

# The Operator Fejér-Riesz Theorem

Michael A. Dritschel and James Rovnyak

*To the memory of Paul Richard Halmos.*

**Abstract.** The Fejér-Riesz theorem has inspired numerous generalizations in one and several variables, and for matrix- and operator-valued functions. This paper is a survey of some old and recent topics that center around Rosenblum's operator generalization of the classical Fejér-Riesz theorem.

**Mathematics Subject Classification (2000).** Primary 47A68; Secondary 60G25, 47A56, 47B35, 42A05, 32A70, 30E99.

**Keywords.** Trigonometric polynomial, Fejer-Riesz theorem, spectral factorization, Schur complement, noncommutative polynomial, Toeplitz operator, shift operator.

## Contents

1. Introduction	2
2. The operator Fejér-Riesz theorem	4
3. Method of Schur complements	7
4. Spectral factorization	11
5. Multivariable theory	18
6. Noncommutative factorization	23
Appendix A. Schur complements	28
References	30

## 1. Introduction

The classical Fejér-Riesz factorization theorem gives the form of a nonnegative trigonometric polynomial on the real line, or, equivalently, a Laurent polynomial that is nonnegative on the unit circle. For the statement, we write  $\mathbb{D} = \{z : |z| < 1\}$  and  $\mathbb{T} = \{\zeta : |\zeta| = 1\}$  for the open unit disk and unit circle in the complex plane.

**Fejér-Riesz Theorem.** *A Laurent polynomial  $q(z) = \sum_{k=-m}^m q_k z^k$  which has complex coefficients and satisfies  $q(\zeta) \geq 0$  for all  $\zeta \in \mathbb{T}$  can be written*

$$q(\zeta) = |p(\zeta)|^2, \quad \zeta \in \mathbb{T},$$

for some polynomial  $p(z) = p_0 + p_1 z + \cdots + p_m z^m$ , and  $p(z)$  can be chosen to have no zeros in  $\mathbb{D}$ .

The original sources are Fejér [22] and Riesz [47]. The proof is elementary and consists in showing that the roots of  $q(z)$  occur in pairs  $z_j$  and  $1/\bar{z}_j$  with  $|z_j| \geq 1$ . Then the required polynomial  $p(z)$  is the product of the factors  $z - z_j$  adjusted by a suitable multiplicative constant  $c$ . Details appear in many places; see e.g. [28, p. 20], [34, p. 235], or [60, p. 26].

The Fejér-Riesz theorem arises naturally in spectral theory, the theory of orthogonal polynomials, prediction theory, moment problems, and systems and control theory. Applications often require generalizations to functions more general than Laurent polynomials, and, more than that, to functions whose values are matrices or operators on a Hilbert space. The spectral factorization problem is to write a given nonnegative matrix- or operator-valued function  $F$  on the unit circle in the form  $F = G^*G$  where  $G$  has an analytic extension to the unit disk (in a suitably interpreted sense). The focal point of our survey is the special case of a Laurent polynomial with operator coefficients.

The operator Fejér-Riesz theorem (Theorem 2.1) obtains a conclusion similar to the classical result for Laurent polynomial whose coefficients are Hilbert space operators: if  $Q_j$ ,  $j = -m, \dots, m$ , are Hilbert space operators such that

$$Q(\zeta) = \sum_{k=-m}^m Q_k \zeta^k \geq 0, \quad \zeta \in \mathbb{T}, \quad (1.1)$$

then there is a polynomial  $P(z) = P_0 + P_1 z + \cdots + P_m z^m$  with operator coefficients such that

$$Q(\zeta) = P(\zeta)^* P(\zeta), \quad \zeta \in \mathbb{T}. \quad (1.2)$$

This was first proved in full generality in 1968 by Marvin Rosenblum [49]. The proof uses Toeplitz operators and a method of Lowdenslager, and it is a fine example of operator theory in the spirit of Paul Halmos. Rosenblum's proof is reproduced in §2.

Part of the fascination of the operator Fejér-Riesz theorem is that it can be stated in a purely algebraic way. The hypothesis (1.1) on  $Q(z)$  is equivalent to the

statement that an associated Toeplitz matrix is nonnegative. The conclusion (1.2) is equivalent to  $2m + 1$  nonlinear equations whose unknowns are the coefficients  $P_0, P_1, \dots, P_m$  of  $P(z)$ . Can it be that this system of equations can be solved by an algebraic procedure? The answer is, yes, and this is a recent development. The iterative procedure uses the notion of a Schur complement and is outlined in §3.

There is a surprising connection between Rosenblum's proof of the operator Fejér-Riesz theorem and spectral factorization. The problem of spectral factorization is formulated precisely in §4, using Hardy class notions. A scalar prototype is Szegő's theorem (Theorem 4.1) on the representation of a positive integrable and log-integrable function  $w$  on the unit circle in the form  $|h|^2$  for some  $H^2$  function  $h$ . The operator and matrix counterparts of Szegő's theorem, Theorems 4.5 and 4.7, have been known for many years and go back to fundamental work in the 1940s and 1950s which was motivated by applications in prediction theory (see the historical notes at the end of §4). We present a proof that is new to the authors and we suspect not widely known. It is based on Theorem 4.3, which traces its origins to Rosenblum's implementation of the Lowdenslager method. In §4 we also state without proof some special results that hold in the matrix case.

The method of Schur complements points the way to an approach to multi-variable factorization problems, which is the subject of §5. Even in the scalar case, the obvious first ideas for multivariable generalizations of the Fejér-Riesz theorem are false by well-known examples. Part of the problem has to do with what one might think are natural restrictions on degrees. In fact, the restrictions on degrees are not so natural after all. When they are removed, we can prove a result, Theorem 5.1, that can be viewed as a generalization of the operator Fejér-Riesz theorem in the strictly positive case. We also look at the problem of outer factorization, at least in some restricted settings.

In recent years there has been increasing interest in noncommutative function theory, especially in the context of functions of freely noncommuting variables. In §6 we consider noncommutative analogues of the  $d$ -torus, and corresponding notions of nonnegative trigonometric polynomials. In the freely noncommutative setting, there is a very nice version of the Fejér-Riesz theorem (Theorem 6.1). In a somewhat more general noncommutative setting, which also happens to cover the commutative case as well, we have a version of Theorem 5.1 for strictly positive polynomials (Theorem 6.2).

Our survey does not aim for completeness in any area. In particular, our bibliography represents only a selection from the literature. The authors regret and apologize for omissions.

## 2. The operator Fejér-Riesz theorem

In this section we give the proof of the operator Fejér-Riesz theorem by Rosenblum [49]. The general theorem had precursors. A finite-dimensional version was given by Rosenblatt [48], an infinite-dimensional special case by Gohberg [26].

We follow standard conventions for Hilbert spaces and operators. If  $A$  is an operator,  $A^*$  is its adjoint. Norms of vectors and operators are written  $\|\cdot\|$ . Except where noted, no assumption is made on the dimension of a Hilbert space, and nonseparable Hilbert spaces are allowed.

**Theorem 2.1 (Operator Fejér-Riesz Theorem).** *Let  $Q(z) = \sum_{k=-m}^m Q_k z^k$  be a Laurent polynomial with coefficients in  $\mathfrak{L}(\mathfrak{G})$  for some Hilbert space  $\mathfrak{G}$ . If  $Q(\zeta) \geq 0$  for all  $\zeta \in \mathbb{T}$ , then*

$$Q(\zeta) = P(\zeta)^* P(\zeta), \quad \zeta \in \mathbb{T}, \quad (2.1)$$

for some polynomial  $P(z) = P_0 + P_1 z + \cdots + P_m z^m$  with coefficients in  $\mathfrak{L}(\mathfrak{G})$ . The polynomial  $P(z)$  can be chosen to be outer.

The definition of an outer polynomial will be given later; in the scalar case, a polynomial is outer if and only if it has no zeros in  $\mathbb{D}$ .

The proof uses (unilateral) shift and Toeplitz operators (see [11] and [29]). By a **shift operator** here we mean an isometry  $S$  on a Hilbert space  $\mathfrak{H}$  such that the unitary component of  $S$  in its Wold decomposition is trivial. With natural identifications, we can write  $\mathfrak{H} = \mathfrak{G} \oplus \mathfrak{G} \oplus \cdots$  for some Hilbert space  $\mathfrak{G}$  and

$$S(h_0, h_1, \dots) = (0, h_0, h_1, \dots)$$

when the elements of  $\mathfrak{H}$  are written in sequence form. Suppose that such a shift  $S$  is chosen and fixed. If  $T, A \in \mathfrak{L}(\mathfrak{H})$ , we say that  $T$  is **Toeplitz** if  $S^* T S = T$ , and that  $A$  is **analytic** if  $AS = SA$ . An analytic operator  $A$  is said to be **outer** if  $\overline{\text{ran } A}$  is a subspace of  $\mathfrak{H}$  of the form  $\mathfrak{F} \oplus \mathfrak{F} \oplus \cdots$  for some closed subspace  $\mathfrak{F}$  of  $\mathfrak{G}$ .

As block operator matrices, Toeplitz and analytic operators have the forms

$$T = \begin{pmatrix} T_0 & T_{-1} & T_{-2} & \cdots \\ T_1 & T_0 & T_{-1} & \ddots \\ T_2 & T_1 & T_0 & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}, \quad A = \begin{pmatrix} A_0 & 0 & 0 & \cdots \\ A_1 & A_0 & 0 & \ddots \\ A_2 & A_1 & A_0 & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}. \quad (2.2)$$

Here

$$T_j = \begin{cases} E_0^* S^{*j} T E_0|_{\mathfrak{G}}, & j \geq 0, \\ E_0^* T S^{|j|} E_0|_{\mathfrak{G}}, & j < 0, \end{cases} \quad (2.3)$$

where  $E_0 g = (g, 0, 0, \dots)$  is the natural embedding of  $\mathfrak{G}$  into  $\mathfrak{H}$ . For examples, consider Laurent and analytic polynomials  $Q(z) = \sum_{k=-m}^m Q_k z^k$  and  $P(z) =$

$P_0 + P_1z + \cdots + P_mz^m$  with coefficients in  $\mathfrak{L}(\mathfrak{G})$ . Set  $Q_j = 0$  for  $|j| > m$  and  $P_j = 0$  for  $j > m$ . Then the formulas

$$T_Q = \begin{pmatrix} Q_0 & Q_{-1} & Q_{-2} & \cdots \\ Q_1 & Q_0 & Q_{-1} & \ddots \\ Q_2 & Q_1 & Q_0 & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}, \quad T_P = \begin{pmatrix} P_0 & 0 & 0 & \cdots \\ P_1 & P_0 & 0 & \ddots \\ P_2 & P_1 & P_0 & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix} \quad (2.4)$$

define bounded operators on  $\mathfrak{H}$ . Boundedness follows from the identity

$$\int_{\mathbb{T}} \langle Q(\zeta)f(\zeta), g(\zeta) \rangle_{\mathfrak{G}} d\sigma(\zeta) = \sum_{k,j=0}^{\infty} \langle Q_{j-k}f_k, g_j \rangle_{\mathfrak{G}}, \quad (2.5)$$

where  $\sigma$  is normalized Lebesgue measure on  $\mathbb{T}$  and  $f(\zeta) = f_0 + f_1\zeta + f_2\zeta^2 + \cdots$  and  $g(\zeta) = g_0 + g_1\zeta + g_2\zeta^2 + \cdots$  have coefficients in  $\mathfrak{G}$ , all but finitely many of which are zero. The operator  $T_Q$  is Toeplitz, and  $T_P$  is analytic. Moreover,

- $Q(\zeta) \geq 0$  for all  $\zeta \in \mathbb{T}$  if and only if  $T_Q \geq 0$ ;
- $Q(\zeta) = P(\zeta)^*P(\zeta)$  for all  $\zeta \in \mathbb{T}$  if and only if  $T_Q = T_P^*T_P$ .

**Definition 2.2.** We say that the polynomial  $P(z)$  is **outer** if the analytic Toeplitz operator  $A = T_P$  is outer.

In view of the example (2.4), the main problem is to write a given nonnegative Toeplitz operator  $T$  in the form  $T = A^*A$ , where  $A$  is analytic. We also want to know that if  $T = T_Q$  for a Laurent polynomial  $Q$ , then we can choose  $A = T_P$  for an outer analytic polynomial of the same degree. Lemmas 2.3 and 2.4 reduce the problem to showing that a certain isometry is a shift operator.

**Lemma 2.3 (Lowdenslager's Criterion).** *Let  $\mathfrak{H}$  be a Hilbert space, and let  $S \in \mathfrak{L}(\mathfrak{H})$  be a shift operator. Let  $T \in \mathfrak{L}(\mathfrak{H})$  be Toeplitz relative to  $S$  as defined above, and suppose that  $T \geq 0$ . Let  $\mathfrak{H}_T$  be the closure of the range of  $T^{1/2}$  in the inner product of  $\mathfrak{H}$ . Then there is an isometry  $S_T$  mapping  $\mathfrak{H}_T$  into itself such that*

$$S_T T^{1/2} f = T^{1/2} S f, \quad f \in \mathfrak{H}.$$

*In order that  $T = A^*A$  for some analytic operator  $A \in \mathfrak{L}(\mathfrak{H})$ , it is necessary and sufficient that  $S_T$  is a shift operator. In this case,  $A$  can be chosen to be outer.*

*Proof.* The existence of the isometry  $S_T$  follows from the identity  $S^*TS = T$ , which implies that  $T^{1/2}Sf$  and  $T^{1/2}f$  have the same norms for any  $f \in \mathfrak{H}$ .

