The Analytic Theory of Matrix Orthogonal Polynomials

David Damanik, Alexander Pushnitski, and Barry Simon

January 30, 2008

Abstract

We survey the analytic theory of matrix orthogonal polynomials.

MSC: 42C05, 47B36, 30C10

keywords: orthogonal polynomials, matrix-valued measures, block Jacobi matrices, block CMV matrices

Contents

Surveys in Approximation Theory
Volume 4, 2008. pp. 1–85.

ISSN 1555-578X

1 Introduction

1.1 Introduction and Overview

Orthogonal polynomials on the real line (OPRL) were developed in the nineteenth century and orthogonal polynomials on the unit circle (OPUC) were initially developed around 1920 by Szegő. Their matrix analogues are of much more recent vintage. They were originally developed in the MOPUC case indirectly in the study of prediction theory [116, 117, 129, 131, 132, 138, 196] in the period 1940–1960. The connection to OPUC in the scalar case was discovered by Krein [131].

Much of the theory since is in the electrical engineering literature [36, 37, 38, 39, 40, 41, 120, 121, 122, 123, 203]; see also [84, 86, 87, 88, 142].

The corresponding real line theory (MOPRL) is still more recent: Following early work of Krein [133] and Berezan'ski [9] on block Jacobi matrices, mainly as applied to self-adjoint extensions, there was a seminal paper of Aptekarev–Nikishin [4] and a flurry of papers since the 1990s [10, 11, 12, 14, 16, 17, 19, 20, 21, 22, 29, 35, 43, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 64, 65, 66, 67, 68, 69, 71, 73, 74, 75, 76, 77, 79, 83, 85, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 137, 139, 140, 143, 144, 145, 148, 149, 150, 155, 156, 157, 154, 161, 162, 179, 186, 198, 200, 201, 202, 204]; see also [7].

There is very little on the subject in monographs — the more classical ones (e.g., [23, 82, 93, 184]) predate most of the subject; see, however, Atkinson [5, Section 6.6]. Ismail [118] has no discussion and Simon [167, 168] has a single section! Because of the use of MOPRL in [33], we became interested in the subject and, in particular, we needed some basic results for that paper which we couldn't find in the literature or which, at least, weren't very accessible. Thus, we decided to produce this comprehensive review that we hope others will find useful.

As with the scalar case, the subject breaks into two parts, conveniently called the analytic theory (general structure results) and the algebraic theory (the set of non-trivial examples). This survey deals entirely with the analytic theory. We note, however, that one of the striking developments in recent years has been the discovery that there are rich classes of genuinely new MOPRL, even at the classical level of Bochner's theorem; see [20, 55, 70, 72, 102, 109, 110, 111, 112, 113, 156, 161] and the forthcoming monograph [63] for further discussion of this algebraic side.

In this introduction, we will focus mainly on the MOPRL case. For scalar OPRL, a key issue is the passage from measure to monic OPRL, then to normalized OPRL, and finally to Jacobi parameters. There are no choices in going from measure to monic OP, $P_n(x)$. They are determined by

$$P_n(x) = x^n + \text{lower order}, \qquad \langle x^j, P_n\rangle = 0 \quad j = 1, \dots, n-1. \tag{1.1}$$

However, the basic condition on the orthonormal polynomials, namely,

$$\langle p_n, p_m\rangle = \delta_{nm} \tag{1.2}$$

does not uniquely determine the $p_n(x)$. The standard choice is

$$p_n(x) = \frac{P_n(x)}{\|P_n\|}.$$

However, if $\theta_0, \theta_1, \dots$ are arbitrary real numbers, then

$$\tilde{p}_n(x) = \frac{e^{i\theta_n} P_n(x)}{\|P_n\|} \tag{1.3}$$

also obey (1.2). If the recursion coefficients (aka Jacobi parameters), are defined via

$$xp_n = a_{n+1}p_{n+1} + b_{n+1}p_n + a_n p_{n-1}, \tag{1.4}$$

then the choice (1.3) leads to

$$\tilde{b}_n = b_n, \qquad \tilde{a}_n = e^{i\theta_n} a_n e^{-i\theta_{n-1}}. \tag{1.5}$$

The standard choice is, of course, most natural here; for example, if

$$p_n(x) = \kappa_n x^n + \text{lower order}, \tag{1.6}$$

then $a_n > 0$ implies $\kappa_n > 0$. It would be crazy to make any other choice.

For MOPRL, these choices are less clear. As we will explain in Section 1.2, there are now two matrix-valued "inner products" formally written as

$$\langle\langle f, g\rangle\rangle_R = \int f(x)^\dagger \, d\mu(x) g(x), \tag{1.7}$$

$$\langle\langle f, g\rangle\rangle_L = \int g(x) \, d\mu(x) f(x)^\dagger, \tag{1.8}$$

where now μ is a matrix-valued measure and † denotes the adjoint, and corresponding two sets of monic OPRL: $P_n^R(x)$ and $P_n^L(x)$. The orthonormal polynomials are required to obey

$$\langle\langle p_n^R, p_m^R\rangle\rangle_R = \delta_{nm}\mathbf{1}. \tag{1.9}$$

The analogue of (1.3) is

$$\tilde{p}_n^R(x) = P_n^R(x)\langle\langle P_n^R, P_n^R\rangle\rangle^{-1/2}\sigma_n \tag{1.10}$$

for a unitary σ_n. For the immediately following, use p_n^R to be the choice $\sigma_n \equiv 1$. For any such choice, we have a recursion relation,

$$xp_n^R(x) = p_{n+1}^R(x)A_{n+1}^\dagger + p_n^R(x)B_{n+1} + p_{n-1}^R(x)A_n \tag{1.11}$$

with the analogue of (1.5) (comparing $\sigma_n \equiv \mathbf{1}$ to general σ_n)

$$\tilde{B}_n = \sigma_n^\dagger B_n \sigma_n \qquad \tilde{A}_n = \sigma_{n-1}^\dagger A_n \sigma_n. \tag{1.12}$$

The obvious analogue of the scalar case is to pick $\sigma_n \equiv \mathbf{1}$, which makes κ_n in

$$p_n^R(x) = \kappa_n x^n + \text{lower order} \tag{1.13}$$

obey $\kappa_n > 0$. Note that (1.11) implies

$$\kappa_n = \kappa_{n+1} A_{n+1}^\dagger \tag{1.14}$$

or, inductively,

$$\kappa_n = (A_n^\dagger \dots A_1^\dagger)^{-1}. \tag{1.15}$$

In general, this choice does not lead to A_n positive or even Hermitian. Alternatively, one can pick σ_n so $\tilde{A}_n$ is positive. Besides these two "obvious" choices, $\kappa_n > 0$ or $A_n > 0$, there is a third that A_n be lower triangular that, as we will see in Section 1.4, is natural. Thus, in the study of MOPRL one needs to talk about equivalent sets of p_n^R and of Jacobi parameters, and this is a major theme of Chapter 2. Interestingly enough for MOPUC, the commonly picked choice equivalent to $A_n > 0$ (namely, $\rho_n > 0$) seems to suffice for applications. So we do not discuss equivalence classes for MOPUC.

Associated to a set of matrix Jacobi parameters is a block Jacobi matrix, that is, a matrix which when written in $l \times l$ blocks is tridiagonal; see (2.29) below.

In Chapter 2, we discuss the basics of MOPRL while Chapter 3 discusses MOPUC. Chapter 4 discusses the Szegő mapping connection of MOPUC and MOPRL. Finally, Chapter 5 discusses the extension of the theory of regular OPs [180] to MOPRL.

While this is mainly a survey, it does have numerous new results, of which we mention:

(a) The clarification of equivalent Jacobi parameters and several new theorems (Theorems 2.8 and 2.9).
(b) A new result (Theorem 2.28) on the order of poles or zeros of $m(z)$ in terms of eigenvalues of J and the once stripped $J^{(1)}$.
(c) Formulas for the resolvent in the MOPRL (Theorem 2.29) and MOPUC (Theorem 3.24) cases.
(d) A theorem on zeros of $\det(\Phi_n^R)$ (Theorem 3.7) and eigenvalues of a cutoff CMV matrix (Theorem 3.10).
(e) A new proof of the Geronimus relations (Theorem 4.2).
(f) Discussion of regular MOPRL (Chapter 5).

There are numerous open questions and conjectures in this paper, of which we mention:

(1) We prove that type 1 and type 3 Jacobi parameters in the Nevai class have $A_n \to \mathbf{1}$ but do not know if this is true for type 2 and, if so, how to prove it.
(2) Determine which monic matrix polynomials, Φ, can occur as monic MOPUC. We know $\det(\Phi(z))$ must have all of its zeros in the unit disk in $\mathbb{C}$, but unlike the scalar case where this is sufficient, we do not know necessary and sufficient conditions.
(3) Generalize Khrushchev theory [125, 126, 101] to MOPUC; see Section 3.13.
(4) Provide a proof of Geronimus relations for MOPUC that uses the theory of canonical moments [43]; see the discussion at the start of Chapter 4.
(5) Prove Conjecture 5.9 extending a result of Stahl–Totik [180] from OPRL to MOPRL.

1.2 Matrix-Valued Measures

Let $\mathcal{M}_l$ denote the ring of all $l \times l$ complex-valued matrices; we denote by $\alpha^\dagger$ the Hermitian conjugate of $\alpha \in \mathcal{M}_l$. (Because of the use of * for Szegő dual in the theory of OPUC, we do not use it for adjoint.) For $\alpha \in \mathcal{M}_l$, we denote by $\|\alpha\|$ its Euclidean norm (i.e., the norm of α as a linear operator on $\mathbb{C}^l$ with the usual Euclidean norm). Consider the set $\mathcal{P}$ of all polynomials in $z \in \mathbb{C}$ with coefficients from $\mathcal{M}_l$. The set $\mathcal{P}$ can be considered either as a right or as a left module over $\mathcal{M}_l$; clearly, conjugation makes the left and right structures isomorphic. For $n = 0, 1, \dots,$ $\mathcal{P}_n$ will denote those polynomials in $\mathcal{P}$ of degree at most n. The set $\mathcal{V}$ denotes the set of all polynomials in $z \in \mathbb{C}$ with coefficients from $\mathbb{C}^l$. The standard inner product in $\mathbb{C}^l$ is denoted by $\langle \cdot, \cdot \rangle_{\mathbb{C}^l}$.

A matrix-valued measure, μ, on $\mathbb{R}$ (or $\mathbb{C}$) is the assignment of a positive semi-definite $l \times l$ matrix $\mu(X)$ to every Borel set X which is countably additive. We will usually normalize it by requiring

$$\mu(\mathbb{R}) = \mathbf{1} \tag{1.16}$$

(or $\mu(\mathbb{C}) = \mathbf{1}$) where $\mathbf{1}$ is the $l \times l$ identity matrix. (We use $\mathbf{1}$ in general for an identity operator, whether in $\mathcal{M}_l$ or in the operators on some other Hilbert space, and $\mathbf{0}$ for the zero operator or matrix.) Normally, our measures for MOPRL will have compact support and, of course, our measures for MOPUC will be supported on all or part of $\partial\mathbb{D}$ ($\mathbb{D}$ is the unit disk in $\mathbb{C}$).

Associated to any such measures is a scalar measure

$$\mu_{\mathrm{tr}}(X) = \mathrm{Tr}(\mu(X)) \tag{1.17}$$

(the trace normalized by $\mathrm{Tr}(\mathbf{1}) = l$). μ_{tr} is normalized by $\mu_{\mathrm{tr}}(\mathbb{R}) = l$.

Applying the Radon–Nikodym theorem to the matrix elements of μ, we see there is a positive semi-definite matrix function $M_{ij}(x)$ so

$$d\mu_{ij}(x) = M_{ij}(x)\, d\mu_{\mathrm{tr}}(x). \tag{1.18}$$

Clearly, by (1.17),

$$\operatorname{Tr}(M(x)) = 1 \tag{1.19}$$

for $d\mu_{\mathrm{tr}}$-a.e. x. Conversely, any scalar measure with $\mu_{\mathrm{tr}}(\mathbb{R}) = l$ and positive semi-definite matrix-valued function M obeying (1.19) define a matrix-valued measure normalized by (1.16).

Given $l \times l$ matrix-valued functions f, g, we define the $l \times l$ matrix $\langle\langle f, g\rangle\rangle_R$ by

$$\langle\langle f, g\rangle\rangle_R = \int f(x)^\dagger M(x) g(x)\, d\mu_{\mathrm{tr}}(x), \tag{1.20}$$

that is, its (j,k) entry is

$$\sum_{nm} \int \overline{f_{nj}(x)}\, M_{nm}(x) g_{mk}(x)\, d\mu_{\mathrm{tr}}(x). \tag{1.21}$$

Since $f^\dagger M f \geq 0$, we see that

$$\langle\langle f, f\rangle\rangle_R \geq 0. \tag{1.22}$$

One might be tempted to think of $\langle\langle f, f\rangle\rangle_R^{1/2}$ as some kind of norm, but that is doubtful. Even if μ is supported at a single point, x_0, with $M = l^{-1}\mathbf{1}$, this "norm" is essentially the absolute value of $A = f(x_0)$, which is known not to obey the triangle inequality! (See [169, Sect. I.1] for an example.)

However, if one looks at

$$\|f\|_R = (\operatorname{Tr}\langle\langle f, f\rangle\rangle_R)^{1/2}, \tag{1.23}$$

one does have a norm (or, at least, a semi-norm). Indeed,

$$\langle f, g\rangle_R = \operatorname{Tr}\langle\langle f, g\rangle\rangle_R \tag{1.24}$$

is a sesquilinear form which is positive semi-definite, so (1.23) is the semi-norm corresponding to an inner product and, of course, one has a Cauchy–Schwarz inequality

$$|\operatorname{Tr}\langle\langle f, g\rangle\rangle_R| \leq \|f\|_R \|g\|_R. \tag{1.25}$$

We have not specified which f's and g's can be used in (1.20). We have in mind mainly polynomials in x in the real case and Laurent polynomials in z in the $\partial\mathbb{D}$ case although, obviously, continuous functions are okay. Indeed, it suffices that f (and g) be measurable and obey

$$\int \operatorname{Tr}(f^\dagger(x) f(x))\, d\mu_{\mathrm{tr}}(x) < \infty \tag{1.26}$$

for the integrals in (1.21) to converge. The set of equivalence classes under $f \sim g$ if $\|f - g\|_R = 0$ defines a Hilbert space, $\mathcal{H}$, and $\langle f, g\rangle_R$ is the inner product on this space.

Instead of (1.20), we use the suggestive shorthand

$$\langle\langle f, g\rangle\rangle_R = \int f(x)^\dagger\, d\mu(x) g(x). \tag{1.27}$$

The use of R here comes from "right" for if $\alpha \in \mathcal{M}_l$,

$$\langle\langle f, g\alpha\rangle\rangle_R = \langle\langle f, g\rangle\rangle_R \alpha, \tag{1.28}$$

$$\langle\langle f\alpha, g\rangle\rangle_R = \alpha^\dagger \langle\langle f, g\rangle\rangle_R, \tag{1.29}$$

but, in general, $\langle\langle f, \alpha g\rangle\rangle_R$ is not related to $\langle\langle f, g\rangle\rangle_R$.

While $(\mathrm{Tr}\langle\langle f, f\rangle\rangle_R)^{1/2}$ is a natural analogue of the norm in the scalar case, it will sometimes be useful to instead consider

$$[\det\langle\langle f, f\rangle\rangle_R]^{1/2}. \tag{1.30}$$

Indeed, this is a stronger "norm" in that $\det > 0 \Rightarrow \mathrm{Tr} > 0$ but not vice-versa.

When $d\mu$ is a "direct sum," that is, each $M(x)$ is diagonal, one can appreciate the difference. In that case, $d\mu = d\mu_1 \oplus \cdots \oplus d\mu_l$ and the MOPRL are direct sums (i.e., diagonal matrices) of scalar OPRL

$$P_n^R(x, d\mu) = P_n(x, d\mu_1) \oplus \cdots \oplus P_n(x, d\mu_l). \tag{1.31}$$

Then

$$\|P_n^R\|_R = \left(\sum_{j=1}^{l} \|P_n(\cdot, d\mu_j)\|^2_{L^2(d\mu_j)}\right)^{1/2}, \tag{1.32}$$

while

$$(\det\langle\langle P_n^R, P_n^R\rangle\rangle_R)^{1/2} = \prod_{j=1}^{l} \|P_n(\cdot, d\mu_j)\|_{L^2(d\mu_j)}. \tag{1.33}$$

In particular, in terms of extending the theory of regular measures [180], $\|P_n^R\|_R^{1/n}$ is only sensitive to $\max\|P_n(\cdot, d\mu_j)\|^{1/2}_{L^2(d\mu_j)}$ while $(\det\langle\langle P_n^R, P_n^R\rangle\rangle_R)^{1/2}$ is sensitive to them all. Thus, det will be needed for that theory (see Chapter 5).

There will also be a left inner product and, correspondingly, two sets of MOPRL and MOPUC. We discuss this further in Sections 2.1 and 3.1.

Occasionally, for $\mathbb{C}^l$ vector-valued functions f and g, we will want to consider the scalar

$$\sum_{k,j} \int \overline{f_k(x)}\, M_{kj}(x) g_j(x)\, d\mu_{\mathrm{tr}}(x), \tag{1.34}$$

which we will denote

$$\int d\langle f(x), \mu(x) g(x)\rangle_{\mathbb{C}^l}. \tag{1.35}$$

We next turn to module Fourier expansions. A set $\{\varphi_j\}_{j=1}^N$ in $\mathcal{H}$ (N may be infinite) is called orthonormal if and only if

$$\langle\langle \varphi_j, \varphi_k\rangle\rangle_R = \delta_{jk}\mathbf{1}. \tag{1.36}$$

This natural terminology is an abuse of notation since (1.36) implies orthogonality in $\langle\cdot,\cdot\rangle_R$ but not normalization, and is much stronger than orthogonality in $\langle\cdot,\cdot\rangle_R$.

Suppose for a moment that $N < \infty$. For any $a_1, \dots, a_N \in \mathcal{M}_l$, we can form $\sum_{j=1}^N \varphi_j a_j$ and, by the right multiplication relations (1.28), (1.29), and (1.36), we have

$$\left\langle\!\!\left\langle \sum_{j=1}^N \varphi_j a_j, \sum_{j=1}^N \varphi_j b_j \right\rangle\!\!\right\rangle_R = \sum_{j=1}^N a_j^\dagger b_j. \tag{1.37}$$

We will denote the set of all such $\sum_{j=1}^N \varphi_j a_j$ by $\mathcal{H}_{(\varphi_j)}$—it is a vector subspace of $\mathcal{H}$ of dimension (over $\mathbb{C}$) Nl^2.

Define for $f \in \mathcal{H}$,

$$\pi_{(\varphi_j)}(f) = \sum_{j=1}^{N} \varphi_j \langle\!\langle \varphi_j, f\rangle\!\rangle_R. \tag{1.38}$$

It is easy to see it is the orthogonal projection in the scalar inner product $\langle \cdot, \cdot \rangle_R$ from $\mathcal{H}$ to $\mathcal{H}_{(\varphi_j)}$.

By the standard Hilbert space calculation, taking care to only multiply on the right, one finds the Pythagorean theorem,

$$\langle\!\langle f, f\rangle\!\rangle_R = \langle\!\langle f - \pi_{(\varphi_j)} f, f - \pi_{(\varphi_j)} f\rangle\!\rangle_R + \sum_{j=1}^{N} \langle\!\langle \varphi_j, f\rangle\!\rangle_R^\dagger \langle\!\langle \varphi_j, f\rangle\!\rangle_R. \tag{1.39}$$

As usual, this proves for infinite N that

$$\sum_{j=1}^{N} \langle\!\langle \varphi_j, f\rangle\!\rangle_R^\dagger \langle\!\langle \varphi_j, f\rangle\!\rangle_R \le \langle\!\langle f, f\rangle\!\rangle_R \tag{1.40}$$

and the convergence of

$$\sum_{j=1}^{N} \varphi_j \langle\!\langle \varphi_j, f\rangle\!\rangle_R \equiv \pi_{(\varphi_j)}(f) \tag{1.41}$$

allowing the definition of $\pi_{(\varphi_j)}$ and of $\mathcal{H}_{(\varphi_j)} \equiv \operatorname{Ran} \pi_{(\varphi_j)}$ for $N = \infty$.

An orthonormal set is called complete if $\mathcal{H}_{(\varphi_j)} = \mathcal{H}$. In that case, equality holds in (1.40) and $\pi_{(\varphi_j)}(f) = f$.

For orthonormal bases, we have the Parseval relation from (1.39)

$$\langle\!\langle f, f\rangle\!\rangle_R = \sum_{j=1}^{\infty} \langle\!\langle \varphi_j, f\rangle\!\rangle_R^\dagger \langle\!\langle \varphi_j, f\rangle\!\rangle_R \tag{1.42}$$

and

$$\|f\|_R^2 = \sum_{j=1}^{\infty} \operatorname{Tr}(\langle\!\langle \varphi_j, f\rangle\!\rangle_R^\dagger \langle\!\langle \varphi_j, f\rangle\!\rangle_R). \tag{1.43}$$

1.3 Matrix Möbius Transformations

Without an understanding of matrix Möbius transformations, the form of the MOPUC Geronimus theorem we will prove in Section 3.10 will seem strange-looking. To set the stage, recall that scalar fractional linear transformations (FLT) are associated to matrices $T = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ with $\det T \neq 0$ via

$$f_T(z) = \frac{az + b}{cz + d}. \tag{1.44}$$

Without loss, one can restrict to

$$\det(T) = 1. \tag{1.45}$$

Indeed, $T \mapsto f_T$ is a 2 to 1 map of $\mathbb{SL}(2, \mathbb{C})$ to maps of $\mathbb{C} \cup \{\infty\}$ to itself. One advantage of the matrix formalism is that the map is a matrix homomorphism, that is,

$$f_{T \circ S} = f_T \circ f_S, \tag{1.46}$$

which shows that the group of FLTs is $\mathbb{SL}(2,\mathbb{C})/\{\mathbf{1},-\mathbf{1}\}$.

While (1.46) can be checked by direct calculation, a more instructive way is to look at the complex projective line. $u,v\in\mathbb{C}^2\setminus\{0\}$ are called equivalent if there is $\lambda\in\mathbb{C}\setminus\{0\}$ so that $u=\lambda v$. Let $[\cdot]$ denote equivalence classes. Except for $[\binom{1}{0}]$, every equivalence class contains exactly one point of the form $\binom{z}{1}$ with $z\in\mathbb{C}$. If $[\binom{1}{0}]$ is associated with ∞, the set of equivalence classes is naturally associated with $\mathbb{C}\cup\{\infty\}$. f_T then obeys

$$\left[T\begin{pmatrix} z \\ 1 \end{pmatrix}\right] = \left[\begin{pmatrix} f_T(z) \\ 1 \end{pmatrix}\right] \tag{1.47}$$

from which (1.46) is immediate.

By Möbius transformations we will mean those FLTs that map $\mathbb{D}$ onto itself. Let

$$J = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.48}$$

Then $[u]=[\binom{z}{1}]$ with $|z|=1$ (resp. $|z|<1$) if and only if $\langle u,Ju\rangle=0$ (resp. $\langle u,Ju\rangle<0$). From this, it is not hard to show that if $\det(T)=1$, then f_T maps $\mathbb{D}$ invertibly onto $\mathbb{D}$ if and only if

$$T^\dagger JT = J. \tag{1.49}$$

If T has the form $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, this is equivalent to

$$|a|^2-|c|^2=1, \qquad |b|^2-|d|^2=-1, \qquad \bar{a}b-\bar{c}d=0. \tag{1.50}$$

The set of T's obeying $\det(T)=1$ and (1.49) is called $\mathbb{SU}(1,1)$. It is studied extensively in [168, Sect. 10.4].

The self-adjoint elements of $\mathbb{SU}(1,1)$ are parametrized by $\alpha\in\mathbb{D}$ via $\rho=(1-|\alpha|^2)^{1/2}$,

$$T_\alpha = \frac{1}{\rho}\begin{pmatrix} 1 & \alpha \\ \bar{\alpha} & 1 \end{pmatrix} \tag{1.51}$$

associated to

$$f_{T_\alpha}(z) = \frac{z+\alpha}{1+\bar{\alpha}z}\,. \tag{1.52}$$

Notice that

$$T_\alpha^{-1} = T_{-\alpha} \tag{1.53}$$

and that

$$\forall z\in\mathbb{D},\ \exists!\,\alpha \text{ such that } T_\alpha(0)=z,$$

namely, $\alpha=z$.

It is a basic theorem that every holomorphic bijection of $\mathbb{D}$ to $\mathbb{D}$ is an f_T for some T in $\mathbb{SU}(1,1)$ (unique up to $\pm\mathbf{1}$).

With this in place, we can turn to the matrix case. Let $\mathcal{M}_l$ be the space of $l\times l$ complex matrices with the Euclidean norm induced by the vector norm $\langle\cdot,\cdot\rangle_{\mathbb{C}^l}^{1/2}$. Let

$$\mathbb{D}_l = \{A\in\mathcal{M}_l : \|A\|<1\}. \tag{1.54}$$

We are interested in holomorphic bijections of $\mathbb{D}_l$ to itself, especially via a suitable notion of FLT. There is a huge (and diffuse) literature on the subject, starting with its use in analytic number theory. It has also been studied in connection with electrical engineering filters and indefinite matrix Hilbert spaces. Among the huge literature, we mention [1, 3, 78, 99, 114, 166]. Especially relevant to MOPUC is the book of Bakonyi–Constantinescu [6].

Consider $\mathcal{M}_l \oplus \mathcal{M}_l = \mathcal{M}_l[2]$ as a right module over $\mathcal{M}_l$. The $\mathcal{M}_l$-projective line is defined by saying $\left[\begin{smallmatrix} X \\ Y \end{smallmatrix}\right] \sim \left[\begin{smallmatrix} X' \\ Y' \end{smallmatrix}\right]$, both in $\mathcal{M}_l[2] \setminus \{\mathbf{0}\}$, if and only if there exists $\Lambda \in \mathcal{M}_l$, Λ invertible so that

$$X = X'\Lambda, \qquad Y = Y'\Lambda. \tag{1.55}$$

Let T be a map of $\mathcal{M}_l[2]$ of the form

$$T = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \tag{1.56}$$

acting on $\mathcal{M}_l[2]$ by

$$T \begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} AX + BY \\ CX + DY \end{bmatrix}. \tag{1.57}$$

Because this acts on the left and Λ equivalence on the right, T maps equivalence classes to themselves. In particular, if $CX + D$ is invertible, T maps the equivalence class of $\left[\begin{smallmatrix} X \\ \mathbf{1} \end{smallmatrix}\right]$ to the equivalence class of $\left[\begin{smallmatrix} f_T[X] \\ \mathbf{1} \end{smallmatrix}\right]$, where

$$f_T[X] = (AX + B)(CX + D)^{-1}. \tag{1.58}$$

So long as $CX + D$ remains invertible, (1.46) remains true. Let J be the $2l \times 2l$ matrix in $l \times l$ block form

$$J = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix}. \tag{1.59}$$

Note that (with $\left[\begin{smallmatrix} X \\ \mathbf{1} \end{smallmatrix}\right]^\dagger = [X^\dagger \mathbf{1}]$)

$$\begin{bmatrix} X \\ \mathbf{1} \end{bmatrix}^\dagger J \begin{bmatrix} X \\ \mathbf{1} \end{bmatrix} \leq \mathbf{0} \Leftrightarrow X^\dagger X \leq \mathbf{1} \Leftrightarrow \|X\| \leq \mathbf{1}. \tag{1.60}$$

Therefore, if we define $\mathbb{SU}(l, l)$ to be those T's with $\det T = 1$ and

$$T^\dagger J T = J, \tag{1.61}$$

then

$$T \in \mathbb{SU}(l, l) \Rightarrow f_T[\mathbb{D}_l] = \mathbb{D}_l \text{ as a bijection.} \tag{1.62}$$

If T has the form (1.56), then (1.61) is equivalent to

$$A^\dagger A - C^\dagger C = D^\dagger D - B^\dagger B = \mathbf{1}, \tag{1.63}$$

$$A^\dagger B = C^\dagger D \tag{1.64}$$

(the fourth relation $B^\dagger A = D^\dagger C$ is equivalent to (1.64)).

This depends on

Proposition 1.1. *If $T = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ obeys* (1.61) *and $\|X\| < 1$, then $CX + D$ is invertible.*

Proof. (1.61) implies that

$$T^{-1} = JT^{\dagger}J \tag{1.65}$$

$$= \begin{pmatrix} A^{\dagger} & -C^{\dagger} \\ -B^{\dagger} & D^{\dagger} \end{pmatrix}. \tag{1.66}$$

Clearly, (1.61) also implies $T^{-1} \in \mathbb{SU}(l,l)$. Thus, by (1.63) for T^{-1},

$$DD^{\dagger} - CC^{\dagger} = \mathbf{1}. \tag{1.67}$$

This implies first that $DD^{\dagger} \geq \mathbf{1}$, so D is invertible, and second that

$$\|D^{-1}C\| \leq 1. \tag{1.68}$$

Thus, $\|X\| < 1$ implies $\|D^{-1}CX\| < 1$ so $\mathbf{1} + D^{-1}CX$ is invertible, and thus so is $D(\mathbf{1} + D^{-1}CX)$. $\square$

It is a basic result of Cartan [18] (see Helgason [114] and the discussion therein) that

Theorem 1.2. *A holomorphic bijection, g, of $\mathbb{D}_l$ to itself is either of the form*

$$g(X) = f_T(X) \tag{1.69}$$

for some $T \in \mathbb{SU}(l,l)$ or

$$g(X) = f_T(X^t). \tag{1.70}$$

Given $\alpha \in \mathcal{M}_l$ with $\|\alpha\| < 1$, define

$$\rho^L = (\mathbf{1} - \alpha^{\dagger}\alpha)^{1/2}, \qquad \rho^R = (\mathbf{1} - \alpha\alpha^{\dagger})^{1/2}. \tag{1.71}$$

Lemma 1.3. *We have*

$$\alpha\rho^L = \rho^R\alpha, \qquad \alpha^{\dagger}\rho^R = \rho^L\alpha^{\dagger}, \tag{1.72}$$

$$\alpha(\rho^L)^{-1} = (\rho^R)^{-1}\alpha, \qquad \alpha^{\dagger}(\rho^R)^{-1} = (\rho^L)^{-1}\alpha^{\dagger}. \tag{1.73}$$

Proof. Let f be analytic in $\mathbb{D}$ with $f(z) = \sum_{n=0}^{\infty} c_n z^n$ its Taylor series at $z = 0$. Since $\|\alpha^{\dagger}\alpha\| < 1$, we have

$$f(\alpha^{\dagger}\alpha) = \sum_{n=0}^{\infty} c_n(\alpha^{\dagger}\alpha)^n \tag{1.74}$$

norm convergent, so $\alpha(\alpha^{\dagger}\alpha)^n = (\alpha\alpha^{\dagger})^n\alpha$ implies

$$\alpha f(\alpha^{\dagger}\alpha) = f(\alpha\alpha^{\dagger})\alpha, \tag{1.75}$$

which implies the first halves of (1.72) and (1.73). The other halves follow by taking adjoints. $\square$

Theorem 1.4. *There is a one-one correspondence between α's in $\mathcal{M}_l$ obeying $\|\alpha\| < 1$ and positive self-adjoint elements of $\mathbb{SU}(l,l)$ via*

$$T_\alpha = \begin{pmatrix} (\rho^R)^{-1} & (\rho^R)^{-1}\alpha \\ (\rho^L)^{-1}\alpha^{\dagger} & (\rho^L)^{-1} \end{pmatrix}. \tag{1.76}$$

Proof. A straightforward calculation using Lemma 1.3 proves that T_α is self-adjoint and $T_\alpha^\dagger J T_\alpha = J$. Conversely, if T is self-adjoint, $T = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ and in $\mathbb{SU}(l,l)$, then $T^\dagger = T \Rightarrow A^\dagger = A$, $B^\dagger = C$, so (1.63) becomes

$$AA^\dagger - BB^\dagger = \mathbf{1}, \tag{1.77}$$

so if

$$\alpha = A^{-1}B, \tag{1.78}$$

then (1.77) becomes

$$A^{-1}(A^{-1})^\dagger + \alpha\alpha^\dagger = \mathbf{1}. \tag{1.79}$$

Since $T \geq 0$, $A \geq 0$ so (1.79) implies $A = (\rho^R)^{-1}$, and then (1.78) implies $B = (\rho^R)^{-1}\alpha$.

