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

Alpay, D.; Dijksma, A.; Langer, H.

Published in:
EPRINTS-BOOK-TITLE

Publication date:
2007

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:

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

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,

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

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:

$$\mathcal{T}_{M_1M_2}(v)=\mathcal{T}_{M_1}(\mathcal{T}_{M_2}(v)),$$

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 $n=\deg\,\Theta$ 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 $\sigma_0, \ldots, \sigma_{n-1}$, 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

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

has a finite number κ of negative squares (here $\text{hol}\,(s)$ 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

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

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

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

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

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

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:

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

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 $w_1, \dots, w_m \in \mathcal{D}$ 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 $z \mapsto K(z, w_i)$, $i = 1, \dots, m$, which has dimension κ and is negative. Let $f_1(z), \dots, f_\kappa(z)$ 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 $A = (a_{ij})_{i,j\ge 0}$ and any integer $k \ge 1$ 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

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

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

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 $i,j = 0,1,\dots$ 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

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

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

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

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

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

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

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

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 $\kappa \geq 0$ 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

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

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, $i = 0, 1, \ldots, 2p - 1$, satisfy the following assumptions:

(1) $|\tau_0| = 1$;
(2) at least one of the numbers $\tau_1, \ldots, \tau_p$ 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,

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

(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

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

(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,

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

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

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

has κ negative squares on $\mathrm{hol}\,(n)$. In this case we also say that $n(z)$ has κ *negative squares* and we write $\mathrm{sq}_-(n) = \kappa$. 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 $\mathrm{Im}\, n(z) \geq 0$ 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:

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

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, $\mathrm{hol}\,(n)$ 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

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

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

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

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 $F(z)$:

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

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

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

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

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

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 $\Theta(z)$. □

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)*:*

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

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 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)*:*

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

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:*

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

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

$$G - A^{*}GA = C^{*}JC \tag{3.19}$$

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

$$\Theta(z) = I_p - (1-zz_0^*)C(zI-A)^{-1}G^{-1}(z_0I-A)^{-*}C^*J, \tag{3.23}$$

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

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

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

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

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

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

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

$$\Theta(z) = \sum_{j=0}^{n} \frac{T_j}{(1 - z z_1^*)^j}.$$

(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, $j = 0, 1, \ldots, n$. 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. $\square$

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. □

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*

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

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}:

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

and the "initial conditions"

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

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

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

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

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

We show that the numbers $\beta_{ij} = \big(D^* D\big)_{ij}$ 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

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

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

$$\binom{\mu}{\nu} - \binom{\mu-1}{\nu} = \binom{\mu-1}{\nu-1} \quad (4.9)$$

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 $j \geq 1$,

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

It follows that all principal submatrices of Γ^0 are positive. $\square$

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

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

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

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

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

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

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

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

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 $k_0(\Gamma)$:

$$k_0(\Gamma) := \min\big\{j \,\big|\, \det\Gamma_j \neq 0\big\}.$$

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

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

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 $k_0(\Gamma) = 2k$. □

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*

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

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))

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

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 $\Gamma = \big(\gamma_{ij}\big)_{i,j=0}^{\infty}$ 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*

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

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

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

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 $L_n(z, w)$:

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

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

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

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 $1 \le k \le \kappa$ and $0 \le q \le \kappa - k$.

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) $s(z) \in \mathbf{S}_\kappa^{z_1} \quad \Longrightarrow \quad \kappa \ge 1$ 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

$$s(z) - \sigma_0 = \frac{\big(1 - |\sigma_0|^2\big) b_c(z)\widetilde{s}(z)}{b_c(z)\sigma_0^* \widetilde{s}(z) + 1}$$

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 $s_0 \neq 0, s_1, \ldots, s_{k-1}$ the polynomial

$$\begin{aligned} q(z) &= q(z; s_0, s_1, \ldots, s_{k-1}) \\ &= c_0 + c_1(z - z_1) + \cdots + c_{k-1}(z - z_1)^{k-1} \\ &\qquad -\big(c_{k-1}^* z^{k+1}(1 - zz_1^*)^{k-1} + c_{k-2}^* z^{k+2}(1 - zz_1^*)^{k-2} + \cdots + c_0^* z^{2k}\big) \end{aligned}$$

of degree $2k$, where the coefficients $c_0, c_1, \ldots, c_{k-1}$ are determined by the formula

$$c_0 s_\ell + \cdots + c_\ell s_0 = \sigma_0 \binom{k}{\ell} (-z_1^*)^\ell \big(1 - |z_1|^2\big)^{k-\ell}, \quad \ell = 0, 1, \ldots, k-1. \tag{5.17}$$

