (1978), 147–165.

[17] D.Z. Djokovic, J. Patera, P. Winternitz, and H. Zassenhaus. Normal forms of elements of classical real and complex Lie and Jordan algebras. *J. Math. Phys.*, **24** (1983), 1363–1374.

[18] D.R. Farenick and B.A.F. Pidkowich. The spectral theorem in quaternions. *Linear Algebra Appl.*, **371** (2003), 75–102.

[19] H. Faßbender, D.S. Mackey, and N. Mackey. Hamilton and Jacobi come full circle: Jacobi algorithms for structured Hamiltonian eigenproblems. *Linear Algebra Appl.*, **332/334** (2001), 37–80.

[20] F.R. Gantmacher. *The Theory of Matrices*. Vols. 1 and 2, Chelsea, New York, 1959. (Translation from Russian.)

[21] F.R. Gantmacher. *Applications of the Theory of Matrices*. Interscience Publishers, New York, 1959. (Translation of part II of the Russian original.)

[22] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials.* Academic Press, New York etc., 1982.

[23] I. Gohberg, P. Lancaster, and L. Rodman. *Invariant Subspaces of Matrices with Applications.* John Wiley, New York etc., 1986.

[24] I. Gohberg, P. Lancaster, and L. Rodman. *Indefinite Linear Algebra and Applications*. Birkhäuser, Boston, 2006.

[25] R.M. Guralnick and L.S. Levy. Presentations of modules when ideals need not be principal. *Illinois J. Math.*, **32** (1988), 593–653.

[26] R.M. Guralnick, L.S. Levy, and C. Odenthal. Elementary divisor theorem for noncommutative PIDs. *Proc. Amer. Math. Soc.*, **103** (1988), 1003–1011.

[27] K. Gürlebeck and W. Sprössig. *Quaternionic and Clifford Calculus for Physicists and Engineers*. Wiley, Chichester, 1997.

[28] F. Gürsey and C.-H. Tze. *On the Role of Division, Jordan and Related Algebras in Particle Physics*. World Scientific Publishing, River Edge, NJ, 1996.

[29] L.K. Hua. On the theory of automorphic functions of a matrix variable, II - The classification of hypercircles under the symplectic group. *Amer. J. Math*, **66** (1944), 531–563.

[30] N. Jacobson. Normal Semi-Linear Transformations. *Amer. J. Math.*, **61** (1939), 45–58.

[31] N. Jacobson. *The Theory of Rings*. American Mathematical Society, Rhode Island, 1943.

[32] M. Karow. Self-adjoint operators and pairs of Hermitian forms over the quaternions. *Linear Algebra and its Applications*, **299** (1999), 101–117.

[33] M. Koecher and R. Remmert. Hamilton's quaternions, in: *Numbers*, Springer, 1991, pp. 189–220.

[34] P. Lancaster and L. Rodman. Canonical forms for hermitian matrix pairs under strict equivalence and congruence. *SIAM Review,* **47** (2005), 407–443.

[35] P. Lancaster and L. Rodman. Canonical forms for symmetric/skew-symmetric real matrix pairs under strict equivalence and congruence. *Linear Algebra and its Applications*, **406** (2005), 1–76.

[36] P. Lancaster and M. Tismenetsky. *The Theory of Matrices.* Second Edition, Academic Press, Orlando, 1985.

[37] H.C. Lee. Eigenvalues and canonical forms of matrices with quaternion coefficients. *Proc. Roy. Irish Acad. Sect. A.*, **52** (1949), 253–260.

[38] Y. Nakayama. A note on the elementary divisor theory in non-commutative domains. *Bull. Amer. Math. Soc.*, **44** (1938), 719–723.

[39] R. Pereira, P. Rocha, and P. Vettori. Algebraic tools for the study of quaternionic behavioral systems. *Linear Algebra Appl.*, **400** (2005), 121–140.

[40] R. von Randow. The involutory antiautomorphisms of the quaternion algebra. *American Mathematical Monthly*, 74 (1967), 699–700.

[41] L. Rodman. Canonical forms for mixed symmeric/skewsymmetric quaternionic matrix pencils. *Linear Algebra Appl.*, to appear.

[42] L. Rodman. Pairs of hermitian and skew-hermitian quaternionic matrices: canonical forms and their applications. *Linear Algebra Appl.*, to appear.

[43] V.V. Sergeĭchuk. Classification problems for systems of forms and linear mappings. (Russian) *Izv. Akad. Nauk SSSR Ser. Mat.*, **51** (1987), 1170–1190, 1358; English translation: *Math. USSR-Izv.*, **31** (1988), 481–501.

[44] V.V. Sergeĭchuk. Classification of sesquilinear forms, pairs of Hermitian forms, self-conjugate and isometric operators over the division ring of quaternions. *Math. Notes* **49** (1991), 409–414.

[45] V.V. Sergeĭchuk. Classification theorems for systems of forms and linear mappings and canonical matrices for isometric operators. Submitted for publication.

[46] R.C. Thompson. Pencils of complex and real symmetric and skew matrices. *Linear Algebra and its Applications*, **147** (1991), 323–371.

[47] Z.-X. Wan. *Geometry of Matrices*. World Scientific Publishing, River Edge, NJ, 1996.

[48] N.A. Wiegmann. Some theorems on matrices with real quaternion entries. *Canadian J. of Math.*, **7** (1955), 191–201.

[49] L.A. Wolff. Similarity of matrices in which the elements are real quaternions. *Bull. Amer. Math. Soc.* **42** (1936), 737–743.

[50] F. Zhang. Quaternions and matrices of quaternions. *Linear Algebra Appl.* **251** (1997), 21–57.

[51] W. Zhang. Jordan canonical form of a partitioned complex matrix and its application to real quaternion matrices. *Comm. Algebra* **29** (2001), 2363–2375.

Leiba Rodman
Department of Mathematics
College of William and Mary
Williamsburg, VA 23187-8795, USA
e-mail: lxrodm@math.wm.edu

Operator Theory:
Advances and Applications, Vol. 176, 255–294

Algorithms to Solve Hierarchically Semi-separable Systems

Zhifeng Sheng, Patrick Dewilde and Shivkumar Chandrasekaran

Abstract. 'Hierarchical Semi-separable' matrices (HSS matrices) form an important class of structured matrices for which matrix transformation algorithms that are linear in the number of equations (and a function of other structural parameters) can be given. In particular, a system of linear equations $Ax = b$ can be solved with linear complexity in the size of the matrix, the overall complexity being linearly dependent on the defining data. Also, LU and ULV factorization can be executed 'efficiently', meaning with a complexity linear in the size of the matrix. This paper gives a survey of the main results, including a proof for the formulas for LU-factorization that were originally given in the thesis of Lyon [1], the derivation of an explicit algorithm for ULV factorization and related Moore-Penrose inversion, a complexity analysis and a short account of the connection between the HSS and the SSS (sequentially semi-separable) case. A direct consequence of the computational theory is that from a mathematical point of view the HSS structure is 'closed' for a number operations. The HSS complexity of a Moore-Penrose inverse equals the HSS complexity of the original, for a sum and a product of operators the HSS complexity is no more than the sum of the individual complexities.

Mathematics Subject Classification (2000). Primary 65F; Secondary 05

Keywords. Hierarchically Semi-separable system, structured matrices, direct solver, ULV factorization

1. Introduction

The term 'semi-separable systems' originated in the work of Gohberg, Kailath and Koltracht [2] where these authors remarked that if an integral kernel is approximated by an outer sum, then the system could be solved with a number of operations essentially determined by the order of the approximation rather than by a power of the number of input and output data. In the same period, Greengard and Rokhlin [3, 4] proposed the 'multipole method' where an integral kernel such as a Green's function is approximated by an outer product resulting in a matrix in which large sub-matrices have low rank. These two theories evolved in parallel in the system theoretical literature and the numerical literature. In the system theoretical literature it was realized that an extension of the semi-separable model (sometimes called ‘quasi-separability’) brings the theory into the realm of time-varying systems, with its rich theory of state realization, interpolation, model order reduction, factorization and embedding [5]. In particular, it was shown in [6] that, based on this theory, a numerically backward stable solver of low complexity can be derived realizing a URV factorization of an operator T, in which U and V are low unitary matrices of state dimensions at most as large as those of T and R is causal, outer and also of state dimensions at most equal those of T. Subsequently, this approach has been refined by a number of authors, a.o. [7, 8, 9].

Although the SSS theory leads to very satisfactory results when applicable, it also became apparent in the late nineties that it is insufficient to cover major physical situations in which it would be very helpful to have system solvers of low complexity – in wiew of the often very large size of the matrices involved. Is it possible to extend the framework of SSS systems so that its major properties remain valid, in particular the fact that the class is closed under system inversion? The HSS theory, pioneered by Chandrasekaran and Gu [10] provides an answer to this question. It is based on a different state space model than the SSS theory, namely a hierarchical rather than a sequential one, but it handles the transition operators very much in the same taste. Based on this, a theory that parallels the basic time-varying theory of [5] can be developed, and remarkably, many results carry over. In the remainder of this paper we recall and derive some major results concerning system inversion, and discuss some further perspectives. The remainder sections of this introduction are devoted to a brief summary of the construction of SSS systems which lay at the basis of the HSS theory. In the numerical literature, the efforts have been concentrated on ‘smooth’ matrices, i.e., matrices in which large sub matrices can be approximated by low rank matrices thanks to the fact that their entries are derived from smooth kernels [11, 12]. Both the SSS and HSS structures are more constrained than the ‘H-matrices’ considered by Hackbusch a.o., but they do have the desirable property that they are closed under inversion and fit naturally in a state space framework. In the sequel we explore in particular the state space structure of HSS systems, other structures such as hierarchical multi-band decomposition have also been considered [13] but are beyond the present scope.

Our basic context is that of block matrices or operators $T = [T_{i,j}]$ with rows of dimensions $\ldots, m_{-1}\ , m_0, m_1, \ldots$ and column dimensions $\ldots, n_{-1}, n_0, n_1, \ldots$. Any of these dimensions may be zero, resulting in an empty row or column (matrix calculus can easily be extended to cover this case, the main rule being that the product of a matrix of dimensions $m \times 0$ with a matrix of dimensions $0 \times n$ results in a zero matrix of dimensions $m \times n$). Concentrating on an upper block matrix (i.e., when $T_{i,j} = 0$ for $i > j$), we define the *the degree of semi-separability* of T as the sequence of ranks $[\delta_i]$ of the matrices H_i where H_i is the sub-matrix corresponding to the row indices $\ldots, n_{i-2}, n_{i-1}$ and the column indices $m_i, m_{i+1}, \ldots$. H_i is called the ith Hankel operator of the matrix T. In case of infinite-dimensional operators, we say that the system is *locally finite* if all H_i have finite dimensions. Corresponding to the local dimension δ_i there are minimal factorizations $H_i = \mathcal{C}_i\mathcal{O}_i$ into what are called the i^{th} *controllability matrix* $\mathcal{C}_i$ and *observability matrix* $\mathcal{O}_i$, of dimensions $(\sum_{k=i-1}^{-\infty} m_k) \times \delta_i$ and $\delta_i \times (\sum_{k=i}^{\infty} n_k)$. Connected to such a system of factorizations there is an indexed realization $\{A_i, B_i, C_i, D_i\}$ of dimensions $\{\delta_i \times \delta_{i+1}, m_i \times \delta_{i+1}, \delta_i \times n_i, m_i \times n_i\}$ constituting a local set of 'small' matrices with the characteristic property of semi-separable realizations for which it holds that

$$\begin{cases} \mathcal{C}_i = \begin{bmatrix} \vdots \\ B_{i-2}A_{i-1} \\ B_{i-1} \end{bmatrix} \quad , \quad \mathcal{O}_i = \begin{bmatrix} C_i & A_iC_{i+1} & A_iA_{i+1}C_{i+2} & \cdots \end{bmatrix} \\ T_{i,j} = D_i \quad \text{for} \quad i = j \\ T_{i,j} = B_iA_{i+1}\cdots A_{j-1}C_j \quad \text{for} \quad i < j. \end{cases} \tag{1.1}$$

The vector-matrix multiplication $y = uT$ can be represented by local state space computations

$$\begin{cases} x_{i+1} &= x_iA_i + u_iB_i \\ y_i &= x_iC_i + u_iD_i. \end{cases} \tag{1.2}$$

The goal of most semi-separable computational theory (as done in [5]) is to perform computations with a complexity linear in the overall dimensions of the matrix, and some function of the degree δ_i, preferably linear, but that is often not achievable (there is still quite some work to do on this topic even in the SSS theory!). The above briefly mentioned realization theory leads to nice representations of the original operator. To this end we only need to introduce a shift operator Z with the characteristic property $Z_{i,i+1} = I$, zero elsewhere, where the dimension of the unit matrix is context dependent, and global representations for the realization as block diagonal operators $\{A = \mathrm{diag}[A_i], B = \mathrm{diag}[B_i], C = \mathrm{diag}[C_i], D = \mathrm{diag}[D_i]\}$. The lower triangular part can of course be dealt with in the same manner as the upper, resulting in the general semi-separable representation of an operator as (the superscript 'H' indicates Hermitian conjugation)

$$T = B_\ell Z^H(I - A_\ell Z^H)^{-1}C_\ell + D + B_uZ(I - A_uZ)^{-1}C_u \tag{1.3}$$

in which the indices refer to the lower, respect. upper semi-separable decomposition. In general we assume that the inverses in this formula do exist and have reasonable bounds, if that is not the case one has to resort to different techniques that go beyond the present exposition. In the finite-dimensional case the matrix $(I-AZ)$ takes the special form when the indexing runs from 0 to n (for orientation the $0,0$ element is boxed in):

$$(I - AZ) = \begin{bmatrix} \boxed{I} & A_0 & & & \\ & I & A_1 & & \\ & & \ddots & \ddots & \\ & & & I & A_n \\ & & & & I \end{bmatrix} \tag{1.4}$$

one may think that this matrix is always invertible, but that is numerically not true, how to deal with numerical instability in this context is also still open territory.

The SSS theory (alias time-varying system theory) has produced many results paralleling the classical LTI theory and translating these results to a matrix context, (see [5] for a detailed account):

- *System inversion*: $T = URV$ in which the unitary matrices U, V and the outer matrix R (outer means: upper and upper invertible) are all semi-separable of degree at most the degree of T;
- *System approximation and model reduction*: sweeping generalizations of classical interpolation theory of the types Nevanlinna-Pick, Caratheodory-Fejer and even Schur-Takagi, resulting in a complete model reduction theory of the 'AAK-type' but now for operators and matrices;
- *Cholesky and spectral factorization*: $T = FF^*$ when T is a positive operator, in which F is semi-separable of the same degree sequence as T – a theory closely related to Kalman filtering;
- and many more results in embedding theory and minimal algebraic realization theory.

2. Hierarchical semi-separable systems

The Hierarchical Semi-Separable representation of a matrix (or operator) A is a layered representation of the multi-resolution type, indexed by the hierarchical level. At the top level 1, it is a 2×2 block matrix representation of the form (notice the redefinition of the symbol A):

$$A = \begin{bmatrix} A_{1;1,1} & A_{1;1,2} \\ A_{1;2,1} & A_{1;2,2} \end{bmatrix} \tag{2.1}$$

in which we implicitely assume that the ranks of the off-diagonal blocks are low so that they can be represented by an 'economical' factorization ('H' indicates Hermitian transposition, for real matrices just transposition), as follows:

$$A = \begin{bmatrix} D_{1;1} & U_{1;1}B_{1;1,2}V_{1;2}^H \\ U_{1;2}B_{1;2,1}V_{1;1}^H & D_{1;2} \end{bmatrix}. \tag{2.2}$$

The second hierarchical level is based on a further but similar decomposition of the diagonal blocks, respect. $D_{1;1}$ and $D_{1;2}$:

$$\begin{aligned} D_{1;1} &= \begin{bmatrix} D_{2:1} & U_{2;1}B_{2;1,2}V_{2;2}^H \\ U_{2;2}B_{2;2,1}V_{2;1}^H & D_{2;2} \end{bmatrix} \\ D_{1;2} &= \begin{bmatrix} D_{2;3} & U_{2;3}B_{2;3,4}V_{2;4}^H \\ U_{2;4}B_{2;4,3}V_{2;3}^H & D_{2;4} \end{bmatrix} \end{aligned} \tag{2.3}$$

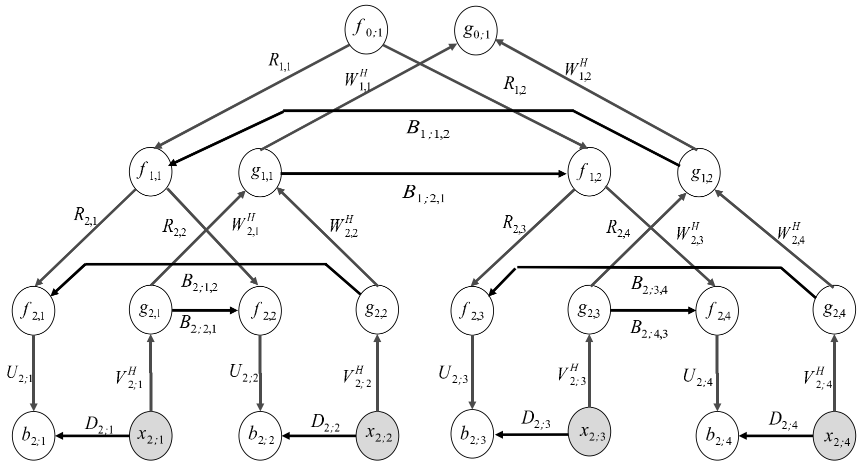

FIGURE 1. HSS Data-flow diagram for a two level hierarchy representing operator-vector multiplication, arrows indicate matrix-vector multiplication of sub-data, nodes correspond to states and are summing incoming data (the top levels f_0 and g_0 are empty).

for which we have the further *level compatibility* assumption (the 'span operator' refers to the column vectors of the subsequent matrix)

$$\operatorname{span}(U_{1;1}) \subset \operatorname{span}\left(\begin{bmatrix} U_{2;1} \\ 0 \end{bmatrix}\right) \oplus \operatorname{span}\left(\begin{bmatrix} 0 \\ U_{2;2} \end{bmatrix}\right), \tag{2.4}$$

$$\operatorname{span}(V_{1;1}) \subset \operatorname{span}\left(\begin{bmatrix} V_{2;1} \\ 0 \end{bmatrix}\right) \oplus \operatorname{span}\left(\begin{bmatrix} 0 \\ V_{2;2} \end{bmatrix}\right) \text{ etc.} \tag{2.5}$$

This spanning property is characteristic for the HSS structure, it allows for a substantial improvement on the numerical complexity for, e.g., matrix-vector multiplication as a multiplication with the higher level structures always can be done using lower level operations, using the *translation operators*

$$U_{1;i} = \begin{bmatrix} U_{2;2i-1}R_{2;2i-1} \\ U_{2;2i}R_{2;2i} \end{bmatrix}, \quad i = 1, 2, \tag{2.6}$$

$$V_{1;i} = \begin{bmatrix} V_{2;2i-1}W_{2;2i-1} \\ V_{2;2i}W_{2;2i} \end{bmatrix}, \quad i = 1, 2. \tag{2.7}$$

Notice the use of indices: at a given level i rows respect. columns are subdivided in blocks indexed by $1, \ldots, i$. Hence the ordered index $(i; k, \ell)$ indicates a block at level i in the position (k, ℓ) in the original matrix. The same kind of subdivision can be used for column vectors, row vectors and bases thereof (as are generally represented in the matrices U and V).

In [14] it is shown how this multilevel structure leads to efficient matrix-vector multiplication and a set of equations that can be solved efficiently as well. For the sake of completeness we review this result briefly. Let us assume that we want to solve the system $Tx = b$ and that T has an HSS representation with deepest hierarchical level K. We begin by accounting for the matrix-vector multiplication Tx. At the leave node $(K;i)$ we can compute

$$g_{K;i} = V_{K;i}^H x_{K;i}.$$

If $(k;i)$ is not a leaf node, we can infer, using the hierarchical relations

$$g_{k;i} = V_{k;i}^H x_{k;i} = W_{k+1;2i-1}^H g_{k+1;2i-1} + W_{k+1;2i}^H g_{k+1;2i}.$$

These operations update a 'hierarchical state' $g_{k;i}$ upward in the tree. To compute the result of the multiplication, a new collection of state variables $\{f_{k;i}\}$ is introduced for which it holds that

$$b_{k;i} = T_{k;i,i} + U_{k;i} f_{k;i}$$

and which can now be computed recursively downward by the equations

$$\begin{bmatrix} f_{k+1;2i-1} \\ f_{k+1;2i} \end{bmatrix} = \begin{bmatrix} B_{k+1;2i-1,2i} g_{k+1;2i} + R_{k+1;2i-1} f_{k,i} \\ B_{k+1;2i,2i-1} g_{k+1;2i-1} + R_{k+1;2i} f_{k;i} \end{bmatrix},$$

the starting point being $f_{0;} = []$, an empty matrix. At the leaf level we can now compute (at least in principle – as we do not know x) the outputs from

$$b_{K;i} = D_{K;i} x_{K;i} + U_{K;i} f_{K;i}.$$

The next step is to represent the multiplication recursions in a compact form using matrix notation and without indices. We fix the maximum order K as before. We define diagonal matrices containing the numerical information, in breadth first order:

$$\mathbf{D} = \mathrm{diag}[D_{K;i}]_{i=1,\ldots,K},\ \mathbf{W} = \mathrm{diag}[(W_{1;i})_{i=1,2}, (W_{2;i})_{i=1\cdots 4}, \ldots],\ \text{etc.}$$

Next, we need two shift operators relevant for the present situation, much as the shift operator Z in time-varying system theory explained above. The first one is the shift-down operator $Z_{\downarrow}$ on a tree. It maps a node in the tree on its children and is a nilpotent operator. The other one is the level exchange operator $Z_{\leftrightarrow}$. At each level it is a permutation that exchanges children of the same node. Finally, we need the leaf projection operator $\mathbf{P}_{\text{leaf}}$ which on a state vector which assembles in breadth first order all the values $f_{k;i}$ produces the values of the leaf nodes (again in breadth first order). The state equations representing the efficient multiplication can now be written as

$$\begin{cases} \mathbf{g} = \mathbf{P}_{\text{leaf}}^H \mathbf{V}^H \mathbf{x} + Z_{\downarrow}^H \mathbf{W}^H \mathbf{g} \\ \mathbf{f} = \mathbf{R} Z_{\downarrow} \mathbf{f} + \mathbf{B} Z_{\leftrightarrow} \mathbf{g} \end{cases} \tag{2.8}$$

while the 'output' equation is given by

$$\mathbf{b} = \mathbf{D}\mathbf{x} + \mathbf{U}\mathbf{P}_{\text{leaf}}\mathbf{f}\ . \tag{2.9}$$

This is the resulting HSS state space representation that parallels the classical SSS state space formulation reviewed above. Written in terms of the hidden state space quantities we find

$$\begin{bmatrix} (I - Z_\downarrow^H \mathbf{W}^H) & 0 \\ -BZ_\leftrightarrow \mathbf{g} & (I - RZ_\downarrow) \end{bmatrix} \begin{bmatrix} \mathbf{g} \\ \mathbf{f} \end{bmatrix} = \begin{bmatrix} \mathbf{P}_{\text{leaf}}^H \mathbf{V}^H \\ 0 \end{bmatrix} \mathbf{x}. \tag{2.10}$$

The state quantities can always be eliminated in the present context as $(I - \mathbf{W}Z_\downarrow)$ and $(I - \mathbf{R}Z_\downarrow)$ are invertible operators due to the fact that $Z_\downarrow$ is nilpotent. We obtain as a representation for the original operator

$$T\mathbf{x} = (\mathbf{D} + \mathbf{U}\mathbf{P}_{\text{leaf}}(I - \mathbf{R}Z_\downarrow)^{-1}\mathbf{B}Z_\leftrightarrow(I - Z_\downarrow^H \mathbf{W}^H)^{-1}\mathbf{P}_{\text{leaf}}^H \mathbf{V}^H)\mathbf{x} = \mathbf{b}. \tag{2.11}$$

3. Matrix operations based on HSS representation

In this section we describe a number of basic matrix operations based on the HSS representation. Matrix operations using the HSS representation are normally much more efficient than operations with plain matrices. Many matrix operations can be done with a computational complexity (or sequential order of basic operations) linear with the dimension of the matrix. These fast algorithms to be described are either collected from other publications [14, 10, 1, 15] or new. We will handle a somewhat informal notation to construct new block diagonals. Suppose, e.g., that $\mathbf{R}_A$ and $\mathbf{R}_B$ are conformal block diagonal matrices from the description given in the preceding section, then the construction operator inter$[\mathbf{R}_A|\mathbf{R}_B]$ will represent a diagonal operator in which the diagonal entries of the two constituents are block-column-wise intertwined:

$$\text{inter}[\mathbf{R}_A|\mathbf{R}_B] = \text{diag}\left[\begin{bmatrix} R_{A;1;1} & R_{B;1;1} \end{bmatrix}, \begin{bmatrix} R_{A;1;2} & R_{B;1;2} \end{bmatrix}, \begin{bmatrix} R_{A;2;1} & R_{B;2;1} \end{bmatrix}, \dots\right].$$

Block-row intertwining

$$\text{inter}[\mathbf{W}_A|\mathbf{W}_B] = \text{diag}\left[\begin{bmatrix} W_{A;1;1} \\ W_{B;1;1} \end{bmatrix}, \begin{bmatrix} W_{A;1;2} \\ W_{B;1;2} \end{bmatrix}, \begin{bmatrix} W_{A;2;1} \\ W_{B;2;1} \end{bmatrix}, \dots\right].$$

matrix intertwining is defined likewise.

3.1. HSS addition

Matrix addition can be done efficiently with HSS representations. The addition algorithm for Sequentially semi-separable representation has been presented in [16]. The addition algorithm for HSS representation which has been studied in [1] is quite similar.

3.1.1. Addition with two commensurately partitioned HSS matrices.

When adding two HSS commensurately partitioned matrices together, the sum will be an HSS matrix with the same partitioning. Let $C = A + B$ where A is defined by sequences U_A, V_A, D_A, R_A, W_A and B_A ; B is defined by sequences U_B, V_B, D_B, R_B, W_B and B_B.

$$\begin{cases} \mathbf{R}_C = \text{inter}\left[\begin{array}{c|c} \mathbf{R}_A & 0 \\ \hline 0 & \mathbf{R}_B \end{array}\right] \\ \mathbf{W}_C = \text{inter}\left[\begin{array}{c|c} \mathbf{W}_A & 0 \\ \hline 0 & \mathbf{W}_B \end{array}\right] \\ \mathbf{B}_C = \text{inter}\left[\begin{array}{c|c} \mathbf{B}_A & 0 \\ \hline 0 & \mathbf{B}_B \end{array}\right] \\ \mathbf{U}_C = \text{inter}\left[\begin{array}{c|c} \mathbf{U}_A & \mathbf{U}_B \end{array}\right] \\ \mathbf{V}_C = \text{inter}\left[\begin{array}{c|c} \mathbf{V}_A & \mathbf{V}_B \end{array}\right] \\ \mathbf{D}_C = \mathbf{D}_A + \mathbf{D}_B. \end{cases} \tag{3.1}$$

The addition can be done in time proportional to the number of entries in the representation. Note that the computed representation of the sum may not be efficient, in the sense that the HSS complexity of the sum increases additively. It is quite possible that the HSS representation is not minimal as well, as is the case when $A = B$. In order to get an efficient HSS representation, we could do *fast model reduction* (described in [15]) or *compression* (to be presented later) on the resulting HSS representation. However, these operations might be too costly to be applied frequently, one could do *model reduction* or *compression* after a number of additions.

3.1.2. Adaptive HSS addition. When two HSS matrices of the same dimensions do not have the same depth, leaf-split or leaf-merge operations described in [15] are needed to make these two HSS representations compatible. Note that we have two choices: we can either split the shallower HSS tree to make it compatible with the deeper one, or we can do leaf-merge on the deeper tree to make it compatible with shallower one. From the point of view of computation complexity, leaf-merge is almost always preferred since it amounts to several matrix multiplications with small matrices (ideally); leaf-split needs several factorization operations which are more costly than matrix multiplications. However, this does not imply leaf-merge should always be used if possible. Keeping in mind the fact that the efficiency of the HSS representation also comes from a deeper HSS tree with smaller translation matrices, the HSS tree should be kept deep enough to capture the low rank off-diagonal blocks. On the other hand, it is obviously impossible to always apply leaf-merge or leaf-split, because one HSS tree may have both a deeper branch and a shallower one than the other HSS tree does.

3.1.3. HSS addition with rank-m matrices. The sum of a level-n hierarchically semi-separable matrix A and a rank-m matrix UBV^H is another level-n hierarchically semi-separable matrix $A' = A + UBV^H$. A rank-m matrix has an almost trivial HSS representation conformal to any hierarchical scheme. With such a representation the HSS addition described in Section 3.1.1 is applicable.

In order to add two matrices together, the rank-m matrix should be represented in a form compatible with the HSS matrix. That is, the rank-m matrix will have to be partitioned recursively according to the partitioning of the HSS matrix A.

Let us first denote U as $U_{0;1}$, V as $V_{0;1}$, UBV^H as $D_{0;1}$. We partition U and V according to the partition of matrix A as follows:
for $k = 0, 1, 2, \ldots, n$ and $i \in 1, 2, \ldots, 2^k$:

$$U_{k;i} = \begin{bmatrix} U_{k+1;2i-1} \\ U_{k+1;2i} \end{bmatrix} \quad V_{k;i} = \begin{bmatrix} V_{k+1;2i-1} \\ V_{k+1;2i} \end{bmatrix}$$

Then at the first level of the partition:

$$U_{0;1}BV_{0;1}^H = \begin{bmatrix} U_{1;1}BV_{1;1}^H & U_{1;1}BV_{1;2}^H \\ U_{1;2}BV_{1;1}^H & U_{1;2}BV_{1;2}^H \end{bmatrix}$$

and following levels are given by:

Theorem 1. *The level-n HSS representation of the rank-m matrix UBV^H is: for $k = 1, 2, \ldots, n$; $i \in 1, 2, \ldots, 2^k$ and $\langle i \rangle = i + 1$ for odd i, $\langle i \rangle = i - 1$ for even i:*

$$\begin{cases} \widehat{D}_{k;i} = U_{k;i}BV_{k;i}^H & \widehat{R}_{k;i} = I \\ \widehat{W}_{k;i} = I & \widehat{B}_{k;i,\langle i \rangle} = B \\ \widehat{U}_{k;i} = U_{k;i} & \widehat{V}_{k;i} = V_{k;i} \end{cases} \tag{3.2}$$

$D_{k;i}$ are again rank-m matrices, assuming recursive correctness of this constructive method, $D_{k;i}$ can also be partitioned and represented recursively.

Other ways of constructing HSS representations for rank-m matrices are possible. One is to firstly form a one-level HSS representation for the rank-m matrix and then use the leaf-split algorithm [15] to compute its HSS representation according to certain partitioning. In principle, this method leads to an efficient HSS tree in the sense that its column bases and row bases are irredundant. However, this method needs much more computations. If m is reasonably small, the method described in this section is recommended.

3.1.4. HSS addition with rank-m matrices with hierarchically semi-separable bases.

In HSS representations, the column bases and row bases of the HSS nodes are not explicitly stored. This means when we compute $\widehat{A} = A + UBV^H$, U and V are probably not explicitly stored, instead, they are implicitly stored with the formulas (2.6) and (2.7).

We can of course compute these row bases and column bases and then construct an HSS representation for UBV^H with the method described in the last subsection. This is not recommended because computing U and V may be costly and not memory efficient.