By Lemma 1.3,

$$C = B^\dagger = \alpha^\dagger(\rho^R)^{-1} = (\rho^L)^{-1}\alpha^\dagger \tag{1.80}$$

and then (by $D = D^\dagger$, $C^\dagger = B$, and (1.63)) $DD^\dagger - CC^\dagger = \mathbf{1}$ plus $D > 0$ implies $D = (\rho^L)^{-1}$. □

Corollary 1.5. *For each $\alpha \in \mathbb{D}_l$, the map*

$$f_{T_\alpha}(X) = (\rho^R)^{-1}(X+\alpha)(\mathbf{1}+\alpha^\dagger X)^{-1}(\rho^L) \tag{1.81}$$

takes $\mathbb{D}_l$ to $\mathbb{D}_l$. Its inverse is given by

$$f_{T_\alpha}^{-1}(X) = f_{T_{-\alpha}}(X) = (\rho^R)^{-1}(X-\alpha)(\mathbf{1}-\alpha^\dagger X)^{-1}(\rho^L). \tag{1.82}$$

There is an alternate form for the right side of (1.81).

Proposition 1.6. *The following identity holds true for any X, $\|X\| \leq 1$:*

$$\rho^R(1+X\alpha^\dagger)^{-1}(X+\alpha)(\rho^L)^{-1} = (\rho^R)^{-1}(X+\alpha)(1+\alpha^\dagger X)^{-1}\rho^L. \tag{1.83}$$

Proof. By the definition of ρ^L and ρ^R, we have

$$X(\rho^L)^{-2}(1-\alpha^\dagger\alpha) = (\rho^R)^{-2}(1-\alpha\alpha^\dagger)X.$$

Expanding, using (1.73) and rearranging, we get

$$X(\rho^L)^{-2} + \alpha(\rho^L)^{-2}\alpha^\dagger X = (\rho^R)^{-2}X + X\alpha^\dagger(\rho^R)^{-2}\alpha.$$

Adding $\alpha(\rho^L)^{-2} + X(\rho^L)^{-2}\alpha^\dagger X$ to both sides and using (1.73) again, we obtain

$$\begin{aligned} X(\rho^L)^{-2} &+ \alpha(\rho^L)^{-2} + X(\rho^L)^{-2}\alpha^\dagger X + \alpha(\rho^L)^{-2}\alpha^\dagger X \\ &= (\rho^R)^{-2}X + (\rho^R)^{-2}\alpha + X\alpha^\dagger(\rho^R)^{-2}X + X\alpha^\dagger(\rho^R)^{-2}\alpha, \end{aligned}$$

which is the same as

$$(X+\alpha)(\rho^L)^{-2}(1+\alpha^\dagger X) = (1+X\alpha^\dagger)(\rho^R)^{-2}(X+\alpha).$$

Multiplying by $(1+X\alpha^\dagger)^{-1}$ and $(1+\alpha^\dagger X)^{-1}$, we get

$$(1+X\alpha^\dagger)^{-1}(X+\alpha)(\rho^L)^{-2} = (\rho^R)^{-2}(X+\alpha)(1+\alpha^\dagger X)^{-1}$$

and the statement follows. □

1.4 Applications and Examples

There are a number of simple examples which show that beyond their intrinsic mathematical interest, MOPRL and MOPUC have wide application.

(a) **Jacobi matrices on a strip**

Let $\Lambda \subset \mathbb{Z}^\nu$ be a subset (perhaps infinite) of the ν-dimensional lattice $\mathbb{Z}^\nu$ and let $\ell^2(\Lambda)$ be square summable sequences indexed by Λ. Suppose a real symmetric matrix α_{ij} is given for all $i, j \in \Lambda$ with $\alpha_{ij} = 0$ unless $|i - j| = 1$ (nearest neighbors). Let β_i be a real sequence indexed by $i \in \Lambda$. Suppose

$$\sup_{i,j} |\alpha_{ij}| + \sup_i |\beta_i| < \infty. \tag{1.84}$$

Define a bounded operator, J, on $\ell^2(\Lambda)$ by

$$(Ju)_i = \sum_j \alpha_{ij} u_j + \beta_i u_i. \tag{1.85}$$

The sum is finite with at most 2ν elements.

The special case $\Lambda = \{1, 2, \dots\}$ with $b_i = \beta_i$, $a_i = \alpha_{i,i+1} > 0$ corresponds precisely to classical semi-infinite tridiagonal Jacobi matrices.

Now consider the situation where $\Lambda' \subset \mathbb{Z}^{\nu-1}$ is a finite set with l elements and

$$\Lambda = \{j \in \mathbb{Z}^\nu : j_1 \in \{1, 2, \dots\};\ (j_2, \dots j_\nu) \in \Lambda'\}, \tag{1.86}$$

a "strip" with cross-section Λ'. J then has a block $l \times l$ matrix Jacobi form where $(\gamma, \delta \in \Lambda')$

$$(B_i)_{\gamma\delta} = b_{(i,\gamma)}, \qquad (\gamma = \delta), \tag{1.87}$$

$$= a_{(i,\gamma)(i,\delta)}, \qquad (\gamma \neq \delta), \tag{1.88}$$

$$(A_i)_{\gamma\delta} = a_{(i,\gamma)(i+1,\delta)}. \tag{1.89}$$

The nearest neighbor condition says $(A_i)_{\gamma\delta} = 0$ if $\gamma \neq \delta$. If

$$a_{(i,\gamma)(i+1,\gamma)} > 0 \tag{1.90}$$

for all i, γ, then A_i is invertible and we have a block Jacobi matrix of the kind described in Section 2.2 below.

By allowing general A_i, B_i, we obtain an obvious generalization of this model—an interpretation of general MOPRL.

Schrödinger operators on strips have been studied in part as approximations to $\mathbb{Z}^\nu$; see [31, 95, 130, 134, 151, 164]. From this point of view, it is also natural to allow periodic boundary conditions in the vertical directions. Furthermore, there is closely related work on Schrödinger (and other) operators with matrix-valued potentials; see, for example, [8, 24, 25, 26, 27, 28, 30, 96, 97, 165].

(b) **Two-sided Jacobi matrices**

This example goes back at least to Nikishin [153]. Consider the case $\nu = 2$, $\Lambda' = \{0,1\} \subset \mathbb{Z}$, and Λ as above. Suppose (1.90) holds, and in addition,

$$a_{(1,0)(1,1)} > 0, \tag{1.91}$$

$$a_{(i,0)(i,1)} = 0, \quad i = 2, 3, \dots. \tag{1.92}$$

Then there are no links between the rungs of the "ladder," $\{1, 2, \dots\} \times \{0, 1\}$ except at the end and the ladder can be unfolded to $\mathbb{Z}$! Thus, a two-sided Jacobi matrix can be viewed as a special kind of one-sided 2×2 matrix Jacobi operator.

It is known that for two-sided Jacobi matrices, the spectral theory is determined by the 2×2 matrix

$$d\mu = \begin{pmatrix} d\mu_{00} & d\mu_{01} \\ d\mu_{10} & d\mu_{11} \end{pmatrix}, \tag{1.93}$$

where $d\mu_{kl}$ is the measure with

$$\langle \delta_k, (J - \lambda)^{-1} \delta_l \rangle = \int \frac{d\mu_{kl}(x)}{x - \lambda}, \tag{1.94}$$

but also that it is very difficult to specify exactly which $d\mu$ correspond to two-sided Jacobi matrices.

This difficulty is illuminated by the theory of MOPRL. By Favard's theorem (see Theorem 2.11), every such $d\mu$ (given by (1.93) and positive definite and non-trivial in a sense we will describe in Lemma 2.1) yields a unique block Jacobi matrix with $A_j > 0$ (positive definite). This $d\mu$ comes from a two-sided Jacobi matrix if and only if

(a) B_j is diagonal for $j = 2, 3, \dots$.
(b) A_j is diagonal for $j = 1, 2, \dots$.
(c) B_j has strictly positive off-diagonal elements.

These are very complicated indirect conditions on $d\mu$!

(c) **Banded matrices**

Classical Jacobi matrices are semi-infinite symmetric tridiagonal matrices, that is,

$$J_{km} = 0 \quad \text{if } |k - m| > 1 \tag{1.95}$$

with

$$J_{km} > 0 \quad \text{if } |k - m| = 1. \tag{1.96}$$

A natural generalization are $(2l + 1)$-diagonal symmetric matrices, that is,

$$J_{km} = 0 \quad \text{if } |k - m| > l, \tag{1.97}$$

$$J_{km} > 0 \quad \text{if } |k - m| = l. \tag{1.98}$$

Such a matrix can be partitioned into $l \times l$ blocks, which is tridiagonal in block. The conditions (1.97) and (1.98) are equivalent to $A_k \in \mathcal{L}$, the set of lower triangular matrices; and conversely, $A_k \in \mathcal{L}$, with A_k, B_k real (and B_k symmetric) correspond precisely to such banded matrices. This is why we introduce type 3 MOPRL.

Banded matrices correspond to certain higher-order difference equations. Unlike the second-order equation (which leads to tridiagonal matrices) where every equation with positive coefficients is equivalent via a variable rescaling to a symmetric matrix, only certain higher-order difference equations correspond to symmetric block Jacobi matrices.

(d) **Magic formula**

In [33], Damanik, Killip, and Simon studied perturbations of Jacobi and CMV matrices with periodic Jacobi parameters (or Verblunsky coefficients). They proved that if Δ is the discriminant of a two-sided periodic J_0, then a bounded two-sided J has $\Delta(J) = S^p + S^{-p}$ ($(Su)_n \equiv u_{n+1}$) if and only if J lies in the isospectral torus of J_0. They call this the magic formula.

This allows the study of perturbations of the isospectral torus by studying $\Delta(J)$ which is a polynomial in J of degree p, and so a $2p+1$ banded matrix. Thus, the study of perturbations of periodic problems is connected to perturbations of $S^p + S^{-p}$ as block Jacobi matrices. Indeed, it was this connection that stimulated our interest in MOPRL, and [33] uses some of our results here.

(e) **Vector-valued prediction theory**

As noted in Section 1.1, both prediction theory and filtering theory use OPUC and have natural MOPUC settings that motivated much of the MOPUC literature.

2 Matrix Orthogonal Polynomials on the Real Line

2.1 Preliminaries

OPRL are the most basic and developed of orthogonal polynomials, and so this chapter on the matrix analogue is the most important of this survey. We present the basic formulas, assuming enough familiarity with the scalar case (see [23, 82, 167, 176, 184, 185]) that we do not need to explain why the objects we define are important.

2.1.1 Polynomials, Inner Products, Norms

Let $d\mu$ be an $l \times l$ matrix-valued Hermitian positive semi-definite finite measure on $\mathbb{R}$ normalized by $\mu(\mathbb{R}) = \mathbf{1} \in \mathcal{M}_l$. We assume for simplicity that μ has a compact support. However, many of the results below do not need the latter restriction and in fact can be found in the literature for matrix-valued measures with unbounded support.

Define (as in (1.20))

$$\langle\langle f, g\rangle\rangle_R = \int f(x)^\dagger \, d\mu(x) \, g(x), \quad \|f\|_R = (\mathrm{Tr}\langle\langle f, f\rangle\rangle_R)^{1/2}, \quad f, g \in \mathcal{P},$$
$$\langle\langle f, g\rangle\rangle_L = \int g(x) \, d\mu(x) \, f(x)^\dagger, \quad \|f\|_L = (\mathrm{Tr}\langle\langle f, f\rangle\rangle_L)^{1/2}, \quad f, g \in \mathcal{P}.$$

Clearly, we have

$$\langle\langle f, g\rangle\rangle_R^\dagger = \langle\langle g, f\rangle\rangle_R, \qquad \langle\langle f, g\rangle\rangle_L^\dagger = \langle\langle g, f\rangle\rangle_L, \tag{2.1}$$
$$\langle\langle f, g\rangle\rangle_L = \langle\langle g^\dagger, f^\dagger\rangle\rangle_R, \qquad \|f\|_L = \|f^\dagger\|_R. \tag{2.2}$$

As noted in Section 1.2, we have the left and right analogues of the Cauchy inequality

$$|\mathrm{Tr}\langle\langle f, g\rangle\rangle_R| \le \|f\|_R \|g\|_R, \quad |\mathrm{Tr}\langle\langle f, g\rangle\rangle_L| \le \|f\|_L \, \|g\|_L.$$

Thus, $\|\cdot\|_R$ and $\|\cdot\|_L$ are semi-norms in $\mathcal{P}$. Indeed, as noted in Section 1.2, they are associated to an inner product. The sets $\{f : \|f\|_R = 0\}$ and $\{f : \|f\|_L = 0\}$ are linear subspaces. Let $\mathcal{P}_R$

be the completion of $\mathcal{P}/\{f : \|f\|_R = 0\}$ (viewed as a right module over $\mathcal{M}_l$) with respect to the norm $\|\cdot\|_R$. Similarly, let $\mathcal{P}_L$ be the completion of $\mathcal{P}/\{f : \|f\|_L = 0\}$ (viewed as a left module) with respect to the norm $\|\cdot\|_L$.

The set $\mathcal{V}$ defined in Section 1.2 is a linear space. Let us introduce a semi-norm in $\mathcal{V}$ by

$$|f| = \left\{ \int d\langle f(x), \mu(x) f(x)\rangle_{\mathbb{C}^l} \right\}^{1/2}. \tag{2.3}$$

Let $\mathcal{V}_0 \subset \mathcal{V}$ be the linear subspace of all polynomials such that $|f| = 0$ and let $\mathcal{V}_\infty$ be the completion of the quotient space $\mathcal{V}/\mathcal{V}_0$ with respect to the norm $|\cdot|$.

Lemma 2.1. *The following are equivalent:*
(1) $\|f\|_R > 0$ *for every non-zero* $f \in \mathcal{P}$.
(2) *For all* n, *the dimension in* $\mathcal{P}_R$ *of the set of all polynomials of degree at most* n *is* $(n+1)l^2$.
(3) $\|f\|_L > 0$ *for every non-zero* $f \in \mathcal{P}$.
(4) *For all* n, *the dimension in* $\mathcal{P}_L$ *of the set of all polynomials of degree at most* n *is* $(n+1)l^2$.
(5) *For every non-zero* $v \in \mathcal{V}$, *we have that* $|v| \neq 0$.
(6) *For all* n, *the dimension in* $\mathcal{V}_\infty$ *of all vector-valued polynomials of degree at most* n *is* $(n+1)l$.
The measure $d\mu$ *is called non-trivial if these equivalent conditions hold.*

Remark. If $l = 1$, these are equivalent to the usual non- triviality condition, that is, $\mathrm{supp}(\mu)$ is infinite. For $l > 1$, we cannot define triviality in this simple way, as can be seen by looking at the direct sum of a trivial and non-trivial measure. In that case, the measure is not non-trivial in the above sense but its support is infinite.

Proof. The equivalences (1) $\Leftrightarrow$ (2), (3) $\Leftrightarrow$ (4), and (5) $\Leftrightarrow$ (6) are immediate. The equivalence (1) $\Leftrightarrow$ (3) follows from (2.2). Let us prove the equivalence (1) $\Leftrightarrow$ (5). Assume that (1) holds and let $v \in \mathcal{V}$ be non-zero. Let $f \in \mathcal{M}_l$ denote the matrix that has v as its leftmost column and that has zero columns otherwise. Then, $0 \neq \|f\|_R^2 = \mathrm{Tr}\langle\langle f, f\rangle\rangle_R = |v|^2$ and hence (5) holds. Now assume that (1) fails and let $f \in \mathcal{P}$ be non-zero with $\|f\|_R = 0$. Then, at least one of the column vectors of f is non-zero. Suppose for simplicity that this is the first column and denote this column vector by v. Let $t \in \mathcal{M}_l$ be the matrix $t_{ij} = \delta_{i1}\delta_{j1}$; then we have

$$\|f\|_R = 0 \Rightarrow \langle\langle f, f\rangle\rangle_R = 0 \Rightarrow 0 = \mathrm{Tr}(t^* \langle\langle f, f\rangle\rangle_R t) = |v|^2$$

and hence (5) fails. $\square$

Throughout the rest of this chapter, we assume the measure $d\mu$ to be non-trivial.

2.1.2 Monic Orthogonal Polynomials

Lemma 2.2. *Let* $d\mu$ *be a non-trivial measure.*
(i) *There exists a unique monic polynomial* P_n^R *of degree* n, *which minimizes the norm* $\|P_n^R\|_R$.
(ii) *The polynomial* P_n^R *can be equivalently defined as the monic polynomial of degree* n *which satisfies*

$$\langle\langle P_n^R, f\rangle\rangle_R = \mathbf{0} \quad \text{for any } f \in \mathcal{P}, \quad \deg f < n. \tag{2.4}$$

(iii) *There exists a unique monic polynomial* P_n^L *of degree* n, *which minimizes the norm* $\|P_n^L\|_L$.

(iv) *The polynomial* P_n^L *can be equivalently defined as the monic polynomial of degree* n *which satisfies*

$$\langle\langle P_n^L, f\rangle\rangle_L = \mathbf{0} \quad \text{for any } f \in \mathcal{P}, \quad \deg f < n. \tag{2.5}$$

(v) *One has* $P_n^L(x) = P_n^R(x)^\dagger$ *for all* $x \in \mathbb{R}$ *and*

$$\langle\langle P_n^R, P_n^R\rangle\rangle_R = \langle\langle P_n^L, P_n^L\rangle\rangle_L. \tag{2.6}$$

Proof. As noted, $\mathcal{P}$ has an inner product $\langle \cdot, \cdot \rangle_R$, so there is an orthogonal projection $\pi_n^{(R)}$ onto $\mathcal{P}_n$ discussed in Section 1.2. Then

$$P_n^R(x) = x^n - \pi_{n-1}^{(R)}(x^n). \tag{2.7}$$

As usual, in inner product spaces, this uniquely minimizes $x^n - Q$ over all $Q \in \mathcal{P}_{n-1}$. It clearly obeys

$$\mathrm{Tr}(\langle\langle P_n^R, f\rangle\rangle_R) = 0 \tag{2.8}$$

for all $f \in \mathcal{P}_{n-1}$. But then for any matrix α,

$$\mathrm{Tr}(\langle\langle P_n^R, f\rangle\rangle_R \alpha) = \mathrm{Tr}(\langle\langle P_n^R, f\alpha\rangle\rangle_R) = 0$$

so (2.4) holds.

This proves (i) and (ii). (iii) and (iv) are similar. To prove (v), note that $P_n^L(x) = P_n^R(x)^\dagger$ follows from the criteria (2.4), (2.5). The identity (2.6) follows from (2.2). $\square$

Lemma 2.3. *Let* μ *be non-trivial. For any monic polynomial* P*, we have* $\det\langle\langle P, P\rangle\rangle_R \neq 0$ *and* $\det\langle\langle P, P\rangle\rangle_L \neq 0$.

Proof. Let P be a monic polynomial of degree n such that $\langle\langle P, P\rangle\rangle_R$ has a non-trivial kernel. Then one can find $\alpha \in \mathcal{M}_l$, $\alpha \neq \mathbf{0}$, such that $\alpha^\dagger \langle\langle P, P\rangle\rangle_R \alpha = \mathbf{0}$. It follows that $\|P\alpha\|_R = 0$. But since P is monic, the leading coefficient of $P\alpha$ is α, so $P\alpha \neq \mathbf{0}$, which contradicts the non-triviality assumption. A similar argument works for $\langle\langle P, P\rangle\rangle_L$. $\square$

By the orthogonality of $Q_n - P_n^R$ to P_n^R for any monic polynomial Q_n of degree n, we have

$$\langle\langle Q_n, Q_n\rangle\rangle_R = \langle\langle Q - P_n^R, Q - P_n^R\rangle\rangle_R + \langle\langle P_n^R, P_n^R\rangle\rangle_R \tag{2.9}$$

and, in particular,

$$\langle\langle P_n^R, P_n^R\rangle\rangle_R \leq \langle\langle Q_n, Q_n\rangle\rangle_R \tag{2.10}$$

with (by non-triviality) equality if and only if $Q_n = P_n^R$. Since Tr and det are strictly monotone on strictly positive matrices, we have the following variational principles ((2.11) restates (i) of Lemma 2.2):

Theorem 2.4. *For any monic* Q_n *of degree* n*, we have*

$$\|Q_n\|_R \geq \|P_n^R\|_R, \tag{2.11}$$

$$\det\langle\langle Q_n, Q_n\rangle\rangle_R \geq \det\langle\langle P_n^R, P_n^R\rangle\rangle_R \tag{2.12}$$

with equality if and only if $P_n^R = Q_n$.

2.1.3 Expansion

Theorem 2.5. *Let $d\mu$ be non-trivial.*

(i) *We have*

$$\langle\!\langle P_k^R, P_n^R \rangle\!\rangle_R = \gamma_n \delta_{kn} \tag{2.13}$$

for some positive invertible matrices γ_n.

(ii) $\{P_k^R\}_{k=0}^n$ *are a right-module basis for $\mathcal{P}_n$; indeed, any $f \in \mathcal{P}_n$ has a unique expansion,*

$$f = \sum_{j=0}^n P_j^R f_j^R. \tag{2.14}$$

Indeed, essentially by (1.38),

$$f_j^R = \gamma_j^{-1} \langle\!\langle P_j^R, f \rangle\!\rangle_R. \tag{2.15}$$

Remark. There are similar formulas for $\langle\!\langle \cdot, \cdot \rangle\!\rangle_L$. By (2.6),

$$\langle\!\langle P_k^L, P_n^L \rangle\!\rangle_L = \gamma_n \delta_{kn} \tag{2.16}$$

(same γ_n, which is why we use γ_n and not γ_n^R).

Proof. (i) (2.13) for $n < k$ is immediate from (2.5) and for $n > k$ by symmetry. $\gamma_n \geq 0$ follows from (1.22). By Lemma 2.3, $\det(\gamma_n) \neq 0$, so γ_n is invertible.

(ii) Map $(\mathcal{M}_l)^{n+1}$ to $\mathcal{P}_n$ by

$$\langle \alpha_0, \dots, \alpha_n \rangle \mapsto \sum_{j=0}^n P_j^R \alpha_j \equiv X(\alpha_0, \dots, \alpha_n).$$

By (2.13),

$$\alpha_j = \gamma_j^{-1} \langle\!\langle P_j^R, X(\alpha_0, \dots, \alpha_n) \rangle\!\rangle$$

so that map is one-one. By dimension counting, it is onto. $\square$

2.1.4 Recurrence Relations for Monic Orthogonal Polynomials

Denote by ζ_n^R (resp. ζ_n^L) the coefficient of x^{n-1} in $P_n^R(x)$ (resp. $P_n^L(x)$), that is,

$$P_n^R(x) = x^n \mathbf{1} + \zeta_n^R x^{n-1} + \text{lower order terms},$$
$$P_n^L(x) = x^n \mathbf{1} + \zeta_n^L x^{n-1} + \text{lower order terms}.$$

Since $P_n^R(x)^\dagger = P_n^L(x)$, we have $(\zeta_n^R)^\dagger = \zeta_n^L$. Using the parameters γ_n of (2.13) and ζ_n^R, ζ_n^L one can write down recurrence relations for $P_n^R(x)$, $P_n^L(x)$.

Lemma 2.6. (i) *We have a commutation relation*

$$\gamma_{n-1}(\zeta_n^R - \zeta_{n-1}^R) = (\zeta_n^L - \zeta_{n-1}^L)\gamma_{n-1}. \tag{2.17}$$

(ii) *We have the recurrence relations*

$$xP_n^R(x) = P_{n+1}^R(x) + P_n^R(x)(\zeta_n^R - \zeta_{n+1}^R) + P_{n-1}(x)\gamma_{n-1}^{-1}\gamma_n, \tag{2.18}$$
$$xP_n^L(x) = P_{n+1}^L(x) + (\zeta_n^L - \zeta_{n+1}^L)P_n^L(x) + \gamma_n\gamma_{n-1}^{-1}P_{n-1}^L(x). \tag{2.19}$$

Proof. (i) We have

$$P_n^R(x) - xP_{n-1}^R(x) = (\zeta_n^R - \zeta_{n-1}^R)x^{n-1} + \text{lower order terms}$$

and so

$$\begin{aligned}
(\zeta_n^L - \zeta_{n-1}^L)\gamma_{n-1} &= (\zeta_n^R - \zeta_{n-1}^R)^\dagger \langle\!\langle P_{n-1}^R, P_{n-1}^R \rangle\!\rangle_R \\
&= (\zeta_n^R - \zeta_{n-1}^R)^\dagger \langle\!\langle x^{n-1}, P_{n-1}^R \rangle\!\rangle_R \\
&= \langle\!\langle P_n^R - xP_{n-1}^R, P_{n-1}^R \rangle\!\rangle_R \\
&= \langle\!\langle P_n^R, P_{n-1}^R \rangle\!\rangle_R - \langle\!\langle xP_{n-1}^R, P_{n-1}^R \rangle\!\rangle_R \\
&= -\langle\!\langle xP_{n-1}^R, P_{n-1}^R \rangle\!\rangle_R \\
&= -\langle\!\langle P_{n-1}^R, xP_{n-1}^R \rangle\!\rangle_R \\
&= \langle\!\langle P_{n-1}^R, P_n^R - xP_{n-1}^R \rangle\!\rangle_R \\
&= \langle\!\langle P_{n-1}^R, x^{n-1}(\zeta_n^R - \zeta_{n-1}^R) \rangle\!\rangle_R \\
&= \langle\!\langle P_{n-1}^R, x^{n-1} \rangle\!\rangle_R (\zeta_n^R - \zeta_{n-1}^R) \\
&= \gamma_{n-1}(\zeta_n^R - \zeta_{n-1}^R).
\end{aligned}$$

(ii) By Theorem 2.5,

$$xP_n^R(x) = P_{n+1}^R(x)C_{n+1} + P_n^R(x)C_n + P_{n-1}^R(x)C_{n-1} + \cdots + P_0^R C_0$$

with some matrices $C_0, \dots, C_{n+1}$. It is straightforward that $C_{n+1} = \mathbf{1}$ and $C_n = \zeta_n^R - \zeta_{n+1}^R$. By the orthogonality property (2.4), we find $C_0 = \cdots = C_{n-2} = \mathbf{0}$. Finally, it is easy to calculate C_{n-1}:

$$\begin{aligned}
\gamma_n &= \langle\!\langle P_n^R, xP_{n-1}^R \rangle\!\rangle_R = \langle\!\langle xP_n^R, P_{n-1}^R \rangle\!\rangle_R \\
&= \langle\!\langle P_{n+1}^R + P_n^R(\zeta_n^R - \zeta_{n+1}^R) + P_{n-1}^R C_{n-1}, P_{n-1}^R \rangle\!\rangle_R \\
&= C_{n-1}^\dagger \gamma_{n-1}
\end{aligned}$$

and so, taking adjoints and using self-adjointness of γ_j, $C_{n-1} = \gamma_{n-1}^{-1}\gamma_n$. This proves (2.18); the other relation (2.19) is obtained by conjugation. □

2.1.5 Normalized Orthogonal Polynomials

We call $p_n^R \in \mathcal{P}$ a right orthonormal polynomial if $\deg p_n^R \le n$ and

$$\langle\!\langle p_n^R, f \rangle\!\rangle_R = \mathbf{0} \text{ for every } f \in \mathcal{P} \text{ with } \deg f < n, \tag{2.20}$$

$$\langle\!\langle p_n^R, p_n^R \rangle\!\rangle_R = \mathbf{1}. \tag{2.21}$$

Similarly, we call $p_n^L \in \mathcal{P}$ a left orthonormal polynomial if $\deg p_n^L \le n$ and

$$\langle\!\langle p_n^L, f \rangle\!\rangle_L = \mathbf{0} \text{ for every } f \in \mathcal{P} \text{ with } \deg f < n, \tag{2.22}$$

$$\langle\!\langle p_n^L, p_n^L \rangle\!\rangle_L = \mathbf{1}. \tag{2.23}$$

Lemma 2.7. *Any orthonormal polynomial has the form*

$$p_n^R(x) = P_n^R(x)\langle\!\langle P_n^R, P_n^R\rangle\!\rangle_R^{-1/2}\sigma_n, \qquad p_n^L(x) = \tau_n\langle\!\langle P_n^L, P_n^L\rangle\!\rangle_L^{-1/2}P_n^L(x) \tag{2.24}$$

where $\sigma_n, \tau_n \in \mathcal{M}_l$ are unitaries. In particular, $\deg p_n^R = \deg p_n^L = n$.

Proof. Let K_n be the coefficient of x^n in p_n^R. Consider the polynomial $q(x) = P_n^R(x)K_n - p_n^R(x)$, where P_n^R is the monic orthogonal polynomial from Lemma 2.2. Then $\deg q < n$ and so from (2.4) and (2.20), it follows that $\langle\!\langle q, q\rangle\!\rangle_R = 0$ and so $q(x)$ vanishes identically. Thus, we have

$$\mathbf{1} = \langle\!\langle p_n^R, p_n^R\rangle\!\rangle_R = K_n^\dagger\langle\!\langle P_n^R, P_n^R\rangle\!\rangle_R K_n \tag{2.25}$$

and so $\det(K_n) \neq 0$. From (2.25) we get $(K_n^\dagger)^{-1}K_n^{-1} = \langle\!\langle P_n^R, P_n^R\rangle\!\rangle_R$, and so $K_nK_n^\dagger = \langle\!\langle P_n^R, P_n^R\rangle\!\rangle_R^{-1}$. From here we get $K_n = \langle\!\langle P_n^R, P_n^R\rangle\!\rangle_R^{-1/2}\sigma_n$ with a unitary σ_n. The proof for p_n^L is similar. $\square$

By Theorem 2.5, the polynomials p_n^R form a right orthonormal module basis in $\mathcal{P}_R$. Thus, for any $f \in \mathcal{P}_R$, we have

$$f(x) = \sum_{m=0}^\infty p_m^R f_m, \qquad f_m = \langle\!\langle p_m^R, f\rangle\!\rangle_R \tag{2.26}$$

and the Parseval identity

$$\sum_{m=0}^\infty \operatorname{Tr}(f_m f_m^\dagger) = \|f\|_R^2 \tag{2.27}$$

holds true. Obviously, since f is a polynomial, there are only finitely many non-zero terms in (2.26) and (2.27).