If  $S_T$  is a shift operator, we can view  $\mathfrak{H}_T$  as a direct sum  $\mathfrak{H}_T = \mathfrak{G}_T \oplus \mathfrak{G}_T \oplus \cdots$  with  $S_T(h_0, h_1, \dots) = (0, h_0, h_1, \dots)$ . Here  $\dim \mathfrak{G}_T \leq \dim \mathfrak{G}$ . To see this, notice that a short argument shows that  $T^{1/2}S_T^*$  and  $S^*T^{1/2}$  agree on  $\mathfrak{H}_T$ , and therefore  $T^{1/2}(\ker S_T^*) \subseteq \ker S^*$ . The dimension inequality then follows because  $T^{1/2}$  is one-to-one on the closure of its range. Therefore we may choose an isometry  $V$  from

$\mathfrak{G}_T$  into  $\mathfrak{G}$ . Define an isometry  $W$  on  $\mathfrak{H}_T$  into  $\mathfrak{H}$  by

$$W(h_0, h_1, \dots) = (Vh_0, Vh_1, \dots).$$

Define  $A \in \mathfrak{L}(\mathfrak{H})$  by mapping  $\mathfrak{H}$  into  $\mathfrak{H}_T$  via  $T^{1/2}$  and then  $\mathfrak{H}_T$  into  $\mathfrak{H}$  via  $W$ :

$$Af = WT^{1/2}f, \quad f \in \mathfrak{H}.$$

Straightforward arguments show that  $A$  is analytic, outer, and  $T = A^*A$ .

Conversely, suppose that  $T = A^*A$  where  $A \in \mathfrak{L}(\mathfrak{H})$  is analytic. Define an isometry  $W$  on  $\mathfrak{H}_T$  into  $\mathfrak{H}$  by  $WT^{1/2}f = Af$ ,  $f \in \mathfrak{H}$ . Then  $WS_T = SW$ , and hence  $S_T^{*n} = W^*S^{*n}W$  for all  $n \geq 1$ . Since the powers of  $S^*$  tend strongly to zero, so do the powers of  $S_T^*$ , and therefore  $S_T$  is a shift operator.  $\square$

**Lemma 2.4.** *In Lemma 2.3, let  $T = T_Q$  be given by (2.4) for a Laurent polynomial  $Q(z)$  of degree  $m$ . If  $T = A^*A$  where  $A \in \mathfrak{L}(\mathfrak{H})$  is analytic and outer, then  $A = T_P$  for some outer analytic polynomial  $P(z)$  of degree  $m$ .*

*Proof.* Let  $Q(z) = \sum_{k=-m}^m Q_k z^k$ . Recall that  $Q_j = 0$  for  $|j| > m$ . By (2.3) applied to  $A$ , what we must show is that  $S^{*j}AE_0 = 0$  for all  $j > m$ . It is sufficient to show that  $S^{*m+1}AE_0 = 0$ . By (2.3) applied to  $T$ , since  $T = A^*A$  and  $A$  is analytic,

$$E_0^*A^*S^{*j}AE_0 = E_0^*S^{*j}TE_0 = Q_j = 0, \quad j > m.$$

It follows that  $\text{ran } S^{*m+1}AE_0 \perp \text{ran } AS^kE_0$  for all  $k \geq 0$ , and therefore

$$\text{ran } S^{*m+1}AE_0 \perp \text{ran } A. \quad (2.6)$$

Since  $A$  is outer,  $\overline{\text{ran } A}$  reduces  $S$ , and so  $\text{ran } S^{*m+1}AE_0 \subseteq \overline{\text{ran } A}$ . Therefore  $S^{*m+1}AE_0 = 0$  by (2.6), and the result follows.  $\square$

The proof of the operator Fejér-Riesz theorem is now easily completed.

*Proof of Theorem 2.1.* Define  $T = T_Q$  as in (2.4). Lemmas 2.3 and 2.4 reduce the problem to showing that the isometry  $S_T$  is a shift operator. It is sufficient to show that  $\|S_T^{*n}f\| \rightarrow 0$  for every  $f$  in  $\mathfrak{H}_T$ .

**Claim:** If  $f = T^{1/2}h$  where  $h \in \mathfrak{H}$  has the form  $h = (h_0, \dots, h_r, 0, \dots)$ , then  $S_T^{*n}f = 0$  for all sufficiently large  $n$ .

For if  $u \in \mathfrak{H}$  and  $n$  is any positive integer, then

$$\left\langle S_T^{*n}f, T^{1/2}u \right\rangle_{\mathfrak{H}_T} = \left\langle f, S_T^n T^{1/2}u \right\rangle_{\mathfrak{H}_T} = \left\langle T^{1/2}h, T^{1/2}S^n u \right\rangle_{\mathfrak{H}} = \langle Th, S^n u \rangle_{\mathfrak{H}}.$$

By the definition of  $T = T_Q$ ,  $Th$  has only a finite number of nonzero entries (depending on  $m$  and  $r$ ), and the first  $n$  entries of  $S^n u$  are zero (irrespective of  $u$ ). The claim follows from the arbitrariness of  $u$ .

In view of the claim,  $\|S_T^{*n}f\| \rightarrow 0$  for a dense set of vectors in  $\mathfrak{H}_T$ , and hence by approximation this holds for all  $f$  in  $\mathfrak{H}_T$ . Thus  $S_T$  is a shift operator, and the result follows.  $\square$

A more general result is proved in the original version of Theorem 2.1 in [49]. There it is only required that  $Q(z)g$  is a Laurent polynomial for a dense set of  $g$  in  $\mathfrak{G}$  (the degrees of these polynomials can be unbounded). We have omitted an accompanying uniqueness statement: the outer polynomial  $P(z)$  in Theorem 2.1 can be chosen such that  $P(0) \geq 0$ , and then it is unique. See [2] and [50].

### 3. Method of Schur complements

We outline now a completely different proof of the operator Fejér-Riesz theorem. The proof is due to Dritschel and Woerdeman [19] and is based on the notion of a Schur complement. The procedure constructs the outer polynomial  $P(z) = P_0 + P_1z + \cdots + P_mz^m$  one coefficient at a time. A somewhat different use of Schur complements in the operator Fejér-Riesz theorem appears in Dritschel [18]. The method in [18] plays a role in the multivariable theory, which is taken up in §5.

We shall explain the main steps of the construction assuming the validity of two lemmas. Full details are given in [19] and also in the forthcoming book [3] by Bakonyi and Woerdeman. The authors thank Mihaly Bakonyi and Hugo Woerdeman for advance copies of key parts of [3], which has been helpful for our exposition. The book [3] includes many additional results not discussed here.

**Definition 3.1.** Let  $\mathfrak{H}$  be a Hilbert space. Suppose  $T \in \mathfrak{L}(\mathfrak{H})$ ,  $T \geq 0$ . Let  $\mathfrak{K}$  be a closed subspace of  $\mathfrak{H}$ , and let  $P_{\mathfrak{K}} \in \mathfrak{L}(\mathfrak{H}, \mathfrak{K})$  be orthogonal projection of  $\mathfrak{H}$  onto  $\mathfrak{K}$ . Then (see Appendix A, Lemma A.2) there is a unique operator  $S \in \mathfrak{L}(\mathfrak{K})$ ,  $S \geq 0$ , such that

- (i)  $T - P_{\mathfrak{K}}^* S P_{\mathfrak{K}} \geq 0$ ;
- (ii) if  $\tilde{S} \in \mathfrak{L}(\mathfrak{K})$ ,  $\tilde{S} \geq 0$ , and  $T - P_{\mathfrak{K}}^* \tilde{S} P_{\mathfrak{K}} \geq 0$ , then  $\tilde{S} \leq S$ .

We write  $S = S(T, \mathfrak{K})$  and call  $S$  the **Schur complement of  $T$  supported on  $\mathfrak{K}$** .

Schur complements satisfy an inheritance property, namely, if  $\mathfrak{K}_- \subseteq \mathfrak{K}_+ \subseteq \mathfrak{H}$ , then  $S(T, \mathfrak{K}_-) = S(S(T, \mathfrak{K}_+), \mathfrak{K}_-)$ . If  $T$  is specified in matrix form,

$$T = \begin{pmatrix} A & B^* \\ B & C \end{pmatrix} : \mathfrak{K} \oplus \mathfrak{K}^\perp \rightarrow \mathfrak{K} \oplus \mathfrak{K}^\perp,$$

then  $S = S(T, \mathfrak{K})$  is the largest nonnegative operator in  $\mathfrak{L}(\mathfrak{K})$  such that

$$\begin{pmatrix} A - S & B^* \\ B & C \end{pmatrix} \geq 0.$$

The condition  $T \geq 0$  is equivalent to the existence of a contraction  $G \in \mathfrak{L}(\mathfrak{K}, \mathfrak{K}^\perp)$  such that  $B = C^{\frac{1}{2}} G A^{\frac{1}{2}}$  (Appendix A, Lemma A.1). In this case,  $G$  can be chosen so that it maps  $\overline{\text{ran } A}$  into  $\overline{\text{ran } C}$  and is zero on the orthogonal complement of  $\overline{\text{ran } A}$ , and then

$$S = A^{\frac{1}{2}} (I - G^* G) A^{\frac{1}{2}}.$$

When  $C$  is invertible, this reduces to the familiar formula  $S = A - B^* C^{-1} B$ .

**Lemma 3.2.** *Let  $M \in \mathfrak{L}(\mathfrak{H})$ ,  $M \geq 0$ , and suppose that*

$$M = \begin{pmatrix} A & B^* \\ B & C \end{pmatrix} : \mathfrak{K} \oplus \mathfrak{K}^\perp \rightarrow \mathfrak{K} \oplus \mathfrak{K}^\perp$$

for some closed subspace  $\mathfrak{K}$  of  $\mathfrak{H}$ .

(1) *If  $S(M, \mathfrak{K}) = P^*P$  and  $C = R^*R$  for some  $P \in \mathfrak{L}(\mathfrak{K})$  and  $R \in \mathfrak{L}(\mathfrak{K}^\perp)$ , then there is a unique  $X \in \mathfrak{L}(\mathfrak{K}, \mathfrak{K}^\perp)$  such that*

$$M = \begin{pmatrix} P^* & X^* \\ 0 & R^* \end{pmatrix} \begin{pmatrix} P & 0 \\ X & R \end{pmatrix} \quad \text{and} \quad \text{ran } X \subseteq \overline{\text{ran } R}. \quad (3.1)$$

(2) *Conversely, if (3.1) holds for some operators  $P, X, R$ , then  $S(M, \mathfrak{K}) = P^*P$ .*

We omit the proof and refer the reader to [3] or [19] for details.

*Proof of Theorem 2.1 using Schur complements.* Let  $Q(z) = \sum_{k=-m}^m Q_k z^k$  satisfy  $Q(\zeta) \geq 0$  for all  $\zeta \in \mathbb{T}$ . We shall recursively construct the coefficients of an outer polynomial  $P(z) = P_0 + P_1 z + \cdots + P_m z^m$  such that  $Q(\zeta) = P(\zeta)^* P(\zeta)$ ,  $\zeta \in \mathbb{T}$ .

Write  $\mathfrak{H} = \mathfrak{G} \oplus \mathfrak{G} \oplus \cdots$  and  $\mathfrak{G}^n = \mathfrak{G} \oplus \cdots \oplus \mathfrak{G}$  with  $n$  summands. As before, set  $Q_k = 0$  for  $|k| > m$ , and define  $T_Q \in \mathfrak{L}(\mathfrak{H})$  by

$$T_Q = \begin{pmatrix} Q_0 & Q_{-1} & Q_{-2} & \cdots \\ Q_1 & Q_0 & Q_{-1} & \ddots \\ Q_2 & Q_1 & Q_0 & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}.$$

For each  $k = 0, 1, 2, \dots$ , define

$$S(k) = S(T_Q, \mathfrak{G}^{k+1}),$$

which we interpret as the Schur complement of  $T_Q$  on the first  $k+1$  summands of  $\mathfrak{H} = \mathfrak{G} \oplus \mathfrak{G} \oplus \cdots$ . Thus  $S(k)$  is a  $(k+1) \times (k+1)$  block operator matrix satisfying

$$S(S(k), \mathfrak{G}^{j+1}) = S(j), \quad 0 \leq j < k < \infty \quad (3.2)$$

by the inheritance property of Schur complements.

**Lemma 3.3.** *For each  $k = 0, 1, 2, \dots$ ,*

$$S(k+1) = \begin{pmatrix} Y_0 & (Y_1 \cdots Y_{k+1}) \\ \begin{pmatrix} Y_1^* \\ \vdots \\ Y_{k+1}^* \end{pmatrix} & S(k) \end{pmatrix}$$

for some operators  $Y_0, Y_1, \dots, Y_{k+1}$  in  $\mathfrak{L}(\mathfrak{G})$ . For  $k \geq m-1$ ,

$$(Y_0 \ Y_1 \ \cdots \ Y_{k+1}) = (Q_0 \ Q_{-1} \ \cdots \ Q_{-k-1}).$$



Again see [3] or [19] for details. Granting Lemmas 3.2 and 3.3, we can proceed with the construction.