Theorem 2. *Suppose U and V are defined in HSS form, the HSS representation of UBV^H is given by the following formulas:*

for $k = 2, 3, \ldots, n$; $i \in 1, 2, \ldots, 2^k$; *and* $\langle i \rangle = i + 1$ *for odd* i, $\langle i \rangle = i - 1$ *for even* i:

$$\begin{cases} \widehat{W}_{1;1} = I & \widehat{W}_{1;2} = I & \widehat{R}_{1;1} = I \\ \widehat{R}_{1;2} = I & \widehat{B}_{1;1,2} = B & \widehat{B}_{1;2,1} = B \\ \widehat{W}_{k;i} = W_{k-1;\lceil \frac{i}{2} \rceil} & \widehat{R}_{k;i} = R_{k-1;\lceil \frac{i}{2} \rceil} & \widehat{B}_{k;i,\langle i \rangle} = R_{k-1;\lceil \frac{i}{2} \rceil} \widehat{B}_{k-1;\lceil \frac{i}{2} \rceil, \langle \lceil \frac{i}{2} \rceil \rangle} W^H_{k-1;\lceil \frac{i}{2} \rceil} \\ \widehat{U}_{n;i} = U_{n;i} R_{n;i} & \widehat{V}_{n;i} = V_{n;i} W_{n;i} & \widehat{D}_{n;i} = U_{n;i} \widehat{B}_{n;\lceil \frac{i}{2} \rceil, \langle \lceil \frac{i}{2} \rceil \rangle} V^H_{n;i}. \end{cases} \tag{3.3}$$

After having the HSS representation of UBV^H, the sum can be computed easily using the HSS addition algorithm described in Section 3.1.1.

3.2. HSS matrix-matrix multiplication

Matrix-matrix multiplication can also be done in time linear with the dimensions of the matrices. The product $C = AB$ is another hierarchically semi-separable matrix.

A is a HSS matrix whose HSS representation is defined by sequences U_A, V_A, D_A, R_A, W_A, and B_A.

B is a HSS matrix whose HSS representation is defined by sequences U_B, V_B, D_B, R_B, W_B, and B_B.

3.2.1. Multiplication of two commensurately partitioned HSS matrices. When two HSS matrices are compatible, that is, they are commensurately partitioned, we can get the HSS representation of the product with the following algorithm. The algorithm was originally given with proof in Lyon's thesis [1].

The notations F and G to be used in following paragraphs represent the intermediate variables representing intermediate states in computing the HSS representation of C. They can be computed using the recursive formulas (3.4) to (3.7).

$F_{k;2i-1}$ represents the F intermediate variable propagated to the left children; similarly, $F_{k;2i}$ represents the intermediate F propagated to the right children. $G_{k;2i-1}$ represents the intermediate variable G coming from the left children; while $G_{k;2i}$ represents the intermediate variable G coming from the right ones. At last, $G_{n;i}$ represents the variable G calculated at leaves.

We first define the intermediate variables recursively via:

Definition 1. *For the multiplication of two level-n HSS matrices the upsweep recursion is defined as:*
for $i \in 1, 2, \ldots, 2^n$:

$$G_{n;i} = V^H_{A;n;i} U_{B;n;i} \tag{3.4}$$

for $k = n, \ldots, 2, 1$ *and* $i \in 1, 2, \ldots, 2^k$:

$$G_{k-1;i} = W^H_{A;k;2i-1} G_{k;2i-1} R_{B;k;2i-1} + W^H_{A;k;2i} G_{k;2i} R_{B;k;2i}. \tag{3.5}$$

Definition 2. *For the multiplication of two level-n HSS matrices the downsweep recursion is defined as:*
for $(i,j) = (1,2)$ *or* $(2,1)$*:*

$$F_{1;i} = B_{A;1;i,j} G_{1;j} B_{B;j,i} \tag{3.6}$$

for $i \in 1,2,\ldots,2^k$, $j = i+1$ *for odd* i, $j = i-1$ *for even* i *and* $k = 2,\ldots,n$*:*

$$F_{k;i} = B_{A;k;i,j} G_{k;j} B_{B;k;j,i} + R_{A;k;i} F_{k-1;\lceil \frac{i}{2} \rceil} W^H_{B;k,i}. \tag{3.7}$$

Theorem 3. *The HSS representation of the product is:*
for $i \in 1,2,\ldots,2^n$

$$\begin{cases} \widehat{D}_{n;i} = D_{A;n;i} D_{B;n;i} + U_{A;n;i} F_{n;i} V^H_{B;n;i} \\ \widehat{U}_{n;i} = \left[\begin{array}{c|c} U_{A;n;i} & D_{A;n;i} U_{B;n;i} \end{array}\right] \\ \widehat{V} = \left[\begin{array}{c|c} D^H_{B;n;i} V_{A;n;i} & V_{B;n;i} \end{array}\right] \end{cases} \tag{3.8}$$

for $k = 1,2,\ldots,n$*;* $i \in 1,2,\ldots,2^k$ *and* $j = i+1$ *for odd* i, $j = i-1$ *for even* i*:*

$$\begin{cases} \widehat{R}_{k;i} = \left[\begin{array}{c|c} R_{A;k;i} & B_{A;k;i,j} G_{k;j} R_{B;k;j} \\ \hline 0 & R_{B;k;i} \end{array}\right] \\ \widehat{W}_{k;i} = \left[\begin{array}{c|c} W_{A;k;i} & 0 \\ \hline B^H_{B;j,i} G^H_{k;j} W_{A;k;j} & W_{B;k;i} \end{array}\right] \\ \widehat{B}_{k;i,j} = \left[\begin{array}{c|c} B_{A;k;i,j} & R_{A;k;i} F_{k-1;\lceil \frac{i}{2} \rceil} W^H_{B;k;j} \\ \hline 0 & B_{B;k;i,j} \end{array}\right]. \end{cases} \tag{3.9}$$

Once again, the complexity of the HSS representation increases additively. *Model reduction* or *compression* may be needed to bring down the complexity. Note that, the algorithm above is given without proof. For a detailed proof and analysis, we refer to [1].

3.2.2. Adaptive HSS Matrix-Matrix multiplication. Adaptive multiplication is needed when two HSS matrices are not completely compatible, then leaf-split and leaf-merge are needed to make them compatible. The comment given in Section (3.1.2) for adaptive addition also applies here.

3.2.3. HSS Matrix-Matrix multiplication with rank-m matrices. A is a level-n HSS matrix whose HSS representation is defined by sequences U_A, V_A, D_A, R_A, W_A, and B_A. UBV^H is a rank-m matrix. The product $AUBV^H$ is also a level-n HSS matrix.

As we mentioned in Section 3.1.3, a rank-m matrix is a hierarchically semi-separable matrix and can be represented with a HSS representation. Then we can easily construct the HSS representation for the rank-m matrix and then perform the HSS Matrix-Matrix multiplication. This is the most straightforward way. However, making use of the fact that the translation matrices (R, W) of the rank-m matrix are identity matrices, the Matrix-Matrix multiplication algorithm can be simplified by substituting the R_B and W_B matrices in Section 3.2.1 with I matrices.

Again, because the complexity has been increased additively, *compression* or *Model reduction* could be helpful.

3.3. HSS matrix transpose

The transpose of a level-n HSS matrix will again be a level-n HSS matrix. Suppose the HSS matrix A is given by sequences B, R, W, U, V, D; it is quite easy to verify that

Theorem 4. *the HSS representation of the transpose A^H is given by the sequences: for $k = 1, 2, \ldots, n$; $i \in 1, 2, \ldots, 2^k$ and $j = i+1$ for odd i, $j = i-1$ for even i:*

$$\begin{cases} \widehat{D}_{k;i} = D^H_{k;i} & \widehat{U}_{k;i} = V_{k;i} \quad \widehat{V}_{k;i} = U_{k;i} \\ \widehat{W}_{k;i} = R_{k;i} & \widehat{R}_{k;i} = W_{k;i} \quad \widehat{B}_{k;i,j} = B^H_{k;j,i}. \end{cases} \tag{3.10}$$

3.4. Generic inversion based on the state space representation

A state space representation for the inverse with the same state complexity can generically be given. We assume the existence of the inverse, the same hierarchical partitioning of the input and output vectors $\mathbf{x}$ and $\mathbf{b}$, and as generic conditions the invertibility of the direct operators $\mathbf{D}$ and $S = (I + \mathbf{B}Z_{\leftrightarrow}\mathbf{P}^H_{\text{leaf}}\mathbf{V}^H\mathbf{D}^{-1}\mathbf{U}\mathbf{P}_{\text{leaf}})$, the latter being a (very) sparse perturbation of the unit operator with a local (that is leaf based) inversion operator. Let $\mathbf{L} = \mathbf{P}^H_{\text{leaf}}\mathbf{V}^H\mathbf{D}^{-1}\mathbf{U}\mathbf{P}_{\text{leaf}}$, then we find

Theorem 5. *Under generic conditions, the inverse system T^{-1} has the following state space representation*

$$\begin{aligned} \begin{bmatrix} \mathbf{g} \\ \mathbf{f} \end{bmatrix} &= \left\{ \begin{bmatrix} I & \\ & 0 \end{bmatrix} - \begin{bmatrix} \mathbf{L} \\ -I \end{bmatrix} S^{-1} \begin{bmatrix} \mathbf{B}Z_{\leftrightarrow} & I \end{bmatrix} \right\} \\ &\cdot \left\{ \begin{bmatrix} Z^H_{\downarrow}\mathbf{W}^H & \\ & \mathbf{R}Z_{\downarrow} \end{bmatrix} \begin{bmatrix} \mathbf{g} \\ \mathbf{f} \end{bmatrix} + \begin{bmatrix} \mathbf{P}^H_{\text{leaf}}\mathbf{V}^H\mathbf{D}^{-1}\mathbf{b} \\ 0 \end{bmatrix} \right\} \end{aligned} \tag{3.11}$$

and the output equation

$$\mathbf{x} = -\mathbf{D}^{-1}\mathbf{U}\mathbf{P}_{\text{leaf}}\mathbf{f} + \mathbf{D}^{-1}\mathbf{b}. \tag{3.12}$$

The proof of the theorem follows from inversion of the output equation which involves the invertibility of the operator $\mathbf{D}$, and replacing the unknown $\mathbf{x}$ in the state equations, followed by a segregation of the terms that are directly dependent on the states and those that are dependent on the shifted states leading to the matrix $\begin{bmatrix} I & \mathbf{L} \\ -\mathbf{B}Z_{\leftrightarrow} & I \end{bmatrix}$ whose inverse is easily computed as the first factor in the right-hand side of the equation above. It should be remarked that this factor only involves operations at the leaf level of the hierarchy tree so that the given state space model can be efficiently executed (actually the inversion can be done using the original hierarchy tree much as is the case for the inversion of upper SSS systems).

Having the theorem, we can derive a closed formula for T^{-1} assuming the generic invertiblity conditions.

$$\begin{aligned} T^{-1} &= \mathbf{D}^{-1} - \mathbf{D}^{-1}\mathbf{U}\mathbf{P}_{\text{leaf}} \cdot \\ &\cdot \left[I - \mathbf{R}Z_{\downarrow} + \mathbf{B}Z_{\leftrightarrow}(I - Z_{\downarrow}^{H}\mathbf{W}^{H})^{-1}\mathbf{P}_{\text{leaf}}^{H}\mathbf{D}^{-1}\mathbf{U}\mathbf{P}_{\text{leaf}} \right]^{-1} \cdot \\ &. \mathbf{B}Z_{\leftrightarrow}(I - Z_{\downarrow}^{H}\mathbf{W}^{H})^{-1}\mathbf{P}_{\text{leaf}}^{H}\mathbf{V}^{H}\mathbf{D}^{-1}. \end{aligned} \tag{3.13}$$

The equation given is a compact diagonal representation of T^{-1}, it also proves that the inverse of an invertible HSS matrix is again a HSS matrix of comparable complexity.

3.5. LU decomposition of HSS matrix

The formulas to compute the L and U factors of a square invertible matrix $T = LU$ in HSS form were originally given without proof in the thesis of Lyon [1] (they were checked computationally and evaluated in the thesis). Here we reproduce the formulas and give proof. The assumptions needed for the existence of the factorization are the same as is in the non-hierarchical case: a hierarchical tree that is n deep, the 2^n (block-) pivots have to be invertible.

The 'generic' situation (which occurs at each level in the HSS LU factorization) is a specialization of the classical Schur inversion theorem as follows:
we are given a matrix with the following 'generic' block decomposition

$$T = \begin{bmatrix} D_A & U_1 B_{12} V_2^H \\ U_2 B_{21} V_1^H & D_B \end{bmatrix} \tag{3.14}$$

in which D_A is a square invertible matrix, D_B is square (but not necessarily invertible), and T is invertible as well. Suppose we dispose of an LU factorization of the 11-block entry $D_A = L_A U_A$ and let us define two new quantities (which in the further proceedings will acquire an important meaning)

$$G_1 = V_1^H D_A^{-1} U_1, \quad F_2 = B_{21} G_1 B_{12}. \tag{3.15}$$

Then the first block step in a LU factorizaton of T is given by

$$T = \begin{bmatrix} L_A & \\ U_2 B_{21} V_1^H U_A^{-1} & I \end{bmatrix} \begin{bmatrix} I & \\ & D_B - U_2 F_2 V_2^H \end{bmatrix} \begin{bmatrix} U_A & L_A^{-1} U_1 B_{12} V_2^H \\ & I \end{bmatrix}. \tag{3.16}$$

The block entry $D_B - U_2 F_2 V_2^H$ is the classical 'Schur-complement' of D_A in the given matrix and it will be invertible if the matrix T is, as we assumed. At this point the first block column of the 'L' factor and the first block row of the 'U' matrix are known (the remainder will follow from an LU-decomposition of the Schur complement $D_B - U_2 F_2 V_2^H$). We see that the 21-entry in L and the 12-entry in U inherit the low rank of the originals with the same U_2, respect. V_2^H entry. In fact, more is true, the hierarchical relations in the first block column of L, respect. block row of U remain valid because $L_A = D_A U_A^{-1}$, respect. $U_A = L_A^{-1} D_A$, with modified row basis, respect. column basis. In the actual HSS computation the

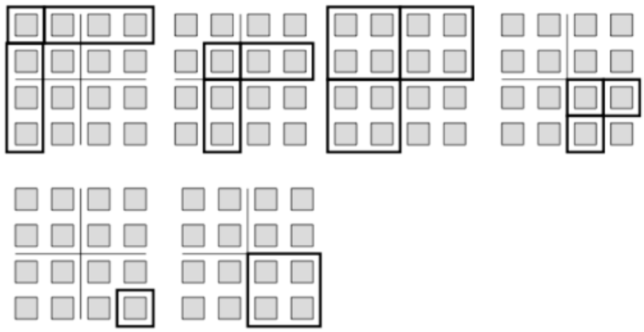

Figure 2. Recursive positioning of the LU first blocks in the HSS post-ordered LU factorization.

Schur complement will not be computed directly – it is lazily evaluated in what is called 'post-order traverse', meaning that each node (k,i) is evaluated only after evaluation of nodes $(k,\ell), \ell < k$ at the same level and its sub nodes $(k+1, 2i-1)$ and $(k+1, 2i)$.

This basic step can be interpreted as a specialization of the LU-factorization algorithm for sequentially separable systems, which reduces here to just two steps. In the first step the F_1 matrix is empty, the LU-factorization of $D_A = L_A U_A$ is done and the $V_1^{LH} = V_1^H U_A^{-1}$, respect. $U_1^U = L_A^{-1} U_1$ are computed. In the second step (in this case there are only two steps), G_1 is computed as $G_1 = V_1^{LH} U_1^U$, with $F_2 = B_{21} G_1 B_{12}$ and finally the Schur complement $D_B - U_2 F_2 V_2^H$ is evaluated (the sequential algorithm would be more complicated if more terms are available).

The HSS LU factorization is executed lazily in post-order traverse (w.r. to the hierarchical ordering of the blocks in the matrix), whereby previously obtained results are used as much as possible. For a tree that is 2 levels deep it goes as in Figure 2.

The collection of information needed to update the Schur complement at each stage of the algorithm is accomplished by an 'upward' movement, represented by the G matrices. Once a certain node (k,i) is reached, the $G_{k,i}$ equals the actual $V_1^H D_A^{-1} U_1$ pertaining to that node and hence subsumes all the data that is needed from previous steps to update the remaining Schur complement. However, the next 'lazy' step in the evaluation does not involve the whole matrix D_B, but only the at that point relevant top left corner matrix, the next candidate for reduction in the ongoing elimination – and determination of the next pivot. This restriction to the relevant top entry is accomplished by the matrix F, which takes information from the G's that are relevant at that level and specializes them to compute the contributions to the Schur-complement update of that specific matrix. Before formulating the algorithm precisely, we make this strategy that leads to efficient computations more precise.

Definition 3. *G propagates the quantity $V_1^H D_A^{-1} U_1$.*

Definition 4. *F propagates the quantity $B_{21} V_1^H D_A^{-1} U_1 B_{12}$ in equation* (3.16).

Updating G. The update situation involves exclusively the upward collection of the $G_{k,i}$. We assume that at some point in the recursion the matrices $G_{k,2i-1}$ and $G_{k,2i}$ are known, the objective is to compute $G_{k-1,i}$. The relevant observation here is that only this recursive data and data from the original matrix are needed to achieve the result. In matrix terms the situation is as follows:

$$\left[\begin{array}{cc|c} D_\ell & U_\ell \widehat{B}_u V_r^H & U_\ell R_\ell [\cdots] \\ U_r \widehat{B}_\ell V_\ell^H & D_r & U_r R_r [\cdots] \\ \hline [\cdots] W_\ell V_\ell^H & [\cdots] W_r^H V_r^H & D_B \end{array}\right] \tag{3.17}$$

where $\widehat{B}_u$ stands for $B_{k;2i-1,2i}$, $\widehat{B}_\ell$ stands for $B_{k;2i,2i-1}$, the subscript 'ℓ' stands for the left branch in the hierarchy for which $G_\ell = G_{k,2i-1} = V_\ell^H D_\ell^{-1} U_\ell$ is known, while the subscript 'r' stands for the right branch, for which $G_r = G_{k,2i} = V_r^H C_r^{-1} U_r$ is known with $C_r = D_r - U_r \widehat{B}_\ell V_\ell^H D_\ell^{-1} U_\ell \widehat{B}_u V_r^H$ the Schur complement of the first block in the left top corner submatrix, the objective being to compute $G = G_{k-1,i}$ given by

$$G = V^H D^{-1} U = \left[\begin{array}{cc} W_\ell^H V_\ell^H & W_r^H V_r^H \end{array}\right] \left[\begin{array}{cc} D_\ell & U_\ell \widehat{B}_u V_r^H \\ U_r \widehat{B}_\ell V_\ell^H & D_r \end{array}\right]^{-1} \left[\begin{array}{c} U_\ell R_\ell \\ U_r R_r \end{array}\right] \tag{3.18}$$

(note that the entries indicated by '$[\cdots]$' in equation (3.17) are irrelevant for this computation, they are taken care of in the F-downdate explained further on, while the $\widehat{B}_u$ and $\widehat{B}_\ell$ subsume the B-data at this level, which are also not relevant at this point of the computation). Computing the inverse of the new Schur complement produces:

$$G = \left[W_\ell^H V_\ell^H \quad W_r^H V_r^H\right] \cdot$$
$$\cdot \begin{bmatrix} D_\ell^{-1} + D_\ell^{-1} U_\ell \widehat{B}_u V_r^H C_r^{-1} U_r \widehat{B}_\ell V_\ell^H D_\ell^{-1} & -D_\ell^{-1} U_\ell \widehat{B}_u V_r^H C_r^{-1} \\ -C_r^{-1} U_r \widehat{B}_\ell V_\ell^H D_\ell^{-1} & C_r^{-1} \end{bmatrix} \begin{bmatrix} U_\ell R_\ell \\ U_r R_r \end{bmatrix}$$
$$G = \left[W_\ell^H \quad W_r^H\right] \cdot$$
$$\cdot \begin{bmatrix} V_\ell^H D_\ell^{-1} U_\ell + V_\ell^H D_\ell^{-1} U_\ell \widehat{B}_u V_r^H C_r^{-1} U_r \widehat{B}_\ell V_\ell^H D_\ell^{-1} U_\ell & -V_\ell^H D_\ell^{-1} U_\ell \widehat{B}_u V_r^H C_r^{-1} U_r \\ -V_r^H C_r^{-1} U_r \widehat{B}_\ell V_\ell^H D_\ell^{-1} U_\ell & V_r^H C_r^{-1} U_r \end{bmatrix} \cdot$$
$$\cdot \begin{bmatrix} R_\ell \\ R_r \end{bmatrix}$$

where $G_\ell = V_\ell^H D_\ell^{-1} U_\ell$ and $G_r = V_r^H C_r^{-1} U_r$ have been introduced. Hence

$$G = \left[\begin{array}{cc} W_\ell^H & W_r^H \end{array}\right] \left[\begin{array}{cc} G_\ell + G_\ell \widehat{B}_u G_r \widehat{B}_\ell G_\ell & -G_\ell \widehat{B}_u G_r \\ -G_r \widehat{B}_\ell G_\ell & G_r \end{array}\right] \left[\begin{array}{c} R_\ell \\ R_r \end{array}\right]. \tag{3.19}$$

Downdating F. The downdate situation can be subsumed as follows. We assume that we have arrived at a stage where the LU factorization has progressed just beyond the (hierarchical) diagonal block D_ℓ in the original matrix, the last block for which the Schur complement data G_ℓ has been updated. The hierarchical diagonal block preceding D_ℓ is subsumed as D_A, covering all the indices preceding those of D_ℓ. For this block, the corresponding G_A is also assumed to be known – these are the recursive assumptions. Let us assume moreover that the next (hierarchical) block to be processed in the post-order is D_r. The relations in the off diagonal entries, using higher level indices as needed are given in the matrix
Let us denote:

$$\left[\begin{array}{cc|cc} D_A & U_A B_u W_\ell^H V_\ell^H & U_A B_u W_r^H V_r^H & \cdots \\ U_\ell R_\ell B_\ell V_A^H & D_\ell & U_\ell B_u' V_r^H & \cdots \\ \hline U_r R_r B_\ell V_A^H & U_r B_\ell' V_\ell^H & D_r & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{array}\right] = \left[\begin{array}{c|cc} A_{11} & A_{12} & \cdots \\ \hline A_{21} & A_{22} & \cdots \\ \vdots & \vdots & \ddots \end{array}\right].$$

The recursive assumptions, expressed in the data of this matrix are the knowledge of $G_A = V_A^H D_A^{-1} V_A^H$ and $G_\ell = V_\ell^H C_\ell^{-1} U_\ell$ in which C_ℓ is the Schur complement of D_A for the diagonal block D_ℓ. then

$$\begin{aligned} A_{21} A_{11}^{-1} A_{12} = {} & \begin{bmatrix} U_r R_r B_\ell V_A^H & U_r B_\ell V_\ell^H \end{bmatrix} \cdot \\ & \cdot \begin{bmatrix} D_A & U_A B_u W_\ell^H V_\ell^H \\ U_\ell R_\ell B_\ell V_A^H & D_\ell \end{bmatrix}^{-1} \begin{bmatrix} U_A B_u W_r^H V_r^H \\ U_\ell B_u V_r^H \end{bmatrix}. \end{aligned} \tag{3.20}$$

With the definition of F and the Schur inverse algorithm, we can rewrite the above formula as:

$$\begin{aligned} & A_{21} A_{11}^{-1} A_{12} = U_r F_r V_r^H \\ & = \begin{bmatrix} U_r R_r B_\ell V_A^H & U_r B_\ell V_\ell^H \end{bmatrix} \cdot \\ & \quad \cdot \begin{bmatrix} D_A^{-1} + D_A^{-1} U_A B_u W_\ell^H V_\ell^H C_\ell^{-1} U_\ell R_\ell B_\ell V_A^H D_A^{-1} & -D_A^{-1} U_A B_u W_\ell^H V_\ell^H C_\ell^{-1} \\ -C_\ell^{-1} U_\ell R_\ell B_\ell V_A^H D_A^{-1} & C_\ell^{-1} \end{bmatrix} \cdot \\ & \quad \cdot \begin{bmatrix} U_A B_u W_r^H V_r^H \\ U_\ell B_u V_r^H \end{bmatrix} \\ & = U_r \begin{bmatrix} R_r B_\ell V_A^H & B_\ell V_\ell^H \end{bmatrix} \cdot \\ & \quad \cdot \begin{bmatrix} D_A^{-1} + D_A^{-1} U_A B_u W_\ell^H V_\ell^H C_\ell^{-1} U_\ell R_\ell B_\ell V_A^H D_A^{-1} & -D_A^{-1} U_A B_u W_\ell^H V_\ell^H C_\ell^{-1} \\ -C_\ell^{-1} U_\ell R_\ell B_\ell V_A^H D_A^{-1} & C_\ell^{-1} \end{bmatrix} \cdot \\ & \quad \cdot \begin{bmatrix} U_A B_u W_r^H \\ U_\ell B_u \end{bmatrix} V_r^H \\ & = U_r \begin{bmatrix} R_r & B_\ell \end{bmatrix} \cdot \\ & \quad \cdot \begin{bmatrix} B_\ell V_A^H D_A^{-1} U_A B_u + Y & -B_\ell V_A^H D_A^{-1} U_A B_u W_\ell^H V_\ell^H C_\ell^{-1} U_\ell \\ -V_\ell^H C_\ell^{-1} U_\ell R_\ell B_\ell V_A^H D_A^{-1} U_A B_u & V_\ell^H C_\ell^{-1} U_\ell \end{bmatrix} \cdot \\ & \quad \cdot \begin{bmatrix} W_r^H \\ B_u \end{bmatrix} V_r^H \end{aligned}$$

where

$$Y = B_\ell V_A^H D_A^{-1} U_A B_u W_\ell^H V_\ell^H C_\ell^{-1} U_\ell R_\ell B_\ell V_A^H D_A^{-1} U_A B_u.$$

As defined, F_r should represent the above term (excluding U_r and V_r). Assuming G_ℓ and $F = F_A$ given, we find

$$G_\ell = V_\ell^H C_\ell^{-1} U_\ell, F = B_\ell V_A^H D_A^{-1} U_A B_u.$$

Finally the update formula for F_r becomes:

$$F_r = \begin{bmatrix} R_r & B_\ell \end{bmatrix} \begin{bmatrix} F + F W_\ell^H G_\ell R_\ell F & -F W_\ell^H G_\ell \\ -G_\ell R_\ell F & G_\ell \end{bmatrix} \begin{bmatrix} W_r^H \\ B_u \end{bmatrix}. \tag{3.21}$$

And F_r again satisfies the definition.

The update formula for F_ℓ can be easily derived from the definition of F. To preserve the definition of F on the left branch, the F from the parent has to be pre-multiplied with R_ℓ and post-multiplied with W_ℓ^H. Thus the update formulas for G and F have been explained and proven.

Modifying B matrices and computing block pivots. To compute the Schur complement $D_{k;i} - U_{k;i} F_{k;i} V_{k;i}^H$ efficiently, we only need to update the B matrices and modify the block pivots. Here we assume that we are moving one level up in the recursion and that the levels below have already been computed. Let

$$S_{k-1;i} = D_{k-1;i} - U_{k-1;i} F_{k-1;i} V_{k-1;i}^H$$

$$S_{k-1;i} = \begin{bmatrix} D_{k;2i-1} & U_{k;2i-1} B_{k;2i-1,2i} V_{k;2i}^H \\ U_{k;2i} B_{k;2i,2i-1} V_{k;2i-1}^H & D_{k;2i} \end{bmatrix} - \begin{bmatrix} U_{k;2i-1} R_{k;2i-1} \\ U_{k;2i} R_{k;2i} \end{bmatrix} F_{k-1;i} \begin{bmatrix} W_{k;2i-1}^H V_{k;2i-1}^H & W_{k;2i}^H V_{k;2i}^H \end{bmatrix}$$

$$S_{k-1;i} = \begin{bmatrix} D_{k;2i-1} & U_{k;2i-1} B_{k;2i-1,2i} V_{k;2i}^H \\ U_{k;2i} B_{k;2i,2i-1} V_{k;2i-1}^H & D_{k;2i} \end{bmatrix} - \begin{bmatrix} U_{k;2i-1} R_{k;2i-1} F_{k-1;i} W_{k;2i-1}^H V_{k;2i-1}^H & U_{k;2i-1} R_{k;2i-1} F_{k-1;i} W_{k;2i}^H V_{k;2i}^H \\ U_{k;2i} R_{k;2i} F_{k-1;i} W_{k;2i-1}^H V_{k;2i-1}^H & U_{k;2i} R_{k;2i} F_{k-1;i} W_{k;2i}^H V_{k;2i}^H \end{bmatrix}$$

$$S_{k-1;i} = \begin{bmatrix} \bar{D}_{k;2i-1} & Y_{k;2i-1,2i} \\ Y_{k;2i,2i-1} & \bar{D}_{k;2i} \end{bmatrix}$$

where

$$Y_{k;2i-1,2i} = U_{k;2i-1}(B_{k;2i-1,2i} - R_{k;2i-1} F_{k-1;i} W_{k;2i}^H) V_{k;2i}^H = U_{k;2i-1} \widehat{B}_{k;2i-1,2i} V_{k;2i}^H$$

$$Y_{k;2i,2i-1} = U_{k;2i}(B_{k;2i,2i-1} - R_{k;2i} F_{k-1;i} W_{k;2i-1}^H) V_{k;2i-1}^H = U_{k;2i} \widehat{B}_{k;2i,2i-1} V_{k;2i-1}^H.$$

Hence

$$\bar{D}_{k;i} = D_{k;i} - U_{k;i} F_{k;i} V_{k;i}^H \tag{3.22}$$

$$\widehat{B}_{k;i,j} = B_{k;i,j} - R_{k;i} F_{k-1;\lceil \frac{i}{2} \rceil} W_{k;j}^H \tag{3.23}$$

and the F for the left branches:

$$F_{k;2i-1} = R_{k;2i-1} F_{k-1;i} W_{k;2i-1}^H. \tag{3.24}$$

Construction formulas for the L and the U matrices. We are now ready to formulate the LU-factorization relations and procedure.