2.2 Block Jacobi Matrices

The study of block Jacobi matrices goes back at least to Krein [133].

2.2.1 Block Jacobi Matrices as Matrix Representations

Suppose that a sequence of unitary matrices $\mathbf{1} = \sigma_0, \sigma_1, \sigma_2, \dots$ is fixed, and p_n^R are defined according to (2.24). As noted above, p_n^R form a right orthonormal basis in $\mathcal{P}_R$.

The map $f(x) \mapsto xf(x)$ can be considered as a right homomorphism in $\mathcal{P}_R$. Consider the matrix J_{nm} of this homomorphism with respect to the basis p_n^R, that is,

$$J_{nm} = \langle\!\langle p_{n-1}^R, xp_{m-1}^R\rangle\!\rangle_R. \tag{2.28}$$

Following Killip–Simon [128] and Simon [167, 168, 176], our Jacobi matrices are indexed with $n = 1, 2, \dots$ but, of course, p_n has $n = 0, 1, 2, \dots$. That is why (2.28) has $n-1$ and $m-1$.

As in the scalar case, using the orthogonality properties of p_n^R, we get that $J_{nm} = \mathbf{0}$ if $|n-m| > 1$. Denote

$$B_n = J_{nn} = \langle\!\langle p_{n-1}^R, xp_{n-1}^R\rangle\!\rangle_R$$

and

$$A_n = J_{n,n+1} = J_{n+1,n}^\dagger = \langle\!\langle p_{n-1}^R, xp_n^R\rangle\!\rangle_R.$$

Then we have

$$J = \begin{pmatrix} B_1 & A_1 & \mathbf{0} & \cdots \\ A_1^\dagger & B_2 & A_2 & \cdots \\ \mathbf{0} & A_2^\dagger & B_3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{2.29}$$

Applying (2.26) to $f(x) = xp_n^R(x)$, we get the recurrence relation

$$xp_n^R(x) = p_{n+1}^R(x)A_{n+1}^\dagger + p_n^R(x)B_{n+1} + p_{n-1}^R(x)A_n, \quad n = 1, 2, \dots. \tag{2.30}$$

If we set $p_{-1}^R(x) = \mathbf{0}$ and $A_0 = \mathbf{1}$, the relation (2.30) also holds for $n = 0$. By (2.2), we can always pick p_n^L so that for x real, $p_n^L(x) = p_n^R(x)^\dagger$, and thus for complex z,

$$p_n^L(z) = p_n^R(\bar{z})^\dagger \tag{2.31}$$

by analytic continuation. By conjugating (2.30), we get

$$xp_n^L(x) = A_{n+1}p_{n+1}^L(x) + B_{n+1}p_n^L(x) + A_n^\dagger p_{n-1}^L(x), \quad n = 0, 1, 2, \dots. \tag{2.32}$$

Comparing this with the recurrence relations (2.18), (2.19), we get

$$A_n = \sigma_{n-1}^\dagger \gamma_{n-1}^{-1/2}\gamma_n^{1/2}\sigma_n, \qquad B_n = \sigma_{n-1}^\dagger \gamma_{n-1}^{1/2}(\zeta_{n-1}^R - \zeta_n^R)\gamma_{n-1}^{-1/2}\sigma_{n-1}. \tag{2.33}$$

In particular, $\det A_n \neq 0$ for all n.

Notice that since σ_n is unitary, $|\det(\sigma_n)| = 1$, so (2.33) implies $\det(\gamma_n^{1/2}) = \det(\gamma_{n-1}^{1/2})|\det(A_n)|$ which, by induction, implies that

$$\det\langle\!\langle P_n^R, P_n^R \rangle\!\rangle = |\det(A_1 \cdots A_n)|^2. \tag{2.34}$$

Any block matrix of the form (2.29) with $B_n = B_n^\dagger$ and $\det A_n \neq 0$ for all n will be called a block Jacobi matrix corresponding to the Jacobi parameters A_n and B_n.

2.2.2 Basic Properties of Block Jacobi Matrices

Suppose we are given a block Jacobi matrix J corresponding to Jacobi parameters A_n and B_n, where $B_n = B_n^\dagger$ and $\det A_n \neq 0$ for each n.

Consider the Hilbert space $\mathcal{H}_v = \ell^2(\mathbb{Z}_+, \mathbb{C}^l)$ (here $\mathbb{Z}_+ = \{1, 2, 3, \dots\}$) with inner product

$$\langle f, g \rangle_{\mathcal{H}_v} = \sum_{n=1}^{\infty} \langle f_n, g_n \rangle_{\mathbb{C}^l}$$

and orthonormal basis $\{e_{k,j}\}_{k \in \mathbb{Z}_+, 1 \le j \le l}$, where

$$(e_{k,j})_n = \delta_{k,n} v_j$$

and $\{v_j\}_{1 \le j \le l}$ is the standard basis of $\mathbb{C}^l$. J acts on $\mathcal{H}_v$ via

$$(Jf)_n = A_{n-1}^\dagger f_{n-1} + B_n f_n + A_n f_{n+1}, \quad f \in \mathcal{H}_v \tag{2.35}$$

(with $f_0 = 0$) and defines a symmetric operator on this space. Note that using invertibility of the A_n's, induction shows

$$\operatorname{span}\{e_{k,j} : 1 \le k \le n,\, 1 \le j \le l\} = \operatorname{span}\{J^{k-1}e_{1,j} : 1 \le k \le n,\, 1 \le j \le l\} \tag{2.36}$$

for every $n \ge 1$. We want to emphasize that elements of $\mathcal{H}_v$ and $\mathcal{H}$ are vector-valued and matrix-valued, respectively. For this reason, we will be interested in both matrix- and vector-valued solutions of the basic difference equations.

We will consider only bounded block Jacobi matrices, that is, those corresponding to Jacobi parameters satisfying

$$\sup_n \operatorname{Tr}(A_n^\dagger A_n + B_n^\dagger B_n) < \infty. \tag{2.37}$$

Equivalently,

$$\sup_n (\|A_n\| + \|B_n\|) < \infty. \tag{2.38}$$

In this case, J is a bounded self-adjoint operator. This is equivalent to μ having compact support.

We call two Jacobi matrices J and $\tilde J$ equivalent if there exists a sequence of unitaries $u_n \in \mathcal{M}_l$, $n \ge 1$, with $u_1 = \mathbf{1}$ such that $\tilde J_{nm} = u_n^\dagger J_{nm} u_m$. From Lemma 2.7 it is clear that if p_n^R, $\tilde p_n^R$ are two sequences of normalized orthogonal polynomials, corresponding to the same measure (but having different normalization), then the Jacobi matrices $J_{nm} = \langle\langle p_{n-1}^R, x p_{m-1}^R \rangle\rangle_R$ and $\tilde J_{nm} = \langle\langle \tilde p_{n-1}^R, x \tilde p_{m-1}^R \rangle\rangle_R$ are equivalent ($u_n = \sigma_{n-1}^\dagger \tilde\sigma_{n-1}$). Thus,

$$\tilde B_n = u_n^\dagger B_n u_n, \qquad \tilde A_n = u_n^\dagger A_n u_{n+1}. \tag{2.39}$$

Therefore, we have a map

$$\begin{aligned}\Phi : \mu \mapsto \{J : J_{mn} = \langle\langle p_{n-1}^R, x p_{m-1}^R \rangle\rangle_R,\ & p_n^R \\ & \text{correspond to } d\mu \text{ for some normalization}\}\end{aligned} \tag{2.40}$$

from the set of all Hermitian positive semi-definite non-trivial compactly supported measures to the set of all equivalence classes of bounded block Jacobi matrices. Below, we will see how to invert this map.

2.2.3 Special Representatives of the Equivalence Classes

Let J be a block Jacobi matrix with the Jacobi parameters A_n, B_n. We say that J is:

- of type 1, if $A_n > \mathbf{0}$ for all n;
- of type 2, if $A_1 A_2 \cdots A_n > \mathbf{0}$ for all n;
- of type 3, if $A_n \in \mathcal{L}$ for all n.

Here, $\mathcal{L}$ is the class of all lower triangular matrices with strictly positive elements on the diagonal. Type 3 is of interest because they correspond precisely to bounded Hermitian matrices with $2l + 1$ non-vanishing diagonals with the extreme diagonals strictly positive; see Section 1.4(c). Type 2 is the case where the leading coefficients of p_n^R are strictly positive definite.

Theorem 2.8. (i) *Each equivalence class of block Jacobi matrices contains exactly one element each of type 1, type 2, or type 3.*

(ii) *Let J be a block Jacobi matrix corresponding to a sequence of polynomials p_n^R as in* (2.24). *Then J is of type 2 if and only if $\sigma_n = \mathbf{1}$ for all n.*

Proof. The proof is based on the following two well-known facts:

(a) For any $t \in \mathcal{M}_l$ with $\det(t) \neq 0$, there exists a unique unitary $u \in \mathcal{M}_l$ such that tu is Hermitian positive semi-definite: $tu \geq 0$.

(b) For any $t \in \mathcal{M}_l$ with $\det(t) \neq 0$, there exists a unique unitary $u \in \mathcal{M}_l$ such that $tu \in \mathcal{L}$.

We first prove that every equivalence class of block Jacobi matrices contains at least one element of type 1. For a given sequence A_n, let us construct a sequence $u_1 = \mathbf{1}, u_2, u_3, \dots$ of unitaries such that $u_n^\dagger A_n u_{n+1} \geq \mathbf{0}$. By the existence part of (a), we find u_2 such that $A_1 u_2 \geq \mathbf{0}$, then find u_3 such that $u_2^\dagger A_2 u_3 \geq \mathbf{0}$, etc. This, together with (2.39), proves the statement. In order to prove the uniqueness part, suppose we have $A_n \geq \mathbf{0}$ and $u_n^\dagger A_n u_{n+1} \geq \mathbf{0}$ for all n. Then, by the uniqueness part of (a), $A_1 \geq \mathbf{0}$ and $A_1 u_2 \geq \mathbf{0}$ imply $u_2 = \mathbf{1}$; next, $A_2 \geq \mathbf{0}$ and $u_2^\dagger A_2 u_3 = A_2 u_3 \geq \mathbf{0}$ imply $u_3 = \mathbf{1}$, etc.

The statement (i) concerning type 3 can be proven in the same way, using (b) instead of (a).

The statement (i) concerning type 2 can be proven similarly. Existence: find u_2 such that $A_1 u_2 \geq \mathbf{0}$, then u_3 such that $(A_1 u_2)(u_2^\dagger A_2 u_3) = A_1 A_2 u_3 \geq \mathbf{0}$, etc. Uniqueness: if $A_1 \dots A_n \geq \mathbf{0}$ and $A_1 \cdots A_n u_{n+1} \geq \mathbf{0}$, then $u_{n+1} = \mathbf{1}$.

By (2.33), we have $A_1 A_2 \cdots A_n = \gamma_n^{1/2} \sigma_n$ and the statement (ii) follows from the positivity of γ_n. ☐

We say that a block Jacobi matrix J belongs to the Nevai class if

$$B_n \to \mathbf{0} \text{ and } A_n^\dagger A_n \to \mathbf{1} \text{ as } n \to \infty.$$

It is clear that J is in the Nevai class if and only if all equivalent Jacobi matrices belong to the Nevai class.

Theorem 2.9. *If J belongs to the Nevai class and is of type 1 or type 3, then $A_n \to 1$ as $n \to \infty$.*

Proof. If J is of type 1, then $A_n^\dagger A_n = A_n^2 \to \mathbf{1}$ clearly implies $A_n \to \mathbf{1}$ since square root is continuous on positive Hermitian matrices.

Suppose J is of type 3. We shall prove that $A_n \to \mathbf{1}$ by considering the rows of the matrix A_n one by one, starting from the lth row. Denote $(A_n)_{jk} = a_{j,k}^{(n)}$. We have

$$(A_n^\dagger A_n)_{ll} = (a_{l,l}^{(n)})^2 \to 1, \text{ and so } a_{l,l}^{(n)} \to 1.$$

Then, for any $k < l$, we have

$$(A_n^\dagger A_n)_{lk} = a_{l,l}^{(n)} a_{l,k}^{(n)} \to 0, \text{ and so } a_{l,k}^{(n)} \to 0.$$

Next, consider the $(l-1)$st row. We have

$$(A_n^\dagger A_n)_{l-1,l-1} = (a_{l-1,l-1}^{(n)})^2 + |a_{l,l-1}^{(n)}|^2 \to 1$$

and so, using the previous step, $a_{l-1,l-1}^{(n)} \to 1$ as $n \to \infty$. Then for all $k < l-1$, we have

$$(A_n^\dagger A_n)_{l-1,k} = \overline{a_{l-1,l-1}^{(n)}}\, a_{l-1,k}^{(n)} + \overline{a_{l,l-1}^{(n)}}\, a_{l,k}^{(n)} \to 0$$

and so, using the previous steps, $a_{l-1,k} \to 0$. Continuing this way, we get $a_{j,k}^{(n)} \to \delta_{j,k}$ as required. ☐

It is an interesting open question if this result also applies to the type 2 case.

2.2.4 Favard's Theorem

Here we construct an inverse of the mapping Φ (defined by (2.40)). Thus, Φ sets up a bijection between non-trivial measures of compact support and equivalence classes of bounded block Jacobi matrices.

Before proceeding to do this, let us prove:

Lemma 2.10. *The mapping Φ is injective.*

Proof. Let μ and $\tilde{\mu}$ be two Hermitian positive semi-definite non-trivial compactly supported measures. Suppose that $\Phi(\mu) = \Phi(\tilde{\mu})$.

Let p_n^R and $\tilde{p}_n^R$ be normalized orthogonal polynomials corresponding to μ and $\tilde{\mu}$. Suppose that the normalization both for p_n^R and for $\tilde{p}_n^R$ has been chosen such that $\sigma_n = \mathbf{1}$ (see (2.24)), that is, type 2. From Lemma 2.8 and the assumption $\Phi(\mu) = \Phi(\tilde{\mu})$ it follows that the corresponding Jacobi matrices coincide, that is, $\langle\!\langle p_n^R, x p_m^R \rangle\!\rangle_R = \langle\!\langle \tilde{p}_n^R, x \tilde{p}_m^R \rangle\!\rangle_R$ for all n and m. Together with the recurrence relation (2.30) this yields $p_n^R = \tilde{p}_n^R$ for all n.

For any $n \geq 0$, we can represent x^n as

$$x^n = \sum_{k=0}^{n} p_k^R(x) C_k^{(n)} = \sum_{k=0}^{n} \tilde{p}_k^R(x) \tilde{C}_k^{(n)}.$$

The coefficients $C_k^{(n)}$ and $\tilde{C}_k^{(n)}$ are completely determined by the coefficients of the polynomials p_n^R and $\tilde{p}_n^R$ and so $C_k^{(n)} = \tilde{C}_k^{(n)}$ for all n and k.

For the moments of the measure μ, we have

$$\int x^n d\mu(x) = \langle\!\langle \mathbf{1}, x^n \rangle\!\rangle_R = \sum_{k=0}^{n} \langle\!\langle \mathbf{1}, p_k^R C_k^{(n)} \rangle\!\rangle_R = \langle\!\langle \mathbf{1}, \mathbf{1} \rangle\!\rangle_R \, C_0^{(n)} = C_0^{(n)}.$$

Since the same calculation is valid for the measure $\tilde{\mu}$, we get

$$\int x^n d\mu(x) = \int x^n d\tilde{\mu}(x)$$

for all n. It follows that

$$\int f(x) d\mu(x) g(x) = \int f(x) d\tilde{\mu}(x) g(x)$$

for all matrix-valued polynomials f and g, and so the measures μ and $\tilde{\mu}$ coincide. $\square$

We can now construct the inverse of the map Φ. Let a block Jacobi matrix J be given. By a version of the spectral theorem for self-adjoint operators with finite multiplicity (see, e.g., [2, Sect. 72]), there exists a matrix-valued measure $d\mu$ with

$$\langle e_{1,j}, f(J) e_{1,k} \rangle_{\mathcal{H}_v} = \int f(x) \, d\mu_{j,k}(x) \tag{2.41}$$

and an isometry

$$R : \mathcal{H}_v \to L^2(\mathbb{R}, d\mu; \mathbb{C}^l)$$

such that (recall that $\{v_j\}$ is the standard basis in $\mathbb{C}^l$)

$$[R e_{1,j}](x) = v_j, \quad 1 \leq j \leq l, \tag{2.42}$$

and, for any $g \in \mathcal{H}_v$, we have

$$(RJg)(x) = x(Rg)(x). \tag{2.43}$$

If the Jacobi matrices J and $\tilde{J}$ are equivalent, then we have $\tilde{J} = U^* J U$ for some $U = \oplus_{n=1}^{\infty} u_n$, $u_1 = \mathbf{1}$. Thus,

$$\langle e_{1,j}, f(\tilde{J}) e_{1,k} \rangle_{\mathcal{H}_v} = \langle U e_{1,j}, f(J) U e_{1,k} \rangle_{\mathcal{H}_v} = \langle e_{1,j}, f(J) e_{1,k} \rangle_{\mathcal{H}_v}$$

and so the measures corresponding to J and $\tilde{J}$ coincide. Thus, we have a map

$$\Psi \colon \{\tilde{J} \colon \tilde{J} \text{ is equivalent to } J\} \mapsto \mu \tag{2.44}$$

from the set of all equivalence classes of bounded block Jacobi matrices to the set of all Hermitian positive semi-definite compactly supported measures.

Theorem 2.11. (i) *All measures in the image of the map Ψ are non- degenerate.*
(ii) $\Phi \circ \Psi = id.$
(iii) $\Psi \circ \Phi = id.$

Proof. (i) To put things in the right context, we first recall that $\|\cdot\|_{\mathcal{H}_v}$ is a norm (rather than a semi-norm), whereas $|\!|\cdot|\!|$ on $\mathcal{V}$ (cf. (2.3)) is, in general, a semi-norm. Using the assumption that $\det(A_k) \neq 0$ for all k (which is included in our definition of a Jacobi matrix), we will prove that $|\!|\cdot|\!|$ is in fact a norm. More precisely, we will prove that $|\!|p|\!| > 0$ for any non-zero polynomial $p \in \mathcal{V}$; by Lemma 2.1 this will imply that μ is non-degenerate.

Let $p \in \mathcal{V}$ be a non-zero polynomial, $\deg p = n$. Notice that (2.42) and (2.43) give

$$[RJ^k e_{1,j}](x) = x^k v_j \tag{2.45}$$

for every $k \geq 0$ and $1 \leq j \leq l$. This shows that p can be represented as $p = Rg$, where $g = \sum_{k=0}^{n} J^k f_k$, and $f_0, \dots, f_n$ are vectors in $\mathcal{H}_v$ such that $\langle f_i, e_{j,k} \rangle_{\mathcal{H}_v} = 0$ for all $i = 0, \dots, n$, $j \geq 2$, $k = 1, \dots, l$ (i.e., the only non-zero components of f_j are in the first $\mathbb{C}^l$ in $\mathcal{H}_v$). Assumption $\deg p = n$ means $f_n \neq 0$.

Since R is isometric, we have $|\!|p|\!| = \|g\|_{\mathcal{H}_v}$, and so we have to prove that $g \neq 0$. Indeed, suppose that $g = 0$. Using the assumption $\det(A_k) \neq 0$ and the tri-diagonal nature of J, we see that $\sum_{k=0}^{n} J^k f_k = 0$ yields $f_n = 0$, contrary to our assumption.

(ii) Consider the elements $Re_{n,k} \in L^2(\mathbb{R}, d\mu; \mathbb{C}^l)$. First note that, by (2.36) and (2.45), $Re_{n,k}$ is a polynomial of degree at most $n-1$. Next, by the unitarity of R, we have

$$\langle Re_{n,k}, Re_{m,j} \rangle_{L^2(\mathbb{R}, d\mu; \mathbb{C}^l)} = \delta_{m,n} \delta_{k,j}. \tag{2.46}$$

Let us construct matrix-valued polynomials $q_n(x)$, using $Re_{n,1}, Re_{n,2}, \dots, Re_{n,l}$ as columns of $q_{n-1}(x)$:

$$[q_{n-1}(x)]_{j,k} = [Re_{n,k}(x)]_j.$$

We have $\deg q_n \leq n$ and $\langle\!\langle q_m, q_n \rangle\!\rangle_R = \delta_{m,n} \mathbf{1}$; the last relation is just a reformulation of (2.46). Hence the q_n's are right normalized orthogonal polynomials with respect to the measure $d\mu$. We find

$$\begin{aligned} J_{nm} &= [\langle e_{n,j}, J e_{m,k} \rangle_{\mathcal{H}_v}]_{1 \leq j,k \leq l} \\ &= [\langle Re_{n,j}, RJe_{m,k} \rangle_{L^2(\mathbb{R}, d\mu; \mathbb{C}^l)}]_{1 \leq j,k \leq l} \end{aligned}$$

$$\begin{aligned}
&= [\langle Re_{n,j}, xRe_{m,k}\rangle_{L^2(\mathbb{R},d\mu;\mathbb{C}^l)}]_{1\le j,k\le l} \\
&= [\langle [q_{n-1}(x)]_{\cdot,j}, x[q_{m-1}(x)]_{\cdot,k}\rangle_{L^2(\mathbb{R},d\mu;\mathbb{C}^l)}]_{1\le j,k\le l} \\
&= \langle\!\langle q_{n-1}, xq_{m-1}\rangle\!\rangle_R
\end{aligned}$$

as required.

(iii) Follows from (ii) and from Lemma 2.10. □

2.3 The m-Function

2.3.1 The Definition of the m-Function

We denote the Borel transform of $d\mu$ by m:

$$m(z) = \int \frac{d\mu(x)}{x-z}, \qquad \operatorname{Im} z > 0. \tag{2.47}$$

It is a matrix-valued Herglotz function, that is, it is analytic and obeys $\operatorname{Im} m(z) > 0$. For information on matrix-valued Herglotz functions, see [98] and references therein. Extensions to operator-valued Herglotz functions can be found in [94].

Lemma 2.12. *Suppose $d\mu$ is given, p_n^R are right normalized orthogonal polynomials, and J is the associated block Jacobi matrix. Then,*

$$m(z) = \langle\!\langle p_0^R, (x-z)^{-1} p_0^R\rangle\!\rangle_R \tag{2.48}$$

and

$$m(z) = \langle e_{1,\cdot}, (J-z)^{-1} e_{1,\cdot}\rangle_{\mathcal{H}_v}. \tag{2.49}$$

Proof. Since $p_0^R = \mathbf{1}$, (2.48) is just a way of rewriting the definition of m. The second identity, (2.49), is a consequence of (2.41) and Theorem 2.11(iii). □

2.3.2 Coefficient Stripping

If J is a block Jacobi matrix corresponding to invertible A_n's and Hermitian B_n's, we denote the k-times stripped block Jacobi matrix, corresponding to $\{A_{k+n}, B_{k+n}\}_{n\ge 1}$, by $J^{(k)}$. That is,

$$J^{(k)} = \begin{pmatrix} B_{k+1} & A_{k+1} & \mathbf{0} & \cdots \\ A_{k+1}^\dagger & B_{k+2} & A_{k+2} & \cdots \\ \mathbf{0} & A_{k+2}^\dagger & B_{k+3} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

The m-function corresponding to $J^{(k)}$ will be denoted by $m^{(k)}$. Note that, in particular, $J^{(0)} = J$ and $m^{(0)} = m$.

Proposition 2.13. *Let J be a block Jacobi matrix with $\sigma_{\mathrm{ess}}(J) \subseteq [a,b]$. Then, for every $\varepsilon > 0$, there is $k_0 \ge 0$ such that for $k \ge k_0$, we have that $\sigma(J^{(k)}) \subseteq [a-\varepsilon, b+\varepsilon]$.*

Proof. This is an immediate consequence of (the proof of) [42, Lemma 1]. □

Proposition 2.14 (Due to Aptekarev–Nikishin [4]). *We have that*

$$m^{(k)}(z)^{-1} = B_{k+1} - z - A_{k+1} m^{(k+1)}(z) A_{k+1}^{\dagger}$$

for $\operatorname{Im} z > 0$ *and* $k \geq 0$.

Proof. It suffices to handle the case $k = 0$. Given (2.49), this is a special case of a general formula for 2×2 block operator matrices, due to Schur [163], that reads

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} (A - BD^{-1}C)^{-1} & -A^{-1}B(D - CA^{-1}B)^{-1} \\ -D^{-1}C(A - BD^{-1}C)^{-1} & (D - CA^{-1}B)^{-1} \end{pmatrix}$$

and which is readily verified. Here $A = B_1 - z$, $B = A_1$, $C = A_1^{\dagger}$, and $D = J^{(1)} - z$. □

2.4 Second Kind Polynomials

Define the second kind polynomials by $q_{-1}^R(z) = -\mathbf{1}$,

$$q_n^R(z) = \int_{\mathbb{R}} d\mu(x) \frac{p_n^R(z) - p_n^R(x)}{z - x}, \quad n = 0, 1, 2, \ldots.$$

As is easy to see, for $n \geq 1$, q_n^R is a polynomial of degree $n - 1$. For future reference, let us display the first several polynomials p_n^R and q_n^R:

$$p_{-1}^R(x) = \mathbf{0}, \qquad p_0^R(x) = \mathbf{1}, \quad p_1^R(x) = (x - B_1)A_1^{-1}, \tag{2.50}$$

$$q_{-1}^R(x) = -\mathbf{1}, \quad q_0^R(x) = \mathbf{0}, \quad q_1^R(x) = A_1^{-1}. \tag{2.51}$$

The polynomials q_n^R satisfy the equation (same form as (2.30))

$$x q_n^R(x) = q_{n+1}^R(x) A_{n+1}^{\dagger} + q_n^R(x) B_{n+1} + q_{n-1}^R(x) A_n, \quad n = 0, 1, 2, \ldots. \tag{2.52}$$

For $n = 0$, this can be checked by a direct substitution of (2.51). For $n \geq 1$, as in the scalar case, this can be checked by taking (2.30) for x and for z, subtracting, dividing by $x - z$, integrating over $d\mu$, and taking into account the orthogonality relation

$$\int d\mu(x) p_n^R(x) = \mathbf{0}, \quad n \geq 1.$$

Finally, let us define

$$\psi_n^R(z) = q_n^R(z) + m(z) p_n^R(z).$$

According to the definition of q_n^R, we have

$$\psi_n^R(z) = \langle\langle f_z, p_n^R \rangle\rangle_R, \quad f_z(x) = (x - \bar{z})^{-1}.$$

By the Parseval identity, this shows that for all $\operatorname{Im} z > 0$, the sequence $\psi_n^R(z)$ is in ℓ^2, that is,

$$\sum_{n=0}^{\infty} \operatorname{Tr}(\psi_n^R(z)^{\dagger} \psi_n^R(z)) < \infty. \tag{2.53}$$

In the same way, we define $q_{-1}^L(z) = -\mathbf{1}$,

$$q_n^L(z) = \int_{\mathbb{R}} \frac{p_n^L(z) - p_n^L(x)}{z - x} d\mu(x), \quad n = 0, 1, 2, \ldots$$

and $\psi_n^L(z) = q_n^L(z) + p_n^L(z) m(z)$.

2.5 Solutions to the Difference Equations

For $\operatorname{Im} z > 0$, consider the solutions to the equations

$$zu_n(z) = \sum_{m=1}^{\infty} u_m(z) J_{mn}, \quad n = 2, 3, \ldots, \tag{2.54}$$

$$zv_n(z) = \sum_{m=1}^{\infty} J_{nm} v_m(z), \quad n = 2, 3, \ldots. \tag{2.55}$$

Clearly, $u_n(z)$ solves (2.54) if and only if $v_n(z) = (u_n(\bar{z}))^\dagger$ solves (2.55). In the above, we normally intend z as a fixed parameter but it then can be used as a variable. That is, z is fixed and $u_n(z)$ is a fixed sequence, not a z-dependent function. A statement like $v_n(z) = (u_n(\bar{z}))^\dagger$ means if u_n is a sequence solving (2.54) for $z = \bar{z}_0$, then v_n obeys (2.55) for $z = z_0$. Of course, if $u_n(z)$ is a function of z in a region, we can apply our estimates to all z in the region. For any solution $\{u_n(z)\}_{n=1}^\infty$ of (2.54), let us define

$$u_0(z) = zu_1(z) - u_1(z)B_1 - u_2(z)A_1^\dagger. \tag{2.56}$$

With this definition, the equation (2.54) for $n = 1$ is equivalent to $u_0(z) = 0$. In the same way, we set

$$v_0(z) = zv_1(z) - B_1 v_1(z) - A_1 v_2(z).$$

Lemma 2.15. *Let $\operatorname{Im} z > 0$ and suppose $\{u_n(z)\}_{n=0}^\infty$ solves (2.54) (for $n \geq 2$) and (2.56) and belongs to ℓ^2. Then*

$$(\operatorname{Im} z) \sum_{n=1}^{\infty} \operatorname{Tr}(u_n(z)^\dagger u_n(z)) = -\operatorname{Im} \operatorname{Tr}(u_1(z) u_0(z)^\dagger). \tag{2.57}$$

In particular, $u_n(z) = \alpha p_{n-1}^R(z)$ is in ℓ^2 only if $\alpha = \mathbf{0}$.

Proof. Denote $s_n = \operatorname{Tr}(u_n(z) A_{n-1}^\dagger u_{n-1}(z)^\dagger)$. Here $A_0 = \mathbf{1}$. Multiplying (2.54) for $n \geq 2$ and (2.56) for $n = 1$ by $u_n(z)^\dagger$ on the right, taking traces, and summing over n, we get

$$z \sum_{n=1}^{N} \operatorname{Tr}(u_n(z) u_n(z)^\dagger) = \sum_{n=1}^{N} s_{n+1} + \sum_{n=1}^{N} \operatorname{Tr}(u_n(z) B_{n+1} u_n(z)^\dagger) + \sum_{n=1}^{N} \overline{s_n}\,.$$

Taking imaginary parts and letting $N \to \infty$, we obtain (2.57) since the middle sum is real and the outer sums cancel up to boundary terms. Applying (2.57) to $u_n(z) = \alpha p_{n-1}^R(z)$, we get zero in the right-hand side:

$$(\operatorname{Im} z) \sum_{n=1}^{\infty} \operatorname{Tr}(\alpha p_{n-1}^R(z) p_{n-1}^R(z)^\dagger \alpha^\dagger) = 0$$

and hence $\alpha = \mathbf{0}$ since $p_0^R = \mathbf{1}$. $\square$

Theorem 2.16. *Let $\operatorname{Im} z > 0$.*

(i) *Any solution $\{u_n(z)\}_{n=0}^\infty$ of (2.54) (for $n \geq 2$) can be represented as*

$$u_n(z) = a p_{n-1}^R(z) + b q_{n-1}^R(z) \tag{2.58}$$

for suitable $a, b \in \mathcal{M}_l$. In fact, $a = u_1(z)$ and $b = -u_0(z)$.

(ii) *A sequence* (2.58) *satisfies* (2.54) *for all* $n \geq 1$ *if and only if* $b = 0$.
(iii) *A sequence* (2.58) *belongs to* ℓ^2 *if and only if* $u_n(z) = c\psi_{n-1}^R(z)$ *for some* $c \in \mathcal{M}_l$. *Equivalently, a sequence* (2.58) *belongs to* ℓ^2 *if and only if* $u_1(z) + u_0(z)m(z) = 0$.