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 $s_0 \neq 0, s_1, \ldots, s_{k-1}$, 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 $q(z) = q(z; s_0, s_1, \ldots, s_{k-1})$ 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

$$\begin{aligned} &\sigma_0^* q(z; s_1, \ldots, s_{k_1})(s(z) - \sigma_0) - (z - z_1)^k (1 - zz_1^*)^k \\ &\qquad = \frac{(z - z_1)^k (1 - zz_1^*)^k (s(z) - \sigma_0)}{\widetilde{s}(z) - \sigma_0} = \mathrm{O}\big((z - z_1)^{2k}\big), \quad z \to z_1. \end{aligned}$$

By comparing this relation with (5.8), we find that

$$q(z; s_0, \ldots, s_{k-1}) = q(z; \sigma_k, \ldots, \sigma_{2k-1})$$

and hence, on account of (5.7) and (5.17), that $\sigma_{k+j} = s_j$, $j = 0, \ldots, k-1$, 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 $\sigma_{2k}, \sigma_{2k+1}, \ldots, \sigma_{2k+q-1}$ of $s(z)$ are determined by $s_0, s_1, \ldots, s_{k-1}$, 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*

$$U = b_c(z_0)^k I_2 + \frac{q(z_0)}{(1-z_0z_1^*)^{2k}}\mathbf{u}\mathbf{u}^*J_c, \quad \mathbf{u} = \begin{pmatrix} 1 \\ \sigma_0^* \end{pmatrix}.$$

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

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

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*

$$\Theta(z) = \left(\frac{b_c(z)}{b_c(z_0)}\right)^k I_2 + \frac{q_1(z)}{(1-zz_1^*)^{2k}}\mathbf{u}\mathbf{u}^*J_c,$$

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*

$$b_c(z_0)^k q_1(z) + b_c(z_0)^{*k} z^{2k} q_1 (1/z^*)^* = 0.$$

(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:

$$s_0(z) = s(z),\; s_1(z) = \mathcal{T}_{\Psi_1(z)^{-1}}(s_0(z)),\; s_2(z) = \mathcal{T}_{\Psi_2(z)^{-1}}(s_1(z)), \dots,$$

$$s_q(z) = \mathcal{T}_{\Psi_q(z)^{-1}}(s_{q-1}(z))$$

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,

$$\begin{aligned}
\begin{pmatrix}1 & -s\end{pmatrix}\mathcal{P}(\Psi) &= \mathcal{P}(s)\\
&= \begin{pmatrix}1 & -s\end{pmatrix}\mathcal{P}(\Psi_1)\oplus(a_1-c_1s)\mathcal{P}(s_1)\\
&\;\;\vdots\\
&= \begin{pmatrix}1 & -s\end{pmatrix}\big(\mathcal{P}(\Psi_1)\oplus\Psi_1\mathcal{P}(\Psi_2)\big)\oplus(a_1-c_1s)(a_2-c_2s_1)\mathcal{P}(s_2)\\
&= \begin{pmatrix}1 & -s\end{pmatrix}\big(\mathcal{P}(\Psi_1)\oplus\Psi_1\mathcal{P}(\Psi_2)\oplus\Psi_1\Psi_2\mathcal{P}(\Psi_3)\oplus\cdots\big)\\
&= \begin{pmatrix}1 & -s\end{pmatrix}\mathcal{P}(\Psi_1\Psi_2\cdots\Psi_q),
\end{aligned}$$

where

$$\Psi_j(z)=\begin{pmatrix}a_j(z) & b_j(z)\\ c_j(z) & d_j(z)\end{pmatrix},\quad j=1,2,\dots.$$

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) $\operatorname{span}\big\{v, T^*v, \ldots, T^{*(k-1)}v\big\}$ *is a* k*-dimensional negative subspace of* $\mathcal{P}$*, so that the space*

$$\widetilde{\mathcal{P}} = \mathcal{P} \ominus \operatorname{span}\big\{v, T^*v, \ldots, T^{*(k-1)}v\big\}$$

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* $\operatorname{span}\{v, T^*v, \dots, T^{*(2k+q-1)}v\}$ *is a* $(2k+q)$*-dimensional Pontryagin subspace of* $\mathcal{P}$ *with negative index equal to* $k+q$*, so that the space*

$$\widetilde{\mathcal{P}} = \mathcal{P} \ominus \operatorname{span}\{v, T^*v, \dots, T^{*(2k+q-1)}v\}$$

and the orthogonal projection P *in* $\mathcal{P}$ *onto* $\widetilde{\mathcal{P}}$ *are well defined, and*

(ii) *with*

$$\begin{aligned} \widetilde{T} &= PTP + \frac{1}{\sigma_k t_{2k+q}} \langle \,\cdot\,, PT^{*(2k+q)}v \rangle Pu, & \widetilde{u} &= \frac{1}{t_{2k+q}} P_{2k+q}u, \\ \widetilde{v} &= \frac{1}{t^*_{2k+q}} PT^{*(2k+q)}v, & \widetilde{\gamma} &= \frac{\sigma_k}{t_{2k+q}} \end{aligned}$$