Theorem 6. *Let a level-n HSS matrix T be given by the sequences R, W, B, U, V and D and assume that the pivot condition for existence of the LU-factorization is satisfied. Then the following relations hold:*
for $i \in 1, 2, \ldots, 2^n$:

$$G_{n;i} = V_{n;i}^H (D_{n;i} - U_{n;i} F_{n;i} V_{n;i}^H)^{-1} U_{n;i} \tag{3.25}$$

for $k = 1, 2, \ldots, n$; $i \in 1, 2, \ldots, 2^k$ and $j = i+1$ for odd i, $j = i-1$ for even i, let

$$\widehat{B}_{k;i,j} = B_{k;i,j} - R_{k;i} F_{k-1;\lceil \frac{i}{2} \rceil} W_{k;j}^H \tag{3.26}$$

for $k = 1, 2, \ldots, n$ and $i \in 1, 2, \ldots, 2^{k-1}$:

$$\begin{aligned} &G_{k-1;i} = \\ &\begin{bmatrix} W_{k;2i-1}^H & W_{k;2i}^H \end{bmatrix} \cdot \\ &\cdot \begin{bmatrix} G_{k;2i-1} + G_{k;2i-1}\widehat{B}_{k;2i-1,2i} G_{k;2i} \widehat{B}_{k;2i,2i-1} G_{k;2i-1} & -G_{k;2i-1}\widehat{B}_{k;2i-1,2i} G_{k;2i} \\ -G_{k;2i}\widehat{B}_{k;2i,2i-1} G_{k;2i-1} & G_{k;2i} \end{bmatrix} \cdot \\ &\cdot \begin{bmatrix} R_{k;2i-1} \\ R_{2i} \end{bmatrix}. \end{aligned} \tag{3.27}$$

Initial value for F is:

$$F_{0;1} = \phi \tag{3.28}$$

left branches F_ℓ are given as:
for $k = 1, 2, \ldots, n$ and $i \in 1, 2, \ldots, 2^{k-1}$:

$$F_{k;2i-1} = R_{k;2i-1} F_{k-1;i} W_{k;2i-1}^H \tag{3.29}$$

right branches F_r are given as:
for $k = 1, 2, \ldots, n$ and $i \in 1, 2, \ldots, 2^{k-1}$:

$$\begin{aligned} F_{k;2i} \;=\; &\begin{bmatrix} R_{k;2i} & B_{k;2i,2i-1} \end{bmatrix} \cdot \\ &\cdot \begin{bmatrix} F_{k-1;i} + F_{k-1;i} W_{k;2i-1}^H G_{k;2i-1} R_{k;2i-1} F_{k-1;i} & -F_{k-1;i} W_{k;2i-1}^H G_{k;2i-1} \\ -G_{k;2i-1} R_{k;2i-1} F_{k;i} & G_{k;2i-1} \end{bmatrix} \cdot \\ &\cdot \begin{bmatrix} W_{k;2i}^H \\ B_{k;2i-1,2i} \end{bmatrix}. \end{aligned} \tag{3.30}$$

The (block) pivots are given by

$$\bar{D}_{n;i} = D_{n;i} - U_{n;i} F_{n;i} V_{n;i}^H. \tag{3.31}$$

Let now the pivots be LU-factored (these are elementary blocks that are not further decomposed): for $i \in 1, 2, \ldots, 2^n$:

$$\bar{L}_{n;i} \bar{U}_{n;i} = \bar{D}_{n;i} = D_{n;i} - U_{n;i} F_{n;i} V_{n;i}^H \tag{3.32}$$

be a LU decomposition at each leaf. Then based on the information generated, the L and U factors are defined as follows:

Theorem 7. *The level-n HSS representation of the L factor will be given as:*
at a non-leaf node:
for $k = 1, 2, \ldots, n$; $i \in 1, 2, \ldots, 2^{k-1}$ *and* $j = 1, 2, \ldots, 2^k$:

$$\begin{cases} \widehat{R}_{k;j} = R_{k;j} & \widehat{W}_{k;2i-1} = W_{k;2i-1} \\ \widehat{W}^H_{k;2i} = W^H_{k;2i} - W^H_{k;2i-1} G_{k;2i-1} B_{k;2i-1,2i} & \\ \widehat{B}_{k;2i,2i-1} = \widehat{B}_{k;2i,2i-1} & \widehat{B}_{k;2i-1,2i} = 0 \end{cases} \tag{3.33}$$

at a leaf:
for $i \in 1, 2, \ldots, 2^n$:

$$\widehat{U}_{n;i} = U_{n;i} \quad \widehat{V}_{n;i} = \bar{U}^{-H}_{n;i} V_{n;i} \quad \widehat{D} = \bar{L}_{n;i}. \tag{3.34}$$

Theorem 8. *The level-n HSS representation of the U factor will be given as:*
at a non-leaf node:
for $k = 1, 2, \ldots, n$; $i \in 1, 2, \ldots, 2^{k-1}$ *and* $j = 1, 2, \ldots, 2^k$:

$$\begin{cases} \widehat{R}_{k;2i-1} = R_{k;2i-1} & \widehat{R}_{k;2i} = R_{k;2i} - \widehat{B}_{k;2i,2i-1} G_{k;2i-1} R_{k;2i-1} & \widehat{W}_{k;j} = W_{k;j} \\ \widehat{B}_{k;2i,2i-1} = 0 & \widehat{B}_{k;2i-1,2i} = \widehat{B}_{k;2i-1,2i} & \end{cases} \tag{3.35}$$

at a leaf:
for $i \in 1, 2, \ldots, 2^n$:

$$\widehat{U}_{n;i} = \bar{L}^{-1}_{n;i} U_{n;i} \quad \widehat{V}_{n;i} = V_{n;i} \quad \widehat{D}_{n;i} = \bar{U}_{n;i}. \tag{3.36}$$

Proof for the traverse. We start with the proof of theorem 6. Given the updating operations on G and downdating operations on F accounted for in the introductory part of this section, it remains to verify that there exists a recursive order to compute all the quantities indicated. Initialization results in the $F_{k,1} = \phi$ for all $k = 1 \cdots n$. In particular, $F_{n,1}$ is now known, and $G_{n,1}$ can be computed. This in turn allows for the computation of $F_{n,2}$ thanks to the F_r downdate formula at level $(k-1, 1)$. Now $G_{n,2}$ can be computed, and next $G_{n-1,1}$ – the first left bottom node is dealt with. We now dispose of enough information to compute $F_{n-1,2}$, since $G_{n-1,1}$ and $F_{n-2,1} = \phi$ are known (this being the beginning of the next step).

The general dependencies in the formulas are as follows. At a leaf: $G_{n;i}$ depends on $F_{n;i}$; at a non-leaf node: $G_{k-1,i}$ is dependent on $G_{k,2i-1}$ and $G_{k,2i}$; $F_{k;2i-1}$ is dependent on $F_{k-1,i}$ and $F_{k,2i}$ is dependent on both $F_{k-1,i}$ and $G_{k,2i-1}$. Considering the closure of data dependencies, the full dependencies at a node are given in Figure 3. With the F matrices on the root initialized, the order in which all the F and G quantities can be computed on a node is $F_{k-1;i} \rightarrow F_{k;2i-1} \rightarrow G_{k;2i-1} \rightarrow F_{k;2i} \rightarrow G_{k;2i} \rightarrow G_{k-1;i}$, or equivalently *parent→left children→right children→parent*. That is: with a post-order traverse on the binary tree (note that: the F on the root is initialized), all unknown Fs and Gs can be filled in.

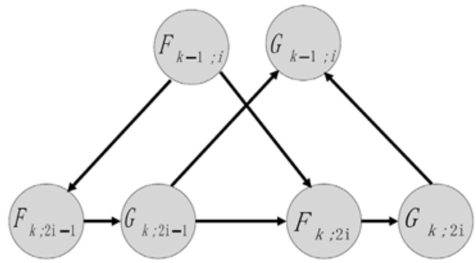

FIGURE 3. The dependencies of the intermediate variables on one no-leaf node.

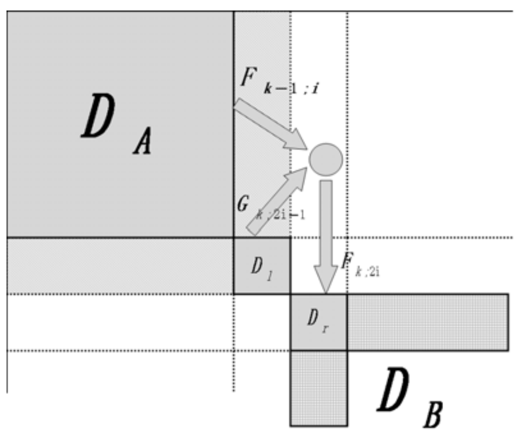

FIGURE 4. The computation of $F_{k;2i}$ with the help of $F_{k-1;i}$ and $G_{k;2i-1}$.

Proof of the formulas for L and U factors. Let now the pivots be LU-factored (these are elementary blocks that are not further decomposed). We may assume that at each step the Schur complements have been computed and updated. To get the $\mathbf{L}$ and $\mathbf{U}$ factors recursively as in formula (3.16), it is obvious that for each leaf of the L factor, $\bar{D} = \widehat{\mathbf{L}}$, $\bar{U} = U$, $\bar{V}^H = V_l^H \widehat{\mathbf{U}}^{-1}$; for each leaf of the U factor, $\bar{D} = \widehat{\mathbf{U}}$, $\bar{U} = \widehat{\mathbf{L}}^{-1} U$, $\bar{V} = V$.

For all left branches, the blocks are updated by modifying B matrices with formula (3.26) to compute the Schur complement $\bar{D}_{k;i} = D_{k;i} - U_{n;i} F_{n;i} V_{n;i}^H$. But for the right branches, updating B matrices with formula (3.26) is not enough because $F_{k-1;i}$ only subsumes the information from its parent. Its left sibling has to be taken into consideration for the update of the Schur complement.

Assuming the correct update has been done for the $\mathbf{D}_A$ block and D_ℓ block (see Figure 4), we may also assume that the Schur complement of D_ℓ has been computed. Hence, we only need to update D_r and the blocks indicated by grids in

Figure 4. That is for the block

$$\begin{bmatrix} \widehat{D}_\ell & U_\ell \widehat{B}_u V_r^H & U_\ell R_\ell B_u \mathbf{V}_B^H & \cdots \\ U_r \widehat{B}_\ell V_\ell^H & D_r & U_r R_r B_u \mathbf{V}_B^H & \cdots \\ \mathbf{U}_B B_\ell W_\ell^H V_\ell^H & \mathbf{U}_B B_\ell W_r^H V_r^H & \mathbf{D}_B & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Hence, only the blocks $\begin{bmatrix} U_r R_r B_u \mathbf{V}_B^H & \cdots \end{bmatrix}$ and $\begin{bmatrix} \mathbf{U}_B B_\ell W_r^H V_r^H \\ \vdots \end{bmatrix}$ have to be updated, other parts of the computation are taken care of by the recursive algorithm. Now, the Schur complement of $\widehat{D}_\ell$ has to be determined. That is:

$$S = \begin{bmatrix} D_r & U_r R_r B_u \mathbf{V}_B^H & \cdots \\ \mathbf{U}_B B_\ell W_r^H V_r^H & \mathbf{D}_B & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} - \begin{bmatrix} U_r \widehat{B}_\ell V_\ell^H \\ \mathbf{U}_B B_\ell W_\ell^H V_\ell^H \\ \vdots \end{bmatrix} \widehat{D}_\ell^{-1} \begin{bmatrix} U_\ell \widehat{B}_u V_r^H & U_\ell R_\ell B_u \mathbf{V}_B^H & \cdots \end{bmatrix}$$

$$S = \begin{bmatrix} D_r - U_r \widehat{B}_\ell V_\ell^H \widehat{D}_\ell^{-1} U_\ell \widehat{B}_u V_r^H & U_r (R_r - \widehat{B}_\ell V_\ell^H D_\ell^{-1} U_\ell R_\ell) B_u \mathbf{V}_B^H & \cdots \\ \mathbf{U}_B B_\ell (W_r^H - W_\ell^H V_\ell^H D_\ell^{-1} U_\ell \widehat{B}_u) V_r^H & \mathbf{D}_B - \mathbf{U}_B B_\ell W_\ell^H V_\ell^H D_\ell^{-1} U_\ell R_\ell B_u \mathbf{V}_B^H & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

Since $G_\ell = V_\ell^H D_\ell^{-1} U_\ell$,

$$S = \begin{bmatrix} D_r - U_r \widehat{B}_l V_\ell^H \widehat{D}_\ell^{-1} U_\ell \widehat{B}_u V_r^H & U_r (R_r - \widehat{B}_\ell G_\ell R_\ell) B_u \mathbf{V}_B^H & \cdots \\ \mathbf{U}_B B_\ell (W_r^H - W_\ell^H G_\ell \widehat{B}_u) V_r^H & \mathbf{D}_B - \mathbf{U}_B B_\ell W_\ell^H V_\ell^H D_\ell^{-1} U_\ell R_\ell B_u \mathbf{V}_B^H & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}.$$

Hence the update of the blocks $\begin{bmatrix} U_r R_r B_u \mathbf{V}_B^H \\ \vdots \end{bmatrix}$ and $\begin{bmatrix} \mathbf{U}_B^H B_\ell W_r^H V_r^H & \cdots \end{bmatrix}$ is given by $\widehat{R}_r = R_r - \widehat{B}_\ell G_\ell R_\ell$ and $\widehat{W}_r^H = W_r^H - W_\ell^H G_\ell \widehat{B}_u$. These prove the update formulas for $\widehat{R}_r$ and $\widehat{W}_r$.

Finally, all the update formulas have been explained, and the whole algorithm consists in recursively applying these formulas which actually compute and update the Schur complement recursively. This will be possible iff the pivot condition is satisfied.

4. Explicit ULV factorization

The LU factorization, however important, has only limited applicability. A backward stable algorithm that can always be applied is 'ULV-factorization'. It factors an arbitrary matrix in three factors, a unitary matrix U, a (generalized) lower triangular L (a non-singular triangular matrix embedded in a possibly larger zero matrix) and another unitary matrix V. In the present section we show that the ULV-factorization for an HSS matrix of order n can be obtained in a special form. Both U and V are again HSS, and the lower triangular factor L has a special HSS form that is extremely sparse (many transfer matrices are zero). The ULV-factorization of A leads directly to the Moore-Penrose inverse for A. One trims the L factor to its purely triangular part, and the U and V factors to the corresponding relevant columns and rows to obtain the so called 'economic ULV factorization' $A = U_e L_e V_e$, the Moore-Penrose inverse then being given as $A^\dagger = V_e^H L_e^{-1} U_e^H$. The determination of the inverse of a lower triangular HSS factor is treated in the following section and gives rise to an HSS matrix of the same order and complexity. In this paper we follow the implicit ULV factorization method presented in [17], and show that the implicit method can be made explicit with some non-trivial modifications. The Moore-Penrose system can be then be solved with the explicit L factor. Alternatively one could follow the method presented in [18] which has similar flavors, but uses a slightly different approach.

For the sake of definiteness and without impairing generality, we assume here that the HSS matrix A has full row rank, and its n-level HSS representation is defined by the sequences U, V, D, B, R, W. Similar to the implicit ULV factorization method, the explicit method involves an upsweep recursion (or equivalently a post-order traverse). We start with the left-most leaf. First, we treat the case in which the HSS representation, which will be recursively reduced, has reached the situation given in equation (4.1). The second block row in that equation has a central purely triangular block $A_{k;i}$ of dimension $\delta_{k;i}$, the goal will be to reduce the matrix further by treating the next block row. Through the steps described in the following treatment this situation will be reached recursively by converting subtrees to leaves, so that the central compression step always happens at the level of a leaf.

4.1. Treatment of a leaf

The situation to be treated in this part of the recursion has the form

$$A = \begin{bmatrix} \ddots & [\vdots]V_{k;i}^{(1)H} & [\vdots]V_{k;i}^{(2)H} & \iddots \\ 0 & A_{k;i} & 0 & 0 \\ U_{k;i}[\cdots] & D_{k;i}^{(1)} & D_{k;i}^{(2)} & U_{k;i}[\cdots] \\ \iddots & [\vdots]V_{k;i}^{(1)H} & [\vdots]V_{k;i}^{(2)H} & \ddots \end{bmatrix}. \tag{4.1}$$

It is assumed at this point that $A_{k;i}$ is already lower triangular and invertible with dimension $\delta_{k;i}$. The next block row stands in line for treatment. The compression step attacks $U_{k;i}$. If $U_{k;i}$ has more rows than columns, it can be compressed by applying QL factorization on it:

$$U_{k;i} = Q_{k;i} \begin{bmatrix} 0 \\ \widehat{U}_{k;i} \end{bmatrix} \tag{4.2}$$

where $\widehat{U}_{k;i}$ is square and has l rows. To keep consistence in the rows, we must apply $Q_{k;i}^H$ to $D_{k;i}$:

$$\widehat{D}_{k;i} = Q_{k;i}^H D_{k;i}. \tag{4.3}$$

Assume that $D_{k;i}$ has m columns. We can partition $D_{k;i}$ as:

$$\begin{array}{c} \begin{array}{cc} \delta_{k;i} & m-\delta_{k;i} \end{array} \\ \begin{bmatrix} \widehat{D}_{k;i}^{(1)} & \widehat{D}_{k;i}^{(2)} \end{bmatrix} \end{array} = \widehat{D}_{k;i} \tag{4.4}$$

Since $A_{k;i}$ is already lower-triangular matrix, to proceed we only have to process the block $\widehat{D}_{k;i}^{(2)}$ so as to obtain a larger upper-triangular reduced block. Hence we LQ factorize $\widehat{D}_{k;i}^{(2)}$ as:

$$\widehat{D}_{k;i}^{(2)} = \begin{bmatrix} \widehat{D}_{k;i;0,0}^{(2)} & 0 \\ \widehat{D}_{k;i;1,0}^{(2)} & \widehat{D}_{k;i;1,1}^{(2)} \end{bmatrix} w_{k;i}, \tag{4.5}$$

where $\widehat{D}_{k;i;0,0}^{(2)}$ is lower triangular and has n columns;$\widehat{D}_{k;i;1,0}^{(2)}$ and $\widehat{D}_{k;i;1,1}^{(2)}$ have l rows. Now to adjust the columns, we must apply $w_{k;i}$ on $V_{k;i}$. Let

$$\begin{array}{c} \delta_{k;i} \\ m-\delta_{k;i} \end{array} \begin{bmatrix} V_{k;i}^{(1)} \\ V_{k;i}^{(2)} \end{bmatrix} = V_{k;i}. \tag{4.6}$$

Apply $w_{k;i}$ on $V_{k;i}^{(2)}$ as

$$\widehat{V}_{k;i}^{(2)} = w_{k;i} V_{k;i}^{(2)} \tag{4.7}$$

let:

$$\begin{bmatrix} \widehat{D}_{k;i}^{(1,1)} \\ \widehat{D}_{k;i}^{(1,2)} \end{bmatrix} = \widehat{D}_{k;i}^{(1)}, \tag{4.8}$$

where $D_{k;i}^{(1,2)}$ has l rows. After these operations, the HSS representation has become

$$\bar{A} = \begin{bmatrix} \ddots & [\vdots]V_{k;i}^{(1)H} & [\vdots]\widehat{V}_{k;i}^{(21)H} & [\vdots]\widehat{V}_{k;i}^{(22)H} & \iddots \\ 0 & A_{k;i} & 0 & 0 & 0 \\ 0 & \widehat{D}_{k;i}^{(11)} & \widehat{D}_{k;i;0,0}^{(2)} & 0 & 0 \\ \widehat{U}_{k;i}[\cdots] & \widehat{D}_{k;i}^{(1,2)} & \widehat{D}_{k;i;1,0}^{(2)} & \widehat{D}_{k;i;1,1}^{(2)} & \widehat{U}_{k;i}[\cdots] \\ \iddots & [\vdots]V_{k;i}^{(1)H} & [\vdots]\widehat{V}_{k;i}^{(21)H} & [\vdots]\widehat{V}_{k;i}^{(22)H} & \ddots \end{bmatrix}. \tag{4.9}$$

The compressed leaf will be returned as:

$$\bar{D}_{k;i} = \begin{bmatrix} \widehat{D}_{k;i}^{(1,2)} & \widehat{D}_{k;i;1,0}^{(2)} & \widehat{D}_{k;i;1,1}^{(2)} \end{bmatrix} \tag{4.10}$$

$$\bar{U}_{k;i} = \widehat{U}_{k;i} \tag{4.11}$$

$$\bar{V}_{k;i} = \begin{bmatrix} V_{k;i}^{(1)} \\ \widehat{V}_{k;i}^{(2)} \end{bmatrix}. \tag{4.12}$$

With

$$\widehat{A}_{k;i} = \begin{bmatrix} A_{k;i} & 0 \\ \widehat{D}^{(1,1)}_{k;i} & \widehat{D}^{(2)}_{k;i;0,0} \end{bmatrix} \tag{4.13}$$

representing the reduced row slices, and

$$\widehat{\delta}_{k;i} = \delta_{k;i} + n. \tag{4.14}$$

Now, the commented HSS representation is exactly the same as the original, except the leaf has become smaller. When $U_{k;i}$ has more columns than rows, nothing can be done to compress in this way. Then a new arrangement has to be created by merging two leaves into a new, integrated leave. This process is treated in the next paragraph.

4.2. Merge

The behavior of this part of the algorithm on a leaf has been specified. If no leaf is available for processing, one can be created by merging. Assume that we are at the node $k;i$, the algorithm works in a post-order traverse way, it proceeds by first calling itself on the left children and then on the right children. When the algorithm comes to the present stage, both the left and the right child are already compressed leaves. They can then be merged by the following explicit procedure.

Before the merge, the HSS representation is, in an obvious notation: Let

$$\begin{aligned} Y^{(1)}_{k+1;2i-1;2i} &= U_{k+1;2i-1} B_{k+1;2i-1;2i} V^{(1)H}_{k+1;2i} \\ Y^{(2)}_{k+1;2i-1;2i} &= U_{k+1;2i-1} B_{k+1;2i-1;2i} V^{(2)H}_{k+1;2i} \\ Y^{(1)}_{k+1;2i;2i-1} &= U_{k+1;2i} B_{k+1;2i;2i-1} V^{(1)H}_{k+1;2i-1} \\ Y^{(2)}_{k+1;2i;2i-1} &= U_{k+1;2i} B_{k+1;2i;2i-1} V^{(2)H}_{k+1;2i-1} \end{aligned}$$

thus $D_{k;i}$ can be represented as:

$$D_{k;i} = \begin{bmatrix} A_{k+1;2i-1} & 0 & 0 & 0 \\ D^{(1)}_{k+1;2i-1} & D^{(2)}_{k+1;2i-1} & Y^{(1)}_{k+1;2i-1;2i} & Y^{(2)}_{k+1;2i-1;2i} \\ 0 & 0 & A_{k+1;2i} & 0 \\ Y^{(1)}_{k+1;2i;2i-1} & Y^{(2)}_{k+1;2i;2i-1} & D^{(1)}_{k+1;2i} & D^{(2)}_{k+1;2i} \end{bmatrix}. \tag{4.15}$$

Next, the rows and columns are moved to put all reduced rows on the top-left. After the reordering, the HSS representation becomes:

$$\widehat{D}_{k;i} = \begin{bmatrix} A_{k+1;2i-1} & 0 & 0 & 0 \\ 0 & A_{k+1;2i} & 0 & 0 \\ D^{(1)}_{k+1;2i-1} & Y^{(1)}_{k+1;2i-1;2i} & D^{(2)}_{k+1;2i-1} & Y^{(2)}_{k+1;2i-1;2i} \\ Y^{(1)}_{k+1;2i;2i-1} & D^{(1)}_{k+1;2i} & Y^{(2)}_{k+1;2i;2i-1} & D^{(2)}_{k+1;2i} \end{bmatrix} \tag{4.16}$$

and the merged leaf now has:

$$\bar{D}_{k;i} = \begin{bmatrix} D^{(1)}_{k+1;2i-1} & Y^{(1)}_{k+1;2i-1;2i} & D^{(2)}_{k+1;2i-1} & Y^{(2)}_{k+1;2i-1;2i} \\ Y^{(1)}_{k+1;2i;2i-1} & D^{(1)}_{k+1;2i} & Y^{(2)}_{k+1;2i;2i-1} & D^{(2)}_{k+1;2i} \end{bmatrix} \tag{4.17}$$

$$\bar{U}_{k;i} = \begin{bmatrix} U_{k+1;2i-1}R_{k+1;2i-1} \\ U_{k+1;2i}R_{k+1;2i} \end{bmatrix}, \bar{V}_{k;i} = \begin{bmatrix} V^{(1)}_{k+1;2i-1}W_{k+1;2i-1} \\ V^{(1)}_{k+1;2i}W_{k+1;2i} \\ V^{(2)}_{k+1;2i-1}W_{k+1;2i-1} \\ V^{(2)}_{k+1;2i}W_{k+1;2i-1} \end{bmatrix}. \tag{4.18}$$

With the intermediate block

$$A_{k;i} = \begin{bmatrix} A_{k+1;2i-1} & 0 \\ 0 & A_{k+1;2i} \end{bmatrix} \tag{4.19}$$

and

$$\delta_{k;i} = \delta_{k+1;2i-1} + \delta_{k+1;2i}. \tag{4.20}$$

Note that now the node has been reduced to a leaf, and the actual HSS system has two fewer leaves. The compression algorithm can then be called on this leaf with $A_{k;i}$ and $\delta_{k;i}$.

4.3. Formal algorithm

Having the above three procedures, we now describe the algorithm formally. Similar to the implicit ULV factorization method, this algorithm is a tree-based recursive algorithm. It involves a post-order traverse of the binary tree of the HSS representation. Let T be the root of the HSS representation.

```
Function: post-order-traverse
Input: an actual HSS node or leaf T;
Output: a compressed HSS leaf)
 1. (node, left-children, right-children) = T;
 2. left-leaf = post-order-traverse left-child;
 3. right-leaf = post-order-traverse right-child;
 4. if left-child is compressible then
    left-leaf = compress left-leaf;
    else
    do nothing;
 5. if right-child is compressible then
    right-leaf = compress right-leaf;
    else
    do nothing;
 6. return compress (Merge(node,left-leaf,right-leaf));
```

```
Function : Explicit-ULV-Factorization
Input: a HSS representation T; Output: the factor L in
sparse matrix format
 1. actual-T = T;
 2. Leaf = post-order-traverse actual-T;
 3. return Leaf.A_{0;1}
```

Once the whole HSS tree is compressed as a leaf and the leaf is further compressed, the L factor has been computed as $L = A_{0;1}$.

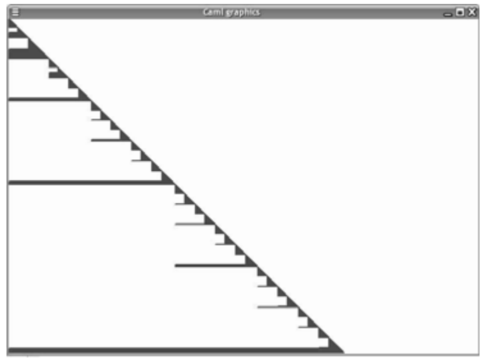

FIGURE 5. The sparsity pattern of L factor of the explicit ULV factorization.

4.4. Results

We show the result of the procedure applied to an HSS matrix A of dimensions 500×700 with full row rank. Its HSS representation is 5 levels deep and balanced. We apply the explicit ULV factorization algorithm on it. Then the sparsity pattern of the L factor will be as in Figure 5. L has 500 rows and 700 columns. Its sparsity is 3.08% (The sparsity depends on the HSS complexity, the lower the complexity is, the sparser the L factor is.). With the assumption that A has full row rank, The non-zero block of L is square and invertible.

4.5. Remarks

- Assume A has full column rank, the algorithm above can be modified to produce the URV factorization (by compressing $V_{k;i}$ instead of $U_{k;i}$).
- The explicit factor shall be kept in sparse matrix form.
- the U and V factors are kept in an implicit form. This is convenient because they can be easily applied to b and x when solve the system $Ax = b$.
- The complexity is higher than the implicit ULV factorization method, but it shall still be linear. It can also be easily be seen that the HSS complexity of the result is the same as of the original (with many transfer matrices reduced to zero).

5. Inverse of triangular HSS matrix

In this section, we will show how a triangular HSS matrix can be inverted efficiently. We shall only present our fast inverse algorithm on upper triangular HSS matrices, since the algorithm for lower triangular matrices is dual and similar. With the combination of the LU factorization algorithm, the inverse algorithm for triangular systems and the matrix-matrix multiplication algorithm, the HSS inverse of a square invertible HSS matrix, of which all block pivots are invertible, can be computed.

Let the level-n HSS representation of the upper triangular matrix A be given by the sequence of R, W, B, U, V and D (where the D's are upper triangular). Assuming all D matrices invertible, the level-n HSS representation of the inverse of A is given by $\widehat{R}$, $\widehat{W}$, $\widehat{B}$, $\widehat{U}$, $\widehat{V}$ and $\widehat{D}$ (where $\widehat{D}$s are again upper triangular) with the formulas given below. We use the following (trivial) fact recursively.

Lemma 1. *The inverse of $D_{k-1;\lceil\frac{i}{2}\rceil}$ (i is an odd number) is given by*

$$\begin{aligned} D^{-1}_{k-1;\lceil\frac{i}{2}\rceil} &= \begin{bmatrix} D_{k;i} & U_{k;i}B_{k;i,i+1}V^H_{k;i+1} \\ 0 & D_{k;i+1} \end{bmatrix}^{-1} \\ &= \begin{bmatrix} D^{-1}_{k;i} & -D^{-1}_{k;i}U_{k;i}B_{k;i,i+1}V^H_{k;i+1}D^{-1}_{k;i+1} \\ 0 & D^{-1}_{k;i+1} \end{bmatrix}. \end{aligned} \tag{5.1}$$

We have

$$\widehat{U}_{k;i} = \begin{bmatrix} D_{k+1;2i-1} & U_{k+1;2i-1}B_{k+1;2i-1,2i}V^H_{k+1;2i} \\ 0 & D_{k+1;2i} \end{bmatrix}^{-1} U_{k;i}$$

$$\widehat{U}_{k;i} = \begin{bmatrix} D_{k+1;2i-1} & U_{k+1;2i-1}B_{k+1;2i-1,2i}V^H_{k+1;2i} \\ 0 & D_{k+1;2i} \end{bmatrix}^{-1} \cdot \begin{bmatrix} U_{k+1;2i-1}R_{k+1;2i-1} \\ U_{k+1;2i}R_{k+1;2i} \end{bmatrix}$$

$$\widehat{U}_{k;i} = \begin{bmatrix} D^{-1}_{k+1;2i-1} & -D^{-1}_{k+1;2i-1}U_{k+1;2i-1}B_{k+1;2i-1,2i}V^H_{k+1;2i}D^{-1}_{k+1;2i} \\ 0 & D^{-1}_{k+1;2i} \end{bmatrix} \cdot \begin{bmatrix} U_{k+1;2i-1}R_{k+1;2i-1} \\ U_{k+1;2i}R_{k+1;2i} \end{bmatrix}$$

$$\widehat{U}_{k;i} = \begin{bmatrix} D^{-1}_{k+1;2i-1}U_{k+1;2i-1}(R_{k+1;2i-1} - B_{k+1;2i-1,2i}V^H_{k+1;2i}D^{-1}_{k+1;2i}U_{k+1;2i}R_{k+1;2i}) \\ D^{-1}_{k+1;2i}U_{k+1;2i}R_{k+1;2i} \end{bmatrix}.$$

Assuming that $\widehat{U}_{k+1;2i-1}$ and $\widehat{U}_{k+1;2i}$ have been updated as $D^{-1}_{k+1;2i-1}\ U_{k+1;2i-1}$ and $D^{-1}_{k+1;2i}U_{k+1;2i}$ the update for $\widehat{U}_{k;i}$ follows from the update $R_{k+1;2i-1}$ as

$$\widehat{R}_{k+1;2i-1} = R_{k+1;2i-1} - B_{k+1;2i-1,2i}V^H_{k+1;2i}\bar{U}^{-1}_{k+1;2i}U_{k+1;2i}R_{k+1;2i}. \tag{5.2}$$

The formulas for $\widehat{V}_{k;i+1}$ become

$$\begin{aligned} \widehat{V}^H_{k;i+1} &= V^H_{k;i+1}\begin{bmatrix} D_{k+1;2i+1} & U_{k+1;2i+1}B_{k+1;2i+1,2i+2}V^H_{k+1;2i+2} \\ 0 & D_{k+1;2i+2} \end{bmatrix}^{-1} \\ \widehat{V}^H_{k;i+1} &= \begin{bmatrix} W^H_{k+1;2i+1}V^H_{k+1;2i+1} & W^H_{k+1;2i+2}V^H_{k+1;2i+2} \end{bmatrix} \cdot \\ &\quad \cdot \begin{bmatrix} D_{k+1;2i+1} & U_{k+1;2i+1}B_{k+1;2i+1,2i+2}V^H_{k+1;2i+2} \\ 0 & D_{k+1;2i+2} \end{bmatrix}^{-1} \\ \widehat{V}^H_{k;i+1} &= \begin{bmatrix} W^H_{k+1;2i+1}V^H_{k+1;2i+1} & W^H_{k+1;2i+2}V^H_{k+1;2i+2} \end{bmatrix} \cdot \\ &\quad \cdot \begin{bmatrix} D^{-1}_{k+1;2i+1} & -D^{-1}_{k+1;2i+1}U_{k+1;2i+1}B_{k+1;2i+1,2i+2}V^H_{k+1;2i+2}D^{-1}_{k+1;2i+2} \\ 0 & D^{-1}_{k+1;2i+2} \end{bmatrix}. \end{aligned}$$

Let

$$\widehat{W}^H_{k+1;2i+2} = W^H_{k+1;2i+2} - W^H_{k+1;2i+1} V^H_{k+1;2i+1} D^{-1}_{k+1;2i+1} U_{k+1;2i+1} B_{k+1;2i+1,2i+2}$$

then

$$\widehat{V}^H_{k;i+1} = \begin{bmatrix} W^H_{k+1;2i+1} V^H_{k+1;2i+1} D^{-1}_{k+1;2i+1} & \widehat{W}^H_{k+1;2i+2} V^H_{k+1;2i+2} D^{-1}_{k+1;2i+2} \end{bmatrix}.$$