Proof. (i) Let $u_n(z)$ be a solution to (2.54). Consider

$$\tilde{u}_n(z) = u_n(z) - u_1(z)p_{n-1}^R(z) + u_0(z)q_{n-1}^R(z).$$

Then $\tilde{u}_n(z)$ also solves (2.54) and $\tilde{u}_0(z) = \tilde{u}_1(z) = 0$. It follows that $\tilde{u}_n(z) = 0$ for all n. This proves (i).

(ii) A direct substitution of (2.58) into (2.54) for $n = 1$ yields the statement.

(iii) We already know that $c\psi_{n-1}^R$ is an ℓ^2 solution. Suppose that $u_n(z)$ is an ℓ^2 solution to (2.54). Rewrite (2.58) as

$$u_n(z) = (a - bm(z))p_{n-1}^R(z) + b\psi_{n-1}^R(z).$$

Since ψ_n^R is in ℓ^2 and cp_n^R is not in ℓ^2, we get $a = bm(z)$, which is equivalent to $u_1(z)+u_0(z)m(z) = 0$ or to $u_n(z) = b\psi_{n-1}^R(z)$. $\square$

By conjugation, we obtain:

Theorem 2.17. *Let* $\operatorname{Im} z > 0$.
(i) *Any solution* $\{v_n(z)\}_{n=0}^\infty$ *of* (2.55) (*for* $n \geq 2$) *can be represented as*

$$v_n(z) = p_{n-1}^L(z)a + q_{n-1}^L(z)b. \tag{2.59}$$

In fact, $a = v_1(z)$ *and* $b = -v_0(z)$.
(ii) *A sequence* (2.59) *satisfies* (2.55) *for all* $n \geq 1$ *if and only if* $b = 0$.
(iii) *A sequence* (2.59) *belongs to* ℓ^2 *if and only if* $v_n(z) = \psi_{n-1}^L(z)c$ *for some* $c \in \mathcal{M}_l$. *Equivalently, a sequence* (2.59) *belongs to* ℓ^2 *if and only if* $v_1(z) + m(z)v_0(z) = 0$.

2.6 Wronskians and the Christoffel–Darboux Formula

For any two $\mathcal{M}_l$-valued sequences u_n, v_n, define the Wronskian by

$$W_n(u, v) = u_n A_n v_{n+1} - u_{n+1} A_n^\dagger v_n. \tag{2.60}$$

Note that $W_n(u, v) = -W_n(v^\dagger, u^\dagger)^\dagger$. If $u_n(z)$ and $v_n(z)$ are solutions to (2.54) and (2.55), then by a direct calculation, we see that $W_n(u(z), v(z))$ is independent of n. Put differently, if both $u_n(z)$ and $v_n(z)$ are solutions to (2.54), then $W_n(u(z), v(\bar{z})^\dagger)$ is independent of n. Or, if both $u_n(z)$ and $v_n(z)$ are solutions to (2.55), then $W_n(u(\bar{z})^\dagger, v(z))$ is independent of n. In particular, by a direct evaluation for $n = 0$, we get

$$\begin{aligned}
W_n(p_{\cdot-1}^R(z), p_{\cdot-1}^R(\bar{z})^\dagger) &= W_n(q_{\cdot-1}^R(z), q_{\cdot-1}^R(\bar{z})^\dagger) = \mathbf{0}, \\
W_n(p_{\cdot-1}^L(\bar{z})^\dagger, p_{\cdot-1}^L(z)) &= W_n(q_{\cdot-1}^L(\bar{z})^\dagger, q_{\cdot-1}^L(z)) = \mathbf{0}, \\
W_n(p_{\cdot-1}^R(z), q_{\cdot-1}^R(\bar{z})^\dagger) &= W_n(p_{\cdot-1}^L(\bar{z})^\dagger, q_{\cdot-1}^L(z)) = \mathbf{1}.
\end{aligned}$$

Let both $u(z)$ and $v(z)$ be solutions to (2.54) of the type (2.58), namely,

$$u_n(z) = ap_{n-1}^R(z) + bq_{n-1}^R(z), \quad v_n(z) = cp_{n-1}^R(z) + dq_{n-1}^R(z).$$

Then the above calculation implies

$$W_n(u(z), v(\bar{z})^\dagger) = ad^\dagger - bc^\dagger.$$

Theorem 2.18 (CD Formula). *For any $x, y \in \mathbb{C}$ and $n \geq 1$, one has*

$$(x-y)\sum_{k=0}^{n} p_k^R(x)p_k^L(y) = -W_{n+1}(p_{\cdot-1}^R(x), p_{\cdot-1}^L(y)). \tag{2.61}$$

Proof. Multiplying (2.30) by $p_n^L(y)$ on the right and (2.32) (with y in place of x) by $p_n^R(x)$ on the left and subtracting, we get

$$(x-y)p_n^R(x)p_n^R(y) = W_n(p_{\cdot-1}^R(x), p_{\cdot-1}^L(y)) - W_{n+1}(p_{\cdot-1}^R(x), p_{\cdot-1}^L(y)).$$

Summing over n and noting that $W_0(p^R(x), p^L(y)) = 0$, we get the required statement. □

2.7 The CD Kernel

The CD kernel is defined for $z, w \in \mathbb{C}$ by

$$K_n(z,w) = \sum_{k=0}^{n} p_k^R(z)p_k^R(\bar{w})^\dagger \tag{2.62}$$

$$= \sum_{k=0}^{n} p_k^L(\bar{z})^\dagger p_k^L(w). \tag{2.63}$$

(2.63) follows from (2.62) and (2.31). Notice that K is independent of the choices σ_n, τ_n in (2.24) and that (2.61) can be written

$$(z-\bar{w})K_n(z,w) = -W_{n+1}(p_{\cdot-1}^R(z), p_{\cdot-1}^R(\bar{w})^\dagger). \tag{2.64}$$

The independence of K_n of σ, τ can be understood by noting that if f_m is given by (2.26), then

$$\int K_n(z,w)\, d\mu(w) f(w) = \sum_{m=0}^{n} p_m^R(z) f_m \tag{2.65}$$

so K is the kernel of the orthogonal projection onto polynomials of degree up to n, and so intrinsic. Similarly, if $f_m^{(L)} = \langle\langle f, p_m^L \rangle\rangle_L$, so

$$f(x) = \sum_{m=0}^{\infty} f_m^{(L)} p_m^L(x), \tag{2.66}$$

then, by (2.63),

$$\int f(z)\, d\mu(z) K_n(z,w) = \sum_{m=0}^{n} f_m^{(L)} p_m^L(w). \tag{2.67}$$

One has

$$\int K_n(z,w)\, d\mu(w) K_n(w,\zeta) = K_n(z,\zeta) \tag{2.68}$$

as can be checked directly and which is an expression of the fact that the map in (2.65) is a projection, and so its own square.

We will let π_n be the map of $L^2(d\mu)$ to itself given by (2.65) (or by (2.67)). (2.64) can then be viewed (for $z, w \in \mathbb{R}$) as an expression of the integral kernel of the commutator $[J, \pi_n]$, which leads to another proof of it [175].

Let $J_{n;F}$ be the finite $nl \times nl$ matrix obtained from J by taking the top leftmost $n2$ blocks. It is the matrix of $\pi_{n-1} M_x \pi_{n-1}$ where M_x is multiplication by x in the $\{p_j^R\}_{j=0}^{n-1}$ basis. For $y \in \mathbb{C}$ and $\gamma \in \mathbb{C}^l$, let $\varphi_{n,\gamma}(y)$ be the vector whose components are

$$(\varphi_{n,\gamma}(y))_j = p_{j-1}^L(y)\gamma \tag{2.69}$$

for $j = 1, 2, \ldots, n$. We claim that

$$[(J_{n;F} - y)\varphi_{n,\gamma}(y)]_j = -\delta_{jn} A_n p_n^L(y)\gamma \tag{2.70}$$

as follows immediately from (2.32).

This is intimately related to (2.61) and (2.64). For recalling J is the matrix in p^R basis, $\varphi_{n,\gamma}(y)$ corresponds to the function

$$\sum_{j=0}^{n-1} p_j^R(x)(\varphi_{n,\gamma}(y))_{j-1} = K_n(x, y)\gamma.$$

As we will see in the next two sections, (2.70) has important consequences.

2.8 Christoffel Variational Principle

There is a matrix version of the Christoffel variational principle (see Nevai [152] for a discussion of uses in the scalar case; this matrix case is discussed by Duran–Polo [76]):

Theorem 2.19. *For any non-trivial $l \times l$ matrix- valued measure, $d\mu$, on $\mathbb{R}$, we have that for any n, any $x_0 \in \mathbb{R}$, and matrix polynomials $Q_n(x)$ of degree at most n with*

$$Q_n(x_0) = \mathbf{1}, \tag{2.71}$$

we have that

$$\langle\!\langle Q_n, Q_n \rangle\!\rangle_R \geq K_n(x_0, x_0)^{-1} \tag{2.72}$$

with equality if and only if

$$Q_n(x) = K_n(x, x_0) K_n(x_0, x_0)^{-1}. \tag{2.73}$$

Remark. (2.72) also holds for $\langle\!\langle \cdot, \cdot \rangle\!\rangle_L$ but the minimizer is then $K_n(x_0, x_0)^{-1} K_n(x, x_0)$.

Proof. Let $Q_n^{(0)}$ denote the right-hand side of (2.73). Then for any polynomial R_n of degree at most n, we have

$$\langle\!\langle Q_n^{(0)}, R_n \rangle\!\rangle_R = K_n(x_0, x_0)^{-1} R_n(x_0) \tag{2.74}$$

because of (2.65). Since $Q_n(x_0) = Q_n^{(0)}(x_0) = \mathbf{1}$, we conclude

$$\langle\!\langle Q_n - Q_n^{(0)}, Q_n - Q_n^{(0)} \rangle\!\rangle_R = \langle\!\langle Q_n, Q_n \rangle\!\rangle_R - K_n(x_0, x_0)^{-1} \tag{2.75}$$

from which (2.72) is immediate and, given the supposed non- triviality, the uniqueness of minimizer. ∎

With this, one easily gets an extension of a result of Máté–Nevai [146] to MOPRL. (They had it for scalar OPUC. For OPUC, it is in Máté–Nevai–Totik [147] on $[-1,1]$ and in Totik [187] for general OPRL. The result below can be proven using polynomial mappings à la Totik [188] or Jost solutions à la Simon [174].)

Theorem 2.20. *Let $d\mu$ be a non-trivial $l \times l$ matrix-valued measure on $\mathbb{R}$ with compact support, E. Let $I = (a,b)$ be an open interval with $I \subset E$. Then for Lebesgue a.e. $x \in I$,*

$$\limsup (n+1) K_n(x,x)^{-1} \leq w(x). \tag{2.76}$$

Remark. This is intended in terms of expectations in any fixed vector.

We state this explicitly since we will need it in Section 5.4, but we note that the more detailed results of Máté–Nevai–Totik [147], Lubinsky [141], Simon [173], and Totik [189] also extend.

2.9 Zeros

We next look at zeros of $\det(P_n^L(z))$, which we will prove soon is also $\det(P_n^R(z))$. Following [66, 177], we will identify these zeros with eigenvalues of $J_{n;F}$. It is not obvious a priori that these zeros are real and, unlike the scalar situation, where the classical arguments on the zeros rely on orthogonality, we do not know how to get reality just from that (but see the remark after Theorem 2.25).

Lemma 2.21. *Let $C(z)$ be an $l \times l$ matrix-valued function analytic near $z = 0$. Let*

$$k = \dim(\ker(C(0))). \tag{2.77}$$

Then $\det(C(z))$ has a zero at $z = 0$ of order at least k.

Remarks. 1. Even in the 1×1 case, where $k = 1$, the zeros can clearly be of higher order than k since $c_{11}(z)$ can have a zero of any order!

2. The temptation to take products of eigenvalues will lead at best to a complicated proof as the cases $C(z) = \left(\begin{smallmatrix} 0 & z \\ 1 & 0 \end{smallmatrix}\right)$ and $C(z) = \left(\begin{smallmatrix} 0 & z^2 \\ 1 & 0 \end{smallmatrix}\right)$ illustrate.

Proof. Let $e_1, \dots, e_l$ be an orthonormal basis with $e_1, \dots, e_k \in \ker(C(0))$. By Hadamard's inequality (see Bhatia [13]),

$$\begin{aligned} |\det(C(z))| &\leq \|C(z)e_1\| \cdots \|C(z)e_l\| \\ &\leq C|z|^k \end{aligned}$$

since $\|C(z)e_j\| \leq C|z|$ if $j = 1, \dots, k$ and $\|C(z)e_j\| \leq d$ for $j = k+1, \dots, l$. □

The following goes back at least to [34]; see also [66, 165, 177, 178].

Theorem 2.22. *We have that*

$$\det\nolimits_{\mathbb{C}^l}(P_n^L(z)) = \det\nolimits_{\mathbb{C}^{nl}}(z - J_{n;F}). \tag{2.78}$$

Proof. By (2.70), if γ is such that $p_n^L(y)\gamma = 0$, then $\varphi_{n,\gamma}(y)$ is an eigenvector for $J_{n;F}$ with eigenvalue y. Conversely, if φ is an eigenvector and γ is defined as that vector in $\mathbb{C}^l$ whose components are the first l components of φ, then a simple inductive argument shows $\varphi = \varphi_{n,\gamma}(y)$ and then, by (2.70) and the fact that A_n is invertible, we see that $p_n^L(y)\gamma = 0$. This shows that for any y,

$$\dim\ker(P_n^L(y)) = \dim\ker(J_{n;F} - y). \tag{2.79}$$

By Lemma 2.21, if y is any eigenvalue of $J_{n;F}$ of multiplicity k, then $\det(P_n^L(z))$ has a zero of order at least k at y. Now let us consider the polynomials in z on the left and right in (2.78). Since $J_{n;F}$ is Hermitian, the sum of the multiplicities of the zeros on the right is nl. Since the polynomial on the left is of degree nl, by a counting argument it has the same set of zeros with the same multiplicities as the polynomial on the right. Since both polynomials are monic, they are equal. □

Corollary 2.23. *All the zeros of* $\det(P_n^L(z))$ *are real. Moreover,*

$$\det(P_n^R(z)) = \det(P_n^L(z)). \tag{2.80}$$

Proof. Since $J_{n;F}$ is Hermitian, all its eigenvalues are real, so (2.78) implies the zeros of $\det(P_n^L(z))$ are real. Thus, since the polynomial is monic,

$$\overline{\det(P_n^L(\bar z))} = \det(P_n^L(z)). \tag{2.81}$$

By Lemma 2.2(v), we have

$$P_n^R(z) = P_n^L(\bar z)^\dagger \tag{2.82}$$

since both sides are analytic and agree if z is real. Thus,

$$\det(P_n^R(z)) = \det(P_n^L(\bar z)^\dagger) = \overline{\det(P_n^L(\bar z))}$$

proving (2.80). □

The following appeared before in [178]; see also [165].

Corollary 2.24. *Let* $\{x_{n,j}\}_{j=1}^{nl}$ *be the zeros of* $\det(P_n^L(x))$ *counting multiplicity ordered by*

$$x_{n,1} \le x_{n,2} \le \cdots \le x_{n,nl}. \tag{2.83}$$

Then

$$x_{n+1,j} \le x_{n,j} \le x_{n+1,j+l}. \tag{2.84}$$

Remarks. 1. This is interlacing if $l = 1$ and replaces it for general l.

2. Using A_n invertible, one can show the inequalities in (2.84) are strict.

Proof. The min-max principle [159] says that

$$x_{n,j} = \max_{\substack{L\subset\mathbb{C}^{nl} \\ \dim(L)\le j-1}} \ \min_{\substack{f\in L^\perp \\ \|f\|=1}} \langle f, J_{n;F} f\rangle_{\mathbb{C}^{nl}}. \tag{2.85}$$

If $P\colon \mathbb{C}^{(n+1)l} \to \mathbb{C}^{nl}$ is the natural projection, then $\langle Pf, J_{n+1;F} Pf\rangle_{\mathbb{C}^{(n+1)l}} = \langle Pf, J_{n;F} Pf\rangle_{\mathbb{C}^{nl}}$ and as L runs through all subspaces of $\mathbb{C}^{(n+1)l}$ dimension at most $j-1$, $P[L]$ runs through all subspaces of dimension at most $j-1$ in $\mathbb{C}^{nl}$, so (2.85) implies $x_{n+1,j} \le x_{n,j}$. Using the same argument on $-J_{n;F}$ and $-J_{n+1;F}$ shows $x_j(-J_{n;F}) \ge x_j(-J_{n+1;F})$. But $x_j(-J_{n;F}) = -x_{nl+1-j}(J_{n;F})$ and $x_j(-J_{n+1;F}) = -x_{(n+1)l+1-j}(J_{n+1;F})$. That yields the other inequality. □

2.10 Lower Bounds on p and the Stieltjes–Weyl Formula for m

Next, we want to exploit (2.70) to prove uniform lower bounds on $\|p_n^L(y)\gamma\|$ when $y \notin \mathrm{cvh}(\sigma(J))$, the convex hull of the support of J, and thereby uniform bounds $\|p_n^L(y)^{-1}\|$. We will then use that to prove that for $z \notin \sigma(J)$, we have

$$m(z) = \lim_{n\to\infty} -p_n^L(z)^{-1} q_n^L(z) \tag{2.86}$$

the matrix analogue of a formula that spectral theorists associate with Weyl's definition of the m-function [193], although for the discrete case, it goes back at least to Stieltjes [181].

We begin by mimicking an argument from [171]. Let $H = \mathrm{cvh}(\sigma(J)) = [c-D, c+D]$ with

$$D = \tfrac{1}{2}\,\mathrm{diam}(H). \tag{2.87}$$

By the definition of A_n,

$$\|A_n\| = \|\langle\langle p_{n-1}^R, (x-c)p_n^R\rangle\rangle_R\| \le D. \tag{2.88}$$

Suppose $y \notin H$ and let

$$d = \mathrm{dist}(y, H). \tag{2.89}$$

By the spectral theorem, for any vector $\varphi \in \mathcal{H}_v$,

$$|\langle \varphi, (J-y)\varphi\rangle_{\mathcal{H}_v}| \ge d\|\varphi\|^2. \tag{2.90}$$

By (2.70), with $\varphi = \varphi_{n,\gamma}(y)$,

$$|\langle \varphi, (J-y)\varphi\rangle| \le \|A_n\|\,\|p_n^L(y)\gamma\|\,\|p_{n-1}^L(y)\gamma\| \tag{2.91}$$

while

$$\|\varphi\|^2 = \sum_{j=0}^{n-1} \|p_j^L(y)\gamma\|^2. \tag{2.92}$$

As in [171, Prop. 2.2], we get:

Theorem 2.25. *If $y \notin H$, for any γ,*

$$\|p_n^L(y)\gamma\| \ge \big(\tfrac{d}{D}\big)\big(1 + \big(\tfrac{d}{D}\big)^2\big)^{(n-1)/2}\|\gamma\|. \tag{2.93}$$

In particular,

$$\|p_n^L(y)^{-1}\| \le \tfrac{D}{d}\,. \tag{2.94}$$

Remark. (2.93) implies $\det(p_n^L(y)) \ne 0$ if $\mathrm{Im}\, y > 0$, providing another proof that its zeros are real.

By the discussion after (2.52), if $\mathrm{Im}\, z > 0$, $q_n^L(z) + p_n^L(z)m(z)$ is in ℓ^2, so goes to zero. Since $p_n^L(z)^{-1}$ is bounded, we obtain:

Corollary 2.26. *For any $z \in \mathbb{C}_+ = \{z : \mathrm{Im}\, z > 0\}$,*

$$m(z) = \lim_{n\to\infty} -p_n^L(z)^{-1} q_n^L(z). \tag{2.95}$$

Remark. This holds for $z \notin H$.

Taking adjoints using (2.82) and $m(z)^\dagger = m(\bar z)$, we see that

$$m(z) = \lim_{n\to\infty} -q_n^R(z) p_n^R(z)^{-1}. \tag{2.96}$$

(2.95) and (2.96) are due to [47], which uses the proof based on the Gauss–Jacobi quadrature formula.

2.11 Wronskians of Vector-Valued Solutions

Let α, β be two vector-valued solutions ($\mathbb{C}^l$) of (2.55) for $n = 2, 3, \dots$. Define their scalar Wronskian as (Euclidean inner product on $\mathbb{C}^l$)

$$W_n(\alpha, \beta) = \langle \alpha_n, A_n \beta_{n+1} \rangle - \langle A_n \alpha_{n+1}, \beta_n \rangle \tag{2.97}$$

for $n = 2, 3, \dots$. One can obtain two matrix solutions by using α or β for one column and 0 for the other columns. The scalar Wronskian is just a matrix element of the resulting matrix Wronskian, so W_n is constant (as can also be seen by direct calculation). Here is an application:

Theorem 2.27. *Let $z_0 \in \mathbb{R} \setminus \sigma_{\rm ess}(J)$. For $k = 0, 1$, let q_k be the multiplicity of z_0 as an eigenvalue of $J^{(k)}$. Then, $q_0 + q_1 \le l$.*

Proof. If $\tilde\beta$ is an eigenfunction for $J^{(1)}$ and we define β by

$$\beta_n = \begin{cases} 0, & n = 1, \\ \tilde\beta_{n-1}, & n \ge 2, \end{cases} \tag{2.98}$$

then β solves (2.55) for $n \ge 2$. If α is an eigenfunction for $J = J^{(0)}$, it also solves (2.55). Since $\alpha_n \to 0$, $\beta_n \to 0$, and A_n is bounded, $W_n(\alpha, \beta) \to 0$ as $n \to \infty$ and so it is identically zero. But since $\beta_1 = 0$,

$$\mathbf{0} = W_1(\alpha, \beta) = \langle \alpha_1, A_1 \beta_2 \rangle = \langle \alpha_1, A_1 \tilde\beta_1 \rangle. \tag{2.99}$$

Let $V^{(k)}$ be the set of values of eigenfunctions of $J^{(k)}$ at $n = 1$. (2.99) says

$$V^{(0)} \subset [A_1 V^{(1)}]^\perp. \tag{2.100}$$

Since $q_k = \dim(V^{(k)})$ and A_1 is non-singular, (2.100) implies that $q_0 \le l - q_1$. $\square$

2.12 The Order of Zeros/Poles of $m(z)$

Theorem 2.28. *Let $z_0 \in \mathbb{R} \setminus \sigma_{\rm ess}(J)$. For $k = 0, 1$, let q_k be the multiplicity of z_0 as an eigenvalue of $J^{(k)}$. If $q_1 - q_0 \ge 0$, then $\det(m(z))$ has a zero at $z = z_0$ of order $q_1 - q_0$. If $q_1 - q_0 < 0$, then $\det(m(z))$ has a pole at $z = z_0$ of order $q_0 - q_1$.*

Remarks. 1. To say $\det(m(z))$ has a zero at $z = z_0$ of order 0 means it is finite and non-vanishing at z_0!

2. Where $d\mu$ is a direct sum of scalar measures, so is $m(z)$, and $\det(m(z))$ is then a product. In the scalar case, $m(z)$ has a pole at z_0 if $J^{(0)}$ has z_0 as eigenvalue and a zero at z_0 if $J^{(1)}$ has z_0 as eigenvalue. In the direct sum case, we see there can be cancellations, which helps explain why $q_1 - q_0$ occurs.

3. Formally, one can understand this theorem as follows. Cramer's rule suggests $\det(m(z)) = \det(J^{(1)} - z)/\det(J^{(0)} - z)$. Even though $\det(J^{(k)} - z)$ is not well-defined in the infinite case, we expect a cancellation of zeros of order q_1 and q_0. For $z_0 \notin H$, the convex hull of $\sigma(J^{(0)})$, one can use (2.95) to prove the theorem following this intuition. Spurious zeros in gaps of $\sigma(J^{(0)})$ make this strategy difficult in gaps.

4. Unlike in Lemma 2.21, we can write m as a product of eigenvalues and analyze that directly because $m(x)$ is self-adjoint for x real, which sidesteps some of the problems associated with non-trivial Jordan normal forms.

5. This proof gives another demonstration that $q_0 + q_1 \le l$.

Proof. $m(z)$ has a simple pole at $z = z_0$ with a residue which is rank q_0 and strictly negative definite on its range. Let

$$f(z) = (z - z_0)m(z).$$

f is analytic near z_0 and self-adjoint for z real and near z_0. Thus, by eigenvalue perturbation theory [124, 159], $f(z)$ has l eigenvalues $\rho_1(z), \dots, \rho_l(z)$ analytic near z_0 with $\rho_1, \rho_2, \dots, \rho_{q_0}$ non-zero at z_0 and $\rho_{q_0+1}, \dots, \rho_l$ zero at z_0.

Thus, $m(z)$ has l eigenvalues near z_0, $\lambda_j(z) = \rho_j(z)/(z - z_0)$, where $\lambda_1, \dots, \lambda_{q_0}$ have simple poles and the others are regular.

By Proposition 2.14, $m(z)^{-1}$ has a simple pole at z_0 with residue of rank q_1 (because A_1 is non-singular), so $m(z)^{-1}$ has q_1 eigenvalues with poles. That means q_1 of $\lambda_{q_0+1}, \dots, \lambda_l$ have simple zeros at z_0 and the others are non-zero. Thus, $\det(m(z)) = \prod_{j=1}^{l} \lambda_j(z)$ has a pole/zero of order $q_0 - q_1$. □

2.13 Resolvent of the Jacobi Matrix

Consider the matrix

$$G_{nm}(z) = \langle\langle p^R_{n-1}, (x - z)^{-1} p^R_{m-1} \rangle\rangle_R.$$

Theorem 2.29. *One has*

$$G_{nm}(z) = \begin{cases} \psi^L_{n-1}(z) p^R_{m-1}(z), & \text{if } n \geq m, \\ p^L_{n-1}(z) \psi^R_{m-1}(z), & \text{if } n \leq m. \end{cases} \tag{2.101}$$

Proof. We have

$$\sum_{m=1}^{\infty} G_{km}(z) J_{mn} = z G_{kn}(z), \quad n \neq k,$$
$$\sum_{m=1}^{\infty} J_{nm} G_{mk}(z) = z G_{nk}(z), \quad n \neq k.$$

Fix $k \geq 0$ and let $u_m(z) = G_{km}(z)$. Then $u_m(z)$ satisfies the equation (2.54) for $n \neq k$, and so we can use Theorem 2.16 to describe this solution. First suppose $k > 1$. As u_m is an ℓ^2 solution and u_m satisfies (2.54) for $n = 1$, we have

$$G_{km}(z) = \begin{cases} a_k(z) p^R_{m-1}(z), & m \leq k, \\ b_k(z) \psi^R_{m-1}(z), & m \geq k. \end{cases} \tag{2.102}$$

If $k = 1$, (2.102) also holds true. For $m \geq k$, this follows by the same argument, and for $m = k = 1$, this is a trivial statement. Next, similarly, let us consider $v_m(z) = G_{mk}(z)$. Then $v_m(z)$ solves (2.55) and so, using Theorem 2.17, we obtain

$$G_{mk}(z) = \begin{cases} p^L_{m-1}(z) c_k(z), & m \leq k, \\ \psi^L_{m-1}(z) d_k(z), & m \geq k. \end{cases} \tag{2.103}$$

Comparing (2.102) and (2.103), we find

$$a_k(z) p^R_{m-1}(z) = \psi^L_{k-1}(z) d_m(z),$$

$$b_k(z)\psi^R_{m-1}(z) = p^L_{k-1}(z)c_m(z).$$

As $p^R_0 = p^L_0 = \mathbf{1}$, it follows that

$$a_k(z) = \psi^L_{k-1}(z)d_1(z), \qquad c_m(z) = b_1(z)\psi^R_{m-1}(z)$$

and so we obtain

$$G_{nm}(z) = \begin{cases} \psi^L_{n-1}(z)d_1(z)p^R_{m-1}(z) & \text{if } n \geq m, \\ p^L_{n-1}(z)b_1(z)\psi^R_{m-1}(z) & \text{if } n \leq m. \end{cases} \tag{2.104}$$

It remains to prove that

$$b_1(z) = d_1(z) = \mathbf{1}. \tag{2.105}$$

Consider the case $m = n = 1$. By the definition of the resolvent,

$$G_{11}(z) = \langle\langle p^R_0, (J-z)^{-1}p^R_0\rangle\rangle_R = \int \frac{d\mu(x)}{x-z} = m(z).$$

On the other hand, by (2.104),

$$G_{11}(z) = \psi^L_0(z)d_1(z)p^R_0(z) = m(z)d_1(z),$$
$$G_{11}(z) = p^L_0(z)b_1(z)\psi^R_0(z) = b_1(z)m(z),$$

which proves (2.105). □

3 Matrix Orthogonal Polynomials on the Unit Circle

3.1 Definition of MOPUC

In this chapter, μ is an $l \times l$ matrix-valued measure on $\partial\mathbb{D}$. $\langle\langle\cdot,\cdot\rangle\rangle_R$ and $\langle\langle\cdot,\cdot\rangle\rangle_L$ are defined as in the MOPRL case. Non-triviality is defined as for MOPRL. We will always assume μ is non-trivial. We define monic matrix polynomials Φ^R_n, Φ^L_n by applying Gram–Schmidt to $\{\mathbf{1}, z\mathbf{1}, \dots\}$, that is, Φ^R_n is the unique matrix polynomial $z^n\mathbf{1}+$ lower order with

$$\langle\langle z^k\mathbf{1}, \Phi^R_n\rangle\rangle_R = 0 \quad k = 0, 1, \dots, n-1. \tag{3.1}$$

We will define the normalized MOPUC shortly. We will only consider the analogue of what we called type 1 for MOPRL because only those appear to be useful. Unlike in the scalar case, the monic polynomials do not appear much because it is for the normalized, but not monic, polynomials that the left and right Verblunsky coefficients are the same.

3.2 The Szegő Recursion

Szegő [184] included the scalar Szegő recursion for the first time. It seems likely that Geronimus had it independently shortly after Szegő. Not knowing of the connection with this work, Levinson [138] rederived the recursion but with matrix coefficients! So the results of this section go back to 1947.