**Construction of  $P_0, P_1$ .** Choose  $P_0 = S(0)^{\frac{1}{2}}$ . Using Lemma 3.3, write

$$S(1) = \begin{pmatrix} Y_0 & Y_1 \\ Y_1^* & S(0) \end{pmatrix}.$$

In Lemma 3.2(1) take  $M = S(1)$  and use the factorizations

$$S(S(1), \mathfrak{G}^1) \stackrel{(3.2)}{=} S(0) = P_0^* P_0 \quad \text{and} \quad S(0) = P_0^* P_0.$$

Choose  $P_1 = X$  where  $X \in \mathfrak{L}(\mathfrak{G})$  is the operator produced by Lemma 3.2(1). Then

$$S(1) = \begin{pmatrix} P_0^* & P_1^* \\ 0 & P_0^* \end{pmatrix} \begin{pmatrix} P_0 & 0 \\ P_1 & P_0 \end{pmatrix} \quad \text{and} \quad \text{ran } P_1 \subseteq \overline{\text{ran } P_0}. \quad (3.3)$$

**Construction of  $P_2$ .** Next use Lemma 3.3 to write

$$S(2) = \begin{pmatrix} Y_0 & (Y_1 \ Y_2) \\ (Y_1^* \\ Y_2^*) & S(1) \end{pmatrix},$$

and apply Lemma 3.2(1) to  $M = S(2)$  with the factorizations

$$S(S(2), \mathfrak{G}^1) \stackrel{(3.2)}{=} S(0) = P_0^* P_0,$$

$$S(1) = \begin{pmatrix} P_0^* & P_1^* \\ 0 & P_0^* \end{pmatrix} \begin{pmatrix} P_0 & 0 \\ P_1 & P_0 \end{pmatrix}.$$

This yields operators  $X_1, X_2 \in \mathfrak{L}(\mathfrak{G})$  such that

$$S(2) = \begin{pmatrix} P_0^* & X_1^* & X_2^* \\ 0 & P_0^* & P_1^* \\ 0 & 0 & P_0^* \end{pmatrix} \begin{pmatrix} P_0 & 0 & 0 \\ X_1 & P_0 & 0 \\ X_2 & P_1 & P_0 \end{pmatrix}, \quad (3.4)$$

$$\text{ran} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \subseteq \overline{\text{ran}} \begin{pmatrix} P_0 & 0 \\ P_1 & P_0 \end{pmatrix}. \quad (3.5)$$

In fact,  $X_1 = P_1$ . To see this, notice that we can rewrite (3.4) as

$$S(2) = \begin{pmatrix} \tilde{P}^* & \tilde{X}^* \\ 0 & \tilde{R}^* \end{pmatrix} \begin{pmatrix} \tilde{P} & 0 \\ \tilde{X} & \tilde{R} \end{pmatrix},$$

$$\tilde{P} = \begin{pmatrix} P_0 & 0 \\ X_1 & P_0 \end{pmatrix}, \quad \tilde{X} = (X_2 \ P_1), \quad \tilde{R} = P_0.$$

By (3.3) and (3.5),  $\text{ran } P_1 \subseteq \overline{\text{ran}} P_0$  and  $\text{ran } X_2 \subseteq \overline{\text{ran}} P_0$ , and therefore  $\text{ran } \tilde{X} \subseteq \overline{\text{ran}} P_0$ . Hence by Lemma 3.2(2),

$$S(S(2), \mathfrak{G}^2) = \tilde{P}^* \tilde{P} = \begin{pmatrix} P_0^* & X_1^* \\ 0 & P_0^* \end{pmatrix} \begin{pmatrix} P_0 & 0 \\ X_1 & P_0 \end{pmatrix}. \quad (3.6)$$

Comparing this with

$$S(S(2), \mathfrak{G}^2) \stackrel{(3.2)}{=} S(1) \stackrel{(3.3)}{=} \begin{pmatrix} P_0^* & P_1^* \\ 0 & P_0^* \end{pmatrix} \begin{pmatrix} P_0 & 0 \\ P_1 & P_0 \end{pmatrix}, \quad (3.7)$$

we get  $P_0^* P_1 = P_0^* X_1$ . By (3.5),  $\text{ran } X_1 \subseteq \overline{\text{ran } P_0}$ , and therefore  $X_1 = P_1$ . Now choose  $P_2 = X_2$  to obtain

$$S(2) = \begin{pmatrix} P_0^* & P_1^* & P_2^* \\ 0 & P_0^* & P_1^* \\ 0 & 0 & P_0^* \end{pmatrix} \begin{pmatrix} P_0 & 0 & 0 \\ P_1 & P_0 & 0 \\ P_2 & P_1 & P_0 \end{pmatrix}, \quad (3.8)$$

$$\text{ran} \begin{pmatrix} P_1 \\ P_2 \end{pmatrix} \subseteq \overline{\text{ran}} \begin{pmatrix} P_0 & 0 \\ P_1 & P_0 \end{pmatrix}. \quad (3.9)$$

**Inductive step.** We continue in the same way for all  $k = 1, 2, 3, \dots$ . At the  $k$ -th stage, the procedure produces operators  $P_0, \dots, P_k$  such that

$$S(k) = \begin{pmatrix} P_0^* & \cdots & P_k^* \\ & \ddots & \vdots \\ 0 & & P_0^* \end{pmatrix} \begin{pmatrix} P_0 & & 0 \\ \vdots & \ddots & \\ P_k & \cdots & P_0 \end{pmatrix}, \quad (3.10)$$

$$\text{ran} \begin{pmatrix} P_1 \\ \vdots \\ P_k \end{pmatrix} \subseteq \overline{\text{ran}} \begin{pmatrix} P_0 & & 0 \\ \vdots & \ddots & \\ P_k & \cdots & P_0 \end{pmatrix}. \quad (3.11)$$

By Lemma 3.3, in the case  $k \geq m$ ,

$$S(k) = \begin{pmatrix} Q_0 & (Q_{-1} & \cdots & Q_{-m} & 0 & \cdots & 0) \\ \left( \begin{array}{c} Q_{-1}^* \\ \vdots \\ Q_{-m}^* \\ 0 \\ \vdots \\ 0 \end{array} \right) & S(k-1) \end{pmatrix}. \quad (3.12)$$

The zeros appear here when  $k > m$ , and their presence leads to the conclusion that  $P_k = 0$  for  $k > m$ . We set then

$$P(z) = P_0 + P_1 z + \cdots + P_m z^m.$$

Comparing (3.10) and (3.12) in the case  $k = m$ , we deduce  $2m + 1$  relations which are equivalent to the identity

$$Q(\zeta) = P(\zeta)^* P(\zeta), \quad \zeta \in \mathbb{T}.$$

**Final step:  $P(z)$  is outer.** Define  $T_P$  as in (2.4). With natural identifications,

$$T_P = \begin{pmatrix} P_0 & (0 & 0 & \cdots) \\ \begin{pmatrix} P_1 \\ P_2 \\ \vdots \end{pmatrix} & T_P \end{pmatrix}. \quad (3.13)$$

The relations (3.11), combined with the fact that  $P_k = 0$  for all  $k > m$ , imply that

$$\text{ran} \begin{pmatrix} P_1 \\ P_2 \\ \vdots \end{pmatrix} \subseteq \overline{\text{ran}} T_P.$$

Hence for any  $g \in \mathfrak{G}$ , a sequence  $f_n$  can be found such that

$$T_P f_n \rightarrow \begin{pmatrix} P_1 \\ P_2 \\ \vdots \end{pmatrix} g.$$

Then by (3.13),

$$T_P \begin{pmatrix} g \\ f_n \end{pmatrix} \rightarrow \begin{pmatrix} P_0 \\ 0 \end{pmatrix} g.$$

It follows that  $\overline{\text{ran}} T_P$  contains every vector  $(P_0 g, 0, 0, \dots)$  with  $g \in \mathfrak{L}(\mathfrak{G})$ , and hence  $\overline{\text{ran}} T_P \supseteq \overline{\text{ran}} P_0 \oplus \overline{\text{ran}} P_0 \oplus \cdots$ . The reverse inclusion holds because by (3.11), the ranges of  $P_1, P_2, \dots$  are all contained in  $\overline{\text{ran}} P_0$ . Thus  $P(z)$  is outer.  $\square$

#### 4. Spectral factorization

The problem of spectral factorization is to write a nonnegative operator-valued function  $F$  on the unit circle in the form  $F = G^* G$  where  $G$  is analytic (in a sense made precise below). The terminology comes from prediction theory, where the nonnegative function  $F$  plays the role of a spectral density for a multidimensional stationary stochastic process. The problem may be viewed as a generalization of a classical theorem of Szegő from Hardy class theory and the theory of orthogonal polynomials (see Hoffman [35, p. 56] and Szegő [62, Chapter X]).

We write  $H^p$  and  $L^p$  for the standard Hardy and Lebesgue spaces for the unit disk and unit circle. See Duren [20]. Recall that  $\sigma$  is normalized Lebesgue measure on the unit circle  $\mathbb{T}$ .

**Theorem 4.1 (Szegő's Theorem).** *Let  $w \in L^1$  satisfy  $w \geq 0$  a.e. on  $\mathbb{T}$  and*

$$\int_{\mathbb{T}} \log w(\zeta) d\sigma > -\infty.$$

*Then  $w = |h|^2$  a.e. on  $\mathbb{T}$  for some  $h \in H^2$ , and  $h$  can be chosen to be an outer function.*

Operator and matrix generalizations of Szegő's theorem are stated in Theorems 4.5 and 4.7 below. Some vectorial function theory is needed to formulate these and other results. We assume familiarity with basic concepts but recall a few definitions. For details, see e.g. [30, 61]) and [50, Chapter 4].

In this section,  $\mathfrak{G}$  denotes a separable Hilbert space. Functions  $f$  and  $F$  on the unit circle with values in  $\mathfrak{G}$  and  $\mathfrak{L}(\mathfrak{G})$ , respectively, are called weakly measurable if  $\langle f(\zeta), v \rangle$  and  $\langle F(\zeta)u, v \rangle$  are measurable for all  $u, v \in \mathfrak{G}$ . Nontangential limits for analytic functions on the unit disk are taken in the strong (norm) topology for vector-valued functions, and in the strong operator topology for operator-valued functions. We fix notation as follows:

- (i) We write  $L_{\mathfrak{G}}^2$  and  $L_{\mathfrak{L}(\mathfrak{G})}^{\infty}$  for the standard Lebesgue spaces of weakly measurable functions on the unit circle with values in  $\mathfrak{G}$  and  $\mathfrak{L}(\mathfrak{G})$ .
- (ii) Let  $H_{\mathfrak{G}}^2$  and  $H_{\mathfrak{L}(\mathfrak{G})}^{\infty}$  be the analogous Hardy classes of analytic functions on the unit disk. We identify elements of these spaces with their nontangential boundary functions, and so the spaces may alternatively be viewed as subspaces of  $L_{\mathfrak{G}}^2$  and  $L_{\mathfrak{L}(\mathfrak{G})}^{\infty}$ .
- (iii) Let  $N_{\mathfrak{L}(\mathfrak{G})}^+$  be the space of all analytic functions  $F$  on the unit disk such that  $\varphi F$  belongs to  $H_{\mathfrak{L}(\mathfrak{G})}^{\infty}$  for some bounded scalar outer function  $\varphi$ . The elements of  $N_{\mathfrak{L}(\mathfrak{G})}^+$  are also identified with their nontangential boundary functions.

A function  $F \in H_{\mathfrak{L}(\mathfrak{G})}^{\infty}$  is called **outer** if  $FH_{\mathfrak{G}}^2$  is dense in  $H_{\mathfrak{F}}^2$  for some closed subspace  $\mathfrak{F}$  of  $\mathfrak{G}$ . A function  $F \in N_{\mathfrak{L}(\mathfrak{G})}^+$  is **outer** if there is a bounded scalar outer function  $\varphi$  such that  $\varphi F \in H_{\mathfrak{L}(\mathfrak{G})}^{\infty}$  and  $\varphi F$  is outer in the sense just defined. The definition of an outer function given here is consistent with the previously defined notion for polynomials in §2.

A function  $A \in H_{\mathfrak{L}(\mathfrak{G})}^{\infty}$  is called **inner** if multiplication by  $A$  on  $\mathfrak{H}_{\mathfrak{G}}^2$  is a partial isometry. In this case, the initial space of multiplication by  $A$  is a subspace of  $\mathfrak{H}_{\mathfrak{G}}^2$  of the form  $\mathfrak{H}_{\mathfrak{F}}^2$  where  $\mathfrak{F}$  is a closed subspace of  $\mathfrak{G}$ . To prove this, notice that both the kernel of multiplication by  $A$  and the set on which it is isometric are invariant under multiplication by  $z$ . Therefore the initial space of multiplication by  $A$  is a reducing subspace for multiplication by  $z$ , and so it has the form  $\mathfrak{H}_{\mathfrak{F}}^2$  where  $\mathfrak{F}$  is a closed subspace of  $\mathfrak{G}$  (see [29, p. 106] and [50, p. p.96]).

Every  $F \in H_{\mathfrak{L}(\mathfrak{G})}^{\infty}$  has an **inner-outer factorization**  $F = AG$ , where  $A$  is an inner function and  $G$  is an outer function. This factorization can be chosen such that the isometric set  $H_{\mathfrak{F}}^2$  for multiplication by  $A$  on  $H_{\mathfrak{G}}^2$  coincides with the closure of the range of multiplication by  $G$ . The inner-outer factorization is extended in an obvious way to functions  $F \in N_{\mathfrak{L}(\mathfrak{G})}^+$ . Details are given, for example, in [50, Chapter 5].

The main problem of this section can now be interpreted more precisely:

**Factorization Problem.** *Given a nonnegative weakly measurable function  $F$  on  $\mathbb{T}$ , find a function  $G$  in  $N_{\mathfrak{L}(\mathfrak{G})}^+$  such that  $F = G^*G$  a.e. on  $\mathbb{T}$ . If such a function exists, we say that  $F$  is **factorable**.*

If a factorization exists, the factor  $G$  can be chosen to be outer by the inner-outer factorization. Moreover, an outer factor  $G$  can be chosen such that  $G(0) \geq 0$ , and then it is unique [50, p. 101]. By the definition of  $N_{\mathfrak{L}(\mathfrak{G})}^+$ , a necessary condition for  $F$  to be factorable is that

$$\int_{\mathbb{T}} \log^+ \|F(\zeta)\| d\sigma < \infty, \quad (4.1)$$

where  $\log^+ x$  is zero or  $\log x$  according as  $0 \leq x \leq 1$  or  $1 < x < \infty$ , and so we only need consider functions which satisfy (4.1). In fact, in proofs we can usually reduce to the bounded case by considering  $F/|\varphi|^2$  for a suitable scalar outer function  $\varphi$ .

The following result is another view of Lowdenslager's criterion, which we deduce from Lemma 2.3. A direct proof is given in [61, pp. 201–203].