Assuming now that $\widehat{V}^H_{k+1;2i+1}$ and $\widehat{V}^H_{k+1;2i+2}$ have been updated as

$$V^H_{k+1;2i+1} D^{-1}_{k+1;2i+1} \qquad \text{and} \qquad V^H_{k+1;2i+2} D^{-1}_{k+1;2i+2},$$

the update for $\widehat{V}_{k;i+1}$ follows from

$$\widehat{W}^H_{k+1;2i+2} = W^H_{k+1;2i+2} - W^H_{k+1;2i+1} V^H_{k+1;2i+1} D^{-1}_{k+1;2i+1} U_{k+1;2i+1} B_{k+1;2i+1,2i+2} \tag{5.3}$$

next the update for $-\widehat{U}_{k;i} B_{k;i,j} \widehat{V}^H_{k;j}$ follows from

$$\widehat{B}_{k;i,j} = -B_{k;i,j}. \tag{5.4}$$

Let the intermediate G be defined as $G_{k;i} = V^H_{k;i} D^{-1}_{k;i} U_{k;i}$, then the above update formulas can be written as

$$\begin{cases} \widehat{W}^H_{k;2i} = W^H_{k;2i} - W^H_{k;2i-1} G_{k;2i-1} B_{k;2i-1,2i} \\ \widehat{W}_{k;2i-1} = W_{k;2i-1} \\ \widehat{R}_{k;2i-1} = R_{k;2i-1} - B_{k;2i-1,2i} G_{k;2i} R_{k;2i} \\ \widehat{R}_{k;2i} = R_{k;2i} \\ \widehat{B}_{k;i,j} = -B_{k;i,j}. \end{cases} \tag{5.5}$$

The recursive formula for the intermediate variable G are as follows. According to the definition of $G_{k-1;i}$:

$$\begin{aligned}
G_{k-1;i} = {} & \begin{bmatrix} W^H_{k;2i-1} V^H_{k;2i-1} & W^H_{k;2i-1} V^H_{k;2i} \end{bmatrix} \cdot \\
& \cdot \begin{bmatrix} D_{k;2i-1} & U_{k;2i-1} B_{k;2i-1,2i} V^H_{k;2i} \\ 0 & D_{k;2i} \end{bmatrix}^{-1} \begin{bmatrix} U_{k;2i-1} R_{k;2i-1} \\ U_{k;2i} R_{k;2i} \end{bmatrix} \\
G_{k-1;i} = {} & \begin{bmatrix} W^H_{k;2i-1} V^H_{k;2i-1} & W^H_{k;2i-1} V^H_{k;2i} \end{bmatrix} \cdot \\
& \cdot \begin{bmatrix} D^{-1}_{k;2i-1} & -D^{-1}_{k;2i-1} U_{k;2i-1} B_{k;2i-1,2i} V^H_{k;2i} D^{-1}_{k;2i} \\ 0 & D^{-1}_{k;2i} \end{bmatrix} \begin{bmatrix} U_{k;2i-1} R_{k;2i-1} \\ U_{k;2i} R_{k;2i} \end{bmatrix} \\
G_{k-1;i} = {} & \begin{bmatrix} W^H_{k;2i-1} & W^H_{k;2i-1} \end{bmatrix} \cdot \\
& \cdot \begin{bmatrix} V^H_{k;2i-1} D^{-1}_{k;2i-1} U_{k;2i-1} & -V^H_{k;2i-1} D^{-1}_{k;2i-1} U_{k;2i-1} B_{k;2i-1,2i} V^H_{k;2i} D^{-1}_{k;2i} U_{k;2i} \\ 0 & V^H_{k;2i} D^{-1}_{k;2i} U_{k;2i} \end{bmatrix} \cdot \\
& \cdot \begin{bmatrix} R_{k;2i-1} \\ R_{k;2i} \end{bmatrix} \\
G_{k-1;i} = {} & \begin{bmatrix} W^H_{k;2i-1} & W^H_{k;2i-1} \end{bmatrix} \begin{bmatrix} G_{k;2i-1} & -G_{k;2i-1} B_{k;2i-1,2i} G_{k;2i} \\ 0 & G_{k;2i} \end{bmatrix} \begin{bmatrix} R_{k;2i-1} \\ R_{k;2i} \end{bmatrix}.
\end{aligned}$$

Summarizing

Definition 5. *Let the intermediate variable G be defined as for $k = 1, 2, \ldots, n$ and $i \in 1, 2, \ldots, 2^{k-1}$:*

$$G_{k;i} = V_{k;i}^H D_{k;i}^{-1} U_{k;i}. \tag{5.6}$$

The upsweep recursion for G is:

$$G_{k-1;i} = \begin{bmatrix} W_{k;2i-1}^H & W_{k;2i}^H \end{bmatrix} \begin{bmatrix} G_{k;2i-1} & -G_{k;2i-1}B_{k;2i-1,2i}G_{k;2i} \\ 0 & G_{k;2i} \end{bmatrix} \begin{bmatrix} R_{k;2i-1} \\ R_{k;2i} \end{bmatrix} \tag{5.7}$$

and hence

Theorem 9. *The level-n HSS representation of the inverse of the upper triangular HSS matrix is given by the following sequence of operations for $k = 1, 2, \ldots, n$; $j \in 1, 2, \ldots, 2^k$ and $i \in 1, 2, \ldots, 2^{k-1}$:*

$$\begin{cases} \widehat{W}_{k;2i}^H = W_{k;2i}^H - W_{k;2i-1}^H G_{k;2i-1} B_{k;2i-1,2i} & \widehat{W}_{k;2i-1} = W_{k;2i-1} \\ \widehat{R}_{k;2i-1} = R_{k;2i-1} - B_{k;2i-1,2i} G_{k;2i} R_{k;2i} & \widehat{R}_{k;2i} = R_{k;2i} \\ \widehat{B}_{k;2i-1,2i} = -B_{k;2i-1,2i} & \widehat{B}_{k;2i,2i-1} = 0 \\ \widehat{U}_{k;j} = D_{k;j}^{-1} U_{k;j} & \widehat{V}_{k;j}^H = V_{k;j}^H D_{k;j}^{-1} \\ \widehat{D}_{k;j}^H = D_{k;j}^{-1}. & \end{cases} \tag{5.8}$$

6. Ancillary operations

In this section, we will discuss various ancillary operations that help to (re-) construct an HSS representation in various circumstances. These operations will help to reduce the HSS complexity or to keep the column base and row base dependencies of the HSS representation.

6.1. Column (row) base insertion

When the off-diagonal blocks have to be changed at the nodes at a higher level, column bases and row bases may have to be changed. To keep the column and row base dependencies, new column (row) bases may have to be added to the lower levels. We might be able to generate these bases from the column (row) bases of the lower level nodes, but this is not guaranteed. Taking a conservative approach we insert column (row) bases into the lower level and then do *compression* to reduce the complexity of HSS representation.

The algorithm combines two sub-algorithms (downsweep base insertion and then a compression). The *compression* procedure is used to eliminate redundant bases and reduce the HSS complexity. *Compression* does not have to be done after every downsweep column (row) base insertion. To save the computation cost, we may do one step of *compression* after a number of steps of bases insertion.

We will present row base insertion in details, while column base insertion is dual and hence similar.

6.1.1. Downsweep row base insertion. Suppose that we need to add a row base represented by a conformal matrix v to an HSS node A without changing the matrix it represents (the column dimension of v should of course be conformal to the row dimension of A.) Let the original HSS node be represented as

$$\begin{bmatrix} D_{1;1} & U_{1;1}B_{1;1,2}V_{1;2}^H \\ U_{1;2}B_{1;2,1}V_{1;1}^H & D_{1;2} \end{bmatrix}.$$

The algorithm works in a downsweep fashion modifying the nodes and leaves in the HSS tree.

- **Row base insertion at a non-leaf node**

  $v_{k;i}$ is split according to the column partition of A at this node: for $k = 1, 2, \ldots, n$ and $i \in 1, 2, \ldots, 2^k$:

$$v_{k;i} = \begin{bmatrix} v_{k+1;2i-1} \\ v_{k+1;2i} \end{bmatrix}$$

  $v_{k+1;2i-1}$ is inserted to the left child, $v_{k+1;2i}$ to the right child recursively. $v_{k+1;2i-1}$ can be generated from $D_{k+1;2i-1}$, and $v_{k+1;2i}$ from $D_{k+1;2i}$. The translation matrices of this node must be modified to make sure that the base insertion does not change the matrix it represents as follows

  for $k = 1, 2, \ldots, n$; $i \in 1, 2, \ldots, 2^k$, and $j = i + 1$ for odd i, $j = i - 1$ for even i:

$$\begin{cases} \widehat{W}_{k;i} = \begin{bmatrix} W_{k;i} & 0 \\ 0 & I \end{bmatrix} \\ \widehat{B}_{k;i,j} = \begin{bmatrix} B_{k;i,j} & 0 \end{bmatrix} \end{cases}. \tag{6.1}$$

- **Row base insertion at a leaf**

  a leaf is reached by recursion, $v_{n;i}$ has to be inserted to the leaf, hence for $k = 1, 2, \ldots, n$ and $i \in 1, 2, \ldots, 2^k$:

$$\widehat{V}_{n;i} = \begin{bmatrix} V_{n;i} & v_{n;i} \end{bmatrix}. \tag{6.2}$$

6.1.2. Compression. After applying downsweep base insertion to A, the row bases v required by the upper level can be generated from A. But the HSS representation we get may have become redundant.

Since only the row base has been modified, we only have to factor $V_{n;i}$ matrices as

for $k = 1, 2, \ldots, n$ and $i \in 1, 2, \ldots, 2^k$:

$$V_{n;i} = \widehat{V}_{n;i} w_{n;i}. \tag{6.3}$$

This should be done by a rank revealing QR or QL factorization, then $\widehat{V}_{n;i}$ will be column orthonormal (and it will surely be column independent). The factor $w_{n;i}$ will then be propagated to the upper level, where the translation matrices $B_{n;i,j}$ and $W_{n;i}$ will be modified by the factor w as follows for $k = 1, 2, \ldots, n$; $i \in 1, 2, \ldots, 2^k$, and $j = i+1$ for odd i, $j = i-1$ for even i:

$$\begin{cases} \widehat{B}_{k;i,j} = B_{k;i,j} w_{k;j}^H \\ \bar{W}_{k;i} = w_{k;i} W_{k;i} \\ \widehat{R}_{k;i} = R_{k;i}. \end{cases} \tag{6.4}$$

Then we do compression on the higher level. Since only row bases have been modified, we only have to factor

$$\begin{bmatrix} \bar{W}_{k;2i-1} \\ \bar{W}_{k;2i} \end{bmatrix} = \begin{bmatrix} \widehat{W}_{k;2i-1} \\ \widehat{W}_{k;2i} \end{bmatrix} w_{k-1;i}. \tag{6.5}$$

Note that $\bar{W}_{k;2i-1}$ and $\bar{W}_{k;2i}$ have been modified by the $w_{k;2i-1}$ and $w_{k;2i}$ factors coming from its children with the formulas (6.4), the factorization should again be rank-revealing. The $w_{k-1;i}$ factor will then be propagated further to the upper level, and the algorithm proceeds recursively.

After applying the whole algorithm to the a HSS node, the new row base v^H will be inserted by appending it to the original row base. Suppose the row base of the original node is given by V^H, the modified node becomes $\begin{bmatrix} V & v \end{bmatrix}^H$. Note that base insertion does not change the HSS matrix, it only modifies its HSS representation.

6.1.3. Column base insertion. The algorithm for column base insertion is similar and dual to the one for row base insertion. Modifications will now be done on U, R instead of V, W. And the B matrices will be modified as $\widehat{B}_{k;i,j} = \begin{bmatrix} B_{k;i,j} \\ 0 \end{bmatrix}$ instead of $\begin{bmatrix} B_{k;i,j} & 0 \end{bmatrix}$. After applying the row bases insertion to a HSS node, the new column bases will be appended after its original column bases. The *compression* algorithm for column base insertion should also be modified accordingly.

6.2. Append a matrix to a HSS matrix

This algorithm appends a thin slice C to a HSS matrix A. This operation is central in the Moore-Penrose HSS inversion treated in [18]. We establish that the result of this operation will still be HSS matrix whose HSS representation can be computed easily. Obviously, we may append the matrix to the top of the HSS matrix, to the left of the HSS matrix, to the right of the HSS matrix or to the bottom of the HSS matrix. Here we just present the method to append matrix to the left of the HSS matrix. Others can be easily derived mutatis mutandis.

6.2.1. Append a rank-k matrix to a HSS matrix. Suppose

$$\widehat{A} = \begin{bmatrix} C & A \end{bmatrix}. \tag{6.6}$$

Matrix B should have the same number of rows as A does, A is HSS matrix whose HSS representation is defined by sequences U_A, V_A, D_A, R_A, W_A and B_A.

Instead of trying to absorb the C matrix into the HSS representation of A matrix, we rewrite the formula (6.6) as:

$$\widehat{A} = \begin{bmatrix} - & - \\ C & A \end{bmatrix} \tag{6.7}$$

where $-$ is a dummy matrix which has no rows. $\widehat{A}$ is an HSS matrix which has one more level than A does.

We then assume that $C = UBV^H$. That is: C is a rank-k matrix. The decomposition of C can be computed by a URV factorization or SVD factorization (in practice, we normally have its decomposition already available).

Then the column base U of C shall be inserted to the HSS representation of A so that U can be generated from the HSS representation of A. This can be done in many different ways. The most straightforward is to insert the column base using the algorithm described in Section 6.1 and followed by a step of *compression* depending on how many columns U has. Suppose that after column bases insertion, the HSS representation of A becomes $\bar{A}$. (Note that: column bases insertion does not change the HSS matrix, it only changes the HSS representation.)

Then $\widehat{A}$ will be represented as

$$\widehat{A} = \begin{bmatrix} - & - \\ UBV^H & \bar{A} \end{bmatrix}. \tag{6.8}$$

It is easy to check that the HSS representation of $\widehat{A}$ will be given as at the top node

$$\begin{cases} B_{1;1,2} = \emptyset & B_{1;2,1} = B & W_{1;1} = | \\ W_{1;2} = \emptyset & R_{1;1} = \emptyset & R_{1;2} = | \end{cases} \tag{6.9}$$

at the left branch:

$$D_{1;1} = - \quad U_{1;1} = \emptyset \quad V_{1;1} = V \tag{6.10}$$

at the right branch:

$$D_{1;2} = \bar{A} \tag{6.11}$$

where $|$ and $-$ represent dummy matrices with no columns respect. no rows. $\emptyset$ represents the dummy matrix without column or or row. The other dimensions of all these should be correct such that the HSS representation is still valid.

6.2.2. Matrix-append when bases are semi-separable. In practice, we almost never compute U and V, since these computations are costly and break the idea of HSS representation. For instance, when a matrix UBV^H needs to be appended to a HSS matrix A, U and V are not explicit stored. They are defined by the formulas (2.6) and (2.7).

In this case, the formulas in the last subsection will have to be modified accordingly, The left branch of $\widehat{A}$ will not be of just one level. Instead, the left child will be a sub-HSS tree defined by the following sequences:
at the root:

$$\begin{cases} B_{1;1,2} = \emptyset & B_{1;2,1} = B & W_{1;1} = | \\ W_{1;2} = \emptyset & R_{1;1} = \emptyset & R_{1;2} = | \end{cases} \tag{6.12}$$

at non-leaf nodes: for $k = 2, 3, \ldots, n$ and $i \in 1, 2, \ldots, 2^{k-1}$:

$$\begin{cases} \widehat{R}_{k;2i-1} = | & \widehat{R}_{k;2i} = | & \widehat{B}_{k;2i-1,2i} = - \\ \widehat{B}_{k;2i,2i-1} = - & \widehat{W}_{k;2i-1} = W_{k;2i-1} & \widehat{W}_{k;2i} = W_{k;2i} \end{cases} \tag{6.13}$$

at the leaves:

$$\widehat{U}_{n;i} = \emptyset \quad \widehat{V}_{n;i} = V_{n;i} \quad \widehat{D}_{n;i} = - \tag{6.14}$$

note that since the column base U is also in a hierarchically semi-separable form, inserting it into A will be somewhat different than that in Section (6.1). The modified formulas for inserting a column base U to A are given by

for $k = 1, 2, \ldots, n$; $i \in 1, 2, \ldots, 2^{k-1}$; $j = i + 1$ for odd i and $j = i - 1$ for even i:

$$\begin{cases} \widehat{B}_{k;i,j} = \begin{bmatrix} B_{A;k;i,j} \\ 0 \end{bmatrix} \\ \widehat{R}_{k;i} = \begin{bmatrix} R_{A;k;i} & 0 \\ 0 & R_{k;i} \end{bmatrix} \\ \widehat{U}_{k;i} = \begin{bmatrix} U_{A;k;i} & U_{k;i} \end{bmatrix}. \end{cases} \tag{6.15}$$

7. Complexity analysis

From the algorithms given, the time complexity of the elementary operations can easily be evaluated together with their effect on the *representation complexity* of the resulting HSS structure. The same matrix can be represented by many different HSS representations, in which some are better than others in terms of computation complexity and space complexity. The *HSS representation complexity* should be defined in such a way that operations on the HSS representation with higher *HSS representation complexity* cost more time and memory than those on HSS representations with lower *HSS representation complexity.* Many indicators can be used. Here, we use a rough measure for the *HSS representation complexity* as follows

Definition 6. *HSS complexity: the total number of free entries in the HSS representation.*

Definition 7. *Free entries: free entries are the entries which can be changed without restriction (For instance, the number of free entries in $n \times n$ diagonal matrix will be n, that in $n \times n$ triangular matrix will be $n(n-1)/2\ldots$ etc.).*

The *HSS complexity* actually indicates the least possible memory needed to store the HSS representation. It also implies the computation complexity, assuming each free entry is accessed once or a small number of times during operations (we may have to account for intermediary representations as well).

Since most of the algorithms given are not so complicated and some have been studied in the literature, we shall limit ourselves to listing a summarizing table for the existing HSS algorithms (including some from the literature). We assume that n is the dimension of the HSS matrix and k is the maximum rank

TABLE 1. Computation complexity analysis table

Operation	**Numerical Complexity**	**Resulting representation complexity**
Matrix-Vector Multiplication [10]	$C_{\mathrm{Matrix}\times\mathrm{Vector}}(n) = O(nk^2)$	A vector of dim. n
Matrix-Matrix Multiplication [10]	$C_{\mathrm{Matrix}\times\mathrm{Matrix}}(n) = O(nk^3)$	Addition
Construct HSS for rank-k matrix	$C_{k\text{-construction}}(n) = O(nk)$	proportional to k
Bases insertion	$C_{\text{Bases-insert}}(n) = O(n)$	Increase by the size of V
Matrix-Append	$C_{\text{Matrix-append}}(n) = O(n)$	Increase by one level
Matrix addition [14]	$C_{\text{Addition}}(n) = O(nk^2)$	Increase additively
Compression	$C_{\text{Compression}}(n) = O(nk^3)$	Does not increase
Model reduction [15]	$C_{\text{Model-reduction}}(n) = O(nk^3)$	Decreases
LU Decomposition [1]	$C_{\mathrm{LU}}(n) = O(nk^3)$	Does not change
Fast solve [10] [14]	$C_{\text{Solve}}(n) = O(nk^3)$	A vector of dim. n
Inverse	$C_{\text{Inverse}}(n) = Onk^3$	Does not change
Transpose	$C_{\text{Transpose}}(n) = O(nk)$	Does not change

of the translation matrices (more accurate formulas can be derived when more detailed information on local rank is available). Table 1 gives a measure of the numerical complexity in terms of n and k, as well as an indication of the HSS complexity of the resulting structure.

We see that in all cases the complexity is linear in the original size of the matrix, and a to be expected power of the size of the translation matrices. Of course, a much more detailed analysis is possible but falls beyond the scope of this paper.

8. Connection between SSS, HSS and the time varying notation

In the earlier papers on SSS [19, 16], efficient algorithms have been developed. Although different algorithms have to be used corresponding to these two seemingly different representations, we would like to show that they are not so different, and we will show how they can be converted to each other. By converting between these two representations, we can take advantages of the fast algorithms for these two different representations.

8.1. From SSS to HSS

In [16], the SSS representation for A is defined as follows: let A be an $N \times N$ matrix satisfying the SSS matrix structure. Then there exist n positive integers $m_1, \ldots, m_n$ with $N = m_1 + \cdots + m_n$ to block-partition A as $A = A_{i,j}$, where $A_{ij} \in C^{m_i \times m_j}$ satisfies

$$A_{ij} = \begin{cases} D_i & \text{if } i = j \\ U_i W_{i+1} \ldots W_{j-1} V_j^H & \text{if } j > i \\ P_i R_{i-1} \ldots R_{j+1} Q_j^H & \text{if } j < i. \end{cases} \tag{8.1}$$

For simplicity, we consider casual operators. For $n = 4$, the matrix A has the form

$$A = \begin{bmatrix} D_1 & U_1 V_2^H & U_1 W_2 V_3^H & U_1 W_2 W_3 V_4^H \\ 0 & D_2 & U_2 V_3^H & U_2 W_3 V_4^H \\ 0 & 0 & D_3 & U_3 V_4^H \\ 0 & 0 & 0 & D_4 \end{bmatrix}. \tag{8.2}$$

Let us first split the SSS matrix as following

$$A = \left[\begin{array}{cc|cc} D_1 & U_1 V_2^H & U_1 W_2 V_3^H & U_1 W_2 W_3 V_4^H \\ 0 & D_2 & U_2 V_3^H & U_2 W_3 V_4^H \\ \hline 0 & 0 & D_3 & U_3 V_4^H \\ 0 & 0 & 0 & D_4 \end{array}\right]. \tag{8.3}$$

The top-left block goes to the left branch of the HSS representation, while the right-bottom block goes to the right branch. The root is defined by setting:

$$\begin{cases} \widehat{B}_{1;1,2} = I & \widehat{B}_{1;2,1} = 0 & \widehat{W}_{1;1} = I \\ \widehat{W}_{1;2} = W_2^H & \widehat{R}_{1;1} = W_3 & \widehat{R}_{1;2} = I. \end{cases} \tag{8.4}$$

Then we construct the left branch with a similar partitioning.

$$\left[\begin{array}{c|c} D_1 & U_1 V_2^H \\ \hline 0 & D_2 \end{array}\right] \tag{8.5}$$

hence

$$\widehat{D}_{2;1} = D_1 \quad \widehat{U}_{2;1} = U_1 \quad \widehat{V}_{2;1} = V_1 \tag{8.6}$$

while for the right child

$$\widehat{D}_{2;2} = D_2 \quad \widehat{U}_{2;2} = U_2 \quad \widehat{V}_{2;2} = V_2. \tag{8.7}$$

In order to keep the HSS representation valid, R and W matrices on the left node should be set properly. That is

$$\begin{cases} \widehat{R}_{2;1} = W_2 & \widehat{R}_{2;2} = I & \widehat{W}_{2;1} = I \\ \widehat{W}_{2;2} = W_1^H & \widehat{B}_{2;2,1} = 0 & B_{2;1,2} = I \end{cases} \tag{8.8}$$

and similarly for the right branch with partitioning as in (8.5)

$$\widehat{D}_{2;3} = D_3 \quad \widehat{U}_{2;3} = U_3 \quad \widehat{V}_{2;3} = V_3, \tag{8.9}$$

$$\widehat{D}_{2;4} = D_4 \quad \widehat{U}_{2;3} = U_4 \quad \widehat{V}_{2;4} = V_4. \tag{8.10}$$

In order to keep the HSS representation valid, R and W matrices on the right node should be set properly. That is

$$\begin{cases} \widehat{R}_{2;3} = W_4 & \widehat{R}_{2;4} = I & \widehat{W}_{2;3} = I \\ \widehat{W}_{2;4} = W_3^H & \widehat{B}_{2;4,3} = 0 & \widehat{B}_{2;3,4} = I. \end{cases} \tag{8.11}$$

Finally the HSS representation can be written as:

$$A = \begin{bmatrix} \widehat{D}_{2;1} & \widehat{U}_{2;1}\widehat{B}_{2;1,2}\widehat{V}_{2;2}^H & \widehat{U}_{2;1}R_{2;1}\widehat{B}_{1;1,2}\widehat{W}_{2;3}^H\widehat{V}_{2;3}^H & \widehat{U}_{2;1}\widehat{R}_{2;1}\widehat{B}_{1;1,2}\widehat{W}_{2;4}^H\widehat{V}_{2;4}^H \\ 0 & \widehat{D}_{2;2} & \widehat{U}_{2;2}\widehat{R}_{2;2}\widehat{B}_{1;1,2}\widehat{W}_{2;3}^H\widehat{V}_{2;3}^H & \widehat{U}_{2;2}\widehat{R}_{2;2}\widehat{B}_{1;1,2}\widehat{W}_{2;4}^H\widehat{V}_{2;4}^H \\ 0 & 0 & \widehat{D}_{2;3} & \widehat{U}_{2;3}\widehat{B}_{2;3,4}\widehat{V}_{2;4}^H \\ 0 & 0 & 0 & \widehat{D}_{2;4} \end{bmatrix} \tag{8.12}$$

with all the translation matrices set in equation (8.4) to (8.11).

The general transformation is then as follows. First we must partition the SSS matrix according to a certain hierarchical partitioning. Then for a current HSS node at k level which should contain the SSS blocks A_{xy} where $i \leq x, y \leq j$ $(1 \leq i < j \leq n)$ and assuming the HSS block is further partitioned at block h $(i < h < j)$ the translation matrices of the current node can be chosen as

$$\begin{cases} \widehat{B}_{k;2i-1,2i} = I & \widehat{B}_{k;2i,2i-1} = 0 & \widehat{W}_{k;2i-1} = I \\ \widehat{W}_{k;2i} = \prod_{x=h}^{i} W_x^H & \widehat{R}_{k;2i-1} = \prod_{x=h+1}^{j} W_x & \widehat{R}_{k;2i} = I \end{cases} \tag{8.13}$$

note that undefined W_x matrices are set equal I (the dimension of I is defined according to context). If $i = h$ or $h + 1 = j$, then one (or two) HSS leaf (leaves) have to be constructed by letting

$$\widehat{D}_{k;i} = D_h \quad \widehat{U}_{k;i} = U_h \quad \widehat{V}_{k;i} = V_h. \tag{8.14}$$

After the HSS node of the current level has been constructed, the same algorithm is applied recursively to construct the HSS node for SSS blocks $A_{xy}, i \leq x, y \leq h$ and for SSS block $A_{xy}, h+1 \leq x, y \leq j$ (the recursion stops when a leaf is constructed.).

Observing the fact that all $\widehat{B}_{k;2i,2i-1}$ matrices are zeros matrices and $W_{k;2i-1}$, $R_{k;2i-1}$ are identity matrices, modifications can be done to get a more efficient HSS representation.

8.2. From HSS to SSS

In this section, we shall consider HSS as recursive SSS using the concise time-varying notation of [5]. We shall first illustrate the algorithm by an example on 8×8 HSS representation. Different partitioning are possible, e.g., those illustrated in Figure 6.

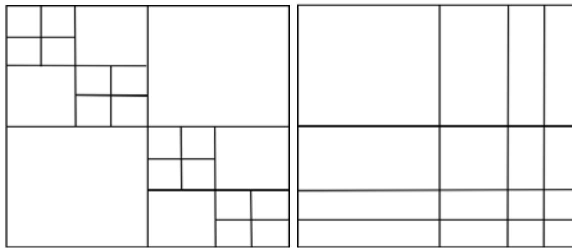

FIGURE 6. HSS partitioning (on the left), SSS partitioning (on the right).

We shall only consider the upper triangular case, as that is the standard case in time-varying system theory. The 4-level balanced HSS representation can be expanded as:

$$A = \begin{bmatrix} D_{1;1} & U_{1;1}B_{1;1,2}W_{2;3}^H V_{2;3}^H & U_{1;1}B_{1;1,2}W_{2;4}^H W_{3;7}^H V_{3;7}^H & U_{1;1}B_{1;1,2}W_{2;4}^H W_{3;8}^H V_{3;8}^H \\ 0 & D_{2,3} & U_{2;3}B_{2;3,4}W_{3;7}^H V_{3;7}^H & U_{2;3}B_{2;3,4}W_{3;8}^H V_{3;8}^H \\ 0 & 0 & D_{3;7} & U_{3;7}B_{3;7,8}V_{3;8}^H \\ 0 & 0 & 0 & D_{3;8}. \end{bmatrix} \tag{8.15}$$

This has to be converted to the time-varying representation for $k = 4$:

$$A = \begin{bmatrix} D_1 & B_1C_2 & B_1A_2C_3 & B_1A_2A_3C_4 \\ 0 & D_2 & B_2C_3 & B_2A_3C_4 \\ 0 & 0 & D_3 & B_3C_4 \\ 0 & 0 & 0 & D_4. \end{bmatrix} \tag{8.16}$$

Representing the time-varying realization matrices as $T_k = \left[\begin{array}{c|c} A_k & C_k \\ \hline B_k & D_k \end{array}\right]$) we obtain

$$T_1 = \left[\begin{array}{c|c} \cdot & \cdot \\ \hline U_{1;1}B_{1;1,2} & D_{1;1} \end{array}\right], T_2 = \left[\begin{array}{c|c} W_{2;4}^H & W_{2;3}^H V_{2;3}^H \\ \hline U_{2;3}B_{2;3,4} & D_{2;3} \end{array}\right] \tag{8.17}$$

$$T_3 = \left[\begin{array}{c|c} W_{3;8}^H & W_{3;7}^H V_{3;7}^H \\ \hline U_{3;7}B_{3;7,8} & D_{3;7} \end{array}\right], T_4 = \left[\begin{array}{c|c} \cdot & V_{3;8}^H \\ \hline \cdot & D_{3;8}. \end{array}\right]. \tag{8.18}$$

More generally, it is easy to see that the realization at k step is given by

$$T_k = \left[\begin{array}{c|c} A_k & C_k \\ \hline B_k & D_k \end{array}\right] = \left[\begin{array}{c|c} W_{k;2^k}^H & W_{k;2^k-1}^H V_{k;2^k-1}^H \\ \hline U_{k;2^k-1}B_{k;2^k-1,2^k} & D_{k;2^k-1}. \end{array}\right]. \tag{8.19}$$

According to the reconfigured partitioning, we see that for step k (indexing the current node) all right children belong to the further steps, while all left children go to $D_{k;2^k-1}$ in the realization of the current step. $W_{k;2^k-1}$, $W_{k;2^k}$ and $B_{k;2^k-1,2^k}$ are the translation matrices of the current node. $U_{k;2^k-1}$ and $V_{k;2^k-1}$ form the column base and row base of the current node, yet they are not explicitly stored.

Note that, according to the HSS definition, they should be generated(recursively) from the left children.

The conversion algorithm should start from the root node and proceed recursively. After constructing the realization on the current step, the algorithm proceeds by setting the right child as the current node and the algorithm goes recursively until it reaches the right bottom where no more right child exist. Then the realization of the last step will be given as:

$$\left[\begin{array}{c|c} \cdot & V^H_{k-1;2^{k-1}} \\ \hline \cdot & D_{k-1;2^{k-1}} \end{array}\right] \tag{8.20}$$

since a leaf does not have a right child.

To show how a HSS tree can be split as time-varying steps, we shall show the partition on an HSS binary tree shown in Figure 7.

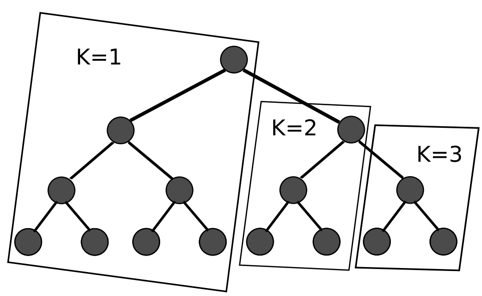

FIGURE 7. Binary tree partitioning.