For a matrix polynomial P_n of degree n, we define the reversed polynomial P^*_n by

$$P^*_n(z) = z^n P_n(1/\bar{z})^\dagger. \tag{3.2}$$

Notice that

$$(P_n^*)^* = P_n \tag{3.3}$$

and for any $\alpha \in \mathcal{M}_l$,

$$(\alpha P_n)^* = P_n^* \alpha^\dagger, \qquad (P_n \alpha)^* = \alpha^\dagger P_n^*. \tag{3.4}$$

Lemma 3.1. *We have*

$$\langle\langle f, g \rangle\rangle_L = \langle\langle g, f \rangle\rangle_L^\dagger, \qquad \langle\langle f, g \rangle\rangle_R = \langle\langle g, f \rangle\rangle_R^\dagger \tag{3.5}$$

and

$$\langle\langle f^*, g^* \rangle\rangle_L = \langle\langle f, g \rangle\rangle_R^\dagger, \quad \langle\langle f^*, g^* \rangle\rangle_R = \langle\langle f, g \rangle\rangle_L^\dagger. \tag{3.6}$$

Proof. The first and second identities follow immediately from the definition. The third identity is derived as follows:

$$\begin{aligned}
\langle\langle f^*, g^* \rangle\rangle_L &= \int e^{in\theta} g(\theta)^\dagger \, d\mu(\theta) \, (e^{in\theta} f(\theta)^\dagger)^\dagger \\
&= \int e^{in\theta} g(\theta)^\dagger \, d\mu(\theta) \, e^{-in\theta} f(\theta) \\
&= \int g(\theta)^\dagger \, d\mu(\theta) \, f(\theta) \\
&= \langle\langle g, f \rangle\rangle_R \\
&= \langle\langle f, g \rangle\rangle_R^\dagger.
\end{aligned}$$

The proof of the last identity is analogous. □

Lemma 3.2. *If P_n has degree n and is left-orthogonal with respect to $z\mathbf{1}, \dots, z^n\mathbf{1}$, then $P_n = c(\Phi_n^R)^*$ for some suitable matrix c.*

Proof. By assumption,

$$\mathbf{0} = \langle\langle P_n, z^j\mathbf{1} \rangle\rangle_L = \langle\langle (z^j\mathbf{1})^*, P_n^* \rangle\rangle_R = \langle\langle z^{n-j}\mathbf{1}, P_n^* \rangle\rangle_R \ \text{ for } 1 \le j \le n.$$

Thus, P_n^* is right-orthogonal with respect to $\mathbf{1}, z\mathbf{1}, \dots, z^{n-1}\mathbf{1}$ and hence it is a right-multiple of Φ_n^R. Consequently, P_n is a left- multiple of $(\Phi_n^R)^*$. □

Let us define normalized orthogonal matrix polynomials by

$$\varphi_0^L = \varphi_0^R = \mathbf{1}, \quad \varphi_n^L = \kappa_n^L \Phi_n^L \quad \text{and} \quad \varphi_n^R = \Phi_n^R \kappa_n^R$$

where the κ's are defined according to the normalization condition

$$\langle\langle \varphi_n^R, \varphi_m^R \rangle\rangle_R = \delta_{nm}\mathbf{1} \quad \langle\langle \varphi_n^L, \varphi_m^L \rangle\rangle_L = \delta_{nm}\mathbf{1}$$

along with (a type 1 condition)

$$\kappa_{n+1}^L (\kappa_n^L)^{-1} > \mathbf{0} \quad \text{and} \quad (\kappa_n^R)^{-1} \kappa_{n+1}^R > \mathbf{0}. \tag{3.7}$$

Notice that κ_n^L are determined by the normalization condition up to multiplication on the left by unitaries; these unitaries can always be uniquely chosen so as to satisfy (3.7).

Now define

$$\rho_n^L = \kappa_n^L (\kappa_{n+1}^L)^{-1} \quad \text{and} \quad \rho_n^R = (\kappa_{n+1}^R)^{-1} \kappa_n^R.$$

Notice that as inverses of positives matrices, $\rho_n^L > 0$ and $\rho_n^R > 0$. In particular, we have that

$$\kappa_n^L = (\rho_{n-1}^L \dots \rho_0^L)^{-1} \quad \text{and} \quad \kappa_n^R = (\rho_0^R \dots \rho_{n-1}^R)^{-1}.$$

Theorem 3.3 (Szegő Recursion). (a) *For suitable matrices $\alpha_n^{L,R}$, one has*

$$z\varphi_n^L - \rho_n^L \varphi_{n+1}^L = (\alpha_n^L)^\dagger \varphi_n^{R,*}, \tag{3.8}$$

$$z\varphi_n^R - \varphi_{n+1}^R \rho_n^R = \varphi_n^{L,*} (\alpha_n^R)^\dagger. \tag{3.9}$$

(b) *The matrices α_n^L and α_n^R are equal and will henceforth be denoted by α_n.*

(c) $\rho_n^L = (\mathbf{1} - \alpha_n^\dagger \alpha_n)^{1/2}$ *and* $\rho_n^R = (\mathbf{1} - \alpha_n \alpha_n^\dagger)^{1/2}$.

Proof. (a) The matrix polynomial $z\varphi_n^L$ has leading term $z^{n+1}\kappa_n^L$. On the other hand, the matrix polynomial $\rho_n^L \varphi_{n+1}^L$ has leading term $z^{n+1}\rho_n^L \kappa_{n+1}^L$. By definition of ρ_n^L, these terms are equal. Consequently, $z\varphi_n^L - \rho_n^L \varphi_{n+1}^L$ is a matrix polynomial of degree at most n. Notice that it is left-orthogonal with respect to $z\mathbf{1}, \dots, z^n\mathbf{1}$ since

$$\langle\langle z\varphi_n^L - \rho_n^L \varphi_{n+1}^L, z^j \mathbf{1} \rangle\rangle_L = \langle\langle \varphi_n^L, z^{j-1}\mathbf{1} \rangle\rangle_L - \langle\langle \rho_n^L \varphi_{n+1}^L, z^j \mathbf{1} \rangle\rangle_L = \mathbf{0} - \mathbf{0} = \mathbf{0}.$$

Now apply Lemma 3.2. The other claim is proved in the same way.

(b) By part (a) and identities established earlier,

$$\begin{aligned}
(\alpha_n^L)^\dagger &= \mathbf{0} + (\alpha_n^L)^\dagger \mathbf{1} \\
&= \langle\langle \varphi_n^{R,*}, \rho_n^L \varphi_{n+1}^L \rangle\rangle_L + (\alpha_n^L)^\dagger \langle\langle \varphi_n^R, \varphi_n^R \rangle\rangle_R \\
&= \langle\langle \varphi_n^{R,*}, \rho_n^L \varphi_{n+1}^L \rangle\rangle_L + (\alpha_n^L)^\dagger \langle\langle \varphi_n^{R,*}, \varphi_n^{R,*} \rangle\rangle_L \quad \text{by (3.6)} \\
&= \langle\langle \varphi_n^{R,*}, \rho_n^L \varphi_{n+1}^L + (\alpha_n^L)^\dagger \varphi_n^{R,*} \rangle\rangle_L \\
&= \langle\langle \varphi_n^{R,*}, z\varphi_n^L \rangle\rangle_L \\
&= \langle\langle z\varphi_n^R, \varphi_n^{L,*} \rangle\rangle_R^\dagger \quad \text{(using the } (n+1)\text{-degree *)} \\
&= \langle\langle \varphi_{n+1}^R \rho_n^R + \varphi_n^{L,*} (\alpha_n^R)^\dagger, \varphi_n^{L,*} \rangle\rangle_R^\dagger \\
&= \langle\langle \varphi_{n+1}^R \rho_n^R, \varphi_n^{L,*} \rangle\rangle_R^\dagger + \langle\langle \varphi_n^{L,*} (\alpha_n^R)^\dagger, \varphi_n^{L,*} \rangle\rangle_R^\dagger \\
&= \mathbf{0} + \langle\langle \varphi_n^{L,*}, \varphi_n^{L,*} (\alpha_n^R)^\dagger \rangle\rangle_R \\
&= \langle\langle \varphi_n^{L,*}, \varphi_n^{L,*} \rangle\rangle_R \, (\alpha_n^R)^\dagger \\
&= \langle\langle \varphi_n^L, \varphi_n^L \rangle\rangle_L \, (\alpha_n^R)^\dagger \\
&= (\alpha_n^R)^\dagger.
\end{aligned}$$

(c) Using parts (a) and (b), we see that

$$\begin{aligned}
\mathbf{1} &= \langle\langle z\varphi_n^L, z\varphi_n^L \rangle\rangle_L \\
&= \langle\langle \rho_n^L \varphi_{n+1}^L + \alpha_n^\dagger \varphi_n^{R,*}, \rho_n^L \varphi_{n+1}^L + \alpha_n^\dagger \varphi_n^{R,*} \rangle\rangle_L
\end{aligned}$$

$$
\begin{aligned}
&= \rho_n^L \langle\langle \varphi_{n+1}^L, \varphi_{n+1}^L \rangle\rangle_L (\rho_n^L)^\dagger + \alpha_n^\dagger \langle\langle \varphi_n^{R,*}, \varphi_n^{R,*} \rangle\rangle_L \, \alpha_n \\
&= (\rho_n^L)2 + \alpha_n^\dagger \langle\langle \varphi_n^R, \varphi_n^R \rangle\rangle_R \, \alpha_n \\
&= (\rho_n^L)2 + \alpha_n^\dagger \alpha_n .
\end{aligned}
$$

A similar calculation yields the other claim. □

The matrices α_n will henceforth be called the Verblunsky coefficients associated with the measure $d\mu$. Since ρ_n^L is invertible, we have

$$\|\alpha_n\| < 1. \tag{3.10}$$

We will eventually see (Theorem 3.12) that any set of α_n's obeying (3.10) occurs as the set of Verblunsky coefficients for a unique non-trivial measure.

Note that the Szegő recursion for the monic orthogonal polynomials is

$$
\begin{aligned}
z\Phi_n^L - \Phi_{n+1}^L &= (\kappa_n^L)^{-1} \alpha_n^\dagger (\kappa_n^R)^\dagger \Phi_n^{R,*}, \\
z\Phi_n^R - \Phi_{n+1}^R &= \Phi_n^{L,*} (\kappa_n^L)^\dagger \alpha_n^\dagger (\kappa_n^R)^{-1},
\end{aligned} \tag{3.11}
$$

so the coefficients in the L and R equations are not equal and depend on all the α_j, $j = 1, \dots, n$.

Let us write the Szegő recursion in matrix form, starting with left-orthogonal polynomials. By Theorem 3.3,

$$
\begin{aligned}
\varphi_{n+1}^L &= (\rho_n^L)^{-1} [z\varphi_n^L - \alpha_n^\dagger \varphi_n^{R,*}], \\
\varphi_{n+1}^R &= [z\varphi_n^R - \varphi_n^{L,*} \alpha_n^\dagger] (\rho_n^R)^{-1},
\end{aligned}
$$

which implies that

$$\varphi_{n+1}^L = z(\rho_n^L)^{-1} \varphi_n^L - (\rho_n^L)^{-1} \alpha_n^\dagger \varphi_n^{R,*}, \tag{3.12}$$

$$\varphi_{n+1}^{R,*} = (\rho_n^R)^{-1} \varphi_n^{R,*} - z(\rho_n^R)^{-1} \alpha_n \varphi_n^L. \tag{3.13}$$

In other words,

$$\begin{pmatrix} \varphi_{n+1}^L \\ \varphi_{n+1}^{R,*} \end{pmatrix} = A^L(\alpha_n, z) \begin{pmatrix} \varphi_n^L \\ \varphi_n^{R,*} \end{pmatrix} \tag{3.14}$$

where

$$A^L(\alpha, z) = \begin{pmatrix} z(\rho^L)^{-1} & -(\rho^L)^{-1}\alpha^\dagger \\ -z(\rho^R)^{-1}\alpha & (\rho^R)^{-1} \end{pmatrix}$$

and $\rho^L = (\mathbf{1} - \alpha^\dagger \alpha)^{1/2}$, $\rho^R = (\mathbf{1} - \alpha\alpha^\dagger)^{1/2}$. Note that, for $z \neq 0$, the inverse of $A^L(\alpha, z)$ is given by

$$A^L(\alpha, z)^{-1} = \begin{pmatrix} z^{-1}(\rho^L)^{-1} & z^{-1}(\rho^L)^{-1}\alpha^\dagger \\ (\rho^R)^{-1}\alpha & (\rho^R)^{-1} \end{pmatrix}$$

which gives rise to the inverse Szegő recursion (first emphasized in the scalar and matrix cases by Delsarte el al. [37])

$$
\begin{aligned}
\varphi_n^L &= z^{-1}(\rho_n^L)^{-1} \varphi_{n+1}^L + z^{-1}(\rho_n^L)^{-1} \alpha_n^\dagger \varphi_{n+1}^{R,*}, \\
\varphi_n^{R,*} &= (\rho_n^R)^{-1} \alpha_n \varphi_{n+1}^L + (\rho_n^R)^{-1} \varphi_{n+1}^{R,*}.
\end{aligned}
$$

For right-orthogonal polynomials, we find the following matrix formulas. By Theorem 3.3,

$$\varphi^R_{n+1} = z\varphi^R_n(\rho^R_n)^{-1} - \varphi^{L,*}_n\alpha^\dagger_n(\rho^R_n)^{-1}, \tag{3.15}$$

$$\varphi^{L,*}_{n+1} = \varphi^{L,*}_n(\rho^L_n)^{-1} - z\varphi^R_n\alpha_n(\rho^L_n)^{-1}. \tag{3.16}$$

In other words,

$$\begin{pmatrix}\varphi^R_{n+1} & \varphi^{L,*}_{n+1}\end{pmatrix} = \begin{pmatrix}\varphi^R_n & \varphi^{L,*}_n\end{pmatrix} A^R(\alpha_n, z)$$

where

$$A^R(\alpha, z) = \begin{pmatrix} z(\rho^R)^{-1} & -z\alpha(\rho^L)^{-1} \\ -\alpha^\dagger(\rho^R)^{-1} & (\rho^L)^{-1}\end{pmatrix}.$$

For $z \neq 0$, the inverse of $A^R(\alpha, z)$ is given by

$$A^R(\alpha, z)^{-1} = \begin{pmatrix} z^{-1}(\rho^R)^{-1} & (\rho^R)^{-1}\alpha \\ z^{-1}(\rho^L)^{-1}\alpha^\dagger & (\rho^L)^{-1}\end{pmatrix}$$

and hence

$$\varphi^R_n = z^{-1}\varphi^R_{n+1}(\rho^R_n)^{-1} + z^{-1}\varphi^{L,*}_{n+1}(\rho^L_n)^{-1}\alpha^\dagger_n, \tag{3.17}$$

$$\varphi^{L,*}_n = \varphi^R_{n+1}(\rho^R_n)^{-1}\alpha_n + \varphi^{L,*}_{n+1}(\rho^L_n)^{-1}. \tag{3.18}$$

3.3 Second Kind Polynomials

In the scalar case, second kind polynomials go back to Geronimus [89, 91, 92]. For $n \geq 1$, let us introduce the second kind polynomials $\psi^{L,R}_n$ by

$$\psi^L_n(z) = \int \frac{e^{i\theta}+z}{e^{i\theta}-z}\,(\varphi^L_n(e^{i\theta}) - \varphi^L_n(z))\,d\mu(\theta), \tag{3.19}$$

$$\psi^R_n(z) = \int \frac{e^{i\theta}+z}{e^{i\theta}-z}\,d\mu(\theta)\,(\varphi^R_n(e^{i\theta}) - \varphi^R_n(z)). \tag{3.20}$$

For $n = 0$, let us set $\psi^L_0(z) = \psi^R_0(z) = \mathbf{1}$. For future reference, let us display the first polynomials of each series:

$$\varphi^L_1(z) = (\rho^L_0)^{-1}(z - \alpha^\dagger_0), \qquad \varphi^R_1(z) = (z - \alpha^\dagger_0)(\rho^R_0)^{-1}, \tag{3.21}$$

$$\psi^L_1(z) = (\rho^L_0)^{-1}(z + \alpha^\dagger_0), \qquad \psi^R_1(z) = (z + \alpha^\dagger_0)(\rho^R_0)^{-1}. \tag{3.22}$$

We will also need formulas for $\psi^{L,*}_n$ and $\psi^{R,*}_n$, $n \geq 1$. These formulas follow directly from the above definition and from

$$\overline{\left(\frac{e^{i\theta}+1/\bar z}{e^{i\theta}-1/\bar z}\right)} = -\frac{e^{i\theta}+z}{e^{i\theta}-z}\,.$$

Indeed, we have

$$\psi^{L,*}_n(z) = z^n \int \frac{e^{i\theta}+z}{e^{i\theta}-z}\,d\mu(\theta)\,(\varphi^L_n(1/\bar z)^\dagger - \varphi^L_n(e^{i\theta})^\dagger), \tag{3.23}$$

$$\psi^{R,*}_n(z) = z^n \int \frac{e^{i\theta}+z}{e^{i\theta}-z}\,(\varphi^R_n(1/\bar z)^\dagger - \varphi^R_n(e^{i\theta})^\dagger)\,d\mu(\theta). \tag{3.24}$$

Proposition 3.4. *The second kind polynomials obey the recurrence relations*

$$\psi^L_{n+1}(z) = (\rho^L_n)^{-1}(z\psi^L_n(z) + \alpha^\dagger_n \psi^{R,*}_n(z)), \tag{3.25}$$
$$\psi^{R,*}_{n+1}(z) = (\rho^R_n)^{-1}(z\alpha_n \psi^L_n(z) + \psi^{R,*}_n(z)) \tag{3.26}$$

and

$$\psi^R_{n+1}(z) = (z\psi^R_n(z) + \psi^{L,*}_n(z)\alpha^\dagger_n)(\rho^R_n)^{-1}, \tag{3.27}$$
$$\psi^{L,*}_{n+1}(z) = (\psi^{L,*}_n(z) + z\psi^R_n(z)\alpha_n)(\rho^L_n)^{-1} \tag{3.28}$$

for $n \geq 0$.

Proof. 1. Let us check (3.25) for $n \geq 1$. Denote the right-hand side of (3.25) by $\tilde\psi^L_{n+1}(z)$. Using the recurrence relations for φ^L_n, $\varphi^{R,*}_n$ and the definition (3.19) of ψ^L_n, we find

$$\psi^L_{n+1}(z) - \tilde\psi^L_{n+1}(z) = \int \frac{e^{i\theta}+z}{e^{i\theta}-z}\, A_n(\theta,z)\, d\mu(\theta)$$

where

$$\begin{aligned} A_n(\theta,z) &= \varphi^L_{n+1}(e^{i\theta}) - \varphi^L_{n+1}(z) \\ &\qquad - (\rho^L_n)^{-1}[z\varphi^L_n(e^{i\theta}) - z\varphi^L_n(z) + \alpha^\dagger_n z^n \varphi^R_n(1/\bar z)^\dagger - \alpha^\dagger_n z^n \varphi^R_n(e^{i\theta})^\dagger] \\ &= (\rho^L_n)^{-1}[e^{i\theta}\varphi^L_n(e^{i\theta}) - \alpha^\dagger_n \varphi^{R,*}_n(e^{i\theta}) - z\varphi^L_n(z) + \alpha^\dagger_n \varphi^{R,*}_n(z) \\ &\qquad - z\varphi^L_n(e^{i\theta}) + z\varphi^L_n(z) - \alpha^\dagger_n\varphi^{R,*}_n(z) + \alpha^\dagger_n z^n \varphi^R_n(e^{i\theta})^\dagger] \\ &= (\rho^L_n)^{-1}[(e^{i\theta}-z)\varphi^L_n(e^{i\theta}) + \alpha^\dagger_n(z^n e^{-in\theta}-1)\varphi^{R,*}_n(e^{i\theta})]. \end{aligned}$$

Using the orthogonality relations

$$\int \varphi^L_n(e^{i\theta})\, d\mu(\theta) e^{-im\theta} = \int \varphi^{R,*}_n(e^{i\theta})\, d\mu(\theta) e^{-i(m+1)\theta} = \mathbf{0}, \tag{3.29}$$

$m = 0, 1, \dots, n-1$, and the formula

$$\frac{e^{in\theta}-z^n}{e^{i\theta}-z} = e^{i(n-1)\theta} + e^{i(n-2)\theta} z + \dots + z^{n-1}$$

we obtain

$$\begin{aligned} \rho^L_n \int \frac{e^{i\theta}+z}{e^{i\theta}-z}\, A_n(\theta,z)\, d\mu(\theta) &= \int (e^{i\theta}\varphi^L_n(e^{i\theta}) - \alpha^\dagger_n \varphi^{R,*}_n(e^{i\theta}))\, d\mu(\theta) \\ &= \rho^L_n \int \varphi^L_{n+1}(e^{i\theta})\, d\mu(\theta) = \mathbf{0}. \end{aligned}$$

2. Let us check (3.26) for $n \geq 1$. Denote the right-hand side of (3.26) by $\tilde\psi^{R,*}_{n+1}(z)$. Similarly to the argument above, we find

$$\psi^{R,*}_{n+1}(z) - \tilde\psi^{R,*}_{n+1}(z) = \int \frac{e^{i\theta}+z}{e^{i\theta}-z}\, B_n(\theta,z)\, d\mu(\theta),$$

where

$$\begin{aligned}B_n(\theta,z) &= z^{n+1}\varphi_{n+1}^R(1/\bar z)^\dagger - z^{n+1}\varphi_{n+1}^R(e^{i\theta})^\dagger\\&\quad - (\rho_n^R)^{-1}[z\alpha_n\varphi_n^L(e^{i\theta}) - z\alpha_n\varphi_n^L(z) + z^n\varphi_n^R(1/\bar z)^\dagger - z^n\varphi_n^R(e^{i\theta})^\dagger]\\&= (\rho_n^R)^{-1}[\varphi_n^{R,*}(z) - \alpha_n z\varphi_n^L(z) - z^{n+1}e^{-i(n+1)\theta}\varphi_n^{R,*}(e^{i\theta}) + z^{n+1}e^{-in\theta}\alpha_n\varphi_n^L(e^{i\theta})\\&\quad - z\alpha_n\varphi_n^L(e^{i\theta}) + z\alpha_n\varphi_n^L(z) - \varphi_n^{R,*}(z) + z^n e^{-in\theta}\varphi_n^{R,*}(e^{i\theta})]\\&= (\rho_n^R)^{-1}[z\alpha_n(z^n e^{-in\theta}-1)\varphi_n^L(e^{i\theta}) + z^n e^{-in\theta}(1-ze^{-i\theta})\varphi_n^{R,*}(e^{i\theta})].\end{aligned}$$

Using the orthogonality relations (3.29), we get

$$\begin{aligned}\rho_n^R\int\frac{e^{i\theta}+z}{e^{i\theta}-z}\,B_n(\theta,z)\,d\mu(\theta) &=\\&= z^{n+1}\int(e^{-i(n+1)\theta}\varphi_n^{R,*}(e^{i\theta}) - \alpha_n e^{-in\theta}\varphi_n^L(e^{i\theta}))\,d\mu(\theta)\\&= z^{n+1}\rho_n^R\int e^{-i(n+1)\theta}\varphi_{n+1}^{R,*}(e^{i\theta})\,d\mu(\theta) = \mathbf{0}.\end{aligned}$$

3. Relations (3.25) and (3.26) can be checked for $n=0$ by a direct substitution of (3.21) and (3.22).

4. We obtain (3.27) and (3.28) from (3.25) and (3.26) by applying the $*$-operation. □

Writing the above recursion in matrix form, we get

$$\begin{pmatrix}\psi_{n+1}^L\\ \psi_{n+1}^{R,*}\end{pmatrix} = A^L(-\alpha_n,z)\begin{pmatrix}\psi_n^L\\ \psi_n^{R,*}\end{pmatrix},\qquad \begin{pmatrix}\psi_0^L\\ \psi_0^{R,*}\end{pmatrix} = \begin{pmatrix}\mathbf{1}\\ \mathbf{1}\end{pmatrix}$$

for left-orthogonal polynomials and

$$\begin{pmatrix}\psi_{n+1}^R & \psi_{n+1}^{L,*}\end{pmatrix} = \begin{pmatrix}\psi_n^R & \psi_n^{L,*}\end{pmatrix}A^R(-\alpha_n,z),\qquad \begin{pmatrix}\psi_0^R & \psi_0^{L,*}\end{pmatrix} = \begin{pmatrix}\mathbf{1} & \mathbf{1}\end{pmatrix}.$$

for right-orthogonal polynomials.

Equivalently,

$$\begin{pmatrix}\psi_{n+1}^L\\ -\psi_{n+1}^{R,*}\end{pmatrix} = A^L(\alpha_n,z)\begin{pmatrix}\psi_n^L\\ -\psi_n^{R,*}\end{pmatrix},\qquad \begin{pmatrix}\psi_0^L\\ -\psi_0^{R,*}\end{pmatrix} = \begin{pmatrix}\mathbf{1}\\ -\mathbf{1}\end{pmatrix}\tag{3.30}$$

and

$$\begin{pmatrix}\psi_{n+1}^R & -\psi_{n+1}^{L,*}\end{pmatrix} = \begin{pmatrix}\psi_n^R & -\psi_n^{L,*}\end{pmatrix}A^R(\alpha_n,z),\qquad \begin{pmatrix}\psi_0^R & -\psi_0^{L,*}\end{pmatrix} = \begin{pmatrix}\mathbf{1} & -\mathbf{1}\end{pmatrix}.$$

In particular, we see that the second kind polynomials $\psi_n^{L,R}$ correspond to Verblunsky coefficients $\{-\alpha_n\}$. We have the following Wronskian-type relations:

Proposition 3.5. *For $n \ge 0$ and $z \in \mathbb{C}$, we have*

$$2z^n\mathbf{1} = \varphi_n^L(z)\psi_n^{L,*}(z) + \psi_n^L(z)\varphi_n^{L,*}(z),\tag{3.31}$$

$$2z^n\mathbf{1} = \psi_n^{R,*}(z)\varphi_n^R(z) + \varphi_n^{R,*}(z)\psi_n^R(z).\tag{3.32}$$

$$\mathbf{0} = \varphi_n^L(z)\psi_n^R(z) - \psi_n^L(z)\varphi_n^R(z),\tag{3.33}$$

$$\mathbf{0} = \psi_n^{R,*}(z)\varphi_n^{L,*}(z) - \varphi_n^{R,*}(z)\psi_n^{L,*}(z).\tag{3.34}$$

Proof. We prove this by induction. The four identities clearly hold for $n = 0$. Suppose (3.31)–(3.34) hold for some $n \geq 0$. Then,

$$\begin{aligned}
&\varphi^L_{n+1}\psi^{L,*}_{n+1} + \psi^L_{n+1}\varphi^{L,*}_{n+1} = \\
&\qquad = \left(\rho^L_n\right)^{-1} [(z\varphi^L_n - \alpha^\dagger_n \varphi^{R,*}_n)(\psi^{L,*}_n + z\psi^R_n \alpha_n) \\
&\qquad\quad + (z\psi^L_n + \alpha^\dagger_n \psi^{R,*}_n)(\varphi^{L,*}_n - z\varphi^R_n \alpha_n)] \left(\rho^L_n\right)^{-1} \\
&\qquad = \left(\rho^L_n\right)^{-1} [z(\varphi^L_n\psi^{L,*}_n + \psi^L_n\varphi^{L,*}_n) - z\alpha^\dagger_n(\psi^{R,*}_n\varphi^R_n + \varphi^{R,*}_n\psi^R_n)\alpha_n \\
&\qquad\quad + \alpha^\dagger_n(\psi^{R,*}_n\varphi^{L,*}_n - \varphi^{R,*}_n\psi^{L,*}_n) + z2(\varphi^L_n\psi^R_n - \psi^L_n\varphi^R_n)\alpha_n] \left(\rho^L_n\right)^{-1} \\
&\qquad = \left(\rho^L_n\right)^{-1} [2z^{n+1}(\mathbf{1} - \alpha^\dagger_n\alpha_n)] \left(\rho^L_n\right)^{-1} = 2z^{n+1}\mathbf{1},
\end{aligned}$$

where we used (3.31)–(3.34) for n in the third step. Thus, (3.31) holds for $n + 1$.

For (3.32), we note that

$$\begin{aligned}
&\psi^{R,*}_{n+1}\varphi^R_{n+1} + \varphi^{R,*}_{n+1}\psi^R_{n+1} = \\
&\qquad = (\rho^R_n)^{-1}[(z\alpha_n\psi^L_n + \psi^{R,*}_n)(z\varphi^R_n - \varphi^{L,*}_n\alpha^\dagger_n) \\
&\qquad\quad + (\varphi^{R,*}_n - z\alpha_n\varphi^L_n)(z\psi^R_n + \psi^{L,*}_n\alpha^\dagger_n)] \left(\rho^R_n\right)^{-1} \\
&\qquad = (\rho^R_n)^{-1}[z(\psi^{R,*}_n\varphi^R_n + \varphi^{R,*}_n\psi^R_n) - z\alpha_n(\psi^L_n\varphi^{L,*}_n + \varphi^L_n\psi^{L,*}_n)\alpha^\dagger_n \\
&\qquad\quad + z2\alpha_n(\psi^L_n\varphi^R_n - \varphi^L_n\psi^R_n) - (\psi^{R,*}_n\varphi^{L,*}_n - \varphi^{R,*}_n\psi^{L,*}_n)\alpha^\dagger_n] \left(\rho^R_n\right)^{-1} \\
&\qquad = (\rho^R_n)^{-1} 2z^{n+1}(\mathbf{1} - \alpha_n\alpha^\dagger_n) \left(\rho^R_n\right)^{-1} = 2z^{n+1}\mathbf{1},
\end{aligned}$$

again using (3.31)–(3.34) for n in the third step.

Next,

$$\begin{aligned}
&\varphi^L_{n+1}\psi^R_{n+1} - \psi^L_{n+1}\varphi^R_{n+1} = \\
&\qquad = (\rho^L_n)^{-1}[(z\varphi^L_n - \alpha^\dagger_n\varphi^{R,*}_n)(z\psi^R_n + \psi^{L,*}_n\alpha^\dagger_n) \\
&\qquad\quad - (z\psi^L_n + \alpha^\dagger_n\psi^{R,*}_n)(z\varphi^R_n - \varphi^{L,*}_n\alpha^\dagger_n)](\rho^R_n)^{-1} \\
&\qquad = (\rho^L_n)^{-1}[z2(\varphi^L_n\psi^R_n - \psi^L_n\varphi^R_n) - \alpha^\dagger_n(\varphi^{R,*}_n\psi^{L,*}_n - \psi^{R,*}_n\varphi^{L,*}_n)\alpha^\dagger_n \\
&\qquad\quad - z\alpha^\dagger_n(\varphi^{R,*}_n\psi^R_n + \psi^{R,*}_n\varphi^R_n) + z(\varphi^L_n\psi^{L,*}_n + \psi^L_n\varphi^{L,*}_n)\alpha^\dagger_n](\rho^R_n)^{-1} \\
&\qquad = \mathbf{0}
\end{aligned}$$

which implies first (3.33) for $n+1$ and then, by applying the $*$-operation of order $2n+2$, also (3.34) for $n + 1$. This concludes the proof of the proposition. $\square$

3.4 Christoffel–Darboux Formulas

Proposition 3.6. (a) (CD)-left orthogonal

$$\begin{aligned}
(1 - \bar{\xi}z)\sum_{k=0}^{n} \varphi^L_k(\xi)^\dagger\varphi^L_k(z) &= \varphi^{R,*}_n(\xi)^\dagger\varphi^{R,*}_n(z) - \bar{\xi}z\varphi^L_n(\xi)^\dagger\varphi^L_n(z) \\
&= \varphi^{R,*}_{n+1}(\xi)^\dagger\varphi^{R,*}_{n+1}(z) - \varphi^L_{n+1}(\xi)^\dagger\varphi^L_{n+1}(z).
\end{aligned}$$

(b) (CD)-right orthogonal

$$
\begin{aligned}
(1-\bar{\xi}z)\sum_{k=0}^{n}\varphi_k^R(z)\varphi_k^R(\xi)^\dagger &= \varphi_n^{L,*}(z)\varphi_n^{L,*}(\xi)^\dagger - \bar{\xi}z\varphi_n^R(z)\varphi_n^R(\xi)^\dagger \\
&= \varphi_{n+1}^{L,*}(z)\varphi_{n+1}^{L,*}(\xi)^\dagger - \varphi_{n+1}^R(z)\varphi_{n+1}^R(\xi)^\dagger.
\end{aligned}
$$

(c) (Mixed CD)-left orthogonal

$$
\begin{aligned}
(1-\bar{\xi}z)\sum_{k=0}^{n}\psi_k^L(\xi)^\dagger\varphi_k^L(z) &= 2\cdot\mathbf{1} - \psi_n^{R,*}(\xi)^\dagger\varphi_n^{R,*}(z) - \bar{\xi}z\psi_n^L(\xi)^\dagger\varphi_n^L(z) \\
&= 2\cdot\mathbf{1} - \psi_{n+1}^{R,*}(\xi)^\dagger\varphi_{n+1}^{R,*}(z) - \psi_{n+1}^L(\xi)^\dagger\varphi_{n+1}^L(z).
\end{aligned}
$$

(d) (Mixed CD)-right orthogonal

$$
\begin{aligned}
(1-\bar{\xi}z)\sum_{k=0}^{n}\varphi_k^R(z)\psi_k^R(\xi)^\dagger &= 2\cdot\mathbf{1} - \varphi_n^{L,*}(z)\psi_n^{L,*}(\xi)^\dagger - \bar{\xi}z\varphi_n^R(z)\psi_n^R(\xi)^\dagger \\
&= 2\cdot\mathbf{1} - \varphi_{n+1}^{L,*}(z)\psi_{n+1}^{L,*}(\xi)^\dagger - \varphi_{n+1}^R(z)\psi_{n+1}^R(\xi)^\dagger.
\end{aligned}
$$

Remark. Since the ψ's are themselves MOPUCs, the analogue of (a) and (b), with all φ's replaced by ψ's, holds.