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*

$$s_0(z) = s(z), \quad s_j(z) = \widehat{s}_{j-1}(z),\ j = 1, 2, \dots.$$

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 $j = 0, 1, \dots,$ 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

$$R(z) = \begin{pmatrix} \dfrac{1}{1 - zz_1^*} & \dfrac{z}{(1 - zz_1^*)^2} & \cdots & \dfrac{z^{k-1}}{(1 - zz_1^*)^k} \end{pmatrix},$$

[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, $\tau_{k+1}, \dots, \tau_{2k-1}$ 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

$$p(z) = z_0(-z_1^*)^k z^{k-1} p(1/z^*)^*.$$

(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 $\leq 2k - 1$ 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* $\mathrm{Im}\ \nu_0 > 0$ *and* $\widehat{\kappa} = \kappa - 1$ *if* $\mathrm{Im}\ \nu_0 < 0$.

(ii) $n(z) \in \mathbf{N}_\kappa^{z_1} \Longrightarrow 1 \leq k \leq \kappa,\ \widehat{n}(z) \in \mathbf{N}_{\kappa-k}$.

(iii) $n(z) \in \mathbf{N}_\kappa$ *and* $n(z)$ *has a pole at* z_1 *of order* $q \geq 1 \Longrightarrow q \leq \kappa$ *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 $\rho_q, \ldots, \rho_{2q-1}$ 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* $\operatorname{Im}\,\nu_0 \neq 0$, *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*

$$\widetilde{\kappa} = \begin{cases} \kappa, & \operatorname{Im}\,\nu_0 > 0, \\ \kappa - 1, & \operatorname{Im}\,\nu_0 < 0. \end{cases}$$

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): $\operatorname{Im}\nu_0 = 0$. 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* $\operatorname{Im}\,\nu_0 = 0$, *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* $\operatorname{Im}\,\nu_0 = 0$, *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*

$$n(z) = \frac{\alpha(z)\widetilde{n}(z) - \nu_0^2}{\widetilde{n}(z) - \beta(z)} \qquad \text{with}$$

$$\alpha(z) = \nu_0 + \frac{(z - z_1)^k(z - z_1^*)^k}{p(z)}, \quad \beta(z) = \nu_0 - \frac{(z - z_1)^k(z - z_1^*)^k}{p(z)}$$

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 $\operatorname{Re} s > 0$ 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 $\mu_{k-1} \ne 0$ 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

$$e_k(z) = \frac{1}{\gamma_k} \det \begin{pmatrix} 0 & 0 & \dots & \mu_{k-1} & \mu_k \\ 0 & 0 & \dots & \mu_k & \mu_{k+1} \\ \vdots & \vdots & & \vdots & \vdots \\ \mu_{k-1} & \mu_k & \dots & \mu_{2k-2} & \mu_{2k-1} \\ 1 & z & \dots & z^{k-1} & z^k \end{pmatrix} \tag{8.5}$$

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 $\mu_{k-1} \ne 0$, 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, $\gamma_{p+1} \ne 0$, 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 $\widetilde{n}(z) = \mathrm{O}(1/z)$, $z = iy,\ y \uparrow \infty$. 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 $\mathbf{r}_z$, $z \in \mathrm{hol}\,(n)$, $z \neq z^*$, defined by

$$\mathbf{r}_z(t) = \frac{1}{t-z}, \quad t \in \mathbb{C}.$$

Equipped with the inner product

$$\langle \mathbf{r}_z, \mathbf{r}_\zeta \rangle = \frac{n(z) - n(\zeta)^*}{z - \zeta^*}, \quad z, \zeta \in \mathrm{hol}\,(n),\ z \neq \zeta^*,$$

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

$$\mathbf{t}_j(t) := t^j, \quad j = 0, 1, \dots, p,$$

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 $u \in \mathrm{dom}\,(A^j)$ and $\mathbf{t}_j = A^j u$, $j = 0, 1, \dots, p$, and the function $n(z)$ admits the representation

$$n(z) = \left\langle (A - z)^{-1} u, u \right\rangle, \quad z \in \mathrm{hol}\,(n).$$

Now let $k\,(\leq p)$ be again the smallest positive integer such that $\mu_{k-1} \neq 0$. We introduce the subspace

$$\mathcal{H}_k = \mathrm{span}\{\mathbf{t}_0, \mathbf{t}_1, \dots, \mathbf{t}_{k-1}\} \tag{8.16}$$

of $\Pi(n)$. It is nondegenerate since $\mu_{k-1} \neq 0$ and its negative index equals

$$\mathrm{ind}_-(\mathcal{H}_k) = \begin{cases} [k/2], & \mu_{k-1} > 0, \\ [(k+1)/2]\,, & \mu_{k-1} < 0. \end{cases}$$

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*

$$n(z) = \left\langle (A - z)^{-1} u, u \right\rangle$$

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`