$D_{k;2^k-1}$ is a potentially large HSS block. Another level of time-varying notation can be used to represent this $D_{k;2^k-1}$ whose realization may again contain sub-blocks represented by the time-varying notation. Since $U_{k;2^k-1}$, $V_{k;2^k-1}$ are not explicitly stored and can be derived locally from the current step, no efficiency is lost by applying recursive time-varying notation.

Here are a number of remarks on the recursive time-varying notation for HSS:

1. $D_{k;2^k-1}$ in the realization is an HSS block which can either be represented in HSS form or by time-varying notation. This suggests a possibly hybrid notation consisting of HSS representations and recursive time-varying notations.
2. $U_{k;2^k-1}$ and $V_{k;2^k-1}$ form HSS bases generated from $D_{k;2^k-1}$. For this recursive time-varying notation, they should not be explicitly stored and can be derived locally in the current step.
3. It is possible to represent general HSS matrices (not just block upper-triangular matrices) with the recursive time-varying notation.
4. All fast HSS algorithms can be interpreted in a recursive time-varying fashion.
5. Some algorithms applied on time-varying notation described in [5] can be extended to the recursive time-varying notation (HSS representation).

9. Final remarks

Although the HSS theory is not yet developed to the same full extent as the sequentially semi-separable theory, the results obtained so far show that the HSS structure has indeed a number of very attractive properties that make it a welcome addition to the theory of structured matrices. Fundamental operations such as matrix-vector multiplication, matrix-matrix multiplication and matrix inversion (including the Moore-Penrose case accounted for in [18]) can all be excuted with a computational complexity linear in the size of the matrix, and additional efficiency induced by the translation operators. A representation in terms of global diagonal and shift operators is also available, very much in the taste of the more restrictive multi-scale theory. These formulas have not yet been exploited fully. The connection with time-varying system theory is also very strong, and it should be possible in the future to transfer a number of its results to the HSS representation, in particular model reduction, interpolation and Hankel norm approximation (i.e., model reduction).

References

[1] W. Lyons, "Fast algorithms with applications to pdes," *PHD thesis*, June 2005.

[2] I. Gohberg, T. Kailath, and I. Koltracht, "Linear complexity algorithms for semiseparable matrices," *Integral Equations and Operator Theory*, vol. 8, pp. 780–804, 1985.

[3] L. Greengard and V. Rokhlin, "A fast algorithm for particle simulations," *J. Comp. Phys.*, vol. 73, pp. 325–348, 1987.

[4] V. Rokhlin, "Applications of volume integrals to the solution of pde's," *J. Comp. Phys.*, vol. 86, pp. 414–439, 1990.

[5] P. Dewilde and A.-J. van der Veen, *Time-varying Systems and Computations.* Kluwer, 1998.

[6] A. van der Veen, "Time-varying lossless systems and the inversion of large structured matrices," *Archiv f. Elektronik u. Übertragungstechnik*, vol. 49, no. 5/6, pp. 372–382, Sept. 1995.

[7] Y. Eidelman and I. Gohberg, "On a new class of structured matrices," *Notes distributed at the 1999 AMS-IMS-SIAM Summer Research Conference*, vol. Structured Matrices in Operator Theory, Numerical Analysis, Control, Signal and Image Processing, 1999.

[8] S. Chandrasekaran, M. Gu, and T. Pals, "A fast and stable solver for smooth recursively semi-separable systems," in *SIAM Annual Conference, San Diego and SIAM Conference of Linear Algebra in Controls, Signals and Systems, Boston*, 2001.

[9] P. Dewilde and A.-J. van der Veen, "Inner-outer factorization and the inversion of locally finite systems of equations," *Linear Algebra and its Applications*, vol. 313, pp. 53–100, 2000.

[10] S. Chandrasekaran, M. Gu, and T. Pals, "Fast and stable algorithms for hierarchically semi-separable representations," in *Technical Report.* University of California at Santa Barbara, April 2004.

[11] W. Hackbusch, "A sparse arithmetic based on $\mathcal{H}$-matrices. part i: Introduction to $\mathcal{H}$-matrices," *Computing*, vol. 64, pp. 21–47, 2000.

[12] T. Pals, "Multipole for scattering computations: Spectral discretization, stabilization, fast solvers," Ph.D. dissertation, Department of Electrical and Computer Engineering, University of California, Santa Barbara, 2004.

[13] P. Dewilde, K. Diepold, and W. Bamberger, "A semi-separable approach to a tridiagonal hierarchy of matrices with application to image flow analysis," in *Proceedings MTNS*, 2004.

[14] S. Chandrasekaran, P. Dewilde, M. Gu, W. Lyons, and T. Pals, "A fast solver for hss representations via sparse matrices," in *Technical Report*. Delft University of Technology, August 2005.

[15] S. Chandrasekaran, Z. Sheng, P. Dewilde, M. Gu, and K. Doshi, "Hierarchically semi-separable representation and dataflow diagrams," vol. Technical document, Nov 2005.

[16] S. Chandrasekaran, P. Dewilde, W. Lyons, T. Pals, and A.-J. van der Veen, "Fast stable solver for sequentially semi-separable linear systems of equations," October 2002.

[17] S. Chandrasekaran, M. Gu, and T. Pals, "A fast ulv decomposition solver for hierachically semiseparable representations," 2004.

[18] P. Dewilde and S. Chandrasekaran, "A hierarchical semi-separable Moore-Penrose equation solver," *Operator Theory: Advances and Applications*, vol. 167, pp. 69–85, Nov 2006, Birkhäuser Verlag.

[19] S. Chandrasekaran, P. Dewilde, M. Gu, W. Lyons, T. Pals, A.-J. van der Veen, and J. Xia, "A fast backward stable solver for sequentially semi-separable matrices," September 2005.

Zhifeng Sheng and Patrick Dewilde
Circuits and Systems Group
Faculty of EEMCS
Delft University of Technology The Netherlands
e-mail: `Z.Sheng@ewi.tudelft.nl`
e-mail: `p.dewilde@ewi.tudelft.nl`

Shivkumar Chandrasekaran
Room 3109, Engineering I
Department of Electrical and Computer Engineering
University of California Santa Barbara
CA 93106-9560, USA
e-mail: `Shiv@ece.ucsb.edu`

Operator Theory:
Advances and Applications, Vol. 176, 295–322

Unbounded Normal Algebras and Spaces of Fractions

Florian-Horia Vasilescu

Abstract. We consider arbitrary families of unbounded normal operators, commuting in a strong sense, in particular algebras consisting of unbounded normal operators, and investigate their connections with some algebras of fractions of continuous functions on compact spaces. New examples and properties of spaces of fractions are also given.

Mathematics Subject Classification (2000). Primary 47L60; Secondary 47B25; 13B99; 47B15; 46G10; 46E15; 46A13.

Keywords. Algebras of unbounded operators; spaces of fractions of continuous functions; spectral measures; Cayley transform.

1. Introduction

The classical Cayley transform $\kappa(t) = (t-i)(t+i)^{-1}$ is a bijective map between the real line $\mathbb{R}$ and the set $\mathbb{T}\setminus\{1\}$, where $\mathbb{T}$ is the unit circle in the complex plane. If $p(t) = \sum_{k=0}^{n} a_k t^k$ is a polynomial of one real variable, generally with complex coefficients, then the function

$$p\circ\kappa^{-1}(z) = \sum_{k=0}^{n} i^k a_k (1+z)^k (1-z)^{-k} = \sum_{k=0}^{n} (-1)^k a_k (\Im z)^k (1-\Re z)^{-k},$$

defined on $\mathbb{T}\setminus\{1\}$, is a sum of fractions with denominators in the family $\{(1-z)^k; k\geq 0\}$, or in the family $\{(1-\Re z)^k; k\geq 0\}$, the latter consisting of positive functions on $\mathbb{T}$. This remark allows us to identify the algebra of polynomials of one real variable with sub-algebras of some algebras of fractions. As a matter of fact, a similar identification can be easily obtained for polynomials in several real variables. We shall use this idea to describe the structure of some algebras of unbounded operators.

We start with some notation and terminology for Hilbert space linear operators. For the properties of the unbounded operators, in particular unbounded self-adjoint and normal operators, and the Cayley transform, we refer to [18], Chapter 13.

Let $\mathcal{H}$ be a complex Hilbert space. For a linear operator T acting in $\mathcal{H}$, we denote by $\mathcal{D}(T)$ its domain of definition. If T is closable, the closure of T will be denoted by $\bar{T}$. If T is densely defined, let T^* be its adjoint. We write $T_1 \subset T_2$ to designate that T_2 extends T_1.

We recall that a densely defined closed operator T is said to be *normal* (resp. *self-adjoint*) if $\mathcal{D}(T) = \mathcal{D}(T^*)$ and $T^*T = TT^*$ (resp. $T = T^*$).

If $\mathcal{D}(T) = \mathcal{D}(T^*)$, the equality $T^*T = TT^*$ is equivalent to $\|Tx\| = \|T^*x\|$ for all $x \in \mathcal{D}(T)$ (see [11], Part II). Clearly, every self-adjoint operator is normal.

Let $\mathcal{D}$ be a dense linear subspace of $\mathcal{H}$, let $\mathcal{L}(\mathcal{D})$ be the algebra of all linear mappings from $\mathcal{D}$ into $\mathcal{D}$, and let, for further use, $\mathcal{B}(\mathcal{H})$ be the algebra of all bounded linear operators on $\mathcal{H}$.

Set

$$\mathcal{L}^{\#}(\mathcal{D}) = \{T \in \mathcal{L}(\mathcal{D}); \mathcal{D}(T^*) \supset \mathcal{D}, T^*\mathcal{D} \subset \mathcal{D}\}.$$

If we put $T^{\#} = T^*|\mathcal{D}$, the set $\mathcal{L}^{\#}(\mathcal{D})$ is an algebra with involution $T \to T^{\#}$.

Let $\mathcal{N}$ be a subalgebra of $\mathcal{L}^{\#}(\mathcal{D})$ such that for every $T \in \mathcal{N}$ one has $T^{\#} \in \mathcal{N}$. We also assume that the identity on $\mathcal{D}$ belongs to $\mathcal{N}$. With the terminology from [20], the algebra $\mathcal{N}$ is an *O^*-algebra*.

As in the bounded case (see [18], Definition 11.24), we say that an O^*-algebra $\mathcal{N}$ is a *normal algebra* if $\mathcal{N}$ consists of (not necessarily bounded) operators whose closures are normal operators.

One of the main aims of this work is to describe normal algebras in terms of algebras of fractions (see Theorem 4.8). We briefly recall the definition of the latter algebras in a commutative framework. In the next section, a more general framework will be presented.

Let $\mathcal{A}$ be a commutative unital complex algebra and let $\mathcal{M} \subset \mathcal{A}$ be a *set of denominators*, that is, a subset closed under multiplication, containing the unit and such that if $ma = 0$ for some $m \in \mathcal{M}$ and $a \in \mathcal{A}$, then $a = 0$. Under these conditions, we can form the algebra of fractions $\mathcal{A}/M$ consisting of the equivalence classes modulo the relation $(a_1, m_1) \sim (a_2, m_2)$ if $a_1m_2 = a_2m_1$ for $a_1, a_2 \in \mathcal{A}$ and $m_1, m_2 \in \mathcal{M}$, endowed with a natural algebraic structure (see [6] or [24] for some details; see also the next section for a more complete discussion). The equivalence class of the pair (a, m) will be denoted by a/m.

If $\mathcal{A}$ is a commutative unital normed complex algebra whose completion $\mathcal{B}$ is a semisimple Banach algebra, and $\mathcal{M} \subset A$ is a set of denominators, the Gelfand representation allows us to replace $\mathcal{A}/\mathcal{M}$ by an algebra of fractions of continuous functions. Specifically, if $\hat{a}$ denotes the Gelfand transform of any $a \in \mathcal{A}$ (computed in $\mathcal{B}$), and $\hat{\mathcal{S}}$ stands for the set $\{\hat{a};\, a \in \mathcal{S}\}$ for a subset $\mathcal{S} \subset A$, then $\hat{\mathcal{M}}$ is a set of denominators in $\hat{\mathcal{A}}$ and the natural assignment $\mathcal{A}/\mathcal{M} \ni a/m \to \hat{a}/\hat{m} \in \hat{\mathcal{A}}/\hat{\mathcal{M}}$ is a unital algebra isomorphism, as one can easily see.

The remark from above shows that the algebras of fractions of continuous functions on compact spaces are of particular interest. Their role in the study of scalar and operator moment problems, and in extension problems, has been already emphasized in [2] and [25]. In the present work, we continue the investigation concerning the algebras of fractions of continuous functions, showing their usefulness in the representation of normal algebras.

We shall briefly describe the contents of this article. In the next section, following some ideas from [6], we present a general construction of spaces of fractions, sufficiently large to allow a partially noncommutative context. Some examples, more or less related to moment problems, are also given.

As mentioned before, algebras of fractions of continuous functions on compact spaces are of particular interest. In the third section, we describe the dual of some spaces of fractions of continuous functions (Theorem 3.1), extending a result from [25], stated as Corollary 3.2. Other examples of algebras of fractions of continuous functions, needed in the next sections, some of them considering functions depending on infinitely many variables, are also presented.

In the fourth section, we introduce and discuss the algebras of unbounded normal operators. To obtain a representation theorem of such an algebra, identifying it with a subalgebra of an algebra of fractions of continuous functions (see Theorem 4.8), we need versions of the spectral theorem for infinitely many unbounded commuting self-adjoint or normal operators (see Theorem 4.2 and Theorem 4.3). This subject has been approached by many authors (see, for instance, [4] and [19]). A recent similar result of this type, which we are aware of, is the main theorem in [16], obtained in the context of semi-groups. Inspired by a result in [18] stated for one operator, our main tool to obtain the versions of the spectral theorem mentioned above is the Cayley transform, leading to an approach seemingly different from other approaches for infinitely many operators.

Some examples of unbounded normal algebras are also given in this section.

The last section contains an extension result of the so-called subnormal families of unbounded operators to normal ones. We use the main result from [2] to get a version of another result from [2], which, unlike in the quoted work, is proved here for families having infinitely many members (see Theorem 5.2), using our Theorem 4.3. The statement of Theorem 5.2 needs another type of an algebra of fractions, which is also included in this section.

Let us mention that fractional transformations have been recently used, in various contexts, in [2], [9], [10], [14], [15], [25] etc. See also [11], [23], [26] for other related results.

2. Spaces of fractions

In this section we present a general setting for a construction of spaces of fractions, associated with real or complex vector spaces, with denominators in appropriate families of linear maps. We adapt some ideas from [6].

Let $\mathbb{K}$ be either the real field $\mathbb{R}$ or the complex one $\mathbb{C}$, let E be a vector space over $\mathbb{K}$, and let $\mathcal{L}(E)$ be the algebra of all linear maps from E into itself.

Definition 2.1. Let $\mathcal{M} \subset \mathcal{L}(E)$ have the following properties:

(1) for all $M_1, M_2 \in \mathcal{M}$ we have $M_1M_2 \in \mathcal{M}$;
(2) the identity map I_E on E belongs to $\mathcal{M}$;
(3) $\mathcal{M}$ is commutative;
(4) every map $M \in \mathcal{M}$ is injective.

Such a family $\mathcal{M} \subset \mathcal{L}(E)$ will be called a *set of denominators*.

Definition 2.2. Let $\mathcal{M} \subset \mathcal{L}(E)$ be a set of denominators. Two elements (x_1, M_1), (x_2, M_2) from $E \times \mathcal{M}$ are said to be *equivalent*, and we write $(x_1, M_1) \sim (x_2, M_2)$, if $M_1x_2 = M_2x_1$.

Remark 2.3. The relation $\sim$ given by Definition 2.2 is clearly reflexive and symmetric. It is also transitive because if $(x_1, M_1) \sim (x_2, M_2)$ and $(x_2, M_2) \sim (x_3, M_3)$, we infer easily that $M_2(M_1x_3 - M_3x_1) = 0$, via condition (3), whence $M_1x_3 = M_3x_1$, by condition (4). Consequently, the relation $\sim$ is an equivalence relation. This allows us to consider the set of equivalence classes $E \times \mathcal{M}/ \sim$, which will be simply denoted by $E/\mathcal{M}$. The equivalence class of the element (x, M) will be denoted x/M.

The set $E/\mathcal{M}$ can be organized as a vector space with the algebraic operations

$$x_1/M_1 + x_2/M_2 = (M_2x_1 + M_1x_2)/M_1M_2,\ x_1, x_2 \in E,\ M_1, M_2 \in \mathcal{M},$$

and

$$\lambda(x/M) = (\lambda x)/M,\ \lambda \in \mathbb{K},\ x \in E,\ M \in \mathcal{M},$$

which are easily seen to be correctly defined.

Definition 2.4. The vector space $E/\mathcal{M}$ will be called the *space of fractions* of E *with denominators in* $\mathcal{M}$.

Note that, if $\mathcal{A}_\mathcal{M}$ is the (commutative) algebra generated by $\mathcal{M}$ in $\mathcal{L}(E)$, the linear space $E/\mathcal{M}$ is actually an $\mathcal{A}_\mathcal{M}$-module, with the action given by

$$N(x/M) = (Nx)/M,\ x \in E,\ N \in \mathcal{A}_\mathcal{M},\ M \in \mathcal{M}.$$

If we regard the multiplication by $M_0 \in \mathcal{M}$ as a linear map on $E/\mathcal{M}$, then M_0 has an inverse on $E/\mathcal{M}$ defined by

$$M_0^{-1}(x/M) = x/(M_0M),\ x \in E,\ M \in \mathcal{M}.$$

We also note that the map $E \ni x \mapsto x/I_E \in E/\mathcal{M}$ is injective, which allows the identification of E with the subspace $\{x/I_E,\ x \in E\}$ of $E/\mathcal{M}$. For this reason, the fraction x/M may be denoted by $M^{-1}x$ for all $x \in E,\ M \in \mathcal{M}$.

We define the subspaces

$$E/M = \{\xi \in E/\mathcal{M};\ M\xi \in E\},\ M \in \mathcal{M}.$$

We clearly have

$$E/\mathcal{M} = \bigcup_{M \in \mathcal{M}} E/M. \tag{2.1}$$

A set of denominators $\mathcal{M} \subset \mathcal{L}(E)$ has a natural division. Namely, if $M', M'' \in \mathcal{M}$, we write $M'|M''$, and say that M' *divides* M'' if there exists $M_0 \in \mathcal{M}$ such that $M'' = M'M_0$.

A subset $\mathcal{M}_0$ of $\mathcal{M}$ is said to be *cofinal* if for every $M \in \mathcal{M}$ we can find an $M_0 \in \mathcal{M}_0$ such that $M|M_0$.

If $M'|M''$, and so $M'' = M'M_0$, the map $E/M' \ni x/M' \mapsto M_0x/M'' \in E/M''$ is the restriction of the identity to E/M', showing that E/M' is a subspace of E/M''. This also shows that the vector space $E/\mathcal{M}$ is the inductive limit of the family of vector spaces $(E/M)_{M\in\mathcal{M}}$.

Remark 2.5. If there is a norm $\| * \|$ on the vector space E, each space E/M can be endowed with the norm

$$\|\xi\|_M = \|M\xi\|,\ \xi \in E/M,\ M \in \mathcal{M}. \tag{2.2}$$

Assuming also that $\mathcal{M}$ consists of bounded operators on E, it is easily seen that

$$\|\xi\|_{M''} \le \|M_0\|\|\xi\|_{M'},\ \xi \in E/M',$$

whenever $M'' = M'M_0$, showing that the inclusion $E/M' \subset E/M''$ is continuous, and so $E/\mathcal{M}$ can be viewed as an inductive limit of normed spaces. If F is a topological vector space and $T : E/\mathcal{M} \mapsto F$ is a linear map, then T is *continuous* if $T|_{E/M}$ is continuous for each $M \in \mathcal{M}$ (see [17] for details).

Remark 2.6. Assume that E is an ordered vector space and let E_+ be the positive cone of E. Let also $\mathcal{M} \subset \mathcal{L}(E)$ be a set of denominators. If one has $M(E_+) \subset E_+$ for all $M \in \mathcal{M}$ (i.e., every $M \in \mathcal{M}$ is a positive operator), we say that $\mathcal{M}$ is a set of *positive* denominators. If $\mathcal{M}$ is a set of positive denominators, we may define a positive cone $(E/M)_+$ in each space E/M by setting

$$(E/M)_+ = \{\xi \in E/M;\ M\xi \in E_+\}.$$

If F is another ordered vector space with the positive cone F_+, a linear map $\phi : E/\mathcal{M} \mapsto F$ is said to be *positive* if $\phi((E/M)_+) \subset F_+$ for all $M \in \mathcal{M}$.

Remark 2.7. Assume that $E = A$ is an algebra. Let $\mathcal{M} \subset \mathcal{L}(A)$ be a set of denominators. Also assume that M is an A-module map of the A-module A for all $M \in \mathcal{M}$. In other words, $M(ab) = aM(b)$ for all $a, b \in A$ and $M \in \mathcal{M}$. Then the fraction space $A/\mathcal{M}$ becomes an algebra, with the multiplication given by the relation

$$(a'/M')(a''/M'') = (a'a'')/(M'M''),\ a', a'' \in A,\ M', M'' \in \mathcal{M}.$$

Particularly, let A be an algebra with unit 1. For each $c \in A$ we set $M_c(a) = ca$, $a \in A$, i.e., the left multiplication map by c on A. A subset $Q \subset A$ is said to be a *set of denominators* if the family $\mathcal{M}_Q = \{M_q;\ q \in Q\} \subset \mathcal{L}(A)$ is a set of denominators. In this case, we identify Q and $\mathcal{M}_Q$ and write $A/\mathcal{M}_Q$ simply A/Q. If Q is in the center of A, then A/Q is an algebra.

Example 2.8. Let $\mathcal{A}$ be a complex $*$-algebra with unit 1, and let $L : \mathcal{A} \mapsto \mathbb{C}$ be a positive form on $\mathcal{A}$. This pair can be associated, in a canonical way, with a certain pre-Hilbert space, via the classical construction due to Gelfand and Naimark. To briefly recall this construction, let $\mathcal{N} = \{a \in \mathcal{A}; L(a^*a) = 0\}$. Since L satisfies the Cauchy-Schwarz inequality, it follows that $\mathcal{N}$ is a left ideal of $\mathcal{A}$. Moreover, the quotient $\mathcal{D} = \mathcal{A}/\mathcal{N}$ is a pre-Hilbert space, whose inner product is given by $\langle a + \mathcal{N}, b + \mathcal{N}\rangle = L(b^*a)$, $a, b \in \mathcal{A}$.

Note also that $\mathcal{D}$ is an $\mathcal{A}$-module. Therefore, we may define a linear map M_a on $\mathcal{D}$ associated to any $a \in \mathcal{A}$, via the relation $M_a(\tilde{x}) = (ax)\tilde{}$, $x \in \mathcal{A}$, where $\tilde{x} = x + \mathcal{N}$.

Let $\mathcal{C}$ be the center of $\mathcal{A}$, and fix $\mathcal{C}_0 \subset \mathcal{C}$ nonempty. The map M_{r_c} is injective if $r_c = 1 + c^*c$ for each $c \in \mathcal{C}_0$. Consequently, the set $\mathcal{M}$ of all maps of the form

$$M_{r_{c_1}}^{\alpha_1} \cdots M_{r_{c_m}}^{\alpha_m},$$

where $\alpha_1, \ldots, \alpha_m$ are arbitrary nonnegative integers and $c_1, \ldots, c_m$ are arbitrary elements from $\mathcal{C}_0$, is a set of denominators in $\mathcal{L}(\mathcal{D})$. This shows that we can consider the space of fractions $\mathcal{D}/\mathcal{M}$.

The following particular case is related to the classical Hamburger moment problem in several variables. Let us denote by $\mathbb{Z}_+^n$ the set of all multi-indices $\alpha = (\alpha_1, \ldots, \alpha_n)$, i.e., $\alpha_j \in \mathbb{Z}_+$ for all $j = 1, \ldots, n$. Let $\mathcal{P}_n$ be the algebra of all polynomial functions on $\mathbb{R}^n$, with complex coefficients, endowed with its natural involution. We denote by t^α the monomial $t_1^{\alpha_1} \cdots t_n^{\alpha_n}$, where $t = (t_1, \ldots, t_n)$ is the current variable in $\mathbb{R}^n$, and $\alpha \in \mathbb{Z}_+^n$.

If an n-sequence $\gamma = (\gamma_\alpha)_{\alpha \in \mathbf{Z}_+^n}$ of real numbers is given, we associate it with the functional $L_\gamma : \mathcal{P}_n \to \mathbb{C}$, where $L_\gamma(t^\alpha) = \gamma_\alpha, \alpha \in \mathbb{Z}_+^n$. Assuming L_γ to be positive, we may perform the construction described above, and obtain a pre-Hilbert space $\mathcal{D}_\gamma = \mathcal{D}$. The spaces of fractions obtained as above will be related to the family $\mathcal{C}_0 = \{t_1, \ldots, t_n\}$. Setting $A_j\tilde{p} = (t_jp)\tilde{}$, $p \in \mathcal{D}_\gamma$, $j = 1, \ldots, n$, the denominator set as above, say $\mathcal{M}_\gamma$, will be given by the family of maps of the form

$$(1 + A_1^2)^{\alpha_1} \cdots (1 + A_n^2)^{\alpha_m},$$

where $\alpha_1, \ldots, \alpha_m$ are arbitrary nonnegative integers. And so, we can form the space of fractions $\mathcal{D}_\gamma/\mathcal{M}_\gamma$, as a particular case of the previous construction.

Example 2.9. Let $C[0, 1]$ (resp. $C[0, 1)$) be the algebra of all complex-valued continuous functions on the interval [0,1] (resp. [0,1)). We consider the Volterra operator $Vf(t) = \int_0^t f(t)dt,\ t \in [0, 1),\ f \in C[0, 1)$, which is an injective map. Let $C_V[0, 1)$ be the subspace of $C[0, 1)$ consisting of those functions f such that $V^nf \in C[0, 1]$ for some integer $n \geq 0$ (depending on f). Let $\mathcal{V} = \{V^n; n \geq 0\}$, which, regarded as a family of linear maps on $C[0, 1]$, is a family of denominators. Therefore, we may form the space of fractions $C[0, 1]/\mathcal{V}$. Note that the space $C_V[0, 1)$ may be identified with a subspace of $C[0, 1]/\mathcal{V}$. Indeed, if $f \in C_V[0, 1)$ and $n \geq 0$ is such that $V^nf \in C[0, 1]$, we identify f with the element $V^nf/V^n \in C[0, 1]/V^n$, and this assignment is linear and injective.

Example 2.10. Let Ω be a compact space and let $C(\Omega)$ be the algebra of all complex-valued continuous functions on Ω, endowed with the natural norm $\|f\|_\infty = \sup_{\omega\in\Omega}|f(\omega)|$, $f \in C(\Omega)$. We consider a collection $\mathcal{P}$ of complex-valued functions p, each defined and continuous on an open set $\Delta_p \subset \Omega$. Let μ be a positive measure on Ω such that $\mu(\Omega \setminus \Delta_p) = 0$, and p (extended with zero on $\Omega \setminus \Delta_p$) is μ-integrable for all $p \in \mathcal{P}$. Via a slight abuse of notation, we may define the numbers $\gamma_p = \int_\Omega p d\mu$, $p \in \mathcal{P}$, which can be called the $\mathcal{P}$-*moments* of μ. A very general, and possibly hopeless moment problem at this level, might be to characterize those families of numbers $(\gamma_p)_{p\in\mathcal{P}}$ which are the $\mathcal{P}$-moments of a certain positive measure.

Let us add some natural supplementary conditions. First of all, assume that $\Omega_0 = \cap_{p\in\mathcal{P}}\Delta_p$ is a dense subset of Ω. Also assume that there exists $\mathcal{R} \subset \mathcal{P}$ a family containing the constant function 1, closed under multiplication in the sense that if $r', r'' \in \mathcal{R}$ then $r'r''$ defined on $\Delta_{r'} \cap \Delta_{r''}$ is in $\mathcal{R}$, and each $r \in \mathcal{R}$ is nonnull on its domain of definition. Finally, we assume that for every function $p \in \mathcal{P}$ there exists a function $r \in \mathcal{R}$ such that the function p/r, defined on $\Delta_p \cap \Delta_r$, has a (unique) continuous extension to Ω. In particular, all functions from the family $\mathcal{Q} = \{1/r; r \in \mathcal{R}\}$ have a continuous extension to Ω. Moreover, the set $\mathcal{Q}$, identified with a family in $C(\Omega)$, is a set of denominators. This allows us to identify each function $p \in \mathcal{P}$ with a fraction from $C(\Omega)/\mathcal{Q}$, namely with h/q, where h is the continuous extension of p/r and $q = 1/r$ for a convenient $r \in \mathcal{R}$. With these conditions, the above $\mathcal{P}$-moment problem can be approached with our methods (see [25]; see also Corollary 3.2).

Summarizing, for a given subspace $\mathcal{P}$ of the algebra of fractions $C(\Omega)/\mathcal{Q}$, and a linear functional ϕ on $\mathcal{P}$, we look for necessary and sufficient conditions on $\mathcal{P}$ and ϕ to insure the existence of a solution, that is, a positive measure μ on Ω such that each p be μ-almost everywhere defined and $\phi(p) = \int_\Omega p d\mu$, $p \in \mathcal{P}$. We may call such a problem a *singular moment problem*, when no data are specified. With this terminology, the classical moment problems of Stieltjes and Hamburger, in one or several variables, are singular moment problems.

3. Spaces of fractions of continuous functions

Let Ω be a compact space and let $C(\Omega)$ be the algebra of all complex-valued continuous functions on Ω, endowed with the natural norm $\|f\|_\infty = \sup_{\omega\in\Omega}|f(\omega)|$, $f \in C(\Omega)$. We denote by $M(\Omega)$ the space of all complex-valued Borel measures on Ω, sometimes identified with the dual of $C(\Omega)$. For an arbitrary function $h \in C(\Omega)$, we set $Z(h) = \{\omega \in \Omega; h(\omega) = 0\}$, which is obviously a compact subset of Ω. If $\mu \in M(\Omega)$, we denote by $|\mu| \in M(\Omega)$ the variation of μ.

We discuss certain spaces of fractions, which were considered in [25]. Let $\mathcal{Q}$ be a family of nonnegative elements of $C(\Omega)$. The set $\mathcal{Q}$ is said to be a *multiplicative family* if (i) $1 \in \mathcal{Q}$, (ii) $q', q'' \in \mathcal{Q}$ implies $q'q'' \in \mathcal{Q}$, and (iii) if $qh = 0$ for some $q \in \mathcal{Q}$ and $h \in C(\Omega)$, then $h = 0$. As in Remark 2.7, a multiplicative family is a set of denominators, and so we can form the algebra of fractions $C(\Omega)/\mathcal{Q}$. To unify the terminology, a multiplicative family $\mathcal{Q} \subset C(\Omega)$ will be called, with no ambiguity, a set of (positive) denominators.

To define a natural topological structure on $C(\Omega)/\mathcal{Q}$, we use Remark 2.5. If

$$C(\Omega)/q = \{f \in C(\Omega)/\mathcal{Q}; qf \in C(\Omega)\},$$

then we have $C(\Omega)/\mathcal{Q} = \cup_{q\in\mathcal{Q}} C(\Omega)/q$. Setting $\|f\|_{\infty,q} = \|qf\|_\infty$ for each $f \in C(\Omega)/q$, the pair $(C(\Omega)/q, \| * \|_{\infty,q})$ becomes a Banach space. Hence, $C(\Omega)/\mathcal{Q}$ is an inductive limit of Banach spaces (see [17], Section V.2).

As in Remark 2.6, in each space $C(\Omega)/q$ we have a positive cone $(C(\Omega)/q)_+$ consisting of those elements $f \in C(\Omega)/q$ such that $qf \geq 0$ as a continuous function.

We use in this text sometimes the notation q^{-1} to designate the element $1/q$ for any $q \in \mathcal{Q}$.