Proof. (a) Write

$$
F_n^L(z) = \begin{pmatrix} \varphi_n^L(z) \\ \varphi_n^{R,*}(z) \end{pmatrix}, \quad J = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix}, \quad \tilde{J} = \begin{pmatrix} \bar{\xi}z\mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix}.
$$

Then,

$$
F_{n+1}^L(z) = A^L(\alpha_n, z)F_n^L(z)
$$

and

$$
\begin{aligned}
&A^L(\alpha,\xi)^\dagger J A^L(\alpha,z) = \\
&\quad = \begin{pmatrix} \bar{\xi}(\rho^L)^{-1} & -\bar{\xi}\alpha^\dagger(\rho^R)^{-1} \\ -\alpha(\rho^L)^{-1} & (\rho^R)^{-1} \end{pmatrix} \begin{pmatrix} z(\rho^L)^{-1} & -(\rho^L)^{-1}\alpha^\dagger \\ z(\rho^R)^{-1}\alpha & -(\rho^R)^{-1} \end{pmatrix} \\
&\quad = \begin{pmatrix} \bar{\xi}z(\rho^L)^{-2} - \bar{\xi}z\alpha^\dagger(\rho^R)^{-2}\alpha & -\bar{\xi}(\rho^L)^{-2}\alpha^\dagger + \bar{\xi}\alpha^\dagger(\rho^R)^{-2} \\ -z\alpha(\rho^L)^{-2} + z(\rho^R)^{-2}\alpha & \alpha(\rho^L)^{-2}\alpha^\dagger - (\rho^R)^{-2} \end{pmatrix} \\
&\quad = \begin{pmatrix} \bar{\xi}z\mathbf{1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{1} \end{pmatrix} = \tilde{J}.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
F_{n+1}^L(\xi)^\dagger J F_{n+1}^L(z) &= F_n^L(\xi)^\dagger A^L(\alpha_n,\xi)^\dagger J A^L(\alpha_n,z)F_n^L(z) \\
&= F_n^L(\xi)^\dagger \tilde{J} F_n^L(z)
\end{aligned}
$$

and hence

$$
\varphi_{n+1}^L(\xi)^\dagger\varphi_{n+1}^L(z) - \varphi_{n+1}^{R,*}(\xi)^\dagger\varphi_{n+1}^{R,*}(z) = \bar{\xi}z\varphi_n^L(\xi)^\dagger\varphi_n^L(z) - \varphi_n^{R,*}(\xi)^\dagger\varphi_n^{R,*}(z)
$$

which shows that the last two expressions in (a) are equal. Denote their common value by $Q_n^L(z,\xi)$. Then,

$$\begin{aligned} Q_n^L(z,\xi) - Q_{n-1}^L(z,\xi) &= \varphi_n^{R,*}(\xi)^\dagger \varphi_n^{R,*}(z) - \bar\xi z \varphi_n^L(\xi)^\dagger \varphi_n^L(z) \\ &\qquad - \varphi_n^{R,*}(\xi)^\dagger \varphi_n^{R,*}(z) + \varphi_n^L(\xi)^\dagger \varphi_n^L(z) \\ &= (1-\bar\xi z)\varphi_n^L(\xi)^\dagger \varphi_n^L(z). \end{aligned}$$

Summing over n completes the proof since $Q_{-1}^L(z,\xi) = 0$.

(b) The proof is analogous to (a): Write $F_n^R(z) = \begin{pmatrix} \varphi_n^R(z) & \varphi_n^{L,*}(z) \end{pmatrix}$. Then, $F_{n+1}^R(z) = F_n^R(z) A^R(\alpha_n, z)$ and $A^R(\alpha,z) J A^R(\alpha,\xi)^\dagger = \tilde J$. Thus,

$$\begin{aligned} F_{n+1}^R(z) J F_{n+1}^R(\xi)^\dagger &= F_n^R(z) A^R(\alpha_n,z) J A^R(\alpha_n,\xi)^\dagger F_n^R(\xi)^\dagger \\ &= F_n^R(z) \tilde J F_n^R(\xi)^\dagger \end{aligned}$$

and hence

$$\varphi_{n+1}^R(z)\varphi_{n+1}^R(\xi)^\dagger - \varphi_{n+1}^{L,*}(z)\varphi_{n+1}^{L,*}(\xi)^\dagger = \bar\xi z \varphi_n^R(z)\varphi_n^R(\xi)^\dagger - \varphi_n^{L,*}(z)\varphi_n^{L,*}(\xi)^\dagger$$

which shows that the last two expressions in (b) are equal. Denote their common value by $Q_n^R(z,\xi)$. Then,

$$Q_n^R(z,\xi) - Q_{n-1}^R(z,\xi) = (1-\bar\xi z)\varphi_n^R(z)\varphi_n^R(\xi)^\dagger$$

and the assertion follows as before.

(c) Write

$$\tilde F_n^L(z) = \begin{pmatrix} \psi_n^L(z) \\ -\psi_n^{R,*}(z) \end{pmatrix}$$

with the second kind polynomials $\psi_n^{L,R}$. As in (a), we see that

$$\begin{aligned} \tilde F_{n+1}^L(\xi)^\dagger J F_{n+1}^L(z) &= \tilde F_n^L(\xi)^\dagger A^L(\alpha_n,\xi)^\dagger J A^L(\alpha_n,z) F_n^L(z) \\ &= \tilde F_n^L(\xi)^\dagger \tilde J F_n^L(z) \end{aligned}$$

and hence

$$\psi_{n+1}^L(\xi)^\dagger \varphi_{n+1}^L(z) + \psi_{n+1}^{R,*}(\xi)^\dagger \varphi_{n+1}^{R,*}(z) = \bar\xi z \psi_n^L(\xi)^\dagger \varphi_n^L(z) + \psi_n^{R,*}(\xi)^\dagger \varphi_n^{R,*}(z).$$

Denote

$$\tilde Q_n^L(z,\xi) = 2\cdot \mathbf{1} - \psi_{n+1}^{R,*}(\xi)^\dagger \varphi_{n+1}^{R,*}(z) - \psi_{n+1}^L(\xi)^\dagger \varphi_{n+1}^L(z).$$

Then,

$$\begin{aligned} \tilde Q_n^L(z,\xi) - \tilde Q_{n-1}^L(z,\xi) &= -\psi_n^{R,*}(\xi)^\dagger \varphi_n^{R,*}(z) - \bar\xi z \psi_n^L(\xi)^\dagger \varphi_n^L(z) \\ &\qquad + \psi_n^{R,*}(\xi)^\dagger \varphi_n^{R,*}(z) + \psi_n^L(\xi)^\dagger \varphi_n^L(z) \\ &= (1-\bar\xi z)\psi_n^L(\xi)^\dagger \varphi_n^L(z) \end{aligned}$$

and the assertion follows as before.

(d) Write $\tilde F_n^R(z) = \begin{pmatrix} \psi_n^R(z) & -\psi_n^{L,*}(z) \end{pmatrix}$. As in (b), we see that

$$\begin{aligned} F_{n+1}^R(z) J \tilde F_{n+1}^R(\xi)^\dagger &= F_n^R(z) A^R(\alpha_n, z) J A^R(\alpha_n, \xi)^\dagger \tilde F_n^R(\xi)^\dagger \\ &= F_n^R(z) \tilde J \tilde F_n^R(\xi)^\dagger \end{aligned}$$

and hence

$$\varphi_{n+1}^R(z)\psi_{n+1}^R(\xi)^\dagger + \varphi_{n+1}^{L,*}(z)\psi_{n+1}^{L,*}(\xi)^\dagger = \bar\xi z \varphi_n^R(z)\psi_n^R(\xi)^\dagger + \varphi_n^{L,*}(z)\psi_n^{L,*}(\xi)^\dagger.$$

With $\tilde Q_n^R(z,\xi) = 2\cdot\mathbf{1} - \varphi_{n+1}^{L,*}(z)\psi_{n+1}^{L,*}(\xi)^\dagger - \varphi_{n+1}^R(z)\psi_{n+1}^R(\xi)^\dagger$, we have

$$\tilde Q_n^R(z,\xi) - \tilde Q_{n-1}^R(z,\xi) = (1-\bar\xi z)\varphi_n^R(z)\psi_n^R(\xi)^\dagger$$

and we conclude as in (c). □

3.5 Zeros of MOPUC

Our main result in this section is:

Theorem 3.7. *All the zeros of* $\det(\varphi_n^R(z))$ *lie in* $\mathbb{D} = \{z : |z| < 1\}$.

We will also prove:

Theorem 3.8. *For each* n,

$$\det(\varphi_n^R(z)) = \det(\varphi_n^L(z)). \tag{3.35}$$

The scalar analogue of Theorem 3.7 has seven proofs in [167]! The simplest is due to Landau [135] and its MOPUC analogue is Theorem 2.13.7 of [167]. There is also a proof in Delsarte et al. [37] who attribute the theorem to Whittle [194]. We provide two more proofs here, not only for their intrinsic interest: our first proof we need because it depends only on the recursion relation (it is related to the proof of Delsarte et al. [37]). The second proof is here since it relates zeros to eigenvalues of a cutoff CMV matrix.

Theorem 3.9. *We have*

(i) *For* $z \in \partial\mathbb{D}$, *all of* $\varphi_n^{R,*}(z)$, $\varphi_n^{L,*}(z)$, $\varphi_n^R(z)$, $\varphi_n^L(z)$ *are invertible.*
(ii) *For* $z \in \partial\mathbb{D}$, $\varphi_n^L(z)(\varphi_n^{R,*}(z))^{-1}$ *and* $(\varphi_n^{*,L}(z))^{-1}\varphi_n^R(z)$ *are unitary.*
(iii) *For* $z \in \mathbb{D}$, $\varphi_n^{R,*}(z)$ *and* $\varphi_n^{L,*}(z)$ *are invertible.*
(iv) *For* $z \in \mathbb{D}$, $\varphi_n^L(z)(\varphi_n^{R,*}(z))^{-1}$ *and* $(\varphi_n^{*,L}(z))^{-1}\varphi_n^R(z)$ *are of norm at most* 1 *and, for* $n \geq 1$, *strictly less than* 1.
(v) *All zeros of* $\det(\varphi_n^{R,*}(z))$ *and* $\det(\varphi_n^{L,*}(z))$ *lie in* $\mathbb{C} \setminus \overline{\mathbb{D}}$.
(vi) *All zeros of* $\det(\varphi_n^R(z))$ *and* $\det(\varphi_n^L(z))$ *lie in* $\mathbb{D}$.

Remark. (vi) is our first proof of Theorem 3.7.

Proof. All these results are trivial for $n = 0$, so we can hope to use an inductive argument. So suppose we have the result for $n-1$.

By (3.13),

$$\varphi_n^{R,*} = (\rho_{n-1}^R)^{-1}(\mathbf{1} - z\alpha_{n-1}\varphi_{n-1}^L(\varphi_{n-1}^{R,*})^{-1})\varphi_{n-1}^{R,*}. \tag{3.36}$$

Since $|\alpha_{n-1}| < 1$, if $|z| \leq 1$, each factor on the right of (3.36) is invertible. This proves (i) and (iii) for $\varphi_n^{R,*}$ and a similar argument works for $\varphi_n^{L,*}$. If $z = e^{i\theta}$, $\varphi_n^R(e^{i\theta}) = e^{in\theta}\varphi_n^{R,*}(e^{i\theta})^\dagger$ is also invertible, so we have (i) and (iii) for n.

Next, we claim that if $z \in \partial\mathbb{D}$, then

$$\varphi_n^{R,*}(z)^\dagger \varphi_n^{R,*}(z) = \varphi_n^L(z)^\dagger \varphi_n^L(z). \tag{3.37}$$

This follows from taking $z = \xi \in \partial\mathbb{D}$ in Proposition 3.6(a). Given that $\varphi_n^{R,*}(z)$ is invertible, this implies

$$\mathbf{1} = (\varphi_n^L(z)\varphi_n^{R,*}(z)^{-1})^\dagger(\varphi_n^L(z)\varphi_n^{R,*}(z)^{-1}) \tag{3.38}$$

proving the first part of (ii) for n. The second part of (ii) is proven similarly by using Proposition 3.6(b).

For $z \in \overline{\mathbb{D}}$, let

$$F(z) = \varphi_n^L(z)\varphi_n^{R,*}(z)^{-1}.$$

Then F is analytic in $\mathbb{D}$, continuous in $\overline{\mathbb{D}}$, and $\|F(z)\| = 1$ on $\partial\mathbb{D}$, so (iv) follows from the maximum principle.

Since $\varphi_n^{R,*}(z)$ is invertible on $\overline{\mathbb{D}}$, its det is non-zero there, proving (v). (vi) then follows from

$$\det(\varphi_n^R(z)) = z^{nl}\, \overline{\det(\varphi_n^{R,*}(1/\bar z))}\,. \tag{3.39}$$

□

Let $\mathcal{V}$ be the $\mathbb{C}^l$-valued functions on $\partial\mathbb{D}$ and $\mathcal{V}_n$ the span of the $\mathbb{C}^l$-valued polynomials of degree at most n, so

$$\dim(\mathcal{V}_n) = \mathbb{C}^{l(n+1)}.$$

Let $\mathcal{V}_\infty$ be the set $\cup_n \mathcal{V}_n$ of all $\mathbb{C}^l$-valued polynomials. Let π_n be the projection onto $\mathcal{V}_n$ in the $\mathcal{V}$ inner product (1.35).

It is easy to see that

$$\mathcal{V}_n \cap \mathcal{V}_{n-1}^\perp = \{\Phi_n^R(z)v : v \in \mathbb{C}^l\} \tag{3.40}$$

since $\langle z^l, \Phi_n^R(z)v\rangle = 0$ for $l = 0, \dots, n-1$ and the dimensions on the left and right of (3.40) coincide. $\mathcal{V}_n \cap \mathcal{V}_{n-1}^\perp$ can also be described as the set of $(v^\dagger \Phi_n^L(z))^\dagger$ for $v \in \mathbb{C}^l$.

We define $M_z\colon \mathcal{V}_{n-1} \to \mathcal{V}_n$ or $\mathcal{V}_\infty \to \mathcal{V}_\infty$ as the operator of multiplication by z.

Theorem 3.10. *For all n, we have*

$$\det\nolimits_{\mathbb{C}^l}(\Phi_n^R(z)) = \det\nolimits_{\mathcal{V}_{n-1}}(z\mathbf{1} - \pi_{n-1}M_z\pi_{n-1}). \tag{3.41}$$

Remarks. 1. Since $\|M_z\| \leq 1$, (3.41) immediately implies zeros of $\det(\varphi_n^R(z))$ lie in $\overline{\mathbb{D}}$, and a small additional argument proves they lie in $\mathbb{D}$. As we will see, this also implies Theorem 3.8.

2. Of course, $\pi_{n-1}M_z\pi_{n-1}$ is a cutoff CMV matrix if written in a CMV basis.

Proof. If $Q \in \mathcal{V}_{n-k}$, then by (3.40),

$$\pi_{n-1}[(z-z_0)^k Q] = 0 \Leftrightarrow (z-z_0)^k Q = \Phi_n^R(z)v \tag{3.42}$$

for some $v \in \mathbb{C}^l$. Thus writing $\det(\Phi_n^R(z)) = \Phi_n^R(z)v_1 \wedge \cdots \wedge \Phi_n^R(z)v_l$ in a Jordan basis for $\Phi_n^R(z)$, we see that the order of the zeros of $\det(\Phi_n^R(z))$ at z_0 is exactly the order of z_0 as an algebraic eigenvalue of $\pi_{n-1}M_z\pi_{n-1}$, that is, the order of z_0 as a zero of the right side of (3.41).

Since both sides of (3.41) are monic polynomials of degree nl and their zeros including multiplicity are the same, we have proven (3.41). □

Proof of Theorem 3.8. On the right side of (3.42), we can put $(\Phi_n^L(z)v^\dagger)^\dagger$ and so conclude (3.41) holds with $\Phi_n^L(z)$ on the left.

This proves (3.35) if φ is replaced by Φ. Since $\alpha_j^*\alpha_j$ and $\alpha_j\alpha_j^*$ are unitarily equivalent, $\det(\rho_j^L) = \det(\rho_j^R)$. Thus, $\det(\kappa_n^L) = \det(\kappa_n^R)$, and we obtain (3.35) for φ. $\square$

It is a basic fact (Theorem 1.7.5 of [167]) that for the scalar case, any set of n zeros in $\mathbb{D}$ are the zeros of a unique OPUC Φ_n and any monic polynomial with all its zeros in $\mathbb{D}$ is a monic OPUC. It is easy to see that any set of nl zeros in $\mathbb{D}$ is the set of zeros of an OPUC Φ_n, but certainly not unique. It is an interesting open question to clarify what matrix monic OPs are monic MOPUCs.

3.6 Bernstein–Szegő Approximation

Given $\{\alpha_j\}_{j=0}^{n-1} \in \mathbb{D}^n$, we use Szegő recursion to define polynomials φ_j^R, φ_j^L for $j = 0, 1, \dots, n$. We define a measure $d\mu_n$ on $\partial\mathbb{D}$ by

$$d\mu_n(\theta) = [\varphi_n^R(e^{i\theta})\varphi_n^R(e^{i\theta})^\dagger]^{-1}\,\frac{d\theta}{2\pi}. \tag{3.43}$$

Notice that (3.37) can be rewritten

$$\varphi_n^R(e^{i\theta})\varphi_n^R(e^{i\theta})^\dagger = \varphi_n^L(e^{i\theta})^\dagger\varphi_n^L(e^{i\theta}). \tag{3.44}$$

We use here and below the fact that the proof of Theorem 3.9 only depends on Szegő recursion and not on the a priori existence of a measure. That theorem also shows the inverse in (3.43) exists. Thus,

$$d\mu_n(\theta) = [\varphi_n^L(e^{i\theta})^\dagger\varphi_n^L(e^{i\theta})]^{-1}\,\frac{d\theta}{2\pi}\,. \tag{3.45}$$

Theorem 3.11. *The measure $d\mu_n$ is normalized (i.e., $\mu_n(\partial\mathbb{D}) = \mathbf{1}$) and its right MOPUC for $j = 0, \dots, n$ are $\{\varphi_j^R\}_{j=0}^n$, and for $j > n$,*

$$\varphi_j^R(z) = z^{j-n}\varphi_n^R(z). \tag{3.46}$$

The Verblunsky coefficients for $d\mu_n$ are

$$\alpha_j(d\mu_n) = \begin{cases} \alpha_j, & j \le n, \\ \mathbf{0}, & j \ge n+1. \end{cases} \tag{3.47}$$

Remarks. 1. In the scalar case, one can multiply by a constant and renormalize, and then prove the constant is 1. Because of commutativity issues, we need a different argument here.

2. Of course, using (3.45), φ_n^L are left MOPUC for $d\mu_n$.

3. Our proof owes something to the scalar proof in [80].

Proof. Let $\langle\langle \cdot, \cdot \rangle\rangle_R$ be the inner product associated with μ_n. By a direct computation, $\langle\langle \varphi_n^R, \varphi_n^R \rangle\rangle_R = \mathbf{1}$, and for $j = 0, 1, \dots, n-1$,

$$\begin{aligned}\langle\langle z^j, \varphi_n^R \rangle\rangle_R &= \frac{1}{2\pi}\int_0^{2\pi} e^{-ij\theta}(\varphi_n^R(e^{i\theta})^\dagger)^{-1}\,d\theta \\ &= \frac{1}{2\pi i}\oint z^{n-j-1}(\varphi_n^{R,*}(z))^{-1}\,dz = \mathbf{0}\end{aligned}$$

by analyticity of $\varphi_n^{R,*}(z)^{-1}$ in $\mathbb{D}$ (continuity in $\overline{\mathbb{D}}$).

This proves φ_n^R is a MOPUC for $d\mu_n$ (and a similar calculation works for the right side of (3.46) if $j \geq n$). By the inverse Szegő recursion and induction downwards, $\{\varphi_j^R\}_{j=0}^{n-1}$ are also OPs, and by the Szegő recursion, they are normalized. In particular, since $\varphi_0^R \equiv \mathbf{1}$ is normalized, $d\mu_n$ is normalized. $\square$

3.7 Verblunsky's Theorem

We can now prove the analogue of Favard's theorem for MOPUC; the scalar version is called Verblunsky's theorem in [167] after [192]. A history and other proofs can be found in [167]. The proof below is essentially the matrix analogue of that of Geronimus [90] (rediscovered in [37, 80]). Delsarte et al. [37] presented their proof in the MOPUC case and they seem to have been the first with a matrix Verblunsky theorem. One can extend the CMV and the Geronimus theorem proofs from the scalar case to get alternate proofs of the theorem below.

Theorem 3.12 (Verblunsky's Theorem for MOPUC). *Any sequence $\{\alpha_j\}_{j=0}^\infty \in \mathbb{D}^\infty$ is the sequence of Verblunsky coefficients of a unique measure.*

Proof. Uniqueness is easy, since the α's determine the φ_j^R's and so the Φ_j^R's which determine the moments.

Given a sequence $\{\alpha_j\}_{j=0}^\infty$, let $d\mu_n$ be the measures of the last section. By compactness of $l \times l$ matrix-valued probability measures on $\partial\mathbb{D}$, they have a weak limit. By using limits, $\{\varphi_j^R\}_{j=0}^\infty$ are the right MOPUC for $d\mu$ and they determine the proper Verblunsky coefficients. $\square$

3.8 Matrix POPUC

Analogously to the scalar case (see [15, 100, 119, 172, 197]), given any unitary β in $\mathcal{M}_l$, we define

$$\varphi_n^R(z;\beta) = z\varphi_{n-1}^R(z) - \varphi_{n-1}^{L,*}(z)\beta^\dagger. \tag{3.48}$$

As in the scalar case, this is related to the secular determinant of unitary extensions of the cutoff CMV matrix. Moreover,

Theorem 3.13. *Fix β. All the zeros of $(\varphi_n(z;\beta))$ lie on $\partial\mathbb{D}$.*

Proof. If $|z| < 1$, $\varphi_{n-1}^{L,*}(z)$ is invertible and

$$\varphi_n^R(z;\beta) = -\varphi_{n-1}^{L,*}(z)\beta^\dagger(\mathbf{1} - z\beta\,\varphi_{n-1}^{L,*}(z)^{-1}\varphi_{n-1}^R(z))$$

is invertible since the last factor differs from $\mathbf{1}$ by a strict contraction. A similar argument shows invertibility if $|z| > 1$. Thus, the only zeros of $\det(\,\cdot\,)$ lie in $\partial\mathbb{D}$. $\square$

3.9 Matrix-Valued Carathéodory and Schur Functions

An analytic matrix-valued function F defined on $\mathbb{D}$ is called a (matrix-valued) Carathéodory function if $F(0) = \mathbf{1}$ and $\operatorname{Re} F(z) \equiv \frac{1}{2}(F(z) + F(z)^\dagger) \geq 0$ for every $z \in \mathbb{D}$. The following result can be found in [44, Thm. 2.2.2].

Theorem 3.14 (Riesz–Herglotz). *If F is a matrix-valued Carathéodory function, then there exists a unique positive semi-definite matrix measure $d\mu$ such that*

$$F(z) = \int \frac{e^{i\theta} + z}{e^{i\theta} - z}\, d\mu(\theta). \tag{3.49}$$

The measure $d\mu$ is given by the unique weak limit of the measures $d\mu_r(\theta) = \operatorname{Re} F(re^{i\theta}) \frac{d\theta}{2\pi}$ as $r \uparrow 1$. Moreover,

$$F(z) = c_0 + 2\sum_{n=1}^{\infty} c_n z^n$$

where

$$c_n = \int e^{-in\theta}\, d\mu(\theta).$$

Conversely, if $d\mu$ is a positive semi-definite matrix measure, then (3.49) *defines a matrix-valued Carathéodory function.*

An analytic matrix-valued function f defined on $\mathbb{D}$ is called a (matrix-valued) Schur function if $f(z)^\dagger f(z) \le \mathbf{1}$ for every $z \in \mathbb{D}$. This condition is equivalent to $f(z)f(z)^\dagger \le \mathbf{1}$ for every $z \in \mathbb{D}$ and to $\|f(z)\| \le 1$ for every $z \in \mathbb{D}$. By the maximum principle, if f is not constant, the inequalities are strict. The following can be found in [167, Prop. 4.5.3]:

Proposition 3.15. *The association*

$$f(z) = z^{-1}(F(z) - \mathbf{1})(F(z) + \mathbf{1})^{-1}, \tag{3.50}$$

$$F(z) = (\mathbf{1} + zf(z))(\mathbf{1} - zf(z))^{-1} \tag{3.51}$$

sets up a one-one correspondence between matrix-valued Carathéodory functions and matrix-valued Schur functions.

Proposition 3.16. *For $z \in \mathbb{D}$, we have*

$$\operatorname{Re} F(z) = (\mathbf{1} - \bar{z} f(z)^\dagger)^{-1}(\mathbf{1} - |z|^2 f(z)^\dagger f(z))(\mathbf{1} - zf(z))^{-1} \tag{3.52}$$

and the non-tangential boundary values $\operatorname{Re} F(e^{i\theta})$ and $f(e^{i\theta})$ exist for Lebesgue almost every θ.

Write $d\mu(\theta) = w(\theta)\frac{d\theta}{2\pi} + d\mu_{\mathrm{s}}$. Then, for almost every θ,

$$w(\theta) = \operatorname{Re} F(e^{i\theta}) \tag{3.53}$$

and for a.e. θ, $\det(w(\theta)) \neq 0$ if and only if $f(e^{i\theta})^\dagger f(e^{i\theta}) < \mathbf{1}$.

Proof. The identity (3.52) follows from (3.51). The existence of the boundary values of f follows by application of the scalar result to the individual entries of f. Then (3.51) gives the boundary values of F. We also used the following fact: Away from a set of zero Lebesgue measure, $\det(\mathbf{1} - zf(z))$ has non-zero boundary values by general H^∞ theory.

(3.53) holds for $\langle \eta, F(z)\eta\rangle_{\mathbb{C}^l}$ and $\langle \eta, d\mu\eta\rangle_{\mathbb{C}^l}$ for any $\eta \in \mathbb{C}^l$ by the scalar result. We get (3.53) by polarization. From

$$w(\theta) = (\mathbf{1} - e^{-i\theta} f(e^{i\theta})^\dagger)^{-1}(\mathbf{1} - f(e^{i\theta})^\dagger f(e^{i\theta}))(\mathbf{1} - e^{i\theta} f(e^{i\theta}))^{-1}$$

it follows immediately that $f(e^{i\theta})^\dagger f(e^{i\theta}) < \mathbf{1}$ implies $\det(w(\theta)) > 0$. Conversely, if $f(e^{i\theta})^\dagger f(e^{i\theta}) \le \mathbf{1}$ but not $f(e^{i\theta})^\dagger f(e^{i\theta}) < 1$, then $\det(\mathbf{1} - f(e^{i\theta})^\dagger f(e^{i\theta})) = 0$ and by our earlier arguments $\det(\mathbf{1} - e^{-i\theta} f(e^{i\theta})^\dagger)^{-1}$ and $\det(\mathbf{1} - e^{i\theta} f(e^{i\theta}))^{-1}$ exist and are finite; hence $\det(w(\theta)) = 0$. All previous statements are true away from suitable sets of zero Lebesgue measure. □

3.10 Coefficient Stripping, the Schur Algorithm, and Geronimus' Theorem

The matrix version of Geronimus' theorem goes back at least to the book of Bakonyi–Constantinescu [6]. Let $F(z)$ be the matrix-valued Carathéodory function (3.49) (with the same measure μ as the one used in the definition of $\langle\langle\cdot,\cdot\rangle\rangle_R$). Let us denote

$$u_n^L(z) = \psi_n^L(z) + \varphi_n^L(z)F(z),$$
$$u_n^R(z) = \psi_n^R(z) + F(z)\varphi_n^R(z).$$

We also define

$$u_n^{L,*}(z) = \psi_n^{L,*}(z) - F(z)\varphi_n^{L,*}(z),$$
$$u_n^{R,*}(z) = \psi_n^{R,*}(z) - \varphi_n^{R,*}(z)F(z).$$

Proposition 3.17. *For any $|z| < 1$, the sequences $u_n^L(z)$, $u_n^R(z)$, $u_n^{L,*}(z)$, $u_n^{R,*}(z)$ are square summable.*

Proof. Denote

$$f(\theta) = \frac{e^{-i\theta} + \bar{z}}{e^{-i\theta} - \bar{z}}, \qquad g(\theta) = \frac{e^{i\theta} + z}{e^{i\theta} - z}.$$

By the definitions (3.19)–(3.24), we have

$$u_n^L(z) = \langle\langle f, \varphi_n^L \rangle\rangle_L,$$
$$-u_n^{R,*}(z) = z^n \langle\langle \varphi_n^R, g \rangle\rangle_R,$$
$$-u_n^{L,*}(z) = z^n \langle\langle \varphi_n^L, g \rangle\rangle_L,$$
$$u_n^R(z) = \langle\langle f, \varphi_n^R \rangle\rangle_R.$$

Using the Bessel inequality and the fact that $|z| < 1$, we obtain the required statements. □

Next we will consider sequences defined by

$$\begin{pmatrix} s_n \\ t_n \end{pmatrix} = A^L(\alpha_{n-1}, z) \cdots A^L(\alpha_0, z) \begin{pmatrix} s_0 \\ t_0 \end{pmatrix} \tag{3.54}$$

where $s_n, t_n \in \mathcal{M}_l$. Similarly, we will consider the sequences

$$(s_n, t_n) = (s_0, t_0) A^R(\alpha_0, z) \cdots A^R(\alpha_{n-1}, z) \tag{3.55}$$

Theorem 3.18. *Let $z \in \mathbb{D}$ and let f be the Schur function associated with $d\mu$ via (3.49) and (3.50). Then:*

(i) *A solution of (3.54) is square summable if and only if the initial condition is of the form*

$$\begin{pmatrix} s_0 \\ t_0 \end{pmatrix} = \begin{pmatrix} c \\ zf(z)c \end{pmatrix}$$

for some matrix c in $\mathcal{M}_l$.