**Lemma 4.2.** *Suppose  $F \in L_{\mathfrak{L}(\mathfrak{G})}^\infty$  and  $F \geq 0$  a.e. on  $\mathbb{T}$ . Let  $\mathfrak{R}_F$  be the closure of  $F^{\frac{1}{2}}H_{\mathfrak{G}}^2$  in  $L_{\mathfrak{G}}^2$ , and let  $S_F$  be the isometry multiplication by  $\zeta$  on  $\mathfrak{R}_F$ . Then  $F$  is factorable if and only if  $S_F$  is a shift operator, that is,*

$$\bigcap_{n=0}^{\infty} \zeta^n \overline{F^{\frac{1}{2}}H_{\mathfrak{G}}^2} = \{0\}. \quad (4.2)$$

*Proof.* In Lemma 2.3 take  $\mathfrak{H} = H_{\mathfrak{G}}^2$  viewed as a subspace of  $L_{\mathfrak{G}}^2$ , and let  $S$  be multiplication by  $\zeta$  on  $\mathfrak{H}$ . Define  $T \in \mathfrak{L}(\mathfrak{H})$  by  $Tf = PFf$ ,  $f \in \mathfrak{H}$ , where  $P$  is the projection from  $L_{\mathfrak{G}}^2$  onto  $H_{\mathfrak{G}}^2$ . One sees easily that  $T$  is a nonnegative Toeplitz operator, and so we can define  $\mathfrak{H}_T$  and an isometry  $S_T$  as in Lemma 2.3. In fact,  $S_T$  is unitarily equivalent to  $S_F$  via the natural isomorphism  $W: \mathfrak{H}_T \rightarrow \mathfrak{R}_F$  such that  $W(T^{\frac{1}{2}}f) = F^{\frac{1}{2}}f$  for every  $f$  in  $\mathfrak{H}$ . Thus  $S_F$  is a shift operator if and only if  $S_T$  is a shift operator, and by Lemma 2.3 this is the same as saying that  $T = A^*A$  where  $A \in \mathfrak{L}(\mathfrak{H})$  is analytic, or equivalently  $F$  is factorable [50, p. 110].  $\square$

We obtain a very useful sufficient condition for factorability.

**Theorem 4.3.** *Suppose  $F \in L_{\mathfrak{L}(\mathfrak{G})}^\infty$  and  $F \geq 0$  a.e. For  $F$  to be factorable, it is sufficient that there exists a function  $\psi$  in  $L_{\mathfrak{L}(\mathfrak{G})}^\infty$  such that*

- (i)  $\psi F \in H_{\mathfrak{L}(\mathfrak{G})}^\infty$ ;
- (ii) for all  $\zeta \in \mathbb{T}$  except at most a set of measure zero,  $\psi(\zeta) \overline{F(\zeta)\mathfrak{G}}$  is one-to-one.

*If these conditions are met and  $F = G^*G$  a.e. with  $G$  outer, then  $\psi G^* \in H_{\mathfrak{L}(\mathfrak{G})}^\infty$ .*

Theorem 4.3 appears in Rosenblum [49] with  $\psi(\zeta) = \zeta^m$  (viewed as an operator-valued function). The case of an arbitrary inner function was proved and applied in a variety of ways by Rosenblum and Rovnyak [50, 51]. V. I. Matsaev first showed that more general functions  $\psi$  can be used. Matsaev's result is

evidently unpublished, but versions were given by D. Z. Arov [1, Lemma to Theorem 4] and A. S. Markus [41, Theorem 34.3 on p. 199]. Theorem 4.3 includes all of these versions.

We do not know if the conditions (i) and (ii) in Theorem 4.3 are necessary for factorability. It is not hard to see that they are necessary in the simple cases  $\dim \mathfrak{G} = 1$  and  $\dim \mathfrak{G} = 2$  (for the latter case, one can use [50, Example 1, p. 125]). The general case, however, is open.

*Proof of Theorem 4.3.* Let  $F$  satisfy (i) and (ii). Define a subspace  $\mathfrak{M}$  of  $L^2_{\mathfrak{G}}$  by

$$\mathfrak{M} = \bigcap_{n=0}^{\infty} \zeta^n \overline{F^{\frac{1}{2}} H^2_{\mathfrak{G}}} = \bigcap_{n=0}^{\infty} \overline{\zeta^n F^{\frac{1}{2}} H^2_{\mathfrak{G}}}.$$

We show that  $\mathfrak{M} = \{0\}$ . By (i),

$$\psi F^{\frac{1}{2}} \mathfrak{M} = \psi F^{\frac{1}{2}} \bigcap_{n=0}^{\infty} \overline{\zeta^n F^{\frac{1}{2}} H^2_{\mathfrak{G}}} \subseteq \bigcap_{n=0}^{\infty} \overline{\zeta^n \psi F H^2_{\mathfrak{G}}} \subseteq \bigcap_{n=0}^{\infty} \zeta^n H^2_{\mathfrak{G}} = \{0\}. \quad (4.3)$$

Thus  $\psi F^{\frac{1}{2}} \mathfrak{M} = \{0\}$ . Now if  $g \in \mathfrak{M}$ , then  $\psi F^{\frac{1}{2}} g = 0$  a.e. by (4.3). Hence  $F^{\frac{1}{2}} g = 0$  a.e. by (ii). By the definition of  $\mathfrak{M}$ ,  $g \in \overline{F^{\frac{1}{2}} H^2_{\mathfrak{G}}}$ , and standard arguments show from this that  $g(\zeta) \in \overline{F(\zeta)^{\frac{1}{2}} \mathfrak{G}}$  a.e. Therefore  $g = 0$  a.e. It follows that  $\mathfrak{M} = \{0\}$ , and so  $F$  is factorable by Lemma 4.2.

Let  $F = G^*G$  a.e. with  $G$  outer. We prove that  $\psi G^* \in H^{\infty}_{\mathfrak{L}(\mathfrak{G})}$  by showing that  $\psi G^* H^2_{\mathfrak{G}} \subseteq H^2_{\mathfrak{G}}$ . Since  $G$  is outer,  $\overline{GH^2_{\mathfrak{G}}} = H^2_{\mathfrak{F}}$  for some closed subspace  $\mathfrak{F}$  of  $\mathfrak{G}$ . By (i),

$$\psi G^*(GH^2_{\mathfrak{G}}) = \psi FH^2_{\mathfrak{G}} \subseteq H^2_{\mathfrak{G}}.$$

Therefore  $\psi G^* H^2_{\mathfrak{F}} \subseteq H^2_{\mathfrak{G}}$ . Suppose  $f \in H^2_{\mathfrak{G} \ominus \mathfrak{F}}$ , and consider any  $h \in L^2_{\mathfrak{G}}$ . Then

$$\langle G^* f, h \rangle_{L^2_{\mathfrak{G}}} = \int_{\mathbb{T}} \langle f(\zeta), G(\zeta)h(\zeta) \rangle_{\mathfrak{G}} d\sigma = 0,$$

because  $\text{ran } G(\zeta) \subseteq \mathfrak{F}$  a.e. Thus  $\psi G^* f = 0$  a.e. It follows that  $\psi G^* H^2_{\mathfrak{G}} \subseteq H^2_{\mathfrak{G}}$ , and therefore  $\psi G^* \in H^{\infty}_{\mathfrak{L}(\mathfrak{G})}$ .  $\square$

For a simple application of Theorem 4.3, suppose that  $F$  is a Laurent polynomial of degree  $m$ , and choose  $\psi$  to be  $\zeta^m I$ . In short order, this yields another proof of the operator Fejér-Riesz theorem (Theorem 2.1).

Another application is a theorem of Sarason [55, p. 198], which generalizes the factorization of a scalar-valued function in  $H^1$  as a product of two functions in  $H^2$  (see [35, p. 56]).

**Theorem 4.4.** *Every  $G$  in  $N^+_{\mathfrak{L}(\mathfrak{G})}$  can be written  $G = G_1 G_2$ , where  $G_1$  and  $G_2$  belong to  $N^+_{\mathfrak{L}(\mathfrak{G})}$  and*

$$G_2^* G_2 = [G^* G]^{1/2} \quad \text{and} \quad G_1^* G_1 = G_2 G_2^* \quad \text{a.e.}$$

*Proof.* Suppose first that  $G \in H_{\mathfrak{L}(\mathfrak{G})}^\infty$ . For each  $\zeta \in \mathbb{T}$ , write

$$G(\zeta) = U(\zeta)[G^*(\zeta)G(\zeta)]^{\frac{1}{2}},$$

where  $U(\zeta)$  is a partial isometry with initial space  $\overline{\text{ran}} [G^*(\zeta)G(\zeta)]^{\frac{1}{2}}$ . It can be shown that  $U$  is weakly measurable. We apply Theorem 4.3 with  $F = [G^*G]^{\frac{1}{2}}$  and  $\psi = U$ . Conditions (i) and (ii) of Theorem 4.3 are obviously satisfied, and so we obtain an outer function  $G_2 \in H_{\mathfrak{L}(\mathfrak{G})}^\infty$  such that

$$G_2^*G_2 = [G^*G]^{1/2} \quad \text{a.e.}$$

and  $UG_2^* \in H_{\mathfrak{L}(\mathfrak{G})}^\infty$ . Set  $G_1 = UG_2^*$ . By construction  $G_1 \in H_{\mathfrak{L}(\mathfrak{G})}^\infty$ ,

$$G = U(G^*G)^{\frac{1}{2}} = (UG_2^*)G_2 = G_1G_2,$$

and  $G_1^*G_1 = G_2U^*UG_2^* = G_2G_2^*$  a.e. The result follows when  $G \in H_{\mathfrak{L}(\mathfrak{G})}^\infty$ .

The general case follows on applying what has just been shown to  $\varphi^2G$ , where  $\varphi$  is a scalar-valued outer function such that  $\varphi^2G \in H_{\mathfrak{L}(\mathfrak{G})}^\infty$ .  $\square$

The standard operator generalization of Szegő's theorem also follows from Theorem 4.3.

**Theorem 4.5.** *Let  $F$  be a weakly measurable function on  $\mathbb{T}$  with values in  $\mathfrak{L}(\mathfrak{G})$  satisfying  $F \geq 0$  a.e. and*

$$\int_{\mathbb{T}} \log^+ \|F(\zeta)\| \, d\sigma < \infty \quad \text{and} \quad \int_{\mathbb{T}} \log^+ \|F(\zeta)^{-1}\| \, d\sigma < \infty.$$

*Then  $F$  is factorable.*

*Proof.* Since  $\log^+ \|F(\zeta)\|$  is integrable, we can choose a scalar outer function  $\varphi_1$  such that

$$F_1 = F/|\varphi_1|^2 \in L_{\mathfrak{L}(\mathfrak{G})}^\infty.$$

Since  $\log^+ \|F(\zeta)^{-1}\|$  is integrable, so is  $\log^+ \|F_1(\zeta)^{-1}\|$ . Hence there is a bounded scalar outer function  $\varphi$  such that

$$\varphi F_1^{-1} \in L_{\mathfrak{L}(\mathfrak{G})}^\infty.$$

We apply Theorem 4.3 to  $F_1$  with  $\psi = \varphi F_1^{-1}$ . Condition (i) is satisfied because  $\psi F_1 = \varphi I$ . Condition (ii) holds because the values of  $\psi$  are invertible a.e. Thus  $F_1$  is factorable, and hence so is  $F$ .  $\square$

Theorem 4.3 has a half-plane version, the scalar inner case of which is given in [50, p. 117]. This has an application to the following generalization of Akhiezer's theorem on factoring entire functions [50, Chapter 6].

**Theorem 4.6.** *Let  $F$  be an entire function of exponential type  $\tau$ , having values in  $\mathfrak{L}(\mathfrak{G})$ , such that  $F(x) \geq 0$  for all real  $x$  and*

$$\int_{-\infty}^{\infty} \frac{\log^+ \|F(t)\|}{1+t^2} dt < \infty.$$

*Then  $F(x) = G(x)^*G(x)$  for all real  $x$  where  $G$  is an entire function with values in  $\mathfrak{L}(\mathfrak{G})$  such that  $\exp(-i\tau z/2)G$  is of exponential type  $\tau/2$  and the restriction of  $G$  to the upper half-plane is an outer function.*

**Matrix case:**

We end this section by quoting a few results for matrix-valued functions. The matrix setting is more concrete, and one can do more. Statements often require invertibility assumptions. We give no details and leave it to the interested reader to consult other sources for further information.

Our previous definitions and results transfer in an obvious way to matrix-valued functions. For this we choose  $\mathfrak{G} = \mathbb{C}^r$  for some positive integer  $r$  and identify operators on  $\mathbb{C}^r$  with  $r \times r$  matrices. The operator norm of a matrix is denoted  $\|\cdot\|$ . We write  $L_{r \times r}^\infty, H_{r \times r}^\infty$  in place of  $L_{\mathfrak{L}(\mathfrak{G})}^\infty, H_{\mathfrak{L}(\mathfrak{G})}^\infty$  and  $\|\cdot\|_\infty$  for the norms on these spaces.

Theorem 4.5 is more commonly stated in a different form for matrix-valued functions.

**Theorem 4.7.** *Suppose that  $F$  is an  $r \times r$  measurable matrix-valued function having invertible values on  $\mathbb{T}$  such that  $F \geq 0$  a.e. and  $\log^+ \|F\|$  is integrable. Then  $F$  is factorable if and only if  $\log \det F$  is integrable.*

Recall that when  $F$  is factorable, there is a unique outer  $G$  such that  $F = G^*G$  and  $G(0) \geq 0$ . It makes sense to inquire about the continuity properties of the mapping  $\Phi: F \rightarrow G$  with respect to various norms. For example, see Jacob and Partington [37]. We cite one recent result in this area.

**Theorem 4.8 (Barclay [5]).** *Let  $F, F_n, n = 1, 2, \dots$ , be  $r \times r$  measurable matrix-valued functions on  $\mathbb{T}$  having invertible values a.e. and integrable norms. Suppose that  $F = G^*G$  and  $F_n = G_n^*G_n$ , where  $G, G_n$  are  $r \times r$  matrix-valued outer functions such that  $G(0) \geq 0$  and  $G_n(0) \geq 0, n = 1, 2, \dots$ . Then*

$$\lim_{n \rightarrow \infty} \int_{\mathbb{T}} \|G(\zeta) - G_n(\zeta)\|^2 d\sigma = 0$$

*if and only if*

- (i)  $\lim_{n \rightarrow \infty} \int_{\mathbb{T}} \|F(\zeta) - F_n(\zeta)\| d\sigma = 0$ , and
- (ii) the family of functions  $\{\log \det F_n\}_{n=0}^\infty$  is uniformly integrable.



A family of functions  $\{\varphi_\alpha\}_{\alpha \in A} \subseteq L^1$  is **uniformly integrable** if for every  $\varepsilon > 0$  there is a  $\delta > 0$  such that  $\int_E |\varphi_\alpha| d\sigma < \varepsilon$  for all  $\alpha \in A$  whenever  $\sigma(E) < \delta$ . See [5] for additional references and similar results in other norms.