Let $\mathcal{Q}_0 \subset \mathcal{Q}$ be nonempty. Let $\mathcal{F} = \sum_{q\in\mathcal{Q}_0} C(\Omega)/q$, that is, the subspace of $C(\Omega)/\mathcal{Q}$ generated by the subspaces $(C(\Omega)/q)_{q\in\mathcal{Q}_0}$, which is itself an inductive limit of Banach spaces. Let also $\psi : \mathcal{F} \to \mathbb{C}$ be linear. As in Remark 2.5, the map ψ is continuous if the restriction $\psi|C(\Omega)/q$ is continuous for all $q \in \mathcal{Q}_0$.

We note that the values of ψ do not depend on the particular representation of the elements of $\mathcal{F}$.

Let us also note that the linear functional $\psi : \mathcal{F} \to \mathbb{C}$ is positive (see Remark 2.6 or [25]) if $\psi|(C(\Omega)/q)_+ \geq 0$ for all $q \in \mathcal{Q}_0$.

The next result, which is an extension of the Riesz representation theorem, describes the dual of a space of fractions, defined as above.

Theorem 3.1. *Let $\mathcal{Q}_0 \subset \mathcal{Q}$ be nonempty, let $\mathcal{F} = \sum_{q\in\mathcal{Q}_0} C(\Omega)/q$, and let $\psi : \mathcal{F} \to \mathbb{C}$ be linear. The functional ψ is continuous if and only if there exists a uniquely determined measure $\mu_\psi \in M(\Omega)$ such that $|\mu_\psi|(Z(q)) = 0$, q^{-1} is $|\mu_\psi|$-integrable for all $q \in \mathcal{Q}_0$ and $\psi(f) = \int_\Omega f d\mu_\psi$ for all $f \in \mathcal{F}$.*

The functional $\psi : \mathcal{F} \to \mathbb{C}$ is positive, if and only if it is continuous and the measure μ_ψ is positive.

Proof. Let $\mu \in M(\Omega)$ be a measure such that $|\mu|(Z(q)) = 0$ and q^{-1} (which is $|\mu|$-almost everywhere defined) is $|\mu|$-integrable for all $q \in \mathcal{Q}_0$. We set $\psi(f) = \int_\Omega f d\mu$ for all $f \in \mathcal{F}$. This definition is correct. Indeed, if $f = \sum_{j\in J} h'_j/q'_j = \sum_{k\in K} h''_k/q''_k$ are two (finite) representations of $f \in \mathcal{F}$, with $h'_j, h''_k \in C(\Omega)$ and $q'_j, q''_k \in \mathcal{Q}_0$, setting $Z = \cup_{j\in J} Z(q'_j) \cup \cup_{k\in K} Z(q''_k)$, we easily derive that $f(\omega) = \sum_{j\in J} h'_j(\omega)/q'_j(\omega) = \sum_{k\in K} h''_k(\omega)/q''_k(\omega)$ for all $\omega \in \Omega \setminus Z$. As $|\mu|(Z) = 0$, we infer that f is a function defined $|\mu|$-almost everywhere and the integral $\psi(f) = \int_\Omega f d\mu$ does not depend on the particular representation of f.

Clearly, the functional ψ is linear. Note also that

$$|\psi(f)| = \left|\int_\Omega hq^{-1} d\mu\right| \leq \int_\Omega |h|q^{-1} d|\mu| \leq \|f\|_{\infty,q} \int_\Omega q^{-1} d|\mu| \tag{3.1}$$

for all $f = h/q \in C(\Omega)/q$, showing the continuity of ψ.

Conversely, let $\psi : \mathcal{F} \to \mathbb{C}$ be linear and continuous. For every $q \in \mathcal{Q}_0$ we set $\theta_q(h) = \psi(h/q)$, $h \in C(\Omega)$. Since the map $C(\Omega) \ni h \to h/q \in C(\Omega)/q$ is continuous (in fact, it is an isometry), the map θ_q is a continuous linear functional on $C(\Omega)$. Therefore, there exists a measure $\nu_q \in M(\Omega)$ such that $\theta_q(h) = \int_\Omega h\, d\nu_q$, $h \in C(\Omega)$, for all $q \in \mathcal{Q}_0$.

Note that $\psi(hq_1/q_1) = \psi(hq_2/q_2)$ for all $q_1, q_2 \in \mathcal{Q}_0$ and $h \in C(\Omega)$, because hq_1/q_1 and hq_2/q_2 are two representations of the same element in $\mathcal{F}$. Therefore, $\theta_{q_1}(hq_1) = \psi(hq_1/q_1) = \psi(hq_2/q_2) = \theta_{q_2}(hq_2)$ for all $q_1, q_2 \in \mathcal{Q}_0$ and $h \in C(\Omega)$, implying the equality $q_1\nu_{q_1} = q_2\nu_{q_2}$. Consequently, there exists a measure μ such that $\mu = q\nu_q$ for all $q \in \mathcal{Q}_0$.

The equality $\mu = q\nu_q$ implies the equality $|\mu| = q|\nu_q|$. This shows the set $Z(q)$ must be $|\mu|$-null. Moreover, the function q^{-1} is $|\mu|$-integrable for all $q \in \mathcal{Q}_0$. Consequently, $\nu_q = q^{-1}\mu$, and the function q^{-1} is μ-integrable for all $q \in \mathcal{Q}_0$.

If $f \in \mathcal{F}$ is arbitrary, then $f = \sum_{j\in J} h_j q_j^{-1}$, with $h_j \in C(\Omega)$, $q_j \in \mathcal{Q}_0$ for all $j \in J$, J finite. We can write

$$\psi(f) = \sum_{j\in J} \theta_{q_j}(h_j) = \sum_{j\in J} \int_\Omega h_j\, d\nu_{q_j} = \int_\Omega f d\mu,$$

giving the desired integral representation for the functional ψ.

As we have $\int_\Omega h d\mu = \psi(hq/q)$ for all $h \in C(\Omega)$ and $q \in \mathcal{Q}_0$, and $\psi(hq/q)$ does not depend on q, it follows that the measure μ is uniquely determined. If we put $\mu = \mu_\psi$, we have the measure whose existence and uniqueness are asserted by the statement.

Let $\psi : \mathcal{F} \to \mathbb{C}$ be linear be linear and positive. Then, as one expects, ψ is automatically continuous. Indeed, if $h \in C(\Omega)$ is real-valued and $q \in \mathcal{Q}_0$, the inequality $-\|h\|_\infty/q \le h/q \le \|h\|_\infty/q$ implies, via the positivity of ψ, the estimate $|\psi(h/q)| \le \|h/q\|_{\infty,q}\psi(1/q)$. If $h \in C(\Omega)$ is arbitrary, the estimate above leads to $|\psi(h/q)| \le 2\|h/q\|_{\infty,q}\psi(1/q)$, showing that ψ is continuous. Finally, the equality $\psi(f) = \int f d\mu_\psi$, $f \in C(\Omega)/q$, $q \in \mathcal{Q}_0$ shows that ψ is positive if and only if μ_ψ is positive. □

The next result is essentially Theorem 3.2 from [25].

Corollary 3.2. *Let $\mathcal{Q}_0 \subset \mathcal{Q}$ be nonempty, let $\mathcal{F} = \sum_{q\in\mathcal{Q}_0} C(\Omega)/q$, and let $\psi : \mathcal{F} \to \mathbb{C}$ be linear. The functional ψ is positive if and only if*

$$\sup\{|\psi(hq^{-1})|;\ h \in C(\Omega),\ \|h\|_\infty \le 1\} = \psi(q^{-1}),\ \ q \in \mathcal{Q}_0. \tag{3.2}$$

Proof. Set $\psi_q = \psi|C(\Omega)/q$ for all $q \in \mathcal{Q}_0$. Then, since $\|f\|_{\infty,q} = \|h\|_\infty$ whenever $f = h/q \in C(\Omega)/q$, we clearly have

$$\|\psi_q\| = \sup\{|\psi(hq^{-1})|;\ h \in C(\Omega),\ \|h\|_\infty \le 1\}.$$

If ψ is positive, and μ_ψ is the associated positive measure given by Theorem 3.1, the estimate (3.1) implies that

$$\|\psi_q\| \le \int_\Omega q^{-1} d\mu_\psi = \psi(q^{-1}) \le \|\psi_q\|,$$

because $\|1/q\|_{\infty,q} = 1$. This shows that (3.2) holds.

Conversely, we use the well-known fact that a linear functional $\theta : C(\Omega) \to \mathbb{C}$ is positive if and only if it is continuous and $\|\theta\| = \theta(1)$

Assuming that (3.2) holds, with the notation of the proof of Theorem 3.1, we have

$$\|\theta_q\| = \|\psi_q\| = \psi(1/q) = \theta_q(1),$$

showing that θ_q is positive for all q. Therefore, the measures ν_q are all positive, implying that the measure $\mu = \mu_\psi$ is positive. Consequently, ψ must be positive. $\square$

Definition 3.3. Let $\mathcal{Q} \subset C(\Omega)$ be a set of denominators. A measure $\mu \in M(\Omega)$ is said to be $\mathcal{Q}$-*divisible* if for every $q \in \mathcal{Q}$ there is a measure $\nu_q \in M(\Omega)$ such that $\mu = q\nu_q$.

Theorem 3.1 shows that a functional on $C(\Omega)/\mathcal{Q}$ is continuous if and only if it has an integral representation via a $\mathcal{Q}$-divisible measure. In addition, Corollary 3.2 asserts that a functional is positive on $C(\Omega)/\mathcal{Q}$ if and only if it is represented by a $\mathcal{Q}$-divisible positive measure μ such that $\mu = q\nu_q$ with $\nu_q \in M(\Omega)$ positive for all $q \in \mathcal{Q}$.

As a matter of fact, the concept given by Definition 3.3 can be considerably extended, as shown by the next example.

Example 3.4. This is a continuation of the discussion started in Example 2.9, whose notation will be kept. We fix an indefinitely differentiable function ϕ, with support in [0,1]. Note the identity

$$\int_0^1 h(t)\phi(t)dt = (-1)^n \int_0^1 V^n h(t)\phi^{(n)}(t)dt, \tag{3.3}$$

valid for all $h \in C[0,1]$ and all integers $n \ge 0$. If we set $d\mu(t) = \phi(t)dt$ and $d\nu_n(t) = (-1)^n\phi^{(n)}(t)dt$, and referring to Definition 3.3, we may say, by (3.3), that the measure μ is $\mathcal{V}$-*divisible*.

It is plausible that the study of $\mathcal{M}$-divisible measures, defined in an appropriate manner for a set of denominators $\mathcal{M}$ consisting of linear and continuous operators on $C(\Omega)$, can be related to the study of continuous linear functionals on the space $C(\Omega)/\mathcal{M}$, via a possible extension of Theorem 3.1.

Example 3.5. Let $\mathcal{S}_1$ be the algebra of polynomials in $z, \bar{z}$, $z \in \mathbb{C}$. We will show that this algebra, which is used to characterize the moment sequences in the complex plane, can be identified with a subalgebra of an algebra of fractions of continuous functions. This example will be extended to infinitely many variables in the last section (similar, yet different examples were considered in [2]). Let $\mathcal{R}_1$ be the set of functions $\{(1+|z|^2)^{-k}; z\in\mathbb{C}, k\in\mathbb{Z}_+\}$, which can be continuously extended to $\mathbb{C}_\infty=\mathbb{C}\cup\{\infty\}$. Identifying $\mathcal{R}_1$ with the set of their extensions in $C(\mathbb{C}_\infty)$, the family $\mathcal{R}_1$ becomes a set of denominators in $C(\mathbb{C}_\infty)$. This will allows us to identify the algebra $\mathcal{S}_1$ with a subalgebra of the algebra of fractions $C(\mathbb{C}_\infty)/\mathcal{R}_1$.

Let $\mathcal{S}_{1,k}^{(1)}$, $k\ge 1$, be the space generated by the monomials $z^j\bar{z}^l$, $0\le j+l<2k$. We put $\mathcal{S}_{1,0}^{(1)}=\mathbb{C}$. Let also $\mathcal{S}_{1,k}^{(2)}$, $k\ge 1$, be the space generated by the monomials $|z|^{2j}$, $0<j\le k$. Put $\mathcal{S}_{1,0}^{(2)}=\{0\}$.

Set $\mathcal{S}_{1,k}=\mathcal{S}_{1,k}^{(1)}+\mathcal{S}_{1,k}^{(2)}$, $k\ge 0$. We clearly have $\mathcal{S}_1=\sum_{k\ge 0}\mathcal{S}_{1,k}$. Since $\mathcal{S}_{1,k}$ may be identified with a subspace of $C(\mathbb{C}_\infty)/r_k$, where $r_k(z)=(1+|z|^2)^{-k}$ for all $k\ge 0$, the space $\mathcal{S}_1$ can be viewed as a subalgebra of the algebra $C(\mathbb{C}_\infty)/\mathcal{R}_1$. Note also that $r_k^{-1}\in\mathcal{S}_{1,k}$ for all $k\ge 1$ and $\mathcal{S}_{1,k}\subset\mathcal{S}_{1,l}$ whenever $k\le l$.

According to Theorem 3.4 from [25], a linear map $\phi:\mathcal{S}_1\mapsto\mathbb{C}$ has a positive extension $\psi: C(\mathbb{C}_\infty)/\mathcal{R}_1\mapsto\mathbb{C}$ with $\|\phi_k\|=\|\psi_k\|$ if and only if $\|\phi_k\|=\phi(r_k^{-1})$, where $\phi_k=\phi|\mathcal{S}_{1,k}$ and $\psi_k=\psi|C(\mathbb{C}_\infty)/r_k$, for all $k\ge 0$ (the norms of the functionals are computed in the sense discussed in Remark 2.5).

This result can be used to characterize the Hamburger moment problem in the complex plane, in the spirit of [25]. Specifically, given a sequence of complex numbers $\gamma=(\gamma_{j,l})_{j\ge 0,l\ge 0}$ with $\gamma_{0,0}=1$, $\gamma_{k,k}\ge 0$ if $k\ge 1$ and $\gamma_{j,l}=\bar\gamma_{l,j}$ for all $j\ge 0, l\ge 0$, the Hamburger moment problem means to find a probability measure on $\mathbb{C}$ such that $\gamma_{j,l}=\int z^j\bar{z}^l d\mu(z)$, $j\ge 0, l\ge 0$.

Defining $L_\gamma:\mathcal{S}_1\mapsto\mathbb{C}$ by setting $L_\gamma(z^j\bar{z}^l)=\gamma_{j,l}$ for all $j\ge 0, l\ge 0$ (extended by linearity), if L_γ has the properties of the functional ϕ above insuring the existence of a positive extension to $C(\mathbb{C}_\infty)/\mathcal{R}_1$, then the measure μ is provided by Corollary 3.2.

For a fixed integer $m\ge 1$, we can state and characterize the existence of solutions for a truncated moment problem (for an extensive study of such problems see [7] and [8]). Specifically, given a finite sequence of complex numbers $\gamma=(\gamma_{j,l})_{j,l}$ with $\gamma_{0,0}=1$, $\gamma_{j,j}\ge 0$ if $1\le j\le m$ and $\gamma_{j,l}=\bar\gamma_{l,j}$ for all $j\ge 0, l\ge 0$, $j\ne l$, $j+l<2m$, find a probability measure on $\mathbb{C}$ such that $\gamma_{j,l}=\int z^j\bar{z}^l d\mu(z)$ for all indices j,l. As in the previous case, a necessary and sufficient condition is that the corresponding map $L_\gamma:\mathcal{S}_{1,m}\mapsto\mathbb{C}$ have the property $\|L_\gamma\|=L_\gamma(r_m^{-1})$, via Theorem 3.4 from [25]. Note also that the actual truncated moment problem is slightly different from the usual one (see [7]).

The similar space $\mathcal{T}_1$, introduced in [2], can be used to characterize the following moment problem: Find a probability measure μ on $\mathbb{C}$ such that the double sequence of the form $\gamma=(\gamma_{j,0},\gamma_{k,k})_{j\ge 0,\,k\ge 1}$ (with $\gamma_{0,0}=1$ and $\gamma_{k,k}\ge 0$ if $k\ge 1$) be a moment sequence in the sense that $\gamma_{j,0}=\int z^j d\mu(z), \gamma_{k,k}=\int |z|^{2k}d\mu(z)$.

Example 3.6. We are particularly interested in some special algebras of fractions of continuous functions, depending on infinitely many variables, necessary for our further discussion.

Let $\mathcal{I}$ be a (nonempty) family of indices. We consider the space $\mathbb{R}^{\mathcal{I}}$, where, as before, $\mathbb{R}$ is the real field. Denote by $t=(t_\iota)_{\iota\in\mathcal{I}}$ the independent variable in

$\mathbb{R}^{\mathcal{I}}$. Let $\mathbb{Z}_+^{(\mathcal{I})}$ be the set of all collections $\alpha = (\alpha_\iota)_{\iota\in\mathcal{I}}$ of nonnegative integers, with finite support. Setting $t^0 = 1$ for $0 = (0)_{\iota\in\mathcal{I}}$ and $t^\alpha = \prod_{\alpha_\iota\neq 0} t_\iota^{\alpha_\iota}$ for $t = (t_\iota)_{\iota\in\mathcal{I}} \in \mathbb{R}^{\mathcal{I}}$, $\alpha = (\alpha_\iota)_{\iota\in\mathcal{I}} \in \mathbb{Z}_+^{(\mathcal{I})}$, $\alpha \neq 0$, we may consider the algebra of complex-valued *polynomial functions* $\mathcal{P}_{\mathcal{I}}$ on $\mathbb{R}^{\mathcal{I}}$, consisting of expressions of the form $\sum_{\alpha\in\mathcal{J}} c_\alpha t^\alpha$, with c_α complex numbers for all $\alpha \in \mathcal{J}$, where $\mathcal{J} \subset \mathbb{Z}_+^{(\mathcal{I})}$ is finite. Note also that the map $\mathcal{P}_{\mathcal{I}} \ni p \to \bar{p} \in \mathcal{P}_{\mathcal{I}}$ is an involution on $\mathcal{P}_{\mathcal{I}}$, where $\bar{p}(t) = \sum_{\alpha\in\mathcal{J}} \bar{c}_\alpha t^\alpha$ if $p(t) = \sum_{\alpha\in\mathcal{J}} c_\alpha t^\alpha$.

When dealing with finite measures, an appropriate framework is the space of all continuous functions on a compact topological space. But neither the space $\mathbb{R}^{\mathcal{I}}$ is compact nor the functions from $\mathcal{P}_{\mathcal{I}}$ are bounded. If we consider the one-point compactification $\mathbb{R}_\infty$ of $\mathbb{R}$, then we can embed $\mathbb{R}^{\mathcal{I}}$ into the compact space $(\mathbb{R}_\infty)^{\mathcal{I}}$. This operation leads us to consider the space $\mathcal{P}_{\mathcal{I}}$ as a subspace of an algebra of fractions derived from the basic algebra $C((\mathbb{R}_\infty)^{\mathcal{I}})$, via a suitable multiplicative family. Specifically, we consider the family $\mathcal{Q}_{\mathcal{I}}$ consisting of all rational functions of the form $q_\alpha(t) = \prod_{\alpha_\iota\neq 0}(1+t_\iota^2)^{-\alpha_\iota}$, $t = (t_\iota)_{\iota\in\mathcal{I}} \in \mathbb{R}^{\mathcal{I}}$, where $\alpha = (\alpha_\iota) \in \mathbb{Z}_+^{(\mathcal{I})}$, $\alpha \neq 0$, is arbitrary (see also [9]). Of course, we set $q_0 = 1$. The function q_α can be continuously extended to $(\mathbb{R}_\infty)^{\mathcal{I}} \setminus \mathbb{R}^{\mathcal{I}}$ for all $\alpha \in \mathbb{Z}_+^{(\mathcal{I})}$. Moreover, the set $\mathcal{Q}_{\mathcal{I}}$ becomes a set of denominators in $C((\mathbb{R}_\infty)^{\mathcal{I}})$. Set also $p_\alpha(t) = q_\alpha(t)^{-1}$, $t \in \mathbb{R}^{\mathcal{I}}$, $\alpha \in \mathbb{Z}_+^{(\mathcal{I})}$.

Let $\mathcal{P}_{\mathcal{I},\alpha}$ be the vector space generated by the monomials t^β, with $\beta_\iota \leq 2\alpha_\iota$, $\iota \in \mathcal{I}$, $\alpha, \beta \in \mathbb{Z}_+^{(\mathcal{I})}$. It is clear that p/p_α can be continuously extended to $(\mathbb{R}_\infty)^{\mathcal{I}} \setminus \mathbb{R}^{\mathcal{I}}$ for every $p \in \mathcal{P}_{\mathcal{I},\alpha}$, and so it can be regarded as an element of $C((\mathbb{R}_\infty)^{\mathcal{I}})$. Therefore, $\mathcal{P}_{\mathcal{I},\alpha}$ is a subspace of $C((\mathbb{R}_\infty)^{\mathcal{I}})/q_\alpha = p_\alpha C((\mathbb{R}_\infty)^{\mathcal{I}})$ for all $\alpha \in \mathbb{Z}_+^{\mathcal{I}}$.

Example 3.7. We give now another example of an algebra of fractions, which will be used in the next section to describe the normal algebras.

As before, let $\mathcal{I}$ be a (nonempty) family of indices. We consider the space $\mathbb{T}^{\mathcal{I}}$, where $\mathbb{T}$ is the unit circle in the complex plane. Denote by $z = (z_\iota)_{\iota\in\mathcal{I}}$ the independent variable in $\mathbb{T}^{\mathcal{I}}$. Let $\mathbb{Z}_+^{(\mathcal{I})}$ be defined as in the previous example. Setting $(\Re z)^0 = 1$ for $0 = (0)_{\iota\in\mathcal{I}}$, $(\Re z)^\alpha = \prod_{\alpha_\iota\neq 0}(\Re z_\iota)^{\alpha_\iota}$ and similar formulas for $(\Im z)^\alpha$ whenever $z = (z_\iota)_{\iota\in\mathcal{I}} \in \mathbb{T}^{\mathcal{I}}$, $\alpha = (\alpha_\iota)_{\iota\in\mathcal{I}} \in \mathbb{Z}_+^{(\mathcal{I})}$, $\alpha \neq 0$, we may consider the algebra of complex-valued functions $\mathcal{R}_{\mathcal{I}}$ on $\mathbb{T}^{\mathcal{I}}$, consisting of expressions of the form $\sum_{\alpha,\beta\in\mathcal{J}} c_{\alpha,\beta}(\Re z)^\alpha(\Im z)^\beta$, with $c_{\alpha,\beta}$ complex numbers for all $\alpha, \beta \in \mathcal{J}$, where $\mathcal{J} \subset \mathbb{Z}_+^{(\mathcal{I})}$ is finite.

We may take in the algebra $\mathcal{R}_{\mathcal{I}}$ a set of denominators $\mathcal{S}_{\mathcal{I}}$ consisting of all functions of the form $s_\alpha(t) = \prod_{\alpha_\iota\neq 0}(1-\Re z_\iota)^{\alpha_\iota}$, $z = (z_\iota)_{\iota\in\mathcal{I}} \in \mathbb{T}^{\mathcal{I}}$, where $\alpha = (\alpha_\iota) \in \mathbb{Z}_+^{(\mathcal{I})}$, $\alpha \neq 0$, is arbitrary. We put $s_0 = 1$. Clearly, $\mathcal{S}_{\mathcal{I}}$ is a set of denominators also in $C((\mathbb{T})^{\mathcal{I}})$.

Note that the map

$$\mathcal{P}_{\mathcal{I}} \ni p \mapsto p \circ \tau \in \mathcal{R}_{\mathcal{I}}/\mathcal{S}_{\mathcal{I}},$$

where $\tau : (\mathbb{T}\setminus\{1\})^{\mathcal{I}} \mapsto \mathbb{R}^{\mathcal{I}}$ is given by $\tau(z)_\iota = -\Im z_\iota/(1-\Re z_\iota)$, $\iota \in \mathcal{I}$, is an injective algebra homomorphism, allowing the identification of $\mathcal{P}_{\mathcal{I}}$ with a subalgebra of $\mathcal{R}_{\mathcal{I}}/\mathcal{S}_{\mathcal{I}}$.

Remark 3.8. Let Ω be a compact Hausdorff space, let A be unital C^*-algebra, and let $C(\Omega, A)$ be the C^*-algebra of all A-valued functions, continuous on Ω. Let $\mathcal{Q}$ be set of denominators in $C(\Omega)$, which can be identified with a set of denominators in $C(\Omega, A)$. Therefore, we may consider the algebra of fractions $C(\Omega, A)/\mathcal{Q}$. Let $\mathcal{Q}_0$ be an arbitrary subset of $\mathcal{Q}$ and let $\mathcal{F} = \sum_{q\in\mathcal{Q}_0} \mathcal{F}_q$ be a subspace of $\sum_{q\in\mathcal{Q}_0} C(\Omega, A)/q$ such that $q^{-1} \in \mathcal{F}_q$ et $\mathcal{F}_q \subset C(\Omega, A)/q$ for all $q \in \mathcal{Q}_0$.

Let $\mathcal{H}$ be a Hilbert space, let $\mathcal{D}$ be a dense linear subspace of $\mathcal{H}$, and denote by $SF(\mathcal{D})$ the space of all sesquilinear forms on $\mathcal{D}$. Let $\phi : \mathcal{F} \mapsto SF(\mathcal{D})$ be linear. Suppose that $\phi(q^{-1})(x, x) > 0$ for all $x \in \mathcal{D}\setminus\{0\}$ and $q \in \mathcal{Q}_0$. Then $\phi(q^{-1})$ induces an inner product on $\mathcal{D}$, and let $\mathcal{D}_q$ be the space $\mathcal{D}$, endowed with the norm given by $\| * \|_q^2 = \phi(q^{-1})(*, *)$. Set $\phi_q = \phi|\mathcal{F}_q$. We may define the quantities

$$\|\phi_q(f)\| = \sup\{|\phi_q(f)(x, y)|; \|x\|_q \leq 1, \|y\|_q \leq 1\},$$

and

$$\|\phi_q\| = \sup\{\|\phi_q(f)\|; \|qf\|_\infty \leq 1\}.$$

We say that the map $\phi : \mathcal{F} \mapsto SF(\mathcal{D})$ is *contractive* if $\|\phi_q\| \leq 1$ for all $q \in \mathcal{Q}_0$.

4. Normal algebras

We keep the notation and terminology from the Introduction.

Let A_1, A_2 be self-adjoint in $\mathcal{H}$. Trying to avoid, at this moment, any involvement of the concept of a spectral measure for unbounded operators, we say that A_1, A_2 *commute* if the bounded operators $(A_1 + iI_{\mathcal{H}})^{-1}$ and $(A_2 + iI_{\mathcal{H}})^{-1}$ commute, where $I_{\mathcal{H}}$ is the identity on $\mathcal{H}$, which is one of the possible (classical) definitions of the commutativity of (unbounded) self-adjoint operators. It is known that the operator N in normal in $\mathcal{H}$ if and only if one has $N = A_1 + iA_2$, where A_1, A_2 are commuting self-adjoint operators (see [11], Part II, or [20]). For this reason, given two normal operators N', N'' having the decompositions $N' = A_1' + iA_2'$, $N'' = A_1'' + iA_2''$, with A_1', A_2', A_1'', A_2'' self-adjoint, we say that N', N'' *commute* if the self-adjoint operators A_1', A_2', A_1'', A_2'' mutually commute.

Let $\mathcal{N} \subset \mathcal{L}^{\#}(\mathcal{D})$ be an O^*-algebra. As mentioned in the Introduction, we say that $\mathcal{N}$ is *normal* if $\bar{N}$ is normal for each $N \in \mathcal{N}$.

A homonymic concept, defined in the framework of bounded operators, can be found in [18]. The aim of this section is to describe the structure of normal algebras, extending Theorem 12.22 from [18]. Let us recall the essential part of that result.

Theorem A. *Let $\mathcal{N}$ be a closed normal algebra of $\mathcal{B}(\mathcal{H})$ containing the identity, and let Δ be the maximal ideal space of $\mathcal{N}$. Then there exists a unique spectral measure E on the Borel subsets of Δ such that $N = \int_\Delta \hat{N} dE$ for every $N \in \mathcal{N}$.*

This result leads us to some versions of the Spectral Theorem, called here Theorem 4.2 and Theorem 4.3, valid for an arbitrary family of not necessarily bounded, commuting self-adjoint or normal operators (see also [4], [16], [19] etc. for various approaches to this topic).

Remark 4.1. Let Ω be a topological space and let E be a spectral measure, defined on the family of all Borel subsets of Ω, with values in $\mathcal{B}(\mathcal{H})$. For every Borel function $f : \Omega \to \mathbb{C}$ we set $\mathcal{D}_f = \{x \in \mathcal{H}; \int_\Omega |f|^2 dE_{x,x} < \infty\}$, where $E_{x,y}(*) = \langle E(*)x, y\rangle$. Then the formula

$$\langle \Psi(f)x, y\rangle = \int_\Omega f dE_{x,y},\ x \in \mathcal{D}_f,\ y \in \mathcal{H}$$

defines a normal operator $\Psi(f)$ for each f, with $\Psi(f)^* = \Psi(\bar{f})$ (see [18], especially Theorem 13.24, for this and other properties to be used in this text).

If f is E-almost everywhere defined (in particular, almost everywhere finite) on Ω, we keep the notation $\int_\Omega f dE_{x,y}$ whenever the integral is well defined.

In the following, for every complex number λ, the operator $\lambda I_{\mathcal{H}}$ will be simply written as λ.

The next statement is a version of Theorem 13.30 from [18] (stated and proved for one self-adjoint operator). We adapt some ideas from the proof of quoted theorem, to get a statement for infinitely many variables.

Theorem 4.2. *Let $(A_\iota)_{\iota\in\mathcal{I}}$ be a commuting family of self-adjoint operators in $\mathcal{H}$. Then there exists a unique spectral measure E on the Borel subsets of $(\mathbb{R}_\infty)^{\mathcal{I}}$ such that each coordinate function $(\mathbb{R}_\infty)^{\mathcal{I}} \ni t \to t_\iota \in \mathbb{R}_\infty$ is E-almost everywhere finite. In addition,*

$$\langle A_\iota x, y\rangle = \int_{(\mathbb{R}_\infty)^{\mathcal{I}}} t_\iota dE_{x,y}(t),\ \ x \in \mathcal{D}(A_\iota),\ y \in \mathcal{H},$$

where

$$\mathcal{D}(A_\iota) = \{x \in \mathcal{H}; \int_{(\mathbb{R}_\infty)^{\mathcal{I}}} |t_\iota|^2 dE_{x,x}(t) < \infty\},$$

for all $\iota \in \mathcal{I}$.

If the set $\mathcal{I}$ is at most countable, then the measure E has support in $\mathbb{R}^{\mathcal{I}}$.

Proof. Let U_ι denote the Cayley transform of the self-adjoint operator A_ι for each $\iota \in \mathcal{I}$. Since

$$U_\iota = (A_\iota - i)(A_\iota + i)^{-1} = 1 - 2i(A_\iota + i)^{-1},\ \iota \in \mathcal{I},$$

the collection $(U_\iota)_{\iota\in\mathcal{I}}$ is a commuting family of unitary operators in $\mathcal{B}(\mathcal{H})$. Moreover, the point 1 does not belong to the point spectrum of $1 - U_\iota$ for all ι, as follows from the general properties of the Cayley transform [18].

Let $\mathcal{B}$ be the closed unital algebra generated by the family $(U_\iota, U_\iota^*)_{\iota\in\mathcal{I}}$ in $\mathcal{B}(\mathcal{H})$, which is a commutative unital C^*-algebra. Let $\Gamma(\mathcal{B})$ be the space of characters of the algebra $\mathcal{B}$.

Standard arguments from the Gelfand theory allow us to assert that the map

$$\Gamma(\mathcal{B}) \ni \gamma \mapsto (\hat{U}_\iota(\gamma))_{\iota\in\mathcal{I}} \in \mathbb{T}^{\mathcal{I}}$$

is a homeomorphism, and so the space $\Gamma(\mathcal{B})$ may be identified with a closed subspace, say Ω, of the compact space $\mathbb{T}^{\mathcal{I}}$.