(ii) *A solution of* (3.55) *is square summable if and only if the initial condition is of the form*

$$(s_0, t_0) = (c, czf(z))$$

for some matrix c.

Proof. We shall prove (i); the proof of (ii) is similar.

1. By Proposition 3.17 and (3.14), (3.30), we have the square summable solution

$$\begin{pmatrix} s_n \\ t_n \end{pmatrix} = \begin{pmatrix} u_n^L(z) & -u_n^{R,*}(z) \end{pmatrix}, \quad \begin{pmatrix} s_0 \\ t_0 \end{pmatrix} = \begin{pmatrix} d \\ zf(z)d \end{pmatrix}, \quad d = (F(z) + \mathbf{1}). \tag{3.56}$$

The matrix $d = F(z) + \mathbf{1}$ is invertible. Thus, multiplying the above solution on the right by $d^{-1}c$ for any given matrix c, we get the "if" part of the theorem.

2. Let us check that $\varphi_n^{R,*}c$ is not square summable for any matrix $c \neq 0$. By the CD formula with $\xi = z$, we have

$$(1 - |z|^2) \sum_{k=1}^{n-1} \varphi_k^L(z)^\dagger \varphi_k^L(z) + \varphi_n^L(z)^\dagger \varphi_n^L(z) = \varphi_n^{R,*}(z)^\dagger \varphi_n^{R,*}(z)$$

and so

$$\varphi_n^{R,*}(z)^\dagger \varphi_n^{R,*}(z) \geq (1 - |z|^2) \varphi_0^L(z)^\dagger \varphi_0^L(z) = (1 - |z|^2)\mathbf{1}.$$

Thus, we get

$$\|\varphi_n^{R,*}c\|_R 2 \geq (1 - |z|^2) \operatorname{Tr} c^\dagger c > 0.$$

3. Let $\binom{s_n}{t_n}$ be any square summable solution to (3.54). Let us write this solution as

$$\begin{pmatrix} s_n \\ t_n \end{pmatrix} = \begin{pmatrix} \varphi_n^L a \\ \varphi_n^{R,*} a \end{pmatrix} + \begin{pmatrix} \psi_n^L b \\ -\psi_n^{R,*} b \end{pmatrix}, \quad a = \frac{s_0 + t_0}{2}, \quad b = \frac{s_0 - t_0}{2}. \tag{3.57}$$

Multiplying the solution (3.56) by b and subtracting from (3.57), we get a square summable solution

$$\begin{pmatrix} \varphi_n^L(z)(a - F(z)b) \\ \varphi_n^{R,*}(a - F(z)b) \end{pmatrix}.$$

It follows that $a = F(z)b$, which proves the "only if" part. $\square$

The Schur function f is naturally associated with $d\mu$ and hence with the Verblunsky coefficients $\alpha_0, \alpha_1, \alpha_2, \ldots$. The Schur functions obtained by coefficient stripping will be denoted by $f_1, f_2, f_3, \ldots$, that is, f_n corresponds to Verblunsky coefficients $\alpha_n, \alpha_{n+1}, \alpha_{n+2}, \ldots$. We also write $f_0 \equiv f$.

Theorem 3.19 (Schur Algorithm and Geronimus' Theorem). *For the Schur functions* $f_0, f_1, f_2, \ldots$ *associated with Verblunsky coefficients* $\alpha_0, \alpha_1, \alpha_2, \ldots$, *the following relations hold:*

$$f_{n+1}(z) = z^{-1}(\rho_n^R)^{-1}[f_n(z) - \alpha_n]\,[\mathbf{1} - \alpha_n^\dagger f_n(z)]^{-1} \rho_n^L, \tag{3.58}$$

$$f_n(z) = (\rho_n^R)^{-1}[zf_{n+1}(z) + \alpha_n]\,[\mathbf{1} + z\alpha_n^\dagger f_{n+1}(z)]^{-1} \rho_n^L. \tag{3.59}$$

Remarks. 1. See (1.81) and Theorem 1.4 to understand this result.

2. (1.83) provides an alternate way to write (3.58) and (3.59).

Proof. It clearly suffices to consider the case $n = 0$. Consider the solution of (3.54) with initial condition
$$\begin{pmatrix} \mathbf{1} \\ zf_0(z) \end{pmatrix}.$$

By Theorem 3.18, there exists a matrix c such that
$$\begin{aligned}\begin{pmatrix} c \\ zf_1(z)c \end{pmatrix} &= A^L(\alpha_0, z) \begin{pmatrix} \mathbf{1} \\ zf_0(z) \end{pmatrix} \\ &= \begin{pmatrix} z(\rho_0^L)^{-1} - z(\rho_0^L)^{-1}\alpha_0^\dagger f_0(z) \\ -z(\rho_0^R)^{-1}\alpha_0 + z(\rho_0^R)^{-1}f_0(z) \end{pmatrix}.\end{aligned}$$

From this we can compute $zf_1(z)$:
$$\begin{aligned} zf_1(z) &= [-z(\rho_0^R)^{-1}\alpha_0 + z(\rho_0^R)^{-1}f_0(z)]\,[z(\rho_0^L)^{-1} - z(\rho_0^L)^{-1}\alpha_0^\dagger f_0(z)]^{-1} \\ &= (\rho_0^R)^{-1}[f_0(z) - \alpha_0]\,[\mathbf{1} - \alpha_0^\dagger f_0(z)]^{-1}\rho_0^L \end{aligned}$$

which is (3.58).

Similarly, we can express f_0 in terms of f_1. From
$$\begin{aligned}\begin{pmatrix} \mathbf{1} \\ zf_0(z) \end{pmatrix} &= A^L(\alpha_0, z)^{-1} \begin{pmatrix} c \\ zf_1(z)c \end{pmatrix} \\ &= \begin{pmatrix} z^{-1}(\rho_0^L)^{-1} & z^{-1}(\rho_0^L)^{-1}\alpha_0^\dagger \\ (\rho_0^R)^{-1}\alpha_0 & (\rho_0^R)^{-1} \end{pmatrix} \begin{pmatrix} c \\ zf_1(z)c \end{pmatrix} \\ &= \begin{pmatrix} z^{-1}(\rho_0^L)^{-1}c + (\rho_0^L)^{-1}\alpha_0^\dagger f_1(z)c \\ (\rho_0^R)^{-1}\alpha_0 c + (\rho_0^R)^{-1}zf_1(z)c \end{pmatrix}\end{aligned}$$

we find that
$$\begin{aligned} zf_0(z) &= [(\rho_0^R)^{-1}\alpha_0 c + (\rho_0^R)^{-1}zf_1(z)c]\,[z^{-1}(\rho_0^L)^{-1}c + (\rho_0^L)^{-1}\alpha_0^\dagger f_1(z)c]^{-1} \\ &= [(\rho_0^R)^{-1}\alpha_0 + (\rho_0^R)^{-1}zf_1(z)]\,[z^{-1}(\rho_0^L)^{-1} + (\rho_0^L)^{-1}\alpha_0^\dagger f_1(z)]^{-1} \\ &= (\rho_0^R)^{-1}[\alpha_0 + zf_1(z)]\,[z^{-1}\mathbf{1} + \alpha_0^\dagger f_1(z)]^{-1}\rho_0^L \end{aligned}$$

which gives (3.59). □

3.11 The CMV Matrix

In this section and the next, we discuss CMV matrices for MOPUC. This was discussed first by Simon in [170], which also has the involved history in the scalar case. Most of the results in this section appear already in [170]; the results of the next section are new here—they parallel the discussion in [167, Sect. 4.4] where these results first appeared in the scalar case.

3.11.1 The CMV basis

Consider the two sequences χ_n, $x_n \in \mathcal{H}$, defined by

$$\chi_{2k}(z) = z^{-k}\varphi_{2k}^{L,*}(z), \qquad \chi_{2k-1}(z) = z^{-k+1}\varphi_{2k-1}^{R}(z),$$
$$x_{2k}(z) = z^{-k}\varphi_{2k}^{R}(z), \qquad x_{2k-1}(z) = z^{-k}\varphi_{2k-1}^{L,*}(z).$$

For an integer $k \geq 0$, let us introduce the following notation: i_k is the $(k+1)$th term of the sequence $0, 1, -1, 2, -2, 3, -3, \dots$, and j_k is the $(k+1)$th term of the sequence $0, -1, 1, -2, 2, -3, 3, \dots$. Thus, for example, $i_1 = 1$, $j_1 = -1$.

We use the right module structure of $\mathcal{H}$. For a set of functions $\{f_k(z)\}_{k=0}^n \subset \mathcal{H}$, its module span is the set of all sums $\sum f_k(z)a_k$ with $a_k \in \mathcal{M}_l$.

Proposition 3.20. (i) *For any $n \geq 1$, the module span of $\{\chi_k\}_{k=0}^n$ coincides with the module span of $\{z^{i_k}\}_{k=0}^n$ and the module span of $\{x_k\}_{k=0}^n$ coincides with the module span of $\{z^{j_k}\}_{k=0}^n$.*
(ii) *The sequences $\{\chi_k\}_{k=0}^\infty$ and $\{x_k\}_{k=0}^\infty$ are orthonormal:*

$$\langle\langle \chi_k, \chi_m \rangle\rangle_R = \langle\langle x_k, x_m \rangle\rangle_R = \delta_{km}. \tag{3.60}$$

Proof. (i) Recall that

$$\varphi_n^R(z) = \kappa_n^R z^n + \text{linear combination of}\{\mathbf{1}, \dots, z^{n-1}\},$$
$$\varphi_n^{L,*}(z) = (\kappa_n^L)^\dagger + \text{linear combination of}\{z, \dots, z^n\},$$

where both κ_n^R and $(\kappa_n^L)^\dagger$ are invertible matrices. It follows that

$$\chi_n(z) = \gamma_n z^{i_n} + \text{linear combination of}\{z^{i_0}, \dots, z^{i_{n-1}}\},$$
$$x_n(z) = \delta_n z^{j_n} + \text{linear combination of}\{z^{j_0}, \dots, z^{j_{n-1}}\},$$

where γ_n, δ_n are invertible matrices. This proves (i).

(ii) By the definition of φ_n^L and φ_n^R, we have

$$\langle\langle \varphi_n^R, \varphi_m^R \rangle\rangle_R = \langle\langle \varphi_n^{L,*}, \varphi_m^{L,*} \rangle\rangle_R = \delta_{nm}, \tag{3.61}$$
$$\langle\langle \varphi_n^R, z^m \rangle\rangle_R = \mathbf{0},\ m = 0, \dots, n-1; \quad \langle\langle \varphi_n^{L,*}, z^m \rangle\rangle_R = \mathbf{0},\ m = 1, \dots, n. \tag{3.62}$$

From (3.61) with $n = m$, we get

$$\langle\langle \chi_n, \chi_n \rangle\rangle_R = \langle\langle x_n, x_n \rangle\rangle_R = \mathbf{1}.$$

Considering separately the cases of even and odd n, it is easy to prove that

$$\langle\langle \chi_n, z^m \rangle\rangle_R = \mathbf{0}, \quad m = i_0, i_1, \dots, i_{n-1}, \tag{3.63}$$
$$\langle\langle x_n, z^m \rangle\rangle_R = \mathbf{0}, \quad m = j_0, j_1, \dots, j_{n-1}. \tag{3.64}$$

For example, for $n = 2k$, $m \in \{i_0, \dots, i_{2k-1}\}$ we have $m + k \in \{1, 2, \dots, 2k\}$ and so, by (3.62),

$$\langle\langle \chi_n, z^m \rangle\rangle_R = \langle\langle z^{-k}\varphi_{2k}^{L,*}, z^m \rangle\rangle_R = \langle\langle \varphi_{2k}^{L,*}, z^{m+k} \rangle\rangle_R = \mathbf{0}.$$

The other three cases are considered similarly. From (3.63), (3.64), and (i), we get (3.60) for $k \neq m$. ☐

From the above proposition and the $\|\cdot\|_\infty$-density of Laurent polynomials in $C(\partial\mathbb{D})$, it follows that $\{\chi_k\}_{k=0}^\infty$ and $\{x_k\}_{k=0}^\infty$ are right orthonormal modula bases in $\mathcal{H}$, that is, any element $f \in \mathcal{H}$ can be represented as a

$$f = \sum_{k=0}^\infty \chi_k \langle\langle \chi_k, f\rangle\rangle_R = \sum_{k=0}^\infty x_k \langle\langle x_k, f\rangle\rangle_R. \tag{3.65}$$

3.11.2 The CMV matrix

Consider the matrix of the right homomorphism $f(z) \mapsto zf(z)$ with respect to the basis $\{\chi_k\}$. Denote $\mathcal{C}_{nm} = \langle\langle \chi_n, z\chi_m\rangle\rangle_R$. The matrix $\mathcal{C}$ is unitary in the following sense:

$$\sum_{k=0}^\infty \mathcal{C}_{kn}^\dagger \mathcal{C}_{km} = \sum_{k=0}^\infty \mathcal{C}_{nk}\mathcal{C}_{mk}^\dagger = \delta_{nm}\mathbf{1}.$$

The proof follows from (3.65):

$$\delta_{nm}\mathbf{1} = \langle\langle z\chi_n, z\chi_m\rangle\rangle_R = \left\langle\!\!\left\langle \sum_{k=0}^\infty \chi_k \langle\langle \chi_k, z\chi_n\rangle\rangle_R, z\chi_m \right\rangle\!\!\right\rangle_R = \sum_{k=0}^\infty \mathcal{C}_{kn}^\dagger \mathcal{C}_{km},$$

$$\delta_{nm}\mathbf{1} = \langle\langle \bar z\chi_n, \bar z\chi_m\rangle\rangle_R = \left\langle\!\!\left\langle \sum_{k=0}^\infty \chi_k \langle\langle \chi_k, \bar z\chi_n\rangle\rangle_R, \bar z\chi_m \right\rangle\!\!\right\rangle_R = \sum_{k=0}^\infty \mathcal{C}_{nk}\mathcal{C}_{mk}^\dagger.$$

We note an immediate consequence:

Lemma 3.21. *Let $|z| \le 1$. Then, for every $m \ge 0$,*

$$\sum_{n=0}^\infty \chi_n(z)\mathcal{C}_{nm} = z\chi_m(z), \qquad \sum_{n=0}^\infty \mathcal{C}_{mn}\chi_n(1/\bar z)^\dagger = z\chi_m(1/\bar z)^\dagger.$$

Proof. First note that the above series contains only finitely many non-zero terms. Expanding $f(z) = z\chi_n$ according to (3.65), we see that

$$z\chi_n(z) = \sum_{k=0}^\infty \chi_k(z)\langle\langle \chi_k, z\chi_n\rangle\rangle_R = \sum_{k=0}^\infty \chi_k(z)\mathcal{C}_{kn}$$

which is the first identity. Next, taking adjoints, we get

$$\bar z\chi_n(z)^\dagger = \sum_{k=0}^\infty \mathcal{C}_{kn}^\dagger \chi_k(z)^\dagger$$

which yields

$$\begin{aligned}
\bar z \sum_{n=0}^\infty \mathcal{C}_{mn}\chi_n(z)^\dagger &= \sum_{n=0}^\infty \mathcal{C}_{mn} \sum_{k=0}^\infty \mathcal{C}_{kn}^\dagger \chi_k(z)^\dagger \\
&= \sum_{k=0}^\infty \Big(\sum_{n=0}^\infty \mathcal{C}_{mn}\, \mathcal{C}_{kn}^\dagger \Big) \chi_k(z)^\dagger \\
&= \sum_{k=0}^\infty \delta_{mk}\chi_k(z)^\dagger = \chi_m(z)^\dagger.
\end{aligned}$$

Replacing z by $1/\bar z$, we get the required statement. $\square$

3.11.3 The $\mathcal{LM}$-representation

Using (3.65) for $f = \chi_m$, we obtain:

$$\mathcal{C}_{nm} = \langle\langle \chi_n, z\chi_m \rangle\rangle_R = \sum_{k=0}^{\infty} \langle\langle \chi_n, zx_k \rangle\rangle_R \langle\langle x_k, \chi_m \rangle\rangle_R = \sum_{k=0}^{\infty} \mathcal{L}_{nk}\mathcal{M}_{km}. \tag{3.66}$$

Denote by $\Theta(\alpha)$ the $2l \times 2l$ unitary matrix

$$\Theta(\alpha) = \begin{pmatrix} \alpha^\dagger & \rho^L \\ \rho^R & -\alpha \end{pmatrix}.$$

Using the Szegő recursion formulas (3.15) and (3.18), we get

$$z\varphi_n^R = \varphi_{n+1}^R \rho_n^R + \varphi_n^{L,*}\alpha_n^\dagger, \tag{3.67}$$
$$\varphi_{n+1}^{L,*} = \varphi_n^{L,*}\rho_n^L - \varphi_{n+1}^R \alpha_n. \tag{3.68}$$

Taking $n = 2k$ and multiplying by z^{-k}, we get

$$zx_{2k} = \chi_{2k}\alpha_{2k}^\dagger + \chi_{2k+1}\rho_{2k}^R,$$
$$zx_{2k+1} = \chi_{2k}\rho_{2k}^L - \chi_{2k+1}\alpha_{2k}.$$

It follows that the matrix $\mathcal{L}$ has the structure

$$\mathcal{L} = \Theta(\alpha_0) \oplus \Theta(\alpha_2) \oplus \Theta(\alpha_4) \oplus \cdots .$$

Taking $n = 2k-1$ in (3.67), (3.68) and multiplying by z^{-k}, we get

$$\chi_{2k-1} = x_{2k-1}\alpha_{2k-1}^\dagger + x_{2k}\rho_{2k-1}^R,$$
$$\chi_{2k} = x_{2k-1}\rho_{2k-1}^L - x_{2k}\alpha_{2k-1}.$$

It follows that the matrix $\mathcal{M}$ has the structure

$$\mathcal{M} = \mathbf{1} \oplus \Theta(\alpha_1) \oplus \Theta(\alpha_3) \oplus \cdots . \tag{3.69}$$

Substituting this into (3.66), we obtain:

$$\mathcal{C} = \begin{pmatrix} \alpha_0^\dagger & \rho_0^L\alpha_1^\dagger & \rho_0^L\rho_1^L & \mathbf{0} & \mathbf{0} & \cdots \\ \rho_0^R & -\alpha_0\alpha_1^\dagger & -\alpha_0\rho_1^L & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{0} & \alpha_2^\dagger\rho_1^R & -\alpha_2^\dagger\alpha_1 & \rho_2^L\alpha_3^\dagger & \rho_2^L\rho_3^L & \cdots \\ \mathbf{0} & \rho_2^R\rho_1^R & -\rho_2^R\alpha_1 & -\alpha_2\alpha_3^\dagger & -\alpha_2\rho_3^L & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \alpha_4^\dagger\rho_3^R & -\alpha_4^\dagger\alpha_3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{3.70}$$

We note that the analogous formula to this in [170], namely, (4.30), is incorrect! The order of the factors below the diagonal is wrong there.

3.12 The Resolvent of the CMV Matrix

We begin by studying solutions to the equations

$$\sum_{k=0}^{\infty} \mathcal{C}_{mk} w_k = z w_m, \quad m \geq 2, \tag{3.71}$$

$$\sum_{k=0}^{\infty} \tilde{w}_k \mathcal{C}_{km} = z \tilde{w}_m, \quad m \geq 1. \tag{3.72}$$

Let us introduce the following functions:

$$\begin{aligned}
\tilde{x}_n(z) &= \chi_n(1/\bar{z})^\dagger, \\
\Upsilon_{2n}(z) &= -z^{-n} \psi_{2n}^{L,*}(z), \\
\Upsilon_{2n-1}(z) &= z^{-n+1} \psi_{2n-1}^{R}(z), \\
y_{2n}(z) &= -\Upsilon_{2n}(1/\bar{z})^\dagger = z^{-n} \psi_{2n}^{L}(z), \\
y_{2n-1}(z) &= -\Upsilon_{2n-1}(1/\bar{z})^\dagger = -z^{-n} \psi_{2n-1}^{R,*}(z), \\
p_n(z) &= y_n(z) + \tilde{x}_n(z) F(z), \\
\pi_n(z) &= \Upsilon_n(z) + F(z) \chi_n(z).
\end{aligned}$$

Proposition 3.22. *Let $z \in \mathbb{D} \setminus \{0\}$.*

(i) *For each $n \geq 0$, a pair of values $(\tilde{w}_{2n}, \tilde{w}_{2n+1})$ uniquely determines a solution $\tilde{w}_n$ to* (3.72). *Also, for any pair of values $(\tilde{w}_{2n}, \tilde{w}_{2n+1})$ in $\mathcal{M}_l$, there exists a solution $\tilde{w}_n$ to* (3.72) *with these values at* $(2n, 2n+1)$.

(ii) *The set of solutions $\tilde{w}_n$ to* (3.72) *coincides with the set of sequences*

$$\tilde{w}_n(z) = a \chi_n(z) + b \pi_n(z) \tag{3.73}$$

where a, b range over $\mathcal{M}_l$.

(iii) *A solution* (3.73) *is in ℓ^2 if and only if $a = 0$.*

(iv) *A solution* (3.73) *obeys* (3.72) *for all $m \geq 0$ if and only if $b = 0$.*

Proposition 3.23. *Let $z \in \mathbb{D} \setminus \{0\}$.*

(i) *For each $n \geq 1$, a pair of values (w_{2n-1}, w_{2n}) uniquely determines a solution w_n to* (3.71). *Also, for any pair of values (w_{2n-1}, w_{2n}) in $\mathcal{M}_l$, there exists a solution w_n to* (3.71) *with these values at* $(2n-1, 2n)$.

(ii) *The set of solutions w_n to* (3.71) *coincides with the set of sequences*

$$w_n(z) = \tilde{x}_n(z) a + p_n(z) b \tag{3.74}$$

where a, b range over $\mathcal{M}_l$.

(iii) *A solution* (3.74) *is in ℓ^2 if and only if $a = 0$.*

(iv) *A solution* (3.74) *obeys* (3.71) *for all $m \geq 0$ if and only if $b = 0$.*

Proof of Proposition 3.22. (i) The matrix $\mathcal{C} - z$ can be written in the form

$$\mathcal{C} - z = \begin{pmatrix} A_0 & B_0 & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{0} & A_1 & B_1 & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{0} & A_2 & B_2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{3.75}$$

where

$$A_0 = \begin{pmatrix} \alpha_0^\dagger - z \\ \rho_0^R \end{pmatrix} \qquad A_n = \begin{pmatrix} \alpha_{2n}^\dagger \rho_{2n-1}^R & -\alpha_{2n}^\dagger \alpha_{2n-1} - z \\ \rho_{2n}^R \rho_{2n-1}^R & -\rho_{2n}^R \alpha_{2n-1} \end{pmatrix}$$

$$B_n = \begin{pmatrix} \rho_{2n}^L \alpha_{2n+1}^\dagger & \rho_{2n}^L \rho_{2n+1}^L \\ -\alpha_{2n}\alpha_{2n+1}^\dagger - z & -\alpha_{2n}\rho_{2n+1}^L \end{pmatrix}.$$

Define $\widetilde{W}_n = (\tilde{w}_{2n}, \tilde{w}_{2n+1})$ for $n = 0, 1, 2, \ldots$. Then (3.72) for $m = 2n+1$, $2n+2$ is equivalent to

$$\widetilde{W}_n B_n + \widetilde{W}_{n+1} A_{n+1} = \mathbf{0}.$$

It remains to prove that the $2l \times 2l$ matrices A_j, B_j are invertible. Suppose that for some $x, y \in \mathbb{C}^l$, $A_n \binom{x}{y} = \binom{0}{0}$. This is equivalent to the system

$$\begin{aligned} \alpha_{2n}^\dagger \rho_{2n-1}^R x - \alpha_{2n}^\dagger \alpha_{2n-1} y - zy &= \mathbf{0}, \\ \rho_{2n}^R \rho_{2n-1}^R x - \rho_{2n}^R \alpha_{2n-1} y &= \mathbf{0}. \end{aligned}$$

The second equation of this system yields $\rho_{2n-1}^R x = \alpha_{2n-1} y$ (since ρ_{2n}^R is invertible), and upon substitution into the first equation, we get $y = x = 0$. Thus, $\ker(A_n) = \{0\}$. In a similar way, one proves that $\ker(B_n) = \{0\}$.

(ii) First note that $\tilde{w}_n = \chi_n$ is a solution to (3.72) by Lemma 3.21. Let us check that $\tilde{w}_n = \Upsilon_n$ is also a solution. If $U_{km} = (-1)^k \delta_{km}$, then $(U\mathcal{C}U)_{km}$ for $m \geq 1$ coincides with the CMV matrix corresponding to the coefficients $\{-\alpha_n\}$. Recall that $\psi_n^{L,R}$ are the orthogonal polynomials $\varphi_n^{L,R}$, corresponding to the coefficients $\{-\alpha_n\}$. Taking into account the minus signs in the definition of Υ_n, we see that $\tilde{w}_n = \Upsilon_n$ solves (3.72) for $m \geq 1$. It follows that any $\tilde{w}_n$ of the form (3.73) is a solution to (3.72).

Let us check that any solution to (3.72) can be represented as (3.73). By (i), it suffices to show that for any $\tilde{w}_0, \tilde{w}_1$, there exist $a, b \in \mathcal{M}_l$ such that

$$\begin{aligned} a\chi_0(z) + b\pi_0(z) &= \tilde{w}_0, \\ a\chi_1(z) + b\pi_1(z) &= \tilde{w}_1. \end{aligned}$$

Recalling that $\chi_0 = 1$, $\Upsilon_0 = -1$, $\Upsilon_1(z) = (z + \alpha_0^\dagger)(\rho_0^R)^{-1}$, $\chi_1(z) = (z - \alpha_0^\dagger)(\rho_0^R)^{-1}$, we see that the above system can be easily solved for a, b if $z \neq 0$.

(iii) Let us prove that the solution π_n is square integrable. We will consider separately the sequences π_{2n} and π_{2n-1} and prove that they both belong to ℓ^2. By (3.20) and (3.23), we have

$$\psi_n^R(z) + F(z)\varphi_n^R(z) = \int \frac{e^{i\theta} + z}{e^{i\theta} - z}\, d\mu(\theta)\varphi_n^R(e^{i\theta}), \tag{3.76}$$

$$\psi_n^{L,*}(z) - F(z)\varphi_n^{L,*}(z) = -z^n \int \frac{e^{i\theta} + z}{e^{i\theta} - z}\, d\mu(\theta)\varphi_n^L(e^{i\theta})^\dagger. \tag{3.77}$$

Taking $n = 2k$ in (3.77) and $n = 2k-1$ in (3.76), we get

$$\pi_{2k}(z) = z^k \int \frac{e^{i\theta} + z}{e^{i\theta} - z}\, d\mu(\theta)\varphi_{2k}^L(e^{i\theta})^\dagger, \tag{3.78}$$

$$\pi_{2k-1}(z) = z^{-k+1} \int \frac{e^{i\theta}+z}{e^{i\theta}-z}\, d\mu(\theta)\varphi^R_{2k-1}(e^{i\theta}). \tag{3.79}$$

As φ^L_{2k} is an orthonormal sequence, using the Bessel inequality, from (3.78) we immediately get that π_{2k} is in ℓ^2.

Consider the odd terms π_{2k-1}. We claim that

$$z^{-k+1} \int \frac{e^{i\theta}+z}{e^{i\theta}-z}\, d\mu(\theta)\varphi^R_{2k-1}(e^{i\theta}) = \int \frac{e^{i\theta}+z}{e^{i\theta}-z}\, d\mu(\theta) e^{i(-k+1)\theta}\varphi^R_{2k-1}(e^{i\theta}). \tag{3.80}$$

Indeed, using the right orthogonality of φ^R_{2k-1} to $e^{im\theta}$, $m = 0, 1, \dots, 2k-2$, we get

$$\begin{aligned}\int \frac{e^{i\theta}+z}{e^{i\theta}-z}\, d\mu(\theta)\varphi^R_{2k-1}(e^{i\theta}) &= \left\langle\!\!\left\langle 1 + 2\sum_{m=1}^{\infty} \bar z^m e^{im\theta}, \varphi^R_{2k-1} \right\rangle\!\!\right\rangle_R \\ &= \left\langle\!\!\left\langle 2\sum_{m=2k-1}^{\infty} \bar z^m e^{im\theta}, \varphi^R_{2k-1} \right\rangle\!\!\right\rangle_R\end{aligned}$$

and

$$\begin{aligned}&\int \frac{e^{i\theta}+z}{e^{i\theta}-z} z^{k-1}\, e^{i(-k+1)\theta}\, d\mu(\theta)\varphi^R_{2k-1}(e^{i\theta}) = \\ &\qquad= \left\langle\!\!\left\langle \bar z^{k-1} e^{i(k-1)\theta}\left(1 + 2\sum_{m=1}^{\infty} \bar z^m e^{im\theta}\right), \varphi^R_{2k-1} \right\rangle\!\!\right\rangle_R \\ &\qquad= \left\langle\!\!\left\langle 2\sum_{m=2k-1}^{\infty} \bar z^m e^{im\theta}, \varphi^R_{2k-1} \right\rangle\!\!\right\rangle_R\end{aligned}$$

which proves (3.80). The identities (3.80) and (3.79) yield

$$\pi_{2k-1}(z) = \left\langle\!\!\left\langle \frac{e^{-i\theta}+\bar z}{e^{-i\theta}-\bar z}, \chi_{2k-1} \right\rangle\!\!\right\rangle_R$$

and, since χ_{2k-1} is a right orthogonal sequence, the Bessel inequality ensures that $\pi_{2k-1}(z)$ is in ℓ^2. Thus, $\pi_k(z)$ is in ℓ^2.

Next, as in the proof of Theorem 3.18, using the CD formula, we check that the sequence $\|\varphi^{L,*}_n(z)\|_R$ is bounded below and therefore the sequence $\chi_{2n}(z)$ is not in ℓ^2. This proves the statement (iii).

(iv) By Lemma 3.21, the solution $\chi_n(z)$ obeys (3.72) for all $m \ge 0$. It is easy to check directly that the solution $\pi_n(z)$ does not obey (3.72) for $m = 0$ if $z \neq 0$. This proves the required statement. $\square$

Proof of Proposition 3.23. (i) For $j = 1, 2, \dots,$ define $W_j = (w_{2j-1}, w_{2j})$. Then, using the block structure (3.75), we can rewrite (3.71) for $m = 2j, 2j+1$ as $A_j W_j + B_j W_{j+1} = \mathbf{0}$. By the proof of Proposition 3.22, the matrices A_j and B_j are invertible, which proves (i).

(ii) Lemma 3.21 ensures that $\tilde x_n(z)$ is a solution of (3.71). As in the proof of Proposition 3.22, by considering the matrix $(U\mathcal{C}U)_{km}$, one checks that $y_n(z)$ is also a solution to (3.71).