A theorem of Bourgain [9] characterizes all functions on the unit circle which are products  $\bar{h}g$  with  $g, h \in H^\infty$ : *A function  $f \in L^\infty$  has the form  $f = \bar{h}g$  where  $g, h \in H^\infty$  if and only if  $\log |f|$  is integrable.* This resolves a problem of Douglas and Rudin [17]. The problem is more delicate than spectral factorization; when  $|f| = 1$  a.e., the factorization cannot be achieved in general with inner functions. Bourgain's theorem was recently generalized to matrix-valued functions.

**Theorem 4.9 (Barclay [4, 6]).** *Suppose  $F \in L_{r \times r}^\infty$  has invertible values a.e. Then  $F$  has the form  $F = H^*G$  a.e. for some  $G, H$  in  $H_{r \times r}^\infty$  if and only if  $\log |\det F|$  is integrable. In this case, for every  $\varepsilon > 0$  such a factorization can be found with*

$$\|G\|_\infty \|H\|_\infty < \|F\|_\infty + \varepsilon.$$

The proof of Theorem 4.9 in [6] is long and technical. In fact, Barclay proves an  $L^p$ -version of this result for all  $p$ ,  $1 \leq p \leq \infty$ .

Another type of generalization is factorization with indices. We quote one result to illustrate this notion.

**Theorem 4.10.** *Let  $F$  be an  $r \times r$  matrix-valued function with rational entries. Assume that  $F$  has no poles on  $\mathbb{T}$  and that  $\det F(\zeta) \neq 0$  for all  $\zeta$  in  $\mathbb{T}$ . Then there exist integers  $\varkappa_1 \leq \varkappa_2 \leq \dots \leq \varkappa_r$  such that*

$$F(z) = F_-(z) \text{diag} \{z^{\varkappa_1}, \dots, z^{\varkappa_r}\} F_+(z),$$

where  $F_\pm$  are  $r \times r$  matrix-valued functions with rational entries such that

- (i)  $F_+(z)$  has no poles for  $|z| \leq 1$  and  $\det F_+(z) \neq 0$  for  $|z| \leq 1$ ;
- (ii)  $F_-(z)$  has no poles for  $|z| \geq 1$  including  $z = \infty$  and  $\det F_-(z) \neq 0$  for  $|z| \geq 1$  including  $z = \infty$ .

The case in which  $F$  is nonnegative on  $\mathbb{T}$  can be handled using the operator Fejér-Riesz theorem (the indices are all zero in this case). The general case is given in Gohberg, Goldberg, and Kaashoek [27, pp. 236–239]. This is a large subject that includes, for example, general theories of factorization in Bart, Gohberg, Kaashoek, and Ran [7] and Clancey and Gohberg [13].

#### Historical remarks:

Historical accounts of spectral factorization appear in [2, 30, 50, 52, 61]. Briefly, the problem of factoring nonnegative matrix-valued functions on the unit circle rose to prominence in the prediction theory of multivariate stationary stochastic processes. The first results of this theory were announced by Zasuhrin [65] without complete proofs; proofs were supplied by M. G. Kreĭn in lectures. Modern accounts of prediction theory and matrix generalizations of Szegő's theorem are based on fundamental papers of Helson and Lowdenslager [31, 32], and Wiener and Masani [63, 64]. The general case of Theorem 4.5 is due to Devinatz [15];

other proofs are given in [16, 30, 50]. For an engineering view and computational methods, see [38, Chapter 8] and [56].

The original source for Lowdenslager's Criterion (Lemmas 2.3 and 4.2) is [40]; an error in [40] was corrected by Douglas [16]. There is a generalization, given by Sz.-Nagy and Foias [61, pp. 201–203], in which the isometry may have a nontrivial unitary component and the shift component yields a maximal factorable summand. Lowdenslager's Criterion is used in the construction of canonical models of operators by de Branges [10]. See also Constantinescu [14] for an adaptation to Toeplitz kernels and additional references.

## 5. Multivariable theory

It is natural to wonder to what extent the results for one variable carry over to several variables. Various interpretations of "several variables" are possible. The most straightforward is to consider Laurent polynomials in complex variables  $z_1, \dots, z_d$  that are nonnegative on the  $d$ -torus  $\mathbb{T}^d$ . The method of Schur complements in §3 suggests an approach to the factorization problem for such polynomials. Care is needed, however, since the first conjectures for a multivariable Fejér-Riesz theorem that might come to mind are false, as explained below. Multivariable generalizations of the Fejér-Riesz theorem are thus necessarily weaker than the one-variable result. One difficulty has to do with degrees, and if the condition on degrees is relaxed, there is a neat result in the strictly positive case (Theorem 5.1).

By a Laurent polynomial in  $z = (z_1, \dots, z_d)$  we understand an expression

$$Q(z) = \sum_{k_1=-m_1}^{m_1} \cdots \sum_{k_d=-m_d}^{m_d} Q_{k_1, \dots, k_d} z_1^{k_1} \cdots z_d^{k_d}. \quad (5.1)$$

We assume that the coefficients belong to  $\mathfrak{L}(\mathfrak{G})$ , where  $\mathfrak{G}$  is a Hilbert space. With obvious interpretations, the scalar case is included. By an analytic polynomial with coefficients in  $\mathfrak{L}(\mathfrak{G})$  we mean an analogous expression, of the form

$$P(z) = \sum_{k_1=0}^{m_1} \cdots \sum_{k_d=0}^{m_d} P_{k_1, \dots, k_d} z_1^{k_1} \cdots z_d^{k_d}. \quad (5.2)$$

The numbers  $m_1, \dots, m_d$  in (5.1) and (5.2) are upper bounds for the degrees of the polynomials in  $z_1, \dots, z_d$ , which we define as the smallest values of  $m_1, \dots, m_d$  that can be used in the representations (5.1) and (5.2).

Suppose that  $Q(z)$  has the form (5.1) and satisfies  $Q(\zeta) \geq 0$  for all  $\zeta \in \mathbb{T}^d$ , that is, for all  $\zeta = (\zeta_1, \dots, \zeta_d)$  with  $|\zeta_1| = \cdots = |\zeta_d| = 1$ . Already in the scalar case, one cannot always find an analytic polynomial  $P(z)$  such that  $Q(\zeta) = P(\zeta)^*P(\zeta)$ ,  $\zeta \in \mathbb{T}^d$ . This was first explicitly shown by Lebow and Schreiber [39]. There are also difficulties in writing  $Q(\zeta) = \sum_{j=1}^r P_j(\zeta)^*P_j(\zeta)$ ,  $\zeta \in \mathbb{T}^d$ , for some finite set of analytic polynomials, at least if one requires that the degrees of the analytic polynomials do not exceed those of  $Q(z)$  as in the one-variable case (see

Naftalevich and Schreiber [44], Rudin [53], and Sakhnovich [54, §3.6]). The example in [44] is based on a Cayley transform of a version of a real polynomial over  $\mathbb{R}^2$  called Motzkin's polynomial, which was the first explicit example of a nonnegative polynomial in  $\mathbb{R}^d$ ,  $d > 1$ , which is not a sum of squares of polynomials. What is not mentioned in these sources is that if we loosen the restriction on degrees, the polynomial in [44] can be written as a sum of squares (see [19]). Nevertheless, for three or more variables, very general results of Scheiderer [57] imply that there exist nonnegative, but not strictly positive, polynomials which cannot be expressed as such finite sums regardless of degrees.

**Theorem 5.1.** *Let  $Q(z)$  be a Laurent polynomial in  $z = (z_1, \dots, z_d)$  with coefficients in  $\mathfrak{L}(\mathfrak{G})$  for some Hilbert space  $\mathfrak{G}$ . Suppose that there is a  $\delta > 0$  such that  $Q(\zeta) \geq \delta I$  for all  $\zeta \in \mathbb{T}^d$ . Then*

$$Q(\zeta) = \sum_{j=1}^r P_j(\zeta)^* P_j(\zeta), \quad \zeta \in \mathbb{T}^d, \quad (5.3)$$

for some analytic polynomials  $P_1(z), \dots, P_r(z)$  in  $z = (z_1, \dots, z_d)$  which have coefficients in  $\mathfrak{L}(\mathfrak{G})$ . Furthermore, for any fixed  $k$ , the representation (5.3) can be chosen such that the degree of each analytic polynomial in  $z_k$  is no more than the degree of  $Q(z)$  in  $z_k$ .

The scalar case of Theorem 5.1 follows by a theorem of Schmüdgen [58], which states that strictly positive polynomials over compact semialgebraic sets in  $\mathbb{R}^n$  (that is, sets which are expressible in terms of finitely many polynomial inequalities) can be written as weighted sums of squares, where the weights are the polynomials used to define the semialgebraic set (see also [12]); the proof is nonconstructive. On the other hand, the proof we sketch using Schur complements covers the operator-valued case, and it gives an algorithm for finding the solution. One can also give estimates for the degrees of the polynomials involved, though we have not stated these.

We prove Theorem 5.1 for the case  $d = 2$ , following Dritschel [18]. The general case is similar. The argument mimics the method of Schur complements, especially in its original form used in [18]. In place of Toeplitz matrices whose entries are operators, in the case of two variables we use Toeplitz matrices whose entries are themselves Toeplitz matrices. The fact that the first level Toeplitz blocks are infinite in size causes problems, and so we truncate these blocks to finite size. Then everything goes through, but instead of factoring the original polynomial  $Q(z)$ , the result is a factorization of polynomials  $Q^{(N)}(z)$  that are close to  $Q(z)$ . When  $Q(\zeta) \geq \delta I$  on  $\mathbb{T}^d$  for some  $\delta > 0$ , there is enough wiggle room to factor  $Q(z)$  itself. We isolate the main steps in a lemma.

**Lemma 5.2.** *Let*

$$Q(z) = \sum_{j=-m_1}^{m_1} \sum_{k=-m_2}^{m_2} Q_{jk} z_1^j z_2^k$$

be a Laurent polynomial with coefficients in  $\mathfrak{L}(\mathfrak{G})$  such that  $Q(\zeta) \geq 0$  for all  $\zeta = (\zeta_1, \zeta_2)$  in  $\mathbb{T}^2$ . Set

$$Q^{(N)}(z) = \sum_{j=-m_1}^{m_1} \sum_{k=-m_2}^{m_2} \frac{N+1-|k|}{N+1} Q_{jk} z_1^j z_2^k.$$

Then for each  $N \geq m_2$ , there are analytic polynomials

$$F_\ell(z) = \sum_{j=0}^{m_1} \sum_{k=0}^N F_{jk}^{(\ell)} z_1^j z_2^k, \quad \ell = 0, \dots, N, \quad (5.4)$$

with coefficients in  $\mathfrak{L}(\mathfrak{G})$  such that

$$Q^{(N)}(\zeta) = \sum_{\ell=0}^N F_\ell(\zeta)^* F_\ell(\zeta), \quad \zeta \in \mathbb{T}^2. \quad (5.5)$$

*Proof.* Write

$$Q(z) = \sum_{j=-m_1}^{m_1} \left( \sum_{k=-m_2}^{m_2} Q_{jk} z_2^k \right) z_1^j = \sum_{j=-m_1}^{m_1} R_j(z_2) z_1^j,$$

and extend all sums to run from  $-\infty$  to  $\infty$  by setting  $Q_{jk} = 0$  and  $R_j(z_2) = 0$  if  $|j| > m_1$  or  $|k| > m_2$ . Introduce a Toeplitz matrix  $T$  whose entries are the Toeplitz matrices  $T_j$  corresponding to the Laurent polynomials  $R_j(z_2)$ , that is,

$$T = \begin{pmatrix} T_0 & T_{-1} & T_{-2} & \cdots \\ T_1 & T_0 & T_{-1} & \ddots \\ T_2 & T_1 & T_0 & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}, \quad T_j = \begin{pmatrix} Q_{j0} & Q_{j,-1} & Q_{j,-2} & \cdots \\ Q_{j1} & Q_{j0} & Q_{j,-1} & \ddots \\ Q_{j2} & Q_{j1} & Q_{j0} & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix},$$

$j = 0, \pm 1, \pm 2, \dots$ . Notice that  $T$  is finitely banded, since  $T_j = 0$  for  $|j| > m_1$ . The identity (2.5) has the following generalization:

$$\langle Th, h \rangle = \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \langle T_{q-p} h_p, h_q \rangle = \int_{\mathbb{T}^2} \langle Q(\zeta) h(\zeta), h(\zeta) \rangle_{\mathfrak{G}} d\sigma_2(\zeta).$$

Here  $\zeta = (\zeta_1, \zeta_2)$  and  $d\sigma_2(\zeta) = d\sigma(\zeta_1) d\sigma(\zeta_2)$ . Also,

$$h(\zeta) = \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} h_{pq} \zeta_1^p \zeta_2^q,$$

where the coefficients are vectors in  $\mathfrak{G}$  and all but finitely many are zero, and

$$h = \begin{pmatrix} h_0 \\ h_1 \\ \vdots \end{pmatrix}, \quad h_p = \begin{pmatrix} h_{p0} \\ h_{p1} \\ \vdots \end{pmatrix}, \quad p = 0, 1, 2, \dots$$

It follows that  $T$  acts as a bounded operator on a suitable direct sum of copies of  $\mathfrak{G}$ . Since  $Q(\zeta) \geq 0$  on  $\mathbb{T}^2$ ,  $T \geq 0$ .