By virtue of the Theorem A above, there exists a spectral measure F defined on the Borel subsets of Ω such that $U_\iota = \int_\Omega z_\iota dF(z)$, where $z = (z_\iota)_{\iota\in\mathcal{I}}$ is the variable in $\mathbb{T}^{\mathcal{I}}$ and the map $\hat{U}_\iota$ has been identified with the coordinate function $z \to z_\iota$ for all ι.

Set $C_\iota = \{z \in \Omega; z_\iota = 1\}$. We want to prove that $F(C_\iota) = 0$ for all ι. Assuming that this is not the case for some ι, we could find a nonnull vector x_ι such that $E(C_\iota)x_\iota = x_\iota$. In addition,

$$x_\iota - U_\iota x_\iota = \int_\Omega (1 - z_\iota)dF(z)x_\iota = \int_{C_\iota} (1 - z_\iota)dF(z)x_\iota = 0,$$

which contradicts the fact that the kernel of $1 - U_\iota$ is null.

In particular, this shows that the function $(1 - z_\iota)^{-1}$ is defined F-almost everywhere for all ι.

For technical reasons, we may extend the set function F to the family of all Borel subsets of $\mathbb{T}^{\mathcal{I}}$ by putting $\tilde{F}(\beta) = F(\beta \cap \Omega)$ for every Borel subset β of $\mathbb{T}^{\mathcal{I}}$. Let $\kappa_{\mathcal{I}} : (\mathbb{R}_\infty)^{\mathcal{I}} \to \mathbb{T}^{\mathcal{I}}$ be the map given by

$$(\mathbb{R}_\infty)^{\mathcal{I}} \ni t = (t_\iota)_{\iota\in\mathcal{I}} \mapsto w = (w_\iota)_{\iota\in\mathcal{I}} \in \mathbb{T}^{\mathcal{I}}$$

where $w_\iota = \kappa_\iota(t) = (t_\iota - i)(t_\iota + i)^{-1}$ if $t_\iota \neq \infty$ and $w_\iota = 1$ if $t_\iota = \infty$. The map $\kappa_{\mathcal{I}}$ is a homeomorphism and the superposition $E = \tilde{F} \circ \kappa_{\mathcal{I}}$ is a spectral measure on the Borel subsets of $(\mathbb{R}_\infty)^{\mathcal{I}}$.

If $D_\iota = \{t \in (\mathbb{R}_\infty)^{\mathcal{I}}; t_\iota = \infty\}$, then $E(D_\iota) = \tilde{F}(\kappa(D_\iota)) = F(C_\iota) = 0$, as noticed above. In other words, the coordinate function $(\mathbb{R}_\infty)^{\mathcal{I}} \ni t \to t_\iota \in \mathbb{R}_\infty$ is E-almost everywhere finite. Moreover, if $\theta_\iota(z) = i(1+z_\iota)(1-z_\iota)^{-1} = -\Im z_\iota(1-\Re z_\iota)^{-1}$, then θ_ι is $\tilde{F}$-almost everywhere defined, and one has

$$\int_{(\mathbb{R}_\infty)^{\mathcal{I}}} t_\iota dE_{x,y}(t) = \int_{\mathbb{T}^{\mathcal{I}}} \theta_\iota(z)d\tilde{F}_{x,y}(z) = \int_\Omega \theta_\iota(z)dF_{x,y}(z),$$

for all $x \in \mathcal{D}_{\theta_\iota}$ and $y \in \mathcal{H}$, via a change of variable and the Remark above. As the function θ_ι is real-valued, the operator $\tilde{A}_\iota$ given by the equality

$$\langle \tilde{A}_\iota x, y\rangle = \int_{(\mathbb{R}_\infty)^{\mathcal{I}}} t_\iota dE_{x,y}(t),\ x \in \mathcal{D}_{\theta_\iota},\ y \in \mathcal{H}$$

is self-adjoint. The arguments from the last part of Theorem 13.30 in [18] show that $\tilde{A}_\iota$ must be precisely A_ι. For the convenience of the reader, we sketch this argument. The equality $(1 - z_\iota)\theta_\iota(z) = i(1 + z_\iota)$ leads to the equality $\tilde{A}_\iota(1 - U_\iota) = i(1 + U_\iota)$. This shows that $\tilde{A}_\iota$ is a self-adjoint extension of the inverse Cayley transform A_ι of U_ι. But any self-adjoint operator is maximally symmetric ([18], Theorem 13.15), and so $A_\iota = \tilde{A}_\iota$.

The equality

$$\mathcal{D}(A_\iota) = \{x \in \mathcal{H}; \int_{(\mathbb{R}_\infty)^{\mathcal{I}}} |t_\iota|^2 dE_{x,x}(t) < \infty\},$$

for all $\iota \in \mathcal{I}$, is a consequence of [18], Theorem 13.24.

To prove the uniqueness of the measure E, we consider another spectral measure G, defined on the Borel subsets of $(\mathbb{R}_\infty)^{\mathcal{I}}$ such that the G-measure of set $D_\iota = \{t \in (\mathbb{R}_\infty)^{\mathcal{I}}; t_\iota = \infty\}$ is null for all ι. In other words, each coordinate function $(\mathbb{R}_\infty)^{\mathcal{I}} \ni t \to t_\iota \in \mathbb{R}_\infty$ is G-almost everywhere finite, and

$$\langle A_\iota x, y\rangle = \int_{(\mathbb{R}_\infty)^{\mathcal{I}}} t_\iota dG_{x,y}(t),\ x \in \mathcal{D}(A_\iota),\ y \in \mathcal{H},\ \iota \in \mathcal{I}.$$

Then the map $H = G \circ \kappa_{\mathcal{I}}^{-1}$ defines a spectral measure on the Borel subsets of the compact space $(\mathbb{R}_\infty)^{\mathcal{I}}$. Because $\kappa_{\mathcal{I}}^{-1} = \theta$, where $\theta(z) = (\theta_\iota(z))_{\iota\in\mathcal{I}}$, we must have

$$U_\iota = \int_{(\mathbb{R}_\infty)^{\mathcal{I}}} \kappa_\iota(t) dG(t) = \int_{\mathbb{T}^{\mathcal{I}}} z_\iota dH(z),$$

where $z = (z_\iota)_{\iota\in\mathcal{I}}$ is the variable in $\mathbb{T}^{\mathcal{I}}$. As the unital C^*-algebra generated by the polynomials in $z_\iota, \bar{z}_\iota$, $\iota \in \mathcal{I}$, is dense in both $C(\Omega)$ and $C(\mathbb{T}^{\mathcal{I}})$, via the Weierstrass-Stone density theorem, the scalar measures $F_{x,y}$ and $H_{x,y}$ are equal for all $x, y \in \mathcal{H}$. Therefore, the spectral measures F and H must be equal too. Consequently, the spectral measures measures E and G obtained from F and H respectively, must be equal, which completes the proof of the theorem.

Finally, if the set $\mathcal{I}$ is at most countable, then the measure E has support in the Borel set

$$(\mathbb{R}_\infty)^{\mathcal{I}} \setminus \bigcup_{\iota\in\mathcal{I}} D_\iota = \mathbb{R}^{\mathcal{I}}. \qquad \square$$

We may consider the one-point compactification $\mathbb{C}_\infty$ of the complex plane $\mathbb{C}$. Then, for every family of indices $\mathcal{I}$, the space $(\mathbb{C}_\infty)^{\mathcal{I}}$ is compact. A version of the previous result, valid for normal operators, is given by the following.

Theorem 4.3. *Let $(N_\iota)_{\iota\in\mathcal{I}}$ be a commuting family of normal operators in $\mathcal{H}$. Then there exists a unique spectral measure G on the Borel subsets of $(\mathbb{C}_\infty)^{\mathcal{I}}$ such that each coordinate function $(\mathbb{C}_\infty)^{\mathcal{I}} \ni z \to z_\iota \in \mathbb{C}_\infty$ is G-almost everywhere finite. In addition,*

$$\langle N_\iota x, y\rangle = \int_{(\mathbb{C}_\infty)^{\mathcal{I}}} z_\iota dG_{x,y}(z),\ x \in \mathcal{D}(N_\iota),\ y \in \mathcal{H},$$

where

$$\mathcal{D}(N_\iota) = \{x \in \mathcal{H}; \int_{(\mathbb{C}_\infty)^{\mathcal{I}}} |z_\iota|^2 dG_{x,x}(z) < \infty\},$$

for all $\iota \in \mathcal{I}$.

If the set $\mathcal{I}$ is at most countable, then the measure G has support in $\mathbb{C}^{\mathcal{I}}$.

Proof. We write $N_\iota = A'_\iota + iA''_\iota$, where $(A'_\iota, A''_\iota)_{\iota\in\mathcal{I}}$ is a commuting family of self-adjoint operators. Set $\mathcal{L} = (\mathcal{I}\times\{0\})\cup(\{0\}\times\mathcal{I})$, $B_\lambda = A'_\iota$ if $\lambda = (\iota, 0)$, and $B_\lambda = A''_\iota$ if $\lambda = (0, \iota)$. We may apply the previous theorem to the family of commuting self-adjoint operators $(B_\lambda)_{\lambda\in\mathcal{L}}$. Therefore, there exists a spectral measure E on the Borel subsets of $(\mathbb{R}_\infty)^{\mathcal{L}}$ such that each coordinate function $\mathbb{R}^{\mathcal{L}} \ni t \to t_\lambda \in \mathbb{R}$ is E-almost everywhere defined. Moreover,

$$\langle B_\lambda x, y\rangle = \int_{(\mathbb{R}_\infty)^{\mathcal{L}}} t_\lambda dE_{x,y}(t),\ x \in \mathcal{D}(B_\lambda),\ y \in \mathcal{H},\ \lambda \in \mathcal{L}.$$

We define a map $\tau : (\mathbb{R}_\infty)^{\mathcal{L}} \mapsto (\mathbb{C}_\infty)^{\mathcal{I}}$ in the following way. If $t = (t_\lambda)_{\lambda\in\mathcal{L}}$ is an arbitrary point in $(\mathbb{R}_\infty)^{\mathcal{I}}$, we put $\tau(t) = (z_\iota)_{\iota\in\mathcal{I}}$, where $z_\iota = t_{(\iota,0)} + it_{(0,\iota)}$ for all $\iota \in \mathcal{I}$. Clearly, we put $z_\iota = \infty$ if either $t_{(\iota,0)} = \infty$ or $t_{(0,\iota)} = \infty$.

We consider the spectral measure given by $G = E\circ\tau^{-1}$, defined on the Borel subsets of $(\mathbb{C}_\infty)^{\mathcal{I}}$. Setting

$$A_\iota = \{z \in (\mathbb{C}_\infty)^{\mathcal{I}}; z_\iota = \infty\},$$

$$B_\iota = \{t \in (\mathbb{R}_\infty)^{\mathcal{L}}; t_{(\iota,0)} = \infty\},\ C_\iota = \{t \in (\mathbb{R}_\infty)^{\mathcal{L}}; t_{(0,\iota)} = \infty\},$$

we obtain $\tau^{-1}(A_\iota) \subset B_\iota \cup C_\iota$ for all $\iota \in \mathcal{I}$. Since $E(B_\iota \cup C_\iota) = 0$, it follows that $G(A_\iota) = 0$, that is, the coordinate function $z \mapsto z_\iota$ is G-almost everywhere finite. In addition,

$$\int_{(\mathbb{C}_\infty)^{\mathcal{I}}} z_\iota dG_{x,y}(z) = \int_{(\mathbb{R}_\infty)^{\mathcal{L}}} (t_{(\iota,0)} + it_{(0,\iota)})dE_{x,y}(z)$$
$$= \langle(A'_\iota + iA''_\iota)x, y\rangle = \langle N_\iota x, y\rangle$$

for all $x \in \mathcal{D}(N_\iota) = \mathcal{D}(A'_\iota)\cap\mathcal{D}(A''_\iota), y \in \mathcal{H}$ and $\iota \in \mathcal{I}$.

We also have

$$\mathcal{D}(N_\iota) = \{x \in \mathcal{H}; \int_{(\mathbb{C}_\infty)^{\mathcal{I}}} |z_\iota|^2 dG_{x,x}(z) < \infty\},$$

for all $\iota \in \mathcal{I}$, as a consequence of [18], Theorem 13.24.

Because $G(A_\iota) = 0$ for all $\iota \in \mathcal{I}$, the last assertion follows as in the proof of Theorem 4.2. $\square$

Example 4.4. Let Ω be a Hausdorff space, let $C(\Omega)$ be the space of all complex-valued continuous functions on Ω, and let E be a spectral measure, defined on the family of all Borel subsets of Ω, with values in $\mathcal{B}(\mathcal{H})$. Assume that the measure E is Radon, that is, the scalar positive measures $E_{x,x}$, $x \in \mathcal{H}$, are all Radon measures (see [5] and [16] for details). We shall see that the set $\mathcal{D}_\infty = \cap_{f\in C(\Omega)}\mathcal{D}_f$ is a dense subspace in $\mathcal{H}$ and

$$\mathcal{N}_\infty = \{\Psi(f)|\mathcal{D}_\infty; f \in C(\Omega)\}$$

is a normal algebra.

That $\mathcal{N}_\infty$ is an O^*-algebra follows easily from [18], Theorem 13.24. To prove that $\mathcal{D}_\infty$ is a dense subspace in $\mathcal{H}$, we adapt some ideas from [16], page 225. We fix a function $f \in C(\Omega)$ and a compact set $K \subset \Omega$. Fix also $y = E(K)x \in E(K)\mathcal{H}$. As the function $f\chi_K$ is bounded, where χ_K is the characteristic function of K, we have $\int_\Omega |f|^2 dE_{y,y} = \int_K |f|^2 dE_{y,y} = \int_K |f|^2 dE_{x,x}$, via the standard properties of the spectral measure E (see [18], Theorem 13.24). In particular, $y \in \mathcal{D}_f$, and so $\cup_{K\in\mathcal{K}(\Omega)} E(K)\mathcal{H} \subset \mathcal{D}_\infty$, where $\mathcal{K}(\Omega)$ is the family of all compact subsets of Ω. To prove the density of $\mathcal{D}_\infty$ in $\mathcal{H}$, it sufficient to prove that the set $\cup_{K\in\mathcal{K}(\Omega)} E(K)\mathcal{H}$ is dense in $\mathcal{H}$.

Fix an $x \in \mathcal{H}$. It follows from [5], Proposition 2.1.7, that

$$\sup_{K\in\mathcal{K}(\Omega)} \|E(K)x\|^2 = \sup_{K\in\mathcal{K}(\Omega)} \int_K dE_{x,x} = \int_\Omega dE_{x,x} = \|x\|^2.$$

In particular, there is a sequence $(K_k)_{k\geq 1}$ of compact subsets in Ω such that $\lim_{k\to\infty} \|E(K_k)x\| = \|x\|$. As the vectors $x - E(K_k)x = E(\Omega \setminus K_k)x$ and $E(K_k)x$ are orthogonal, we infer that

$$\|x - E(K_k)x\|^2 = \|x\|^2 - \|E(K_k)x\|^2,$$

for all $k \geq 1$. Therefore, $\lim_{k\to\infty} E(K_k)x = x$, showing the desired density.

We have only to show that the closure of the operator $\Psi(f)|\mathcal{D}_\infty$ equals $\Psi(f)$ for each $f \in C(\Omega)$. With $f \in C(\Omega)$ and $x \in \mathcal{D}_f$, we refine a previous argument. Because we have

$$\sup_{K\in\mathcal{K}(\Omega)} \int_K dE_{x,x} = \int_\Omega dE_{x,x}, \quad \sup_{K\in\mathcal{K}(\Omega)} \int_K |f|^2 dE_{x,x} = \int_\Omega |f|^2 dE_{x,x},$$

(via [5], Proposition 2.1.7), and left side integrals depend increasingly on the compact set K, there is a sequence $(K_k)_{k\geq 1}$ of compact subsets in Ω such that $\lim_{k\to\infty} \|E(K_k)x\| = \|x\|$, and $\lim_{k\to\infty} \|\Psi(f)E(K_k)x\| = \|\Psi(f)x\|$. The orthogonality of the vectors $x - E(K_k)x$ and $E(K_k)x$ on one side, and that of the vectors $\Psi(f)x - \Psi(f)E(K_k)x$ and $\Psi(f)E(K_k)x$ on the other side show, as above, that $\lim_{k\to\infty} E(K_k)x = x$ and $\lim_{k\to\infty} \Psi(f)E(K_k)x = \Psi(f)x$, which implies that the closure of the operator $\Psi(f)|\mathcal{D}_\infty$ is equal $\Psi(f)$.

Consequently, $\mathcal{N}_\infty$ is a normal algebra.

Example 4.5. Let Ω be a topological space of the form $\Omega = \cup_{n\geq 1}\Omega_n$, where $(\Omega_n)_{n\geq 1}$ is an increasing sequence of Borel subsets. Let also $\mathcal{A}$ be an algebra of Borel functions on Ω, containing the constant functions and the complex conjugate of every given function from $\mathcal{A}$. Also assume that $f|_{\Omega_n}$ is bounded for all $f \in \mathcal{A}$ and $n \geq 1$.

Let E be a spectral measure (not necessarily Radon), defined on the family of all Borel subsets of Ω, with values in $\mathcal{B}(\mathcal{H})$. Setting $\mathcal{D} = \cap_{f\in\mathcal{A}}\mathcal{D}_f$, then $\mathcal{D}$ is dense in $\mathcal{H}$ and $\mathcal{N} = \{\Psi(f)|\mathcal{D}; f \in \mathcal{A}\}$ is a normal algebra. Indeed, if $f \in \mathcal{A}$, and so $f|_{\Omega_n}$ is bounded for all $n \geq 1$, we have $E(\Omega_n)x \in \mathcal{D}_f$. In addition, $\lim_{n\to\infty} E(\Omega_n)x = x$ for all $x \in \mathcal{H}$, showing that $\mathcal{D}$ is dense in $\mathcal{H}$. Moreover, if $x \in \mathcal{D}_f$, then the sequence $\Psi(f)E(\Omega_n)x$ is convergent to $\Psi(f)x$, implying $\overline{\Psi(f)|\mathcal{D}} = \Psi(f)$ for each $f \in \mathcal{A}$. It is clear that $\mathcal{N}$ is an O^*-algebra, and therefore, $\mathcal{N}_\infty$ is a normal algebra.

Lemma 4.6. *If $\mathcal{A}$ is a normal algebra and N_1, N_2 are arbitrary elements, then $\bar{N}_1, \bar{N}_2$ commute.*

Proof. We first note that if $S \in \mathcal{N}$ is symmetric, then $\bar{S}$ is self-adjoint, which is clear because $\bar{S}$ is normal. Note also that for an arbitrary $N \in \mathcal{N}$ we have $NN^{\#} = N^{\#}N$, as a consequence of the fact that $\bar{N}(\bar{N})^* = (\bar{N})^*\bar{N}$.

Now, let $S_1, S_2 \in \mathcal{N}$ be symmetric. Note that

$$2i(S_1S_2 - S_2S_1)x = (S_1 - iS_2)(S_1 + iS_2)x - (S_1 + iS_2)(S_1 - iS_2)x$$
$$= (\overline{S_1 + iS_2})^*\overline{(S_1 + iS_2)}x - \overline{(S_1 + iS_2)}(\overline{S_1 + iS_2})^*x = 0,$$

for all $x \in \mathcal{D}$, since $\overline{(S_1 + iS_2)}$ is normal and $(\overline{S_1 + iS_2})^* = (S_1 + iS_2)^* \supset (S_1 - iS_2)$. It follows from Proposition 7.1.3 in [20] that $\bar{S}_1, \bar{S}_2$ commute.

If $N_1, N_2 \in \mathcal{N}$ are arbitrary, we write $N_1 = S_{11} + iS_{12}$, $N_2 = S_{21} + iS_{22}$, with S_{jk} symmetric. The previous argument shows that the self-adjoint operators $\bar{S}_{11}$, $\bar{S}_{12}$, $\bar{S}_{21}$, $\bar{S}_{22}$ commute. Therefore, the normal operators $\bar{N}_1, \bar{N}_2$ also commute. □

Lemma 4.7. *Let $A_1, \ldots, A_n$ be commuting self-adjoint operators and let $U_1, \ldots, U_n$ be the Cayley transforms of $A_1, \ldots, A_n$ respectively. Then we have:*

$$\prod_{j=1}^{n}(2 - U_j - U_j^*)^{k_j} = \prod_{j=1}^{n} 4^{k_j}(1 + A_j^2)^{-k_j},$$

for all integers $k_1 \geq 1, \ldots, k_n \geq 1$.

In particular, the operator $\prod_{j=1}^{n}(2 - U_j - U_j^*)^{k_j}$ is positive and injective for all integers $k_1 \geq 1, \ldots, k_n \geq 1$.

Proof. We have the equality $U_j = (A_j - i)(A_j + i)^{-1}$ for all indices j, by the definition of Cayley transform, whence $1 - U_j = 2i(A_j + i)^{-1}$. Because one has $(2i(A_j + i)^{-1})^* = -2i(A_j - i)^{-1}$, via the self-adjointness of A_j, we infer $1 - U_j^* = -2i(A_j - i)^{-1}$. Therefore,

$$2 - U_j - U_j^* = 2i(A_j + i)^{-1}((A_j - i) - (A_j + i))(A_j - i)^{-1} = 4(A_j^2 + 1)^{-1}.$$

Taking into account that the operators $(A_j^2 + 1)^{-1}$, $j = 1, \ldots, n$, are all bounded and commute, we deduce easily the stated formula. The remaining assertions are direct consequences of that formula. □

Theorem 4.8. *Let $\mathcal{N} \subset \mathcal{L}^{\#}(\mathcal{L})$ be a normal algebra. Then there exists a family of indices $\mathcal{I}$, a compact subspace $\Omega \subset \mathbb{T}^{\mathcal{I}}$, a set of denominators $\mathcal{M} \subset C(\Omega)$, and an injective $*$-homomorphism $\mathcal{N} \ni N \mapsto \phi_N \in C(\Omega)/\mathcal{M}$.*

In addition, there exists a uniquely determined spectral measure F on the Borel subsets of Ω such that ϕ_N is F-almost everywhere defined and

$$\langle Nx, y\rangle = \int_{\Omega} \phi_N(z)dF_{x,y}(z),\ x \in \mathcal{D},\ y \in \mathcal{H}.$$

Proof. Let $A = (A_\iota)_{\iota\in\mathcal{I}}$ be a family of Hermitian generators of the algebra $\mathcal{N}$. Such a family obviously exists because $N = (N + N^{\#})/2 + i(N - N^{\#})/2i$ and both $(N + N^{\#})/2$, $(N - N^{\#})/2i$ are Hermitian. Every A_ι can be associated with the Cayley transform U_ι of $\bar{A}_\iota$. As in the proof of Theorem 4.2, since the self-adjoint operators $(\bar{A}_\iota)_{\iota\in\mathcal{I}}$ mutually commute, the corresponding Cayley transforms, as well as their adjoints, mutually commute. Let $\mathcal{B}$ be the closed unital algebra generated by the family $(U_\iota, U_\iota^*)_{\iota\in\mathcal{I}}$ in $\mathcal{B}(\mathcal{H})$, which is a commutative unital C^*-algebra.

It follows from Lemma 4.6 that each operator $2-U_\iota-U_\iota^*$ is injective. If $\tilde{\mathcal{M}}$ is the family of all possible finite products of operators of the form $2-U_\iota-U_\iota^*$, $\iota\in\mathcal{I}$, and the identity 1, then $\tilde{\mathcal{M}}$ is a set of denominators consisting of positive operators.

We define a map from $\mathcal{N}$ to $\mathcal{B}/\tilde{\mathcal{M}}$ by associating to each generator A_ι the fraction $i(U_\iota-U_\iota^*)/(2-U_\iota-U_\iota^*)$, and extending this assignment to $\mathcal{N}$ by linearity and multiplicativity. Let us do this operation properly.

Setting $A^0 = 1$ for $0=(0)_{\iota\in\mathcal{I}}$ and $A^\alpha = \prod_{\alpha_\iota\neq 0} A_\iota^{\alpha_\iota}$ for $\alpha=(\alpha_\iota)_{\iota\in\mathcal{I}}\in\mathbb{Z}_+^{(\mathcal{I})}$, $\alpha\neq 0$, we may define $p(A)=\sum_{\alpha\in\mathcal{J}} c_\alpha A^\alpha$ for every polynomial $p(t)=\sum_{\alpha\in\mathcal{J}} c_\alpha t^\alpha$ in $\mathcal{P}_\mathcal{I}$. As a matter of fact, we have the equality $\mathcal{N}=\{p(A);\, p\in\mathcal{P}_\mathcal{I}\}$ because A is family of generators of $\mathcal{N}$. If we put $W_\iota = i(U_\iota-U_\iota^*)/(2-U_\iota-U_\iota^*)$ and $W=(W_\iota)_{\iota\in\mathcal{I}}$, then $\Phi(p(A))=p(W)$ for each polynomial p. To show that this definition does not depend on a particular representation of $p(A)$, it suffices to show that $p(W)=0$ implies $p(A)=0$ for a fixed polynomial $p\in\mathcal{P}_\mathcal{I}$ such that $p(A)=\sum_{\alpha\in\mathcal{J}} c_\alpha A^\alpha$. Let $\beta\in\mathbb{Z}_+^{(\mathcal{I})}$ be such that $\alpha_\iota\le\beta_\iota$ for all $\iota\in\mathcal{I}$ and $\alpha\in\mathcal{J}$. Also set $V_\gamma=\prod_\iota(2-U_\iota-U_\iota^*)^{\gamma_\iota}$ and $D_\alpha=\prod_\iota(i(U_\iota-U_\iota^*))^{\alpha_\iota}$ for all γ and α in $\mathbb{Z}_+^{(\mathcal{I})}$. Then

$$V_\beta p(W)=\sum_{\alpha\in\mathcal{J}} c_\alpha V_\beta W^\alpha=\sum_{\alpha\in\mathcal{J}} c_\alpha V_{\beta-\alpha}D_\alpha=0.$$

It follows from Lemma 4.7 and the fact that $\mathcal{D}$ is invariant under A^α for all α, that the operator V_β^{-1} is defined on $\mathcal{D}$ and leaves it invariant, for all β. Consequently, for an arbitrary $x\in\mathcal{D}$ we have:

$$p(A)x=\sum_{\alpha\in\mathcal{J}} c_\alpha D_\alpha V_\alpha^{-1}x=\sum_{\alpha\in\mathcal{J}} c_\alpha V_{\beta-\alpha}D_\alpha V_\beta^{-1}x=0.$$

This allows us to define correctly an injective unital $*$-homomorphism from the algebra $\mathcal{N}$ into the algebra $\mathcal{B}/\tilde{\mathcal{M}}$.

As in the proof of Theorem 4.2, the space $\Gamma(\mathcal{B})$ of characters of the algebra $\mathcal{B}$ may be identified with a closed subspace Ω of the compact space $\mathbb{T}^\mathcal{I}$. Then $\mathcal{B}$ is identified with $C(\Omega)$, and the function $\hat{U}_\iota$ with the coordinate function $\mathbb{T}^\mathcal{I}\ni z\to z_\iota\in\mathbb{C}$ for all ι. Hence, the set of denominators $\mathcal{M}$ corresponding to $\tilde{\mathcal{M}}$ will be the set of all possible finite products of functions of the form $2-z_\iota-\bar{z}_\iota$, $\iota\in\mathcal{I}$, and the constant function 1. As noticed in the Introduction, the algebra of fractions $\mathcal{B}/\tilde{\mathcal{M}}$ can be identified with the algebra of fractions $C(\Omega)/\mathcal{M}$. The image of the algebra $\mathcal{N}$ in $C(\Omega)/\mathcal{M}$ will be the unital algebra generated by the fractions $\theta_\iota(z)=-\Im z_\iota/(1-\Re z_\iota)$, $\iota\in\mathcal{I}$. Specifically, if $N\in\mathcal{N}$ is arbitrary, and if $p_N\in\mathcal{P}_\mathcal{I}$ is a polynomial such that $N=p_N(A)$, then, setting $\phi_N=p_N\circ\theta$, where $\theta=(\theta_\iota)_{\iota\in\mathcal{I}}$, the map $\mathcal{M}\ni N\mapsto\phi_N\in C(\Omega)/\mathcal{M}$ is a $*$-homomorphism.

As in the proof of Theorem 4.2, there exists a spectral measure F defined on the Borel subsets of Ω such that $U_\iota=\int_\Omega z_\iota dF(z)$ for all ι. Moreover, if $C_\iota=\{z\in\Omega;\, z_\iota=1\}$, we have $F(C_\iota)=0$ for all ι. This shows that the function ϕ_N, which is not defined on a finite union of the sets C_ι, is almost everywhere defined.

Fix an $N \in \mathcal{N}$, $N = p_N(A)$, with $p_N(t)) = \sum_{\alpha \in \mathcal{J}} c_\alpha t^\alpha$. Because we have

$$\langle \bar{A}_\iota x, y\rangle = \int_\Omega \theta_\iota(z) dF_{x,y}(z),\ x \in \mathcal{D}(\bar{A}_\iota),\ y \in \mathcal{H},$$

we can write the following equations:

$$\begin{aligned}
\int_\Omega \phi_N(z) dF_{x,y}(z) &= \int_\Omega (p_N \circ \theta)(z) dF_{x,y}(z) \\
&= \sum_{\alpha \in \mathcal{J}} c_\alpha \int_\Omega \prod_{\alpha_\iota \neq 0} \theta_\iota^{\alpha_\iota}(z) dF_{x,y}(z) \\
&= \sum_{\alpha \in \mathcal{J}} c_\alpha \int_\Omega \prod_{\alpha_\iota \neq 0} A_\iota^{\alpha_\iota}(z) dF_{x,y}(z) \\
&= \langle p_N(A) x, y\rangle \\
&= \langle N x, y\rangle,
\end{aligned}$$

for all $x \in \mathcal{D}$ and $y \in \mathcal{H}$, via the usual properties of the unbounded functional calculus (see [18]).

The measure F is uniquely determined on Ω, via the uniqueness in Theorem 4.2. $\square$

Remark 4.9. We may apply Theorem 4.3 to the family $\{\bar{N}; N \in \mathcal{N}\}$, $\mathcal{N}$ a normal algebra, which is a commuting family of normal operators in $\mathcal{H}$. According to this result, there exists a unique spectral measure E on the Borel subsets of $(\mathbb{C}_\infty)^{\mathcal{N}}$ such that each coordinate function $(\mathbb{C}_\infty)^{\mathcal{N}} \ni z \to z_N \in \mathbb{C}_\infty$ is E-almost everywhere finite, and

$$\langle \bar{N} x, y\rangle = \int_{(\mathbb{C}_\infty)^{\mathcal{N}}} z_N dE_{x,y}(z),\ x \in \mathcal{D}(\bar{N}),\ y \in \mathcal{H},\ N \in \mathcal{N}.$$

5. Normal extensions

In this section we present a version of a Theorem 3.4 in [2], concerning the existence of normal extensions. Unlike in [2], we prove it here for infinitely many operators.

Fix a Hilbert space $\mathcal{H}$, a dense subspace $\mathcal{D}$ of $\mathcal{H}$, and a compact Hausdorff space Ω. As before, we denote by $SF(\mathcal{D})$ the space of all sesquilinear forms on $\mathcal{D}$.

For the convenience of the reader, we shall reproduce some statements from [2], to be used in the sequel. We start by recalling some terminology from [2].

Let $\mathcal{Q} \subset C(\Omega)$ be a set of positive denominators. Fix a $q \in \mathcal{Q}$. A linear map $\psi : C(\Omega)/q \to SF(\mathcal{D})$ is called *unital* if $\psi(1)(x, y) = \langle x, y\rangle$, $x, y \in \mathcal{D}$. We say that ψ is *positive* if $\psi(f)$ is positive semidefinite for all $f \in (C(\Omega)/q)_+$.