Let us prove that any solution to (3.71) can be represented in the form (3.74). By (i), it suffices to show that for any w_1, w_2, there exist $a, b \in \mathcal{M}_l$ such that

$$\begin{aligned}\tilde{x}_1(z)a + p_1(z)b &= w_1,\\ \tilde{x}_2(z)a + p_2(z)b &= w_2.\end{aligned}$$

We claim that this system of equations can be solved for a, b. Here are the main steps. Substituting the definitions of $\tilde{x}_n(z)$ and $p_n(z)$, we rewrite this system as

$$\begin{aligned}\varphi_1^{R,*}(a + F(z)b) - \psi_1^{R,*}(z)b &= zw_1,\\ \varphi_2^{L}(a + F(z)b) + \psi_2^{L}(z)b &= zw_2.\end{aligned}$$

Using Szegő recurrence, we can substitute the expressions for φ_2^L, ψ_2^L, which helps rewrite our system as

$$\begin{aligned}\varphi_1^{R,*}(a + F(z)b) - \psi_1^{R,*}(z)b &= zw_1,\\ \varphi_1^{L}(a + F(z)b) + \psi_1^{L}(z)b &= \rho_1^L w_2 + \alpha_1^\dagger w_1.\end{aligned}$$

Substituting explicit formulas for $\varphi_1^{R,*}$, φ_1^L, $\psi_1^{R,*}$, ψ_1^L, and expressing $F(z)$ in terms of $f(z)$, we can rewrite this as

$$\begin{aligned}(\rho_0^R)^{-1}(1 - \alpha_0 z)a + 2z(\rho_0^R)^{-1}(f(z) - \alpha_0)(\mathbf{1} - zf(z))^{-1}b &= zw_1,\\ (\rho_0^L)^{-1}(z - \alpha_0^\dagger)a + 2z(\rho_0^L)^{-1}(\mathbf{1} - \alpha_0^\dagger f(z))(\mathbf{1} - zf(z))^{-1}b &= \rho_1^L w_2 + \alpha_1^\dagger w_1.\end{aligned}$$

Denote

$$\begin{aligned}a_1 &= (\rho_0^R)^{-1}(1 - \alpha_0 z)a,\\ b_1 &= 2z(\rho_0^L)^{-1}(\mathbf{1} - \alpha_0^\dagger f(z))(\mathbf{1} - zf(z))^{-1}b.\end{aligned}$$

Then in terms of a_1, b_1, our system can be rewritten as

$$\begin{aligned}a_1 + X_1 b_1 &= zw_1,\\ X_2 a_1 + b_1 &= \rho_1^L w_2 + \alpha_1^\dagger w_1,\end{aligned}$$

where

$$\begin{aligned}X_1 &= (\rho_0^R)^{-1}(f(z) - \alpha_0)(\mathbf{1} - \alpha_0 f(z))^{-1}\rho_0^L,\\ X_2 &= (\rho_0^L)^{-1}(z - \alpha_0^\dagger)(\mathbf{1} - \alpha_0 z)^{-1}\rho_0^R.\end{aligned}$$

Since $\|f(z)\| < 1$ and $|z| < 1$, we can apply Corollary 1.5, which yields $\|X_1\| < 1$ and $\|X_2\| < 1$. It follows that our system can be solved for a_1, b_1.

(iii) As $p_n(z) = -\pi_n(1/\bar{z})^\dagger$, by Proposition 3.22, we get that $p_n(z)$ is in ℓ^2. In the same way, as $\tilde{x}_n(z) = \chi_n(1/\bar{z})^\dagger$, we get that $\tilde{x}_n(z)$ is not in ℓ^2.

(iv) By Lemma 3.21, the solution $\tilde{x}_n(z)$ obeys (3.71) for all $m \geq 0$. Using the explicit formula for $y_n(z)$, one easily checks that the solution $y_n(z)$ does not obey (3.71) for $m = 0, 1$. □

Theorem 3.24. *We have for $z \in \mathbb{D}$,*

$$[(\mathcal{C}-z)^{-1}]_{k,l} = \begin{cases} (2z)^{-1}\tilde{x}_k(z)\pi_l(z), & l > k \text{ or } k = l \text{ even,} \\ (2z)^{-1}p_k(z)\chi_l(z), & k > l \text{ or } k = l \text{ odd.} \end{cases}$$

Proof. Fix $z \in \mathbb{D}$. Write $G_{k,l}(z) = [(\mathcal{C}-z)^{-1}]_{k,l}$. Then $G_{\cdot,l}(z)$ is equal to $(\mathcal{C}-z)^{-1}\delta_l$, which means that $G_{k,l}(z)$ solves (3.71) for $m \neq l$. Since $G_{\cdot,l}(z)$ is ℓ^2 at infinity and obeys the equation at $m=0$, we see that it is a right-multiple of p for large k and a right-multiple of $\tilde{x}$ for small k. Thus,

$$G_{k,l}(z) = \begin{cases} \tilde{x}_k(z)a_l(z), & k < l \text{ or } k = l \text{ even,} \\ p_k(z)b_l(z), & k > l \text{ or } k = l \text{ odd.} \end{cases}$$

Similarly,

$$G_{k,l}(z) = \begin{cases} \tilde{b}_k(z)\pi_l(z), & k < l \text{ or } k = l \text{ even,} \\ \tilde{a}_k(z)\chi_l(z), & k > l \text{ or } k = l \text{ odd.} \end{cases}$$

Equating the two expressions, we find

$$\tilde{x}_k(z)a_l(z) = \tilde{b}_k(z)\pi_l(z) \quad k < l \text{ or } k = l \text{ even,} \tag{3.81}$$
$$p_k(z)b_l(z) = \tilde{a}_k(z)\chi_l(z) \quad k > l \text{ or } k = l \text{ odd.} \tag{3.82}$$

Putting $k=0$ in (3.81) and setting $\tilde{b}_0(z) = c_1(z)$, we find $a_l(z) = c_1(z)\pi_l(z)$. Putting $l=0$ in (3.82) and setting $b_0(z) = c_2(z)$, we find $p_k(z)c_2(z) = \tilde{a}_k(z)$. Thus,

$$G_{k,l}(z) = \begin{cases} \tilde{x}_k(z)c_1(z)\pi_l(z), & k < l \text{ or } k = l \text{ even,} \\ p_k(z)c_2(z)\chi_l(z), & k > l \text{ or } k = l \text{ odd.} \end{cases}$$

We claim that $c_1(z) = c_2(z) = (2z)^{-1}\mathbf{1}$. Consider the case $k=l=0$. Then, on the one hand, by the definition,

$$\begin{aligned} G_{0,0}(z) &= \int \frac{1}{e^{i\theta}-z}\,d\mu(e^{i\theta}) \\ &= \int (2z)^{-1}\left[\frac{e^{i\theta}+z}{e^{i\theta}-z} - 1\right] d\mu(e^{i\theta}) \\ &= (2z)^{-1}(F(z)-\mathbf{1}) \end{aligned} \tag{3.83}$$

and on the other hand,

$$G_{0,0}(z) = \tilde{x}_0(z)c_1(z)\pi_0(z) = c_1(z)(F(z)-\mathbf{1}).$$

This shows $c_1(z) = (2z)^{-1}\mathbf{1}$. Next, consider the case $k=1$, $l=0$. Then, on the one hand, by the definition,

$$G_{1,0}(z) = \langle\!\langle \chi_1, (e^{i\theta}-z)^{-1}\chi_0 \rangle\!\rangle_R$$

and on the other hand,

$$G_{1,0}(z) = p_1(z)c_2(z)\chi_0(z).$$

Let us calculate the expressions on the right-hand side. We have

$$p_1(z)c_2(z)\chi_0(z) = (\rho_0^R)^{-1}(-z^{-1} - \alpha_0 + (z^{-1} - \alpha_0)F(z))c_2(z) \tag{3.84}$$

and

$$\begin{aligned}
&\langle\!\langle \chi_1, (e^{i\theta} - z)^{-1}\chi_0\rangle\!\rangle_R = \\
&= (\rho_0^R)^{-1} \int (e^{-i\theta} - \alpha_0) d\mu(\theta)(e^{i\theta} - z)^{-1} \\
&= (\rho_0^R)^{-1} \int [z^{-1}(e^{i\theta} - z)^{-1} - z^{-1}e^{-i\theta} - \alpha_0(e^{i\theta} - z)^{-1}]\, d\mu(\theta) \\
&= (\rho_0^R)^{-1} \left[\frac{1}{2z2}\,(F(z) - \mathbf{1}) - \frac{1}{2z}\,\alpha_0(F(z) - \mathbf{1}) - \frac{1}{z}\int e^{-i\theta}\, d\mu(\theta)\right].
\end{aligned}$$

Taking into account the identity

$$\int e^{-i\theta} d\mu(\theta) = \alpha_0$$

(which can be obtained, e.g., by expanding $\langle\!\langle \varphi_1^R, \varphi_0^R\rangle\!\rangle_R = 0$), we get

$$\langle\!\langle \chi_1, (e^{i\theta} - z)^{-1}\chi_0\rangle\!\rangle_R = \frac{1}{2z}\,(\rho_0^R)^{-1}(-z^{-1} - \alpha_0 + (z^{-1} - \alpha_0)F(z)).$$

Comparing this with (3.84), we get $c_2(z) = (2z)^{-1}\mathbf{1}$. □

As an immediate corollary, evaluating the kernel on the diagonal for even and odd indices, we obtain the formulas

$$\int \varphi_{2n}^L(e^{i\theta})\,\frac{d\mu(\theta)}{e^{i\theta} - z}\,\varphi_{2n}^L(e^{i\theta})^\dagger = -\frac{1}{2z^{2n+1}}\,\varphi_{2n}^L(z)u_{2n}^{L,*}(z), \tag{3.85}$$

$$\int \varphi_{2n-1}^R(e^{i\theta})^\dagger\,\frac{d\mu(\theta)}{e^{i\theta} - z}\,\varphi_{2n-1}^R(e^{i\theta}) = -\frac{1}{2z^{2n}}\,u_{2n-1}^{R,*}(z)\varphi_{2n-1}^R(z). \tag{3.86}$$

Combining this with (3.31) and (3.32), we find

$$u_n^L(z)\varphi_n^{L,*}(z) + \varphi_n^L(z)u_n^{L,*}(z) = 2z^n, \tag{3.87}$$

$$\varphi_n^{R,*}(z)u_n^R(z) + u_n^{R,*}(z)\varphi_n^R(z) = 2z^n. \tag{3.88}$$

3.13 Khrushchev Theory

Among the deepest and most elegant methods in OPUC are those of Khrushchev [125, 126, 101]. We have not been able to extend them to MOPUC! We regard their extension as an important open question; we present the first very partial steps here.

Let

$$\Omega = \{\theta : \det w(\theta) > 0\}.$$

Theorem 3.25. *For every $n \geq 0$,*

$$\{\theta : f_n(e^{i\theta})^\dagger f_n(e^{i\theta}) < \mathbf{1}\} = \Omega$$

up to a set of zero Lebesgue measure.

Consequently,

$$\int \|f_n(e^{i\theta})\|\,\frac{d\theta}{2\pi} \geq 1 - \frac{|\Omega|}{2\pi}\,. \tag{3.89}$$

Proof. Recall that, by Proposition 3.16, up to a set of zero Lebesgue measure,

$$\{\theta : f_0(e^{i\theta})^\dagger f_0(e^{i\theta}) < \mathbf{1}\} = \{\theta :\ \det w(\theta) > 0\}$$

so, by induction, it suffices to show that, up to a set of zero Lebesgue measure,

$$\{\theta : f_0(e^{i\theta})^\dagger f_0(e^{i\theta}) < \mathbf{1}\} = \{\theta : f_1(e^{i\theta})^\dagger f_1(e^{i\theta}) < \mathbf{1}\}.$$

This in turn follows from the fact that the Schur algorithm, which relates the two functions, preserves the property $g^\dagger g < \mathbf{1}$.

Notice that away from Ω, $f_n(e^{i\theta})$ has norm one and therefore,

$$\int \|f_n(e^{i\theta})\| \, \frac{d\theta}{2\pi} \geq \int_{\Omega^c} \|f_n(e^{i\theta})\| \, \frac{d\theta}{2\pi} = \int_{\Omega^c} 1 \, \frac{d\theta}{2\pi}$$

which yields (3.89). □

Define

$$b_n(z; d\mu) = \varphi_n^L(z; d\mu)\varphi_n^{R,*}(z; d\mu)^{-1}.$$

Proposition 3.26. (a) $b_{n+1} = (\rho_n^L)^{-1}(zb_n - \alpha_n^\dagger)(\mathbf{1} - z\alpha_n b_n)^{-1}\rho_n^R$.
(b) *The Verblunsky coefficients of* b_n *are* $(-\alpha_{n-1}^\dagger, -\alpha_{n-2}^\dagger, \dots, -\alpha_0^\dagger, \mathbf{1})$.

Proof. (a) By the Szegő recursion, we have that

$$\begin{aligned}
b_{n+1} &= \varphi_{n+1}^L(\varphi_{n+1}^{R,*})^{-1} \\
&= ((\rho_n^L)^{-1} z\varphi_n^L - (\rho_n^L)^{-1}\alpha_n^\dagger\varphi_n^{R,*})((\rho_n^R)^{-1}\varphi_n^{R,*} - z(\rho_n^R)^{-1}\alpha_n\varphi_n^L)^{-1} \\
&= (\rho_n^L)^{-1}(z\varphi_n^L - \alpha_n^\dagger\varphi_n^{R,*})(\varphi_n^{R,*} - z\alpha_n\varphi_n^L)^{-1}\rho_n^R \\
&= (\rho_n^L)^{-1}(z\varphi_n^L(\varphi_n^{R,*})^{-1} - \alpha_n^\dagger\varphi_n^{R,*}(\varphi_n^{R,*})^{-1}) \\
&\qquad\qquad (\varphi_n^{R,*}(\varphi_n^{R,*})^{-1} - z\alpha_n\varphi_n^L(\varphi_n^{R,*})^{-1})^{-1}\rho_n^R \\
&= (\rho_n^L)^{-1}(zb_n - \alpha_n^\dagger)(\mathbf{1} - z\alpha_n b_n)^{-1}\rho_n^R.
\end{aligned}$$

(b) It follows from part (a) that the first Verblunsky coefficient of b_n is $-\alpha_{n-1}^\dagger$ and that its first Schur iterate is b_{n-1}; compare Theorem 3.19. This gives the claim by induction and the fact that $b_0 = \mathbf{1}$. □

4 The Szegő Mapping and the Geronimus Relations

In this chapter, we present the matrix analogue of the Szegő mapping and the resulting Geronimus relations. This establishes a correspondence between certain matrix-valued measures on the unit circle and matrix-valued measures on the interval $[-2, 2]$ and, consequently, a correspondence between Verblunsky coefficients and Jacobi parameters. Throughout this chapter, we will denote measures on the circle by $d\mu_C$ and measures on the interval by $d\mu_I$.

The scalar versions of these objects are due to Szegő [182] and Geronimus [90]. There are four proofs that we know of: the original argument of Geronimus [90] based on Szegő's formula in [182], a proof of Damanik–Killip [32] using Schur functions, a proof of Killip–Nenciu [127] using CMV matrices, and a proof of Faybusovich–Gekhtman [81] using canonical moments.

The matrix version of these objects was studied by Yakhlef–Marcellán [199] who proved Theorem 4.2 below using the Geronimus–Szegő approach. Our proof uses the Killip–Nenciu–CMV approach. In comparing our formula with [199], one needs the following dictionary (their objects on the left of the equal sign and ours on the right):

$$\begin{aligned} H_n &= -\alpha_{n+1}^\dagger, \\ D_n &= A_n, \\ E_n &= B_{n+1}. \end{aligned}$$

Dette–Studden [43] have extended the theory of canonical moments from OPRL to MOPRL. It would be illuminating to use this to extend the proof that Faybusovich–Gekhtman [81] gave of Geronimus relations for scalar OPUC to MOPUC.

Suppose $d\mu_C$ is a non-trivial positive semi-definite Hermitian matrix measure on the unit circle that is invariant under $\theta \mapsto -\theta$ (i.e., $z \mapsto \bar{z} = z^{-1}$). Then we define the measure $d\mu_I$ on the interval $[-2, 2]$ by

$$\int f(x)\, d\mu_I(x) = \int f(2\cos\theta)\, d\mu_C(\theta)$$

for f measurable on $[-2, 2]$. The map

$$\mathrm{Sz} : d\mu_C \mapsto d\mu_I$$

is called the Szegő mapping.

The Szegő mapping can be inverted as follows. Suppose $d\mu_I$ is a non-degenerate positive semi-definite matrix measure on $[-2, 2]$. Then we define the measure $d\mu_C$ on the unit circle which is invariant under $\theta \mapsto -\theta$ by

$$\int g(\theta)\, d\mu_C(\theta) = \int g\left(\arccos(x/2)\right)\, d\mu_I(x)$$

for g measurable on $\partial\mathbb{D}$ with $g(\theta) = g(-\theta)$.

We first show that for the measures on the circle of interest in this section, the Verblunsky coefficients are always Hermitian.

Lemma 4.1. *Suppose $d\mu_C$ is a non-trivial positive semi-definite Hermitian matrix measure on the unit circle. Denote the associated Verblunsky coefficients by $\{\alpha_n\}$. Then, $d\mu_C$ is invariant under $\theta \mapsto -\theta$ if and only if $\alpha_n^\dagger = \alpha_n$ for every n.*

Proof. For a polynomial P, denote $\tilde{P}(z) = P(\bar{z})^\dagger$.

1. Suppose that $d\mu_C$ is invariant under $\theta \mapsto -\theta$. Then we have

$$\langle\!\langle f, g \rangle\!\rangle_L = \langle\!\langle \tilde{g}, \tilde{f} \rangle\!\rangle_R$$

for all f, g. Inspecting the orthogonality conditions which define Φ_n^L and Φ_n^R, we see that

$$\tilde{\Phi}_n^L = \Phi_n^R \text{ and } \langle\!\langle \Phi_n^L, \Phi_n^L \rangle\!\rangle_L = \langle\!\langle \Phi_n^R, \Phi_n^R \rangle\!\rangle_R. \tag{4.1}$$

Next, we claim that

$$\kappa_n^L = \kappa_n^{R,\dagger}. \tag{4.2}$$

Indeed, recall the definition of κ_n^L, κ_n^R:

$$\kappa_n^L = u_n \langle\!\langle \Phi_n^L, \Phi_n^L \rangle\!\rangle_L^{-1/2}, \ u_n \text{ is unitary}, \ \kappa_{n+1}^L (\kappa_n^L)^{-1} > 0, \quad \kappa_0^L = 1,$$
$$\kappa_n^R = \langle\!\langle \Phi_n^R, \Phi_n^R \rangle\!\rangle_R^{-1/2} v_n, \ v_n \text{ is unitary}, \ (\kappa_n^R)^{-1} \kappa_{n+1}^R > 0, \quad \kappa_0^R = 1.$$

Using this definition and (4.1), one can easily prove by induction that $v_n = u_n^\dagger$ and therefore (4.2) holds true.

Next, taking $z = 0$ in (3.11), we get

$$\alpha_n = -(\kappa_n^R)^{-1} \Phi_{n+1}^L(0)^\dagger (\kappa_n^L)^\dagger,$$
$$\alpha_n = -(\kappa_n^R)^\dagger \Phi_{n+1}^R(0)^\dagger (\kappa_n^L)^{-1}.$$

From here and (4.1), (4.2), we get $\alpha_n = \alpha_n^\dagger$.

2. Assume $\alpha_n^\dagger = \alpha_n$ for all n. Then, by Theorem 3.3(c), we have $\rho_n^L = \rho_n^R$. It follows that in this case the Szegő recurrence relation is invariant with respect to the change $\varphi_n^L \mapsto \tilde{\varphi}_n^R$, $\varphi_n^R \mapsto \tilde{\varphi}_n^L$. It follows that $\varphi_n^L = \tilde{\varphi}_n^R$, $\varphi_n^R = \tilde{\varphi}_n^L$. In particular, we get

$$\langle\!\langle \varphi_n^L, \varphi_m^L \rangle\!\rangle_L = \langle\!\langle \tilde{\varphi}_m^R, \tilde{\varphi}_n^R \rangle\!\rangle_R. \tag{4.3}$$

Now let f and g be any polynomials; we have

$$f(z) = \sum_n f_n \varphi_n^L(z), \quad \tilde{f}(z) = \sum_n \tilde{\varphi}_n^L(z) f_n^\dagger,$$

and a similar expansion for g. Using these expansions and (4.3), we get

$$\langle\!\langle f, g \rangle\!\rangle_L = \langle\!\langle \tilde{g}, \tilde{f} \rangle\!\rangle_R \text{ for all polynomials } f,\ g.$$

From here it follows that the measure $d\mu_C$ is invariant under $\theta \mapsto -\theta$. $\square$

Now consider two measures $d\mu_C$ and $d\mu_I = \mathrm{Sz}(d\mu_C)$ and the associated CMV and Jacobi matrices. What are the relations between the parameters of these matrices?

Theorem 4.2. *Given $d\mu_C$ and $d\mu_I = \mathrm{Sz}(d\mu_C)$ as above, the coefficients of the associated CMV and Jacobi matrices satisfy the Geronimus relations:*

$$B_{k+1} = \sqrt{\mathbf{1} - \alpha_{2k-1}}\, \alpha_{2k} \sqrt{\mathbf{1} - \alpha_{2k-1}} - \sqrt{\mathbf{1} + \alpha_{2k-1}}\, \alpha_{2k-2} \sqrt{\mathbf{1} + \alpha_{2k-1}}\,, \tag{4.4}$$
$$A_{k+1} = \sqrt{\mathbf{1} - \alpha_{2k-1}} \sqrt{\mathbf{1} - \alpha_{2k}^2} \sqrt{\mathbf{1} + \alpha_{2k+1}}\,. \tag{4.5}$$

Remarks. 1. For these formulas to hold for $k = 0$, we set $\alpha_{-1} = -\mathbf{1}$.

2. There are several proofs of the Geronimus relations in the scalar case. We follow the proof given by Killip and Nenciu in [127].

3. These A's are, in general, not type 1 or 2 or 3.

Proof. For a Hermitian $l \times l$ matrix α with $\|\alpha\| < 1$, define the unitary $2l \times 2l$ matrix $S(\alpha)$ by

$$S(\alpha) = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{\mathbf{1} - \alpha} & -\sqrt{\mathbf{1} + \alpha} \\ \sqrt{\mathbf{1} + \alpha} & \sqrt{\mathbf{1} - \alpha} \end{pmatrix}.$$

Since $\alpha^\dagger = \alpha$, the associated ρ^L and ρ^R coincide and will be denoted by ρ. We therefore have

$$\Theta(\alpha) = \begin{pmatrix} \alpha & \rho \\ \rho & -\alpha \end{pmatrix}$$

and hence, by a straightforward calculation,

$$\begin{aligned} S(\alpha)\Theta(\alpha)S(\alpha)^{-1} &= \frac{1}{2} \begin{pmatrix} \sqrt{\mathbf{1} - \alpha} & -\sqrt{\mathbf{1} + \alpha} \\ \sqrt{\mathbf{1} + \alpha} & \sqrt{\mathbf{1} - \alpha} \end{pmatrix} \begin{pmatrix} \alpha & \rho \\ \rho & -\alpha \end{pmatrix} \begin{pmatrix} \sqrt{\mathbf{1} - \alpha} & \sqrt{\mathbf{1} + \alpha} \\ -\sqrt{\mathbf{1} + \alpha} & \sqrt{\mathbf{1} - \alpha} \end{pmatrix} \\ &= \begin{pmatrix} -\mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix}. \end{aligned}$$

Thus, if we define

$$\mathcal{S} = \mathbf{1} \oplus S(\alpha_1) \oplus S(\alpha_3) \oplus \dots$$

and

$$\mathcal{R} = \mathbf{1} \oplus (-\mathbf{1}) \oplus \mathbf{1} \oplus (-\mathbf{1}) \oplus \dots,$$

it follows that (see (3.69))

$$\mathcal{S}\mathcal{M}\mathcal{S}^\dagger = \mathcal{R}.$$

The matrix $\mathcal{L}\mathcal{M} + \mathcal{M}\mathcal{L}$ is unitarily equivalent to

$$\mathcal{A} = \mathcal{S}(\mathcal{L}\mathcal{M} + \mathcal{M}\mathcal{L})\mathcal{S}^\dagger = \mathcal{S}\mathcal{L}\mathcal{S}^\dagger\mathcal{R} + \mathcal{R}\mathcal{S}\mathcal{L}\mathcal{S}^\dagger.$$

Observe that $\mathcal{A}$ is the direct sum of two block Jacobi matrices. Indeed, it follows quickly from the explicit form of $\mathcal{R}$ that the even-odd and odd-even block entries of $\mathcal{A}$ vanish. Consequently, $\mathcal{A}$ is the direct sum of its odd-odd and even-even block entries. We will call these two block Jacobi matrices J and $\tilde{J}$, respectively.

Consider $\mathcal{C}$ as an operator on $\mathcal{H}_v$. Then $d\mu_C$ is the spectral measure of $\mathcal{C}$ in the following sense:

$$[\mathcal{C}^m]_{0,0} = \int e^{im\theta} d\mu_C(\theta);$$

see (3.83). Then, by the spectral theorem, $d\mu_C(-\theta)$ is the spectral measure of $\mathcal{C}^{-1}$ and so $d\mu_I$ is the spectral measure of $\mathcal{C} + \mathcal{C}^{-1} = \mathcal{L}\mathcal{M} + (\mathcal{L}\mathcal{M})^{-1} = \mathcal{L}\mathcal{M} + \mathcal{M}\mathcal{L}$. Since $\mathcal{S}$ leaves $\begin{pmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots \end{pmatrix}^t$ invariant, we see that $d\mu_I$ is the spectral measure of J.

To determine the block entries of J, we only need to compute the odd-odd block entries of $\mathcal{A}$. For $k \geq 0$, we have

$$\begin{aligned} \mathcal{A}_{2k+1,2k+1} &= \begin{pmatrix} \sqrt{\mathbf{1} + \alpha_{2k-1}} & \sqrt{\mathbf{1} - \alpha_{2k-1}} \end{pmatrix} \begin{pmatrix} -\alpha_{2k-2} & \mathbf{0} \\ \mathbf{0} & \alpha_{2k} \end{pmatrix} \begin{pmatrix} \sqrt{\mathbf{1} + \alpha_{2k-1}} \\ \sqrt{\mathbf{1} - \alpha_{2k-1}} \end{pmatrix} \\ &= \sqrt{\mathbf{1} - \alpha_{2k-1}}\, \alpha_{2k} \sqrt{\mathbf{1} - \alpha_{2k-1}} - \sqrt{\mathbf{1} + \alpha_{2k-1}}\, \alpha_{2k-2} \sqrt{\mathbf{1} + \alpha_{2k-1}} \end{aligned}$$

and

$$\begin{aligned}\mathcal{A}_{2k+1,2k+3} &= \begin{pmatrix}\sqrt{\mathbf{1}+\alpha_{2k-1}} & \sqrt{\mathbf{1}-\alpha_{2k-1}}\end{pmatrix}\begin{pmatrix}\mathbf{0} & \mathbf{0}\\ \rho_{2k} & \mathbf{0}\end{pmatrix}\begin{pmatrix}\sqrt{\mathbf{1}+\alpha_{2k+1}}\\ \sqrt{\mathbf{1}-\alpha_{2k+1}}\end{pmatrix}\\ &= \sqrt{\mathbf{1}-\alpha_{2k-1}}\sqrt{\mathbf{1}-\alpha_{2k}^2}\sqrt{\mathbf{1}+\alpha_{2k+1}}\,.\end{aligned}$$

The result follows. □

As in [127, 168], one can also use this to get Geronimus relations for the second Szegő map.

5 Regular MOPRL

The theory of regular (scalar) OPs was developed by Stahl–Totik [180] generalizing a definition of Ullman [190] for $[-1,1]$. (See Simon [171] for a review and more about the history.) Here we develop the basics for MOPRL; it is not hard to do the same for MOPUC.

5.1 Upper Bound and Definition

Theorem 5.1. *Let $d\mu$ be a non-trivial $l\times l$ matrix-valued measure on $\mathbb{R}$ with $E=\operatorname{supp}(d\mu)$ compact. Then (with $C(E)$ = logarithmic capacity of E)*

$$\limsup_{n\to\infty}|\det(A_1\cdots A_n)|^{1/n}\leq C(E)^l. \tag{5.1}$$

Remarks. 1. $|\det(A_1\cdots A_n)|$ is constant over equivalent Jacobi parameters.

2. For the scalar case, this is a result of Widom [195] (it might be older) whose proof extends to the matrix case.

Proof. Let T_n be the Chebyshev polynomials for E (see [171, Appendix B] for a definition) and let $T_n^{(l)}$ be $T_n\otimes\mathbf{1}$, that is, the $l\times l$ matrix polynomial obtained by multiplying $T_n(x)$ by $\mathbf{1}$. $T_n^{(l)}$ is monic so, by (2.12) and (2.34),

$$\begin{aligned}|\det(A_1\cdots A_n)|^{1/n} &\leq \left|\det\biggl(\int|T_n^{(l)}(x)|^2\,d\mu(x)\biggr)\right|^{1/2n}\\ &\leq \sup_n|T_n(x)|^{l/n}.\end{aligned}$$

By a theorem of Szegő [183], $\sup_n|T_n(x)|^{1/n}\to C(E)$, so (5.1) follows. □

Definition. Let $d\mu$ be a non-trivial $l\times l$ matrix-valued measure with $E=\operatorname{supp}(d\mu)$ compact. We say μ is *regular* if equality holds in (5.1).

5.2 Density of Zeros

The following is a simple extension of the scalar results (see [74] or [171, Sect. 2]):

Theorem 5.2. *Let $d\mu$ be a regular measure with $E = \operatorname{supp}(d\mu)$. Let $d\nu_n$ be the zero counting measure for $\det(p_n^L(x))$, that is, if $\{x_j^{(n)}\}_{j=1}^{nl}$ are the zeros of this determinant (counting degenerate zeros multiply), then*

$$d\nu_n = \frac{1}{nl}\sum_{j=1}^{nl}\delta_{x_j^{(n)}}. \tag{5.2}$$

Then $d\nu_n$ converges weakly to $d\rho_E$, the equilibrium measure for E.

Remark. For a discussion of capacity, equilibrium measure, quasi-every (q.e.), etc., see [171, Appendix A] and [115, 136, 158].

Proof. By (2.80) and (2.93),

$$|\det(p_n^R(x))| \geq \bigl(\tfrac{d}{D}\bigr)^l\bigl(1+\bigl(\tfrac{d}{D}\bigr)^2\bigr)^{(n-1)l/2} \tag{5.3}$$

so, in particular,

$$\liminf_n |\det(p_n^R(x))|^{1/nl} \geq \bigl(1+\bigl(\tfrac{d}{D}\bigr)^2\bigr)^{1/2} \geq 1. \tag{5.4}$$

But

$$|\det(p_n^R(x))| = |\det(A_1\cdots A_n)|^{-1}\exp(-nl\Phi_{\nu_n}(x)) \tag{5.5}$$

where Φ_ν is the potential of the measure ν. Let ν_∞ be a limit point of ν_n and use equality in (5.1) and (5.4) to conclude, for $x \notin \operatorname{cvh}(E)$,

$$\exp(-\Phi_{\nu_\infty}(x)) \geq C(E)$$

which, as in the proof of Theorem 2.4 of [171], implies that $