Fix  $N \geq m_2$ . Set

$$T' = \begin{pmatrix} T'_0 & T'_{-1} & T'_{-2} & \cdots \\ T'_1 & T'_0 & T'_{-1} & \ddots \\ T'_2 & T'_1 & T'_0 & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix},$$

where  $T'_j$  is the upper  $(N+1) \times (N+1)$  block of  $T_j$  with a normalizing factor:

$$T'_j = \frac{1}{N+1} \begin{pmatrix} Q_{j0} & Q_{j,-1} & \cdots & Q_{j,-N} \\ Q_{j1} & Q_{j0} & \cdots & Q_{j,-N+1} \\ \cdots & \cdots & \cdots & \cdots \\ Q_{jN} & Q_{j,N-1} & \cdots & Q_{j0} \end{pmatrix}, \quad j = 0, \pm 1, \pm 2, \dots$$

Then  $T'$  is the Toeplitz matrix corresponding to the Laurent polynomial

$$\Psi(w) = \sum_{j=-m_1}^{m_1} T'_j w^j.$$

Moreover,  $T' \geq 0$  since it is a positive constant multiple of a compression of  $T$ . Thus  $\Psi(w) \geq 0$  for  $|w| = 1$ . By the operator Fejér-Riesz theorem (Theorem 2.1),

$$\Psi(w) = \Phi(w)^* \Phi(w), \quad |w| = 1, \quad (5.6)$$

for some analytic polynomial  $\Phi(w) = \sum_{j=0}^{m_1} \Phi_j w^j$  whose coefficients are  $(N+1) \times (N+1)$  matrices with entries in  $\mathfrak{L}(\mathfrak{G})$ . Write

$$\Phi_j = (\Phi_{jN} \quad \Phi_{j,N-1} \quad \cdots \quad \Phi_{j0}),$$

where  $\Phi_{jk}$  is the  $k$ -th column in  $\Phi_j$ . Set

$$\tilde{F}(z) = \sum_{j=0}^{m_1} \sum_{k=0}^N \Phi_{jk} z_1^j z_2^k.$$

The identity (5.6) is equivalent to  $2m_1 + 1$  relations for the coefficients of  $\Psi(w)$ . The coefficients of  $\Psi(w)$  are constant on diagonals, there being  $N+1-k$  terms in the  $k$ -th diagonal above the main diagonal, and similarly below. If these terms are summed, the result gives  $2m_1 + 1$  relations equivalent to the identity

$$Q^{(N)}(\zeta) = \tilde{F}(\zeta)^* \tilde{F}(\zeta), \quad \zeta \in \mathbb{T}^2. \quad (5.7)$$

We omit the calculation, which is straightforward but laborious. To convert (5.7) to the form (5.5), write

$$\Phi_{jk} = \begin{pmatrix} F_{jk}^{(0)} \\ F_{jk}^{(1)} \\ \vdots \\ F_{jk}^{(N)} \end{pmatrix}, \quad j = 0, \dots, m_1 \text{ and } k = 0, \dots, N.$$

Then

$$\tilde{F}(z) = \begin{pmatrix} F_0(z) \\ F_1(z) \\ \vdots \\ F_N(z) \end{pmatrix},$$

where  $F_0(z), \dots, F_N(z)$  are given by (5.4), and so (5.7) takes the form (5.5).  $\square$

*Proof of Theorem 5.1 for the case  $d = 2$ .* Suppose  $N \geq m_2$ , and set

$$\tilde{Q}(z) = \sum_{j=-m_1}^{m_1} \sum_{k=-m_2}^{m_2} \frac{N+1}{N+1-|k|} Q_{jk} z_1^j z_2^k.$$

The values of  $\tilde{Q}(z)$  are selfadjoint on  $\mathbb{T}^2$ , and  $\tilde{Q}(z) = Q(z) + S(z)$ , where

$$S(z) = \sum_{j=-m_1}^{m_1} \sum_{k=-m_2}^{m_2} \frac{|k|}{N+1-|k|} Q_{jk} z_1^j z_2^k.$$

Now choose  $N$  large enough that  $\|S(\zeta)\| < \delta$ ,  $\zeta \in \mathbb{T}^2$ . Then  $\tilde{Q}(\zeta) \geq 0$  on  $\mathbb{T}^2$ , and the result follows on applying Lemma 5.2 to  $\tilde{Q}(z)$ .  $\square$

Further details can be found in [18], and a variation on this method yielding good numerical results is given in Geronimo and Lai [23].

While, as we mentioned, there is in general little hope of finding a factorization of a positive trigonometric polynomial in two or more variables in terms of one or more analytic polynomials of the same degree, it happens that there are situations where the existence of such a factorization is important. In particular, Geronimo and Woerdeman consider this question in the context of the autoregressive filter problem [24, 25], with the first paper addressing the scalar case and the second the operator-valued case, both in two variables. They show that for scalar-valued polynomials in this setting there exists a factorization in terms of a single **stable** (so invertible in the bidisk  $\mathbb{D}^2$ ) analytic polynomial of the same degree if and only if a full rank condition holds for certain submatrices of the associated Toeplitz matrix ([24, Theorem 1.1.3]). The condition for operator-valued polynomials is similar, but more complicated to state. We refer the reader to the original papers for details.

Stable scalar polynomials in one variable are by definition outer, so the Geronimo and Woerdeman results can be viewed as a statement about outer factorizations in two variables. In [19], a different notion of outerness is considered. As we saw in §3, in one variable outer factorizations can be extracted using Schur complements. The same Schur complement method in two or more variables gives rise to a version of “outer” factorization which in general does not agree with that coming from stable polynomials. In [19], this Schur complement version of outerness is used when considering outer factorizations for polynomials in two or more variables. As in the Geronimo and Woerdeman papers, it is required that the factorization be in terms of a single analytic polynomial of the same degree as the polynomial being factored. Then necessary and sufficient conditions for such an outer factorization under these constraints are found ([19, Theorem 4.1]).

The problem of spectral factorization can also be considered in the multivariable setting. Blower [8] has several results along these lines for bivariate matrix-valued functions, including a matrix analogue of Szegő’s theorem similar to Theorem 4.7. His results are based on a two-variable matrix version of Theorem 5.1, and the arguments he gives coupled with Theorem 5.1 can be used to extend these results to polynomials in  $d > 2$  variables as well.

## 6. Noncommutative factorization

We now present some noncommutative interpretations of the notion of “several variables,” starting with the one most frequently considered, and for which there is an analogue of the Fejér-Riesz theorem. It is due to Scott McCullough and comes very close to the one-variable result. Further generalizations have been obtained by Helton, McCullough and Putinar in [33]. For a broad overview of the area, two nice survey articles have recently appeared by Helton and Putinar [34] and Schmüdgen [59] covering noncommutative real algebraic geometry, of which the noncommutative analogues of the Fejér-Riesz theorem are one aspect.

In keeping with the assumptions made in [42], all Hilbert spaces in this section are taken to be separable. Fix Hilbert spaces  $\mathfrak{G}$  and  $\mathfrak{H}$ , and assume that  $\mathfrak{H}$  is infinite dimensional.

Let  $S$  be the free semigroup with generators  $a_1, \dots, a_d$ . Thus  $S$  is the set of words

$$w = a_{j_1} \cdots a_{j_k}, \quad j_1, \dots, j_k \in \{1, \dots, d\}, \quad k = 0, 1, 2, \dots, \quad (6.1)$$

with the binary operation concatenation. The empty word is denoted  $e$ . The length of the word (6.1) is  $|w| = k$  (so  $|e| = 0$ ). Let  $S_m$  be the set of all words (6.1) of length at most  $m$ . The cardinality of  $S_m$  is  $\ell_m = 1 + d + d^2 + \cdots + d^m$ .

We extend  $S$  to a free group  $G$ . We can think of the elements of  $G$  as words in  $a_1, \dots, a_d, a_1^{-1}, \dots, a_d^{-1}$ , with two such words identified if one can be obtained from the other by cancelling adjacent terms of the form  $a_j$  and  $a_j^{-1}$ . The binary operation in  $G$  is also concatenation. Words in  $G$  of the form  $h = v^{-1}w$  with

$v, w \in S$  play a special role and are called **hereditary**. Notice that a hereditary word  $h$  has many representations  $h = v^{-1}w$  with  $v, w \in S$ . Let  $H_m$  be the set of hereditary words  $h$  which have at least one representation in the form  $h = v^{-1}w$  with  $v, w \in S_m$ .

We can now introduce the noncommutative analogues of Laurent and analytic polynomials. A hereditary polynomial is a formal expression

$$Q = \sum_{h \in H_m} h \otimes Q_h, \quad (6.2)$$

where  $Q_h \in \mathfrak{L}(\mathfrak{G})$  for all  $h$ . Analytic polynomials are hereditary polynomials of the special form

$$P = \sum_{w \in S_m} w \otimes P_w, \quad (6.3)$$

where  $P_w \in \mathfrak{L}(\mathfrak{G})$  for all  $w$ . The identity

$$Q = P^*P$$

is defined to mean that

$$Q_h = \sum_{\substack{v, w \in S_m \\ h = v^{-1}w}} P_v^* P_w, \quad h \in H_d.$$

Next we give meaning to the expressions  $Q(U)$  and  $P(U)$  for hereditary and analytic polynomials (6.2) and (6.3) and any tuple  $U = (U_1, \dots, U_d)$  of unitary operators on  $\mathfrak{H}$ . First define  $U^w \in \mathfrak{L}(\mathfrak{H})$  for any  $w \in S$  by writing  $w$  in the form (6.1) and setting

$$U^w = U_{j_1} \cdots U_{j_k}.$$

By convention,  $U^e = I$  is the identity operator on  $\mathfrak{H}$ . If  $h \in \mathfrak{G}$  is a hereditary word, set

$$U^h = (U^v)^* U^w$$

for any representation  $h = v^{-1}w$  with  $v, w \in S$ ; this definition does not depend on the choice of representation. Finally, define  $Q(U), P(U) \in \mathfrak{L}(\mathfrak{H} \otimes \mathfrak{G})$  by

$$Q(U) = \sum_{h \in H_m} U^h \otimes Q_h, \quad P(U) = \sum_{w \in S_m} U^w \otimes P_w.$$

The reader is referred to, for example, Murphy [43, §6.3] for the construction of tensor products of Hilbert spaces and algebras, or Palmer, [45, §1.10] for a more detailed account.



**Theorem 6.1 (McCullough [42]).** *Let*

$$Q = \sum_{h \in H_m} h \otimes Q_h$$

*be a hereditary polynomial with coefficients in  $\mathfrak{L}(\mathfrak{G})$  such that  $Q(U) \geq 0$  for every tuple  $U = (U_1, \dots, U_d)$  of unitary operators on  $\mathfrak{H}$ . Then for some  $\ell \leq \ell_m$ , there exist analytic polynomials*

$$P_j = \sum_{w \in S_m} w \otimes P_{j,w}, \quad j = 1, \dots, \ell,$$

*with coefficients in  $\mathfrak{L}(\mathfrak{G})$  such that*

$$Q = P_1^* P_1 + \dots + P_\ell^* P_\ell.$$

*Moreover, for any tuple  $U = (U_1, \dots, U_d)$  of unitary operators on  $\mathfrak{H}$ ,*

$$Q(U) = P_1(U)^* P_1(U) + \dots + P_\ell(U)^* P_\ell(U).$$

*In these statements, when  $\mathfrak{G}$  is infinite dimensional, we can choose  $\ell = 1$ .*

As noted by McCullough, when  $d = 1$ , Theorem 6.1 gives a weaker version of Theorem 2.1. However, Theorem 2.1 can be deduced from this by a judicious use of Beurling's theorem and an inner-outer factorization.

McCullough's theorem uses one of many possible choices of noncommutative spaces on which some form of trigonometric polynomials can be defined. We place this, along with the commutative versions, within a general framework, which we now explain.

The complex scalar-valued trigonometric polynomials in  $d$  variables form a unital  $*$ -algebra  $\mathfrak{P}$ , the involution taking  $z^n$  to  $z^{-n}$ , where for  $n = (n_1, \dots, n_d)$ ,  $-n = (-n_1, \dots, -n_d)$ . If instead the coefficients are in the algebra  $\mathfrak{C} = \mathfrak{L}(\mathfrak{G})$  for some Hilbert space  $\mathfrak{G}$ , then the unital involutive algebra of trigonometric polynomials with coefficients in  $\mathfrak{C}$  is  $\mathfrak{P} \otimes \mathfrak{C}$ . The unit is  $1 \otimes 1$ . A representation of  $\mathfrak{P} \otimes \mathfrak{C}$  is a unital algebra  $*$ -homomorphism from  $\mathfrak{P} \otimes \mathfrak{C}$  into  $\mathfrak{L}(\mathfrak{H})$  for a Hilbert space  $\mathfrak{H}$ . The key thing here is that  $z_1, \dots, z_d$  generate  $\mathfrak{C}$ , and so assuming we do not mess with the coefficient space, a representation  $\pi$  is determined by specifying  $\pi(z_k)$ ,  $k = 1, \dots, d$ .

First note that since  $z_k^* z_k = 1$ ,  $\pi(z_k)$  is isometric, and since  $z_k^* = z_k^{-1}$ , we then have that  $\pi(z_k)$  is unitary. Assuming the variables commute, the  $z_k$ s generate a commutative group  $G$  which we can identify with  $\mathbb{Z}^d$  under addition, and the irreducible representations of commutative groups are one dimensional. This essentially follows from the spectral theory for normal operators (see, for example, Edwards [21, p. 718]). However, the one-dimensional representations are point evaluations on  $\mathbb{T}^d$ . Discrete groups with the discrete topology are examples of locally compact groups. Group representations of locally compact groups extend naturally to the algebraic group algebra, which in this case is  $\mathfrak{P}$ , and then on to the algebra  $\mathfrak{P} \otimes \mathfrak{C}$  by tensoring with the identity representation of  $\mathfrak{C}$ . So a seemingly more complex way of stating that a commutative trigonometric polynomial

$P$  in several variables is positive / strictly positive is to say that for each (topologically) irreducible unitary representation  $\pi$  of  $G$ , the extension of  $\pi$  to a unital  $*$ -representation of the algebra  $\mathfrak{P} \otimes \mathfrak{C}$ , also called  $\pi$ , has the property that  $\pi(P) \geq 0$  /  $\pi(P) > 0$ . By the way, since  $\mathbb{T}^d$  is compact,  $\pi(P) > 0$  implies the existence of some  $\epsilon > 0$  such that  $\pi(P - \epsilon 1 \otimes 1) = \pi(P) - \epsilon 1 \geq 0$ .

What is gained through this perspective is that we may now define noncommutative trigonometric polynomials over a finitely generated discrete (so locally compact) group  $G$  in precisely the same manner. These are the elements of the algebraic group algebra  $\mathfrak{P}$  generated by  $G$ ; that is, formal complex linear combinations of elements of  $G$  endowed with pointwise addition and a convolution product (see Palmer [45, section 1.9]). Then a trigonometric polynomial in  $\mathfrak{P} \otimes \mathfrak{C}$  is formally a finite sum over  $G$  of the form  $P = \sum_g g \otimes P_g$  where  $P_g \in \mathfrak{C}$  for all  $g$ .