More generally, let $\mathcal{Q}_0 \subset \mathcal{Q}$ be nonempty. Let $\mathcal{C} = \sum_{q \in \mathcal{Q}_0} C(\Omega)/q$, and let $\psi : \mathcal{C} \to SF(\mathcal{D})$ be linear. The map ψ is said to be *unital* (resp. *positive*) if $\psi|C(\Omega)/q$ is unital (resp. positive) for all $q \in \mathcal{Q}_0$.

We start with a part from Theorem 2.2 in [2].

Theorem B. *Let $\mathcal{Q}_0 \subset \mathcal{Q}$ be nonempty, let $\mathcal{C} = \sum_{q\in\mathcal{Q}_0} C(\Omega)/q$, and let $\psi : \mathcal{C} \to SF(\mathcal{D})$ be linear and unital. The map ψ is positive if and only if*

$$\sup\{|\psi(hq^{-1})(x,x)|;\, h \in C(\Omega),\, \|h\|_\infty \le 1\} = \psi(q^{-1})(x,x),\ q \in \mathcal{Q}_0,\ x \in \mathcal{D}.$$

Let again $\mathcal{Q}_0 \subset \mathcal{Q}$ be nonempty and let $\mathcal{F} = \sum_{q\in\mathcal{Q}_0} \mathcal{F}_q$, where $q^{-1} \in \mathcal{F}_q$ and $\mathcal{F}_q$ is a vector subspace of $C(\Omega)/q$ for all $q \in \mathcal{Q}_0$. Let $\phi : \mathcal{F} \mapsto SF(\mathcal{D})$ be linear. Suppose that $\phi(q^{-1})(x,x) > 0$ for all $x \in \mathcal{D}\setminus\{0\}$ and $q \in \mathcal{Q}_0$. Then $\phi(q^{-1})$ induces an inner product on $\mathcal{D}$, and let $\mathcal{D}_q$ be the space $\mathcal{D}$, endowed with the norm given by $\| * \|_q^2 = \phi(q^{-1})(*,*)$.

Let $M_n(\mathcal{F}_q)$ (resp. $M_n(\mathcal{F})$) denote the space of $n \times n$-matrices with entries in $\mathcal{F}_q$ (resp. in $\mathcal{F}$). Note that $M_n(\mathcal{F}) = \sum_{q\in\mathcal{Q}_0} M_n(\mathcal{F}_q)$ may be identified with a subspace of the algebra of fractions $C(\Omega, M_n)/\mathcal{Q}$, where M_n is the C^*-algebra of $n \times n$-matrices with entries in $\mathbb{C}$. Moreover, the map ϕ has a natural extension $\phi^n : M_n(\mathcal{F}) \mapsto SF(\mathcal{D}^n)$, given by

$$\phi^n(\mathbf{f})(\mathbf{x},\mathbf{y}) = \sum_{j,k=1}^{n} \phi(f_{j,k})(x_k, y_j),$$

for all $\mathbf{f} = (f_{j,k}) \in M_n(\mathcal{F})$ and $\mathbf{x} = (x_1,\ldots,x_n), \mathbf{y} = (y_1,\ldots,y_n) \in \mathcal{D}^n$.

Let $\phi_q^n = \phi^n \mid M_n(\mathcal{F}_q)$. Endowing the Cartesian product $\mathcal{D}^n$ with the norm $\|\mathbf{x}\|_q^2 = \sum_{j=1}^n \phi(q^{-1})(x_j,x_j)$ if $\mathbf{x} = (x_1,\ldots,x_n) \in \mathcal{D}^n$, and denoting it by $\mathcal{D}_q^n$, we are in the conditions of Remark 3.8, with M_n for A and $\mathcal{D}_q^n$ for $\mathcal{D}_q$. Hence we say that the map ϕ^n is contractive if $\|\phi_q^n\| \le 1$ for all $q \in \mathcal{Q}_0$. Using the standard norm $\| * \|_n$ in the space of M_n, the space $M_n(\mathcal{F}_q)$ is endowed with the norm $\|(qf_{j,k})\|_{n,\infty} = \sup_{\omega\in\Omega} \|(q(\omega)f_{j,k}(\omega))\|_n$, for all $(f_{j,k}) \in M_n(\mathcal{F}_q)$.

Following [3] and [13] (see also [12]), we shall say that the map $\phi : \mathcal{F} \mapsto SF(\mathcal{D})$ is *completely contractive* if the map $\phi^n : M_n(\mathcal{F}) \mapsto SF(\mathcal{D}^n)$ is contractive for all integers $n \ge 1$.

Note that a linear map $\phi : \mathcal{F} \mapsto SF(\mathcal{D})$ with the property $\phi(q^{-1})(x,x) > 0$ for all $x \in \mathcal{D}\setminus\{0\}$ and $q \in \mathcal{Q}_0$ is completely contractive if and only if for all $q \in \mathcal{Q}_0$, $n \in \mathbb{N}$, $x_1,\ldots,x_n, y_1,\ldots,y_n \in \mathcal{D}$ with

$$\sum_{j=1}^{n} \phi(q^{-1})(x_j,x_j) \le 1, \quad \sum_{j=1}^{n} \phi(q^{-1})(y_j,y_j) \le 1,$$

and for all $(f_{j,k}) \in M_n(\mathcal{F}_q)$ with $\|(qf_{j,k})\|_{n,\infty} \le 1$, we have

$$\left| \sum_{j,k=1}^{n} \phi(f_{j,k})(x_k,y_j) \right| \le 1. \tag{5.1}$$

Let us now recall the main result of [2], namely Theorem 2.5, in a shorter form (see also [21] and [25] for some particular cases).

Theorem C. *Let Ω be a compact space and let $\mathcal{Q} \subset C(\Omega)$ be a set of positive denominators. Let also $\mathcal{Q}_0$ be a cofinal subset of $\mathcal{Q}$, with $1 \in \mathcal{Q}_0$. Consider $\mathcal{F} = \sum_{q\in\mathcal{Q}_0} \mathcal{F}_q$, where $\mathcal{F}_q$ is a vector subspace of $C(\Omega)/q$ such that $r^{-1} \in \mathcal{F}_r \subset \mathcal{F}_q$ for all $r \in \mathcal{Q}_0$ and $q \in \mathcal{Q}_0$, with $r|q$. Let also $\phi : \mathcal{F} \to SF(\mathcal{D})$ be linear and unital, and set $\phi_q = \phi|\mathcal{F}_q$, $\phi_{q,x}(*) = \phi_q(*)(x,x)$ for all $q \in \mathcal{Q}_0$ and $x \in \mathcal{D}$.*

The following two statements are equivalent:

(a) *The map ϕ extends to a unital, positive, linear map ψ on $C(\Omega)/\mathcal{Q}$ such that, for all $x \in \mathcal{D}$ and $q \in \mathcal{Q}_0$, we have:*

$$\|\psi_{q,x}\| = \|\phi_{q,x}\|, \quad \textit{where} \quad \psi_q = \psi|C(\Omega)/q,\ \psi_{q,x}(*) = \psi_q(*)(x,x).$$

(b) (i) *$\phi(q^{-1})(x,x) > 0$ for all $x \in \mathcal{D} \setminus \{0\}$ and $q \in \mathcal{Q}_0$.*
(ii) *The map ϕ is completely contractive.*

Remark. A "minimal" subspace of $C(\Omega)/\mathcal{Q}$ to apply Theorem C is obtained as follows. If $\mathcal{Q}_0$ is a cofinal subset of $\mathcal{Q}$ with $1 \in \mathcal{Q}_0$, we define $\mathcal{F}_q$ for some $q \in \mathcal{Q}_0$ to be the vector space generated by all fractions of the form r/q, where $r \in \mathcal{Q}_0$ and $r|q$. It is clear that the subspace $\mathcal{F} = \sum_{q\in\mathcal{Q}_0} \mathcal{F}_q$ has the properties required to apply Theorem C.

We also need Corollary 2.7 from [2].

Corollary D. *Suppose that, with the hypotheses of Theorem* B, *condition* (b) *is satisfied. Then there exists a positive $B(\mathcal{H})$-valued measure F on the Borel subsets of Ω such that*

$$\phi(f)(x,y) = \int_\Omega f\, dF_{x,y},\ \ f \in \mathcal{F},\ \ x,y \in \mathcal{D}. \tag{5.2}$$

For every such measure F and every $q \in \mathcal{Q}_0$, we have $F(Z(q)) = 0$.

Example 5.1. We extend to infinitely many variables the Example 3.5. Let $\mathcal{I}$ be a (nonempty) family of indices. Denote by $z = (z_\iota)_{\iota\in\mathcal{I}}$ the independent variable in $\mathbb{C}^{\mathcal{I}}$. Let also $\bar{z} = (\bar{z}_\iota)_{\iota\in\mathcal{I}}$. As before, let $\mathbb{Z}_+^{(\mathcal{I})}$ be the set of all collections $\alpha = (\alpha_\iota)_{\iota\in\mathcal{I}}$ of nonnegative integers, with finite support. Setting $z^0 = 1$ for $0 = (0)_{\iota\in\mathcal{I}}$ and $z^\alpha = \prod_{\alpha_\iota\neq 0} z_\iota^{\alpha_\iota}$ for $z = (z_\iota)_{\iota\in\mathcal{I}} \in \mathbb{C}^{\mathcal{I}}$, $\alpha = (\alpha_\iota)_{\iota\in\mathcal{I}} \in \mathbb{Z}_+^{(\mathcal{I})}$, $\alpha \neq 0$, we may consider the algebra of complex-valued functions $\mathcal{S}_{\mathcal{I}}$ on $\mathbb{C}^{\mathcal{I}}$, consisting of expressions of the form $\sum_{\alpha,\beta\in\mathcal{J}} c_{\alpha,\beta} z^\alpha \bar{z}^\beta$, with $c_{\alpha,\beta}$ complex numbers for all $\alpha, \beta \in \mathcal{J}$, where $\mathcal{J} \subset \mathbb{Z}_+^{(\mathcal{I})}$ is finite.

We can embed the space $\mathcal{S}_{\mathcal{I}}$ into the algebra of fractions derived from the basic algebra $C((\mathbb{C}_\infty)^{\mathcal{I}})$, using a suitable set of denominators. Specifically, we consider the family $\mathcal{R}_{\mathcal{I}}$ consisting of all rational functions of the form $r_\alpha(t) = \prod_{\alpha_\iota\neq 0}(1+|z_\iota|^2)^{-\alpha_\iota}$, $z = (z_\iota)_{\iota\in\mathcal{I}} \in \mathbb{C}^{\mathcal{I}}$, where $\alpha = (\alpha_\iota) \in \mathbb{Z}_+^{(\mathcal{I})}$, $\alpha \neq 0$, is arbitrary. Of course, we set $r_0 = 1$. The function r_α can be continuously extended to $(\mathbb{C}_\infty)^{\mathcal{I}} \setminus \mathbb{C}^{\mathcal{I}}$ for all $\alpha \in \mathbb{Z}_+^{(\mathcal{I})}$. In fact, actually the function $f_{\beta,\gamma}(z) = z^\beta \bar{z}^\gamma r_\alpha(z)$ can be continuously extended to $(\mathbb{C}_\infty)^{\mathcal{I}} \setminus \mathbb{C}^{\mathcal{I}}$ whenever $\beta_\iota + \gamma_\iota < 2\alpha_\iota$, and $\beta_\iota = \gamma_\iota = 0$ if $\alpha_\iota = 0$, for all $\iota \in \mathcal{I}$ and $\alpha, \beta, \gamma \in \mathbb{Z}_+^{(\mathcal{I})}$. Moreover, the family $\mathcal{R}_{\mathcal{I}}$ becomes a set of denominators in $C((\mathbb{C}_\infty)^{\mathcal{I}})$. This shows that the space $\mathcal{S}_{\mathcal{I}}$ can be embedded into the algebra of fractions $C((\mathbb{C}_\infty)^{\mathcal{I}})/\mathcal{R}_{\mathcal{I}}$.

To be more specific, for all $\alpha \in \mathbb{Z}_+^{(\mathcal{I})}$, $\alpha \neq 0$, we denote by $\mathcal{S}_{\mathcal{I},\alpha}^{(1)}$ the linear spaces generated by the monomials $z^\beta \bar{z}^\gamma$, with $\beta_\iota + \gamma_\iota < 2\alpha_\iota$ whenever $\alpha_\iota > 0$, and $\beta_\iota = \gamma_\iota = 0$ if $\alpha_\iota = 0$. Put $\mathcal{S}_{\mathcal{I},0}^{(1)} = \mathbb{C}$.

We also define $\mathcal{S}_{\mathcal{I},\alpha}^{(2)}$, for $\alpha \in \mathbb{Z}_+^{(\mathcal{I})}$, $\alpha \neq 0$, to be the linear space generated by the monomials $|z|^{2\beta} = \prod_{\beta_\iota \neq 0}(z_\iota \bar{z}_\iota)^{\beta_\iota}$, $0 \neq \beta$, $\beta_\iota \leq \alpha_\iota$ for all $\iota \in \mathcal{I}$ and $|z| = (|z_\iota|)_{\iota\in\mathcal{I}}$. We define $\mathcal{S}_{\mathcal{I},0}^{(2)} = \{0\}$.

Set $\mathcal{S}_{\mathcal{I},\alpha} = \mathcal{S}_{\mathcal{I},\alpha}^{(1)} + \mathcal{S}_{\mathcal{I},\alpha}^{(2)}$ for all $\alpha \in \mathbb{Z}_+^{(\mathcal{I})}$. Note that, if $f \in \mathcal{S}_{\mathcal{I},\alpha}$, the function $r_\alpha f$ extends continuously to $(\mathbb{C}_\infty)^{\mathcal{I}}$ and that $\mathcal{S}_{\mathcal{I},\alpha} \subset \mathcal{S}_{\mathcal{I},\beta}$ if $\alpha_\iota \leq \beta_\iota$ for all $\iota \in \mathcal{I}$.

It is now clear that the algebra $\mathcal{S}_{\mathcal{I}} = \sum_{\alpha\in\mathbb{Z}_+^{(\mathcal{I})}} \mathcal{S}_{\mathcal{I},\alpha}$ can be identified with a subalgebra of $C((\mathbb{C}_\infty)^{\mathcal{I}})/\mathcal{R}_{\mathcal{I}}$. This algebra has the properties of the space $\mathcal{F}$ appearing in the statement of Theorem C.

Let now $T = (T_\iota)_{\iota\in\mathcal{I}}$ be a family of linear operators defined on a dense subspace $\mathcal{D}$ of a Hilbert space $\mathcal{H}$ such that $T_\iota(\mathcal{D}) \subset \mathcal{D}$ and $T_\iota T_\kappa x = T_\kappa T_\iota x$ for all $\iota, \kappa \in \mathcal{I}$, $x \in \mathcal{D}$.

Setting T^α as in the case of complex monomials, which is possible because of the commutativity of the family T on $\mathcal{D}$, we may define a unital linear map $\phi_T : \mathcal{S}_{\mathcal{I}} \to SF(\mathcal{D})$ by

$$\phi_T(z^\alpha \bar{z}^\beta)(x, y) = \langle T^\alpha x, T^\beta y\rangle\ , \quad x, y \in \mathcal{D},\ \alpha, \beta \in \mathbb{Z}_+^{(\mathcal{I})}, \tag{5.3}$$

which extends by linearity to the subspace $\mathcal{S}_{\mathcal{I}}$ generated by these monomials.

An easy proof shows that, for all α, β in $\mathbb{Z}_+^{(\mathcal{I})}$ with $\beta - \alpha \in \mathbb{Z}_+^{(\mathcal{I})}$, and $x \in \mathcal{D} \setminus \{0\}$, we have

$$0 < \langle x, x\rangle \leq \phi_T(r_\alpha^{-1})(x, x) \leq \phi_T(r_\beta^{-1})(x, x)\,. \tag{5.4}$$

The polynomial r_α^{-1} will be denoted by s_α for all $\alpha \in \mathbb{Z}_+^{(\mathcal{I})}$.

The family $T = (T_\iota)_{\iota\in\mathcal{I}}$ is said to have a *normal extension* if there exist a Hilbert space $\mathcal{K} \supset \mathcal{H}$ and a family $N = (N_\iota)_{\iota\in\mathcal{I}}$ consisting of commuting normal operators in $\mathcal{K}$ such that $\mathcal{D} \subset \mathcal{D}(N_\iota)$ and $N_\iota x = T_\iota x$ for all $x \in \mathcal{D}$ and $\iota \in \mathcal{I}$.

A family $T = (T_\iota)_{\iota\in\mathcal{I}}$ having a normal extension is also called a *subnormal family* (see, for instance, [1]).

The following result is a version of Theorem 3.4 from [2], valid for an arbitrary family of operators. We mention that the basic space of fractions from [2] is slightly modified.

Theorem 5.2. *Let $T = (T_\iota)_{\iota\in\mathcal{I}}$ be a family of linear operators defined on a dense subspace $\mathcal{D}$ of a Hilbert space $\mathcal{H}$. Assume that $\mathcal{D}$ is invariant under T_ι for all $\iota \in \mathcal{I}$ and that T is a commuting family on $\mathcal{D}$. The family T admits a normal* *extension if and only if the map* $\phi_T : \mathcal{S}_\mathcal{I} \mapsto SF(\mathcal{D})$ *has the property that for all* $\alpha \in \mathbb{Z}_+^{(\mathcal{I})}$, $m \in \mathbb{N}$ *and* $x_1, \ldots, x_m, y_1, \ldots, y_m \in \mathcal{D}$ *with*

$$\sum_{j=1}^m \phi_T(s_\alpha)(x_j, x_j) \leq 1, \quad \sum_{j=1}^m \phi_T(s_\alpha)(y_j, y_j) \leq 1,$$

and for all $p = (p_{j,k}) \in M_m(\mathcal{S}_{\mathcal{I},\alpha})$ *with* $\sup_z \|r_\alpha(z)p(z)\|_m \leq 1$*, we have*

$$\left| \sum_{j,k=1}^m \phi_T(p_{j,k})(x_k, y_j) \right| \leq 1. \tag{5.5}$$

Proof. We follow the lines of the proof of Theorem 3.4 in [2].

If the condition of the theorem is fulfilled, and so we have a linear and unital map $\phi_T : \mathcal{S}_\mathcal{I} \to SF(\mathcal{D})$ induced by (5.3), then conditions (i) (by (5.4)) and (ii) of Theorem C are satisfied for ϕ_T. Hence, by that theorem and Corollary D, there exists a positive $\mathcal{B}(\mathcal{H})$-valued measure F on the Borel sets of $\Omega = (\mathbb{C}_\infty)^\mathcal{I}$, such that (5.2) holds for ϕ_T. Because ϕ_T is unital, $F(\Omega)$ is the identity operator on $\mathcal{H}$. By the classical Naimark dilation theorem (see, for instance [12]), there exists a Hilbert space $\mathcal{K}$ containing $\mathcal{H}$ as a closed subspace and a spectral measure E on the Borel subsets of Ω with values in $\mathcal{B}(\mathcal{K})$, such that $F(*) = PE(*)|\mathcal{H}$, where P denotes the orthogonal projection from $\mathcal{K}$ onto $\mathcal{H}$. As in Remark 4.1, for each $\iota \in \mathcal{I}$, let N_ι be the normal operator with domain

$$\mathcal{D}(N_\iota) = \left\{ x \in \mathcal{K}\,;\, \int_{K_\iota} |z_\iota|^2 dE_{x,x}(z) \} < \infty \right\}$$

and

$$N_\iota x = \int_{K_\iota} z_\iota dE(z)x\,, \quad x \in \mathcal{D}(N_\iota)\,,$$

where $K_\iota = \{z \in \Omega; z_\iota \neq \infty\}$. For all $x, y \in \mathcal{D}$, $\iota \in \mathcal{I}$, we have

$$\langle PN_\iota x, y\rangle = \langle N_\iota x, y\rangle = \int_{K_\iota} z_\iota dE_{x,y}(z)$$

$$= \int_\Omega z_\iota dF_{x,y}(z) = \phi_T(z_\iota)(x,y) = \langle T_\iota x, y\rangle\,,$$

because $F(K_\iota) = F(\Omega)$. Indeed, $F(\Omega \setminus K_\iota) = F(\{z \in \Omega; z_\iota = \infty\}) = F(Z((1+|z_\iota|^2)^{-1}) = 0$, by Corollary D. Hence, $PN_\iota x = T_\iota x$ for all $x \in \mathcal{D}$, $\iota \in \mathcal{I}$. Note also that

$$\|T_\iota x\|^2 = \phi_T(|z_\iota|^2)(x,x) = \int_\Omega |z_\iota|^2 dF_{x,x}(z) =$$

$$\int_{K_\iota} |z_\iota|^2 dE_{x,x}(z) = \|N_\iota x\|^2.$$

for all $x \in \mathcal{D}$, $\iota \in \mathcal{I}$, which shows that $N = (N_\iota)_{\iota \in \mathcal{I}}$ is a normal extension of $T = (T_\iota)_{\iota \in \mathcal{I}}$, via the following:

Remark. Let $S : \mathcal{D}(S) \subset \mathcal{H} \mapsto \mathcal{H}$ be an arbitrary linear operator. If $B : \mathcal{D}(B) \subset \mathcal{K} \mapsto \mathcal{K}$ is a normal operator such that $\mathcal{H} \subset \mathcal{K}$, $\mathcal{D}(S) \subset \mathcal{D}(B)$, $Sx = PBx$ and $\|Sx\| = \|Bx\|$ for all $x \in \mathcal{D}(S)$, where P is the projection of $\mathcal{K}$ onto $\mathcal{H}$, then we have $Sx = Bx$ for all $x \in \mathcal{D}(S)$. Indeed, $\langle Sx, Sx\rangle = \langle Sx, Bx\rangle$ and $\langle Bx, Sx\rangle = \langle PBx, Sx\rangle = \langle Sx, Sx\rangle = \langle Bx, Bx\rangle$. Hence, we have $\|Sx - Bx\| = 0$ for all $x \in \mathcal{D}(S)$ (see [2], Remark 3.1).

We continue the proof of Theorem 5.2.

Conversely, if $T = (T_\iota)_{\iota\in\mathcal{I}}$ admits a normal extension $N = (N_\iota)_{\iota\in\mathcal{I}}$, the latter has a spectral measure E with support in $\Omega = (\mathbb{C}_\infty)^{\mathcal{I}}$, via Theorem 4.3. Then for all $\alpha \in \mathbb{Z}_+^{(\mathcal{I})}$, the space $\mathcal{D}$ is contained in

$$\mathcal{D}(T^\alpha) \subset \mathcal{D}(N^\alpha) = \left\{x \in \mathcal{K}\,;\ \int_\Omega |z^\alpha|^2 dE_{x,x}(z) < \infty\right\}.$$

It follows that, for all $f \in C(\Omega)/r_\alpha$, the function f is integrable on Ω with respect to the positive scalar measure $E_{x,x}$. Using the decomposition $4E_{x,y} = E_{x+y,x+y} - E_{x-y,x-y} + iE_{x+iy,x+iy} - iE_{x-iy,x-iy}$, we see that $\psi : C(\Omega)/\mathcal{R}_{\mathcal{I}} \mapsto SF(\mathcal{D})$, defined by

$$\psi(f)(x,y) = \int_\Omega f(z)dE_{x,y}(z),\ x,y \in \mathcal{D},\ f \in C(\Omega)/\mathcal{R}_{\mathcal{I}},$$

is a linear map which is obviously unital and positive. Moreover,

$$\psi(z^\alpha \bar{z}^\beta)(x,y) = \langle N^\alpha x, N^\beta y\rangle = \langle T^\alpha x, T^\beta y\rangle = \phi_T(z^\alpha \bar{z}^\beta)(x,y),$$

for all $x, y \in \mathcal{D}$ and $\alpha \in \mathbb{Z}_+^{(\mathcal{I})}$, because N^α extends T^α, showing that ψ is an extension of ϕ_T.

Setting $\phi = \phi_T$, $\psi_{r_\alpha} = \psi|C(\Omega)/r_\alpha$, $\psi_{r_\alpha,x}(*) = \psi_{r_\alpha}(*)(x,x)$, $\phi_{r_\alpha} = \phi|\mathcal{S}_{r_\alpha}$, $\phi_{r_\alpha,x}(*) = \psi_{r_\alpha}(*)(x,x)$ for all $\alpha \in \mathbb{Z}_+^{(\mathcal{I})}$ and $x \in \mathcal{D}$, we have:

$$\phi(r_\alpha^{-1})(x,x) = \psi(r_\alpha^{-1})(x,x) = \|\psi_{r_\alpha,x}\| \ge \|\phi_{r_\alpha,x}\| \ge \phi(r_\alpha^{-1})(x,x),$$

via Theorem B. This shows that the map $\phi : \mathcal{S}_{\mathcal{I}} \to SF(\mathcal{D})$ satisfies condition (a) in Theorem C. We infer that the condition in the actual statement, derived from condition (b) in Theorem C, should be also satisfied. This completes the proof of Theorem 5.2. □

Remark 5.3. Let $T = (T_\iota)_{\iota\in\mathcal{I}}$ be a family of linear operators defined on a dense subspace $\mathcal{D}$ of a Hilbert space $\mathcal{H}$. Assume that $\mathcal{D}$ is invariant under T_ι and that T is a commuting family on $\mathcal{D}$. If the map $\phi_T : \mathcal{S}_{\mathcal{I}} \mapsto SF(\mathcal{D})$ has the property (5.5), the family has a proper quasi-invariant subspace. In other words, there exists a proper Hilbert subspace $\mathcal{L}$ of the Hilbert space $\mathcal{H}$ such that the subspace $\{x \in \mathcal{D}(T_\iota) \cap \mathcal{L}; Tx \in \mathcal{L}\}$ is dense in in $\mathcal{L}$ for each $\iota \in \mathcal{I}$. This is a consequence of Theorem 5.3 and Theorem 11 from [1].

We use Example 3.5 for the particular case of a single operator. We take $\mathcal{I} = \{1\}$ and put $\mathcal{S}_{\mathcal{I}} = \mathcal{S}_1$, which is the set of all polynomials in z and $\bar{z}$, $z \in \mathbb{C}$. The set $\mathcal{R}_{\mathcal{I}} = \mathcal{R}_1$ consists of all functions of the form $r_k(z) = (1 + |z|^2)^{-k}$, with $z \in \mathbb{C}$ and $k \in \mathbb{Z}_+$.

Considering a single operator S, we may define a unital linear map $\phi_S : \mathcal{S}_1 \to SF(\mathcal{D})$ by

$$\phi_S(z^j \bar{z}^k)(x,y) = \langle S^j x, S^k y\rangle \ , \quad x, y \in \mathcal{D},\ j \in \mathbb{Z}_+,$$

extended by linearity to the subspace $\mathcal{S}_1$. The next result is a version of Corollary 3.5 from [2] (stated for a different basic space of fraction).

Corollary 5.4. *Let $S : \mathcal{D}(S) \subset \mathcal{H} \mapsto \mathcal{H}$ be a densely defined linear operator such that $S\mathcal{D}(S) \subset \mathcal{D}(S)$. The operator S admits a normal extension if and only if for all $m \in \mathbb{Z}_+$, $n \in \mathbb{N}$ and $x_1, \dots, x_n, y_1, \dots, y_n \in \mathcal{D}(S)$ with*

$$\sum_{j=1}^{n}\sum_{k=0}^{m} \frac{m!}{k!(m-k)!}\langle S^k x_j, S^k x_j\rangle \le 1, \quad \sum_{j=1}^{n}\sum_{k=0}^{m} \frac{m!}{k!(m-k)!}\langle S^k y_j, S^k y_j\rangle \le 1,$$

and for all $p = (p_{j,k}) \in M_n(\mathcal{S}_m)$, with $\sup_{z\in\mathbb{C}} \|(1+|z|^2)^{-m} p(z)\|_n \le 1$, we have

$$\left| \sum_{j,k=1}^{n} \langle \phi_S(p_{j,k}) x_k, y_j\rangle \right| \le 1.$$

Corollary 5.3 is a direct consequence of Theorem 5.2.

The case of one operator, covered by our Corollary 5.3, is also studied in [22], via a completely different approach.

References

[1] E. Albrecht and F.-H. Vasilescu, *Invariant subspaces for some families of unbounded subnormal operators*, Glasgow Math. J. 45 (2003), 53–67.

[2] E. Albrecht and F.-H. Vasilescu, *Unbounded extensions and operator moment problems*, preprint, 2004.

[3] W.B. Arveson, *Subalgebras of C^*-algebras*, Acta. Math. 123 (1969), 141–224.

[4] Ju.M. Berezanskii, *Self-adjoint operators in spaces of functions of an infinite number of variables* (Russian), Naukova Dumka, Kiev, 1978.

[5] C. Berg, J.P.R. Christensen, P. Ressel, *Harmonic analysis on semigroups*, Springer-Verlag, New York/Berlin/Heidelberg/Tokyo, 1984.

[6] N. Bourbaki, *Algèbre commutative*, Hermann, Paris, 1961.

[7] R.E. Curto and L.A. Fialkow, *Solution of the truncated complex moment problem for flat data*, Memoirs of the A. M. S. Number 568, 1996.

[8] R.E. Curto and L.A. Fialkow, *Flat extensions of positive moment matrices: Recursively generated relations*, Memoirs of the A. M. S. Number 648, 1998.

[9] O. Demanze and A. Mouze, *Dénominateurs universels pour la décomposition de polynômes positifs sur un espace de Hilbert*, J. Operator Theory 56 (2006), 273–290.

[10] M.A. Dritschel, *On factorization on trigonometric polynomials*, Integr. equ. oper. theory 49 (2004), 11–42.

[11] N. Dunford and J.T. Schwartz, *Linear Operators*, Part I, Interscience Publishers, New York, London, 1958; Part II, Interscience Publishers John Wiley and sons, New York, London, 1963; Part III, Wiley-Interscience, New York, 1971.

[12] V. Paulsen, *Completely bounded maps and operator algebras*, Cambridge University Press, Cambridge, 2002.

[13] R.T. Powers, *Selfadjoint algebras of unbounded operators.* II, Trans. Amer. Math. Soc., 187 (1974), 261–293.

[14] M. Putinar and F.-H. Vasilescu, *Solving moment problems by dimensional extension*, Annals of Math. 149 (1999), 1087–1107.

[15] M. Putinar and F.-H. Vasilescu, *A uniqueness criterion in the multivariate moment problem*, Math. Scand., 92 (2003), 295–300.

[16] P. Ressel and W.J. Ricker, *An integral representation for semi-groups of unbounded normal operators*, J. Operator Theory, 47 (2002), 219–243.

[17] A.P. Robertson and W.J. Robertson, *Topological vector spaces*, Cambridge University Press, Cambridge, 1973.

[18] W. Rudin, *Functional Analysis*, McGraw-Hill Book Company, New York, 1973.

[19] Yu.S. Samoilenko, *The spectral theory of sets of selfadjoint operators* (Russian), Naukova Dumka, Kiev, 1984.

[20] K. Schmüdgen, *Unbounded Operator Algebras and Representation Theory*, Operator Theory: Advances and Applications Vol. 37, Birkhäuser Verlag, Basel/Boston/Berlin, 1990.

[21] J. Stochel, *Solving the truncated moment problem solves the full moment problem*, Glasgow Math. J. 43 (2001), 335–341.

[22] J. Stochel and F.H. Szafraniec, *The complex moment problem and subnormality: A polar decomposition approach*, J. Functional Analysis 159 (1998), 432–491.

[23] J. Stochel and F.H. Szafraniec, *Domination of unbounded operators and commutativity*, J. Math. Soc. Japan 55 (2003), no. 2, 405–437.

[24] B.L. van der Waerden, *Algebra*, Vol. 2, Frederick Ungar Publ. Co., New York, 1970.

[25] F.-H. Vasilescu, *Spaces of fractions and positive functionals*, Math. Scand. 96 (2005), 257–279.

[26] M. Woytylak, *Algebras dense in L^2 spaces: an operator approach*, Glasgow Math. J. 47 (2005), 155–165.

Florian-Horia Vasilescu
Laboratoire Paul Painlevé
U.F.R de Mathématiques
Université de Lille 1
F-59655 Villeneuve d'Ascq Cedex, France
e-mail: `fhvasil@math.univ-lille1.fr`