We also introduce an involution by setting  $g^* = g^{-1}$  for  $g \in G$ . A trigonometric polynomial  $P$  is **self adjoint** if for all  $g$ ,  $P_{g^*} = P_g^*$ . There is an order structure on selfadjoint elements defined by saying that a selfadjoint polynomial  $P$  is **positive** / **strictly positive** if for every irreducible unital  $*$ -representation  $\pi$  of  $G$ , the extension as above of  $\pi$  to the algebra  $\mathfrak{P} \otimes \mathfrak{C}$  (again called  $\pi$ ), satisfies  $\pi(P) \geq 0$  /  $\pi(P) > 0$ ; where by  $\pi(P) > 0$  we mean that there exists some  $\epsilon > 0$  independent of  $\pi$  such that  $\pi(P - \epsilon(1 \otimes 1)) \geq 0$ . Letting  $\Omega$  represent the set of such irreducible representations, we can in a manner suggestive of the Gel'fand transform define  $\hat{P}(\pi) = \pi(P)$ , and in this way think of  $\Omega$  as a sort of noncommutative space on which our polynomial is defined. The Gel'fand-Raïkov theorem (see, for example, Palmer [46, Theorem 12.4.6]) ensures the existence of sufficiently many irreducible representations to separate  $G$ , so in particular,  $\Omega \neq \emptyset$ .

For a finitely generated discrete group  $G$  with generators  $\{a_1, \dots, a_d\}$ , let  $S$  be a fixed unital subsemigroup of  $G$  containing the generators. The most interesting case is when  $S$  is the subsemigroup generated by  $e$  (the group identity) and  $\{a_1, \dots, a_d\}$ . As an example of this, if  $G$  is the noncommutative free group in  $d$  generators, then the unital subsemigroup generated by  $\{a_1, \dots, a_d\}$  consists of group elements  $w$  of the form  $e$  (for the empty word) and those which are an arbitrary finite product of positive powers of the generators, as in (6.1).

We also need to address the issue of what should play the role of Laurent and analytic trigonometric polynomials in the noncommutative setting. The **hereditary** trigonometric polynomials are defined as those polynomials of the form  $P = \sum_j w_{j1}^* w_{j2} \otimes P_j$ , where  $w_{j1}, w_{j2} \in S$ . We think of these as the Laurent polynomials. Trigonometric polynomials over  $S$  are referred to as **analytic** polynomials. The **square** of an analytic polynomial  $Q$  is the hereditary trigonometric polynomial  $Q^*Q$ . Squares are easily seen to be positive. As a weak analogue of the Fejér-Riesz theorem, we prove a partial converse below.

We refer to those hereditary polynomials which are selfadjoint as real hereditary polynomials, and denote the set of such polynomials by  $H$ . While these polynomials do not form an algebra, they are clearly a vector space. Those which

are finite sums of squares form a cone  $C$  in  $H$  (that is,  $C$  is closed under sums and positive scalar multiplication). Any real polynomial is the sum of terms of the form  $1 \otimes A$  or  $w_2^* w_1 \otimes B + w_1^* w_2 \otimes B^*$ , where  $w_1, w_2 \in S$  and  $A$  is selfadjoint. The first of these is obviously the difference of squares. Using  $w^* w = 1$  for any  $w \in G$ , we also have

$$w_2^* w_1 \otimes B + w_1^* w_2 \otimes B^* = (w_1 \otimes B + w_2 \otimes 1)^* (w_1 \otimes B + w_2 \otimes 1) - 1 \otimes (1 + B^* B).$$

Hence  $H = C - C$ .

For  $A, B \in \mathfrak{L}(\mathfrak{H})$  and  $w_1, w_2 \in S$ ,

$$\begin{aligned} 0 &\leq (w_1 \otimes A + w_2 \otimes B)^* (w_1 \otimes A + w_2 \otimes B) \\ &\leq (w_1 \otimes A + w_2 \otimes B)^* (w_1 \otimes A + w_2 \otimes B) \\ &\quad + (w_1 \otimes A - w_2 \otimes B)^* (w_1 \otimes A - w_2 \otimes B) \\ &= 2(1 \otimes A^* A + 1 \otimes B^* B) \\ &\leq (\|A\|^2 + \|B\|^2)(1 \otimes 1). \end{aligned}$$

Applying this iteratively, we see that for any  $P \in H$ , there is some constant  $0 \leq \alpha < \infty$  such that  $\alpha 1 \pm P \in C$ . In other words, the cone  $C$  is **archimedean**. In particular,  $1 \otimes 1$  is in the algebraic interior of  $C$ , meaning that if  $P \in H$ , then there is some  $0 < t_0 \leq 1$  such that for all  $0 < t < t_0$ ,  $t(1 \otimes 1) + (1 - t)P \in C$ .

**Theorem 6.2.** *Let  $G$  be a finitely generated discrete group,  $P$  a strictly positive trigonometric polynomial over  $G$  with coefficients in  $\mathfrak{L}(\mathfrak{G})$ . Then  $P$  is a sum of squares of analytic polynomials.*

*Proof.* The proof uses a standard GNS construction and separation argument. Suppose that for some  $\epsilon > 0$ ,  $P - \epsilon(1 \otimes 1) \geq 0$  but that  $P \notin C$ . Since  $C$  has nonempty algebraic interior, it follows from the Edelmet-Kakutani theorem<sup>1</sup> that there is a nonconstant linear functional  $\lambda : H \rightarrow \mathbb{R}$  such that  $\lambda(C) \geq 0$  and  $\lambda(P) \leq 0$ . Since  $\lambda$  is nonzero, there is some  $R \in H$  with  $\lambda(R) > 0$ , and so since the cone  $C$  is archimedean, there exists  $\alpha > 0$  such that  $\alpha(1 \otimes 1) - R \in C$ . From this we see that  $\lambda(1 \otimes 1) > 0$ , and so by scaling, we may assume  $\lambda(1 \otimes 1) = 1$ .

We next define a nontrivial scalar product on  $H$  by setting

$$\langle w_1 \otimes A, w_2 \otimes B \rangle = \lambda(w_2^* w_1 \otimes B^* A)$$

and extending linearly to all of  $H$ . It is easily checked that this satisfies all of the properties of an inner product, except that  $\langle w \otimes A, w \otimes A \rangle = 0$  may not necessarily imply that  $w \otimes A = 0$ . Even so, such scalar products satisfy the Cauchy-Schwarz inequality, and so  $N = \{w \otimes A : \langle w \otimes A, w \otimes A \rangle = 0\}$  is a vector subspace of  $H$ . Therefore this scalar product induces an inner product on  $H/N$ , and the completion  $\mathfrak{H}$  of  $H/N$  with respect to the associated norm makes  $H/N$  into a Hilbert space.

<sup>1</sup>[36] Holmes, Corollary, §4B. *Let  $A$  and  $B$  be nonempty convex subsets of  $X$ , and assume the algebraic interior of  $A$ ,  $\text{cor}(A)$  is nonempty. Then  $A$  and  $B$  can be separated if and only if  $\text{cor}(A) \cap B = \emptyset$ .*

We next define a representation  $\pi : H \rightarrow \mathfrak{L}(\mathfrak{H})$  by the left regular representation; that is,  $\pi(P)[w \otimes A] = [P(w \otimes A)]$ , where  $[\cdot]$  indicates an equivalence class in  $H/N$ . Since  $P \geq \epsilon(1 \otimes 1) \geq 0$  for some  $\epsilon > 0$ ,  $P - \epsilon/2(1 \otimes 1) > 0$ . Suppose that  $P \notin C$ . Then

$$\lambda((P - \epsilon/2(1 \otimes 1)) + \epsilon/2(1 \otimes 1)) = \lambda(P - \epsilon/2(1 \otimes 1)) + \epsilon/2 \leq 0.$$

Hence

$$\langle \pi(P - \epsilon/2(1 \otimes 1))[1 \otimes 1], [1 \otimes 1] \rangle \leq -\epsilon/2,$$

and so  $\pi(P - \epsilon/2(1 \otimes 1)) \not\geq 0$ . The representation  $\pi$  obviously induces a unitary representation of  $G$  via  $\pi(a_i) = \pi(a_i \otimes 1)$ , where  $a_i$  is a generator of  $G$ . The (irreducible) representations of  $G$  are in bijective correspondence with the essential unital  $*$ -representations of the group  $C^*$ -algebra  $C^*(G)$  (Palmer, [46, Theorem 12.4.1]), which then restrict back to representations of  $H$ . Since unitary representations of  $G$  are direct integrals of irreducible unitary representations (see, for example, Palmer, [46, p. 1386]), there is an irreducible unitary representation  $\pi'$  of  $G$  such that the corresponding representation of  $H$  has the property that  $\pi'(P - \epsilon/2(1 \otimes 1)) \not\geq 0$ , giving a contradiction.  $\square$

The above could equally well have been derived using any  $C^*$ -algebra in the place of  $\mathfrak{L}(\mathfrak{H})$ . One could also further generalize to non-discrete locally compact groups, replacing the trigonometric polynomials by functions of compact support.

We obtain Theorem 5.1 as a corollary if we take  $G$  to be the free group in  $d$  commuting letters. On the other hand, if  $G$  is the noncommutative free group on  $d$  letters, it is again straightforward to specify the irreducible representations of  $G$ . These take the generators  $(a_1, \dots, a_d)$  to irreducible  $d$ -tuples  $(U_1, \dots, U_d)$  of (noncommuting) unitary operators, yielding a weak form of McCullough's theorem.

As mentioned earlier, it is known by results of Scheiderer [57] that when  $G$  is the free group in  $d$  commuting letters,  $d \geq 3$ , there are positive polynomials which cannot be expressed as sums of squares of analytic polynomials, so no statement along the lines of Theorem 6.1 can be true for trigonometric polynomials if it is to hold for all finitely generated discrete groups. Just what can be said in various special cases is still largely unexplored.

## Appendix A. Schur complements

We prove the existence and uniqueness of Schur complements for Hilbert space operators as required in Definition 3.1.

**Lemma A.1.** *Let  $T \in \mathfrak{L}(\mathfrak{H})$ , where  $\mathfrak{H}$  is a Hilbert space. Let  $\mathfrak{K}$  be a closed subspace of  $\mathfrak{H}$ , and write*

$$T = \begin{pmatrix} A & B^* \\ B & C \end{pmatrix} : \mathfrak{K} \oplus \mathfrak{K}^\perp \rightarrow \mathfrak{K} \oplus \mathfrak{K}^\perp.$$

Then  $T \geq 0$  if and only if  $A \geq 0$ ,  $C \geq 0$ , and  $B = C^{\frac{1}{2}}GA^{\frac{1}{2}}$  for some contraction  $G \in \mathfrak{L}(\mathfrak{K}, \mathfrak{K}^\perp)$ . The operator  $G$  can be chosen so that it maps  $\overline{\text{ran}} A$  into  $\overline{\text{ran}} C$  and is zero on the orthogonal complement of  $\overline{\text{ran}} A$ , and then it is unique.

*Proof.* If  $B = C^{\frac{1}{2}}GA^{\frac{1}{2}}$  where  $G \in \mathfrak{L}(\mathfrak{K}, \mathfrak{K}^\perp)$  is a contraction, then

$$T = \begin{pmatrix} A^{\frac{1}{2}} & 0 \\ C^{\frac{1}{2}}G & C^{\frac{1}{2}}(I - GG^*)^{\frac{1}{2}} \end{pmatrix} \begin{pmatrix} A^{\frac{1}{2}} & G^*C^{\frac{1}{2}} \\ 0 & (I - GG^*)^{\frac{1}{2}}C^{\frac{1}{2}} \end{pmatrix} \geq 0.$$

Conversely, if  $T \geq 0$ , it is trivial that  $A \geq 0$  and  $C \geq 0$ . Set

$$N = T^{\frac{1}{2}} = \begin{pmatrix} N_1 \\ N_2 \end{pmatrix} : \mathfrak{H} \rightarrow \mathfrak{H} \oplus \mathfrak{K}.$$

Then  $A = N_1N_1^*$  and  $C = N_2N_2^*$ , and so there exist partial isometries  $V_1 \in \mathfrak{L}(\mathfrak{K}, \mathfrak{H})$  and  $V_2 \in \mathfrak{L}(\mathfrak{K}^\perp, \mathfrak{H})$  with initial spaces  $\overline{\text{ran}} A$  and  $\overline{\text{ran}} C$  such that  $N_1^* = V_1A^{\frac{1}{2}}$  and  $N_2^* = V_2C^{\frac{1}{2}}$ . Thus  $B = N_2N_1^* = C^{\frac{1}{2}}GA^{\frac{1}{2}}$ , where  $G = V_2^*V_1$  is a contraction. By construction  $G$  has the properties in the last statement, and clearly such an operator is unique.  $\square$

**Lemma A.2.** Let  $\mathfrak{H}$  be a Hilbert space, and suppose  $T \in \mathfrak{L}(\mathfrak{H})$ ,  $T \geq 0$ . Let  $\mathfrak{K}$  be a closed subspace of  $\mathfrak{H}$ , and write

$$T = \begin{pmatrix} A & B^* \\ B & C \end{pmatrix} : \mathfrak{K} \oplus \mathfrak{K}^\perp \rightarrow \mathfrak{K} \oplus \mathfrak{K}^\perp.$$

Then there is a largest operator  $S \geq 0$  in  $\mathfrak{L}(\mathfrak{K})$  such that

$$\begin{pmatrix} A - S & B^* \\ B & C \end{pmatrix} \geq 0. \quad (\text{A.1})$$

It is given by  $S = A^{\frac{1}{2}}(I - G^*G)A^{\frac{1}{2}}$ , where  $G \in \mathfrak{L}(\mathfrak{K}, \mathfrak{K}^\perp)$  is a contraction which maps  $\overline{\text{ran}} A$  into  $\overline{\text{ran}} C$  and is zero on the orthogonal complement of  $\overline{\text{ran}} A$ .

*Proof.* By Lemma A.1, we may define  $S = A^{\frac{1}{2}}(I - G^*G)A^{\frac{1}{2}}$  with  $G$  as in the last statement of the lemma. Then

$$\begin{pmatrix} A - S & B^* \\ B & C \end{pmatrix} = \begin{pmatrix} A^{\frac{1}{2}}G^*GA^{\frac{1}{2}} & A^{\frac{1}{2}}G^*C^{\frac{1}{2}} \\ C^{\frac{1}{2}}GA^{\frac{1}{2}} & C \end{pmatrix} = \begin{pmatrix} A^{\frac{1}{2}}G^* \\ C^{\frac{1}{2}} \end{pmatrix} (GA^{\frac{1}{2}} \quad C^{\frac{1}{2}}) \geq 0.$$

Consider any  $X \geq 0$  in  $\mathfrak{L}(\mathfrak{K})$  such that

$$\begin{pmatrix} A - X & B^* \\ B & C \end{pmatrix} \geq 0.$$

Since  $A \geq X \geq 0$ , we can write  $X = A^{\frac{1}{2}}KA^{\frac{1}{2}}$  where  $K \in \mathfrak{L}(\mathfrak{K})$  and  $0 \leq K \leq I$ . We can choose  $K$  so that it maps  $\overline{\text{ran}} A$  into itself and is zero on  $(\overline{\text{ran}} A)^\perp$ . Then

$$\begin{aligned} \begin{pmatrix} A - X & B^* \\ B & C \end{pmatrix} &= \begin{pmatrix} A - A^{\frac{1}{2}}KA^{\frac{1}{2}} & A^{\frac{1}{2}}G^*C^{\frac{1}{2}} \\ C^{\frac{1}{2}}GA^{\frac{1}{2}} & C \end{pmatrix} \\ &= \begin{pmatrix} A^{\frac{1}{2}} & 0 \\ 0 & C^{\frac{1}{2}} \end{pmatrix} \begin{pmatrix} I - K & G^* \\ G & I \end{pmatrix} \begin{pmatrix} A^{\frac{1}{2}} & 0 \\ 0 & C^{\frac{1}{2}} \end{pmatrix}. \end{aligned}$$

By our choices  $G$  and  $K$ , we deduce that

$$\begin{pmatrix} I - K & G^* \\ G & I \end{pmatrix} \geq 0.$$

By Lemma A.1,  $G = G_1(I - K)^{\frac{1}{2}}$  where  $G_1 \in \mathfrak{L}(\mathfrak{K}, \mathfrak{K}^\perp)$  is a contraction. Therefore  $G^*G \leq I - K$ , and so

$$X = A^{\frac{1}{2}}KA^{\frac{1}{2}} \leq A^{\frac{1}{2}}(I - G^*G)A^{\frac{1}{2}} = S.$$

This shows  $S$  is maximal with respect to the property (A.1).  $\square$

## References

1. D. Z. Arov, *Stable dissipative linear stationary dynamical scattering systems*, J. Operator Theory **2** (1979), no. 1, 95–126, English Transl. with appendices by the author and J. Rovnyak, Oper. Theory Adv. Appl., vol. 134, Birkhäuser Verlag, Basel, 2002, 99–136.
2. M. Bakonyi and T. Constantinescu, *Schur's algorithm and several applications*, Pitman Research Notes in Mathematics Series, vol. 261, Longman Scientific & Technical, Harlow, 1992.
3. M. Bakonyi and H. J. Woerdeman, *Matrix completions, moment problems, and factorizations*, book manuscript, in preparation, 2008.
4. S. Barclay, *A solution to the Douglas-Rudin problem for matrix-valued functions*, preprint, 2007.
5. ———, *Continuity of the spectral factorization mapping*, J. London Math. Soc. (2) **70** (2004), no. 3, 763–779.
6. ———, *Banach spaces of analytic vector-valued functions*, Ph.D. thesis, University of Leeds, 2007.
7. H. Bart, I. Gohberg, M. A. Kaashoek, and A. C. M. Ran, *Factorization of matrix and operator functions: the state space method*, Oper. Theory Adv. Appl., vol. 178, Birkhäuser Verlag, Basel, 2008.
8. G. Blower, *On analytic factorization of positive Hermitian matrix functions over the disc*, Linear Algebra Appl. **295** (1999), no. 1-3, 149–158.
9. J. Bourgain, *A problem of Douglas and Rudin on factorization*, Pacific J. Math. **121** (1986), no. 1, 47–50.
10. L. de Branges, *The expansion theorem for Hilbert spaces of entire functions*, Entire Functions and Related Parts of Analysis (Proc. Sympos. Pure Math., La Jolla, Calif., 1966), Amer. Math. Soc., Providence, RI, 1968, pp. 79–148.
11. A. Brown and P. R. Halmos, *Algebraic properties of Toeplitz operators*, J. Reine Angew. Math. **213** (1963/1964), 89–102.
12. G. Cassier, *Problème des moments sur un compact de  $\mathbf{R}^n$  et décomposition de polynômes à plusieurs variables*, J. Funct. Anal. **58** (1984), no. 3, 254–266.
13. K. F. Clancey and I. Gohberg, *Factorization of matrix functions and singular integral operators*, Oper. Theory Adv. Appl., vol. 3, Birkhäuser Verlag, Basel, 1981.

14. T. Constantinescu, *Factorization of positive-definite kernels*, Topics in operator theory: Ernst D. Hellinger memorial volume, Oper. Theory Adv. Appl., vol. 48, Birkhäuser Verlag, Basel, 1990, pp. 245–260.
15. A. Devinatz, *The factorization of operator valued functions*, Ann. of Math. (2) **73** (1961), 458–495.
16. R. G. Douglas, *On factoring positive operator functions*, J. Math. Mech. **16** (1966), 119–126.
17. R. G. Douglas and W. Rudin, *Approximation by inner functions*, Pacific J. Math. **31** (1969), 313–320.
18. M. A. Dritschel, *On factorization of trigonometric polynomials*, Integral Equations Operator Theory **49** (2004), no. 1, 11–42.
19. M. A. Dritschel and H. J. Woerdeman, *Outer factorizations in one and several variables*, Trans. Amer. Math. Soc. **357** (2005), no. 11, 4661–4679.
20. P. L. Duren, *Theory of  $H^p$  spaces*, Academic Press, New York, 1970; Dover reprint, Mineola, New York, 2000.
21. R. E. Edwards, *Functional analysis. Theory and applications*, Holt, Rinehart and Winston, New York, 1965; Dover reprint, Mineola, New York, 1995.
22. L. Fejér, *Über trigonometrische polynome*, J. Reine Angew. Math. **146** (1916), 53–82.
23. J. S. Geronimo and Ming-Jun Lai, *Factorization of multivariate positive Laurent polynomials*, J. Approx. Theory **139** (2006), no. 1-2, 327–345.
24. J. S. Geronimo and H. J. Woerdeman, *Positive extensions, Fejér-Riesz factorization and autoregressive filters in two variables*, Ann. of Math. (2) **160** (2004), no. 3, 839–906.
25. ———, *The operator valued autoregressive filter problem and the suboptimal Nehari problem in two variables*, Integral Equations Operator Theory **53** (2005), no. 3, 343–361.
26. I. Gohberg, *The factorization problem for operator functions*, Izv. Akad. Nauk SSSR Ser. Mat. **28** (1964), 1055–1082, Amer. Math. Soc. Transl. (2) **49** 130–161.
27. I. Gohberg, S. Goldberg, and M. A. Kaashoek, *Classes of linear operators. Vol. I*, Oper. Theory Adv. Appl., vol. 49, Birkhäuser Verlag, Basel, 1990.
28. U. Grenander and G. Szegő, *Toeplitz forms and their applications*, California Monographs in Mathematical Sciences, University of California Press, Berkeley, 1958.
29. P. R. Halmos, *Shifts on Hilbert spaces*, J. Reine Angew. Math. **208** (1961), 102–112.
30. H. Helson, *Lectures on invariant subspaces*, Academic Press, New York, 1964.
31. H. Helson and D. Lowdenslager, *Prediction theory and Fourier series in several variables*, Acta Math. **99** (1958), 165–202.
32. ———, *Prediction theory and Fourier series in several variables. II*, Acta Math. **106** (1961), 175–213.
33. J. W. Helton, S. A. McCullough, and M. Putinar, *Matrix representations for positive noncommutative polynomials*, Positivity **10** (2006), no. 1, 145–163.
34. J. W. Helton and M. Putinar, *Positive polynomials in scalar and matrix variables, the spectral theorem, and optimization*, Operator theory, structured matrices, and dilations, Theta Ser. Adv. Math., vol. 7, Theta, Bucharest, 2007, pp. 229–306.

35. K. Hoffman, *Banach spaces of analytic functions*, Prentice-Hall Inc., Englewood Cliffs, N. J., 1962; Dover reprint, Mineola, New York, 1988.
36. R. B. Holmes, *Geometric functional analysis and its applications*, Graduate Texts in Mathematics, No. 24, Springer-Verlag, New York, 1975.
37. B. Jacob and J. R. Partington, *On the boundedness and continuity of the spectral factorization mapping*, SIAM J. Control Optim. **40** (2001), no. 1, 88–106.
38. T. Kailath, A. H. Sayed, and B. Hassibi, *Linear estimation*, Prentice Hall, Englewood Cliffs, NJ, 1980.
39. A. Lebow and M. Schreiber, *Polynomials over groups and a theorem of Fejér and Riesz*, Acta Sci. Math. (Szeged) **44** (1982), no. 3-4, 335–344 (1983).
40. D. Lowdenslager, *On factoring matrix valued functions*, Ann. of Math. (2) **78** (1963), 450–454.
41. A. S. Markus, *Introduction to the spectral theory of polynomial operator pencils*, Translations of Mathematical Monographs, vol. 71, Amer. Math. Soc., Providence, RI, 1988.
42. S. McCullough, *Factorization of operator-valued polynomials in several non-commuting variables*, Linear Algebra Appl. **326** (2001), no. 1-3, 193–203.
43. G. J. Murphy,  *$C^*$ -algebras and operator theory*, Academic Press Inc., Boston, MA, 1990.
44. A. Naftalevich and M. Schreiber, *Trigonometric polynomials and sums of squares*, Number theory (New York, 1983–84), Lecture Notes in Math., vol. 1135, Springer-Verlag, Berlin, 1985, pp. 225–238.
45. T. W. Palmer, *Banach algebras and the general theory of  $*$ -algebras. Vol. I*, Encyclopedia of Mathematics and its Applications, vol. 49, Cambridge University Press, Cambridge, 1994.
46. ———, *Banach algebras and the general theory of  $*$ -algebras. Vol. 2*, Encyclopedia of Mathematics and its Applications, vol. 79, Cambridge University Press, Cambridge, 2001.
47. F. Riesz, *Über ein Problem des Herrn Carathéodory*, J. Reine Angew. Math. **146** (1916), 83–87.
48. M. Rosenblatt, *A multi-dimensional prediction problem*, Ark. Mat. **3** (1958), 407–424.
49. M. Rosenblum, *Vectorial Toeplitz operators and the Fejér-Riesz theorem*, J. Math. Anal. Appl. **23** (1968), 139–147.
50. M. Rosenblum and J. Rovnyak, *Hardy classes and operator theory*, Oxford University Press, New York, 1985; Dover reprint, Mineola, New York, 1997.
51. ———, *The factorization problem for nonnegative operator valued functions*, Bull. Amer. Math. Soc. **77** (1971), 287–318.
52. Yu. A. Rozanov, *Stationary random processes*, Holden-Day Inc., San Francisco, Calif., 1967.
53. W. Rudin, *The extension problem for positive-definite functions*, Illinois J. Math. **7** (1963), 532–539.
54. L. A. Sakhnovich, *Interpolation theory and its applications*, Kluwer, Dordrecht, 1997.
55. D. Sarason, *Generalized interpolation in  $H^\infty$* , Trans. Amer. Math. Soc. **127** (1967), 179–203.



56. A. H. Sayed and T. Kailath, *A survey of spectral factorization methods*, Numer. Linear Algebra Appl. **8** (2001), no. 6-7, 467–496, Numerical linear algebra techniques for control and signal processing.
57. C. Scheiderer, *Sums of squares of regular functions on real algebraic varieties*, Trans. Amer. Math. Soc. **352** (2000), no. 3, 1039–1069.
58. K. Schmüdgen, *The  $K$ -moment problem for compact semi-algebraic sets*, Math. Ann. **289** (1991), no. 2, 203–206.
59. ———, *Noncommutative real algebraic geometry - some basic concepts and first ideas*, Emerging Applications of Algebraic Geometry, The IMA Volumes in Mathematics and its Applications, vol. 149, Springer-Verlag, Berlin, 2009, pp. 325–350.
60. B. Simon, *Orthogonal polynomials on the unit circle. Part 1*, Amer. Math. Soc. Colloq. Publ., vol. 54, Amer. Math. Soc., Providence, RI, 2005.
61. B. Sz.-Nagy and C. Foias, *Harmonic analysis of operators on Hilbert space*, North-Holland Publishing Co., Amsterdam, 1970.
62. G. Szegő, *Orthogonal polynomials*, fourth ed., Amer. Math. Soc. Colloq. Publ., vol. 23, Amer. Math. Soc., Providence, RI, 1975.
63. N. Wiener and P. Masani, *The prediction theory of multivariate stochastic processes. I. The regularity condition*, Acta Math. **98** (1957), 111–150.
64. ———, *The prediction theory of multivariate stochastic processes. II. The linear predictor*, Acta Math. **99** (1958), 93–137.
65. V. Zasuĥin, *On the theory of multidimensional stationary random processes*, C. R. (Doklady) Acad. Sci. URSS (N.S.) **33** (1941), 435–437.

Michael A. Dritschel

School of Mathematics and Statistics, Herschel Building, University of Newcastle, Newcastle upon Tyne NE1 7RU, UK

e-mail: [m.a.dritschel@ncl.ac.uk](mailto:m.a.dritschel@ncl.ac.uk)

James Rovnyak

University of Virginia, Department of Mathematics, P. O. Box 400137, Charlottesville, VA 22904–4137

e-mail: [rovnyak@virginia.edu](mailto:rovnyak@virginia.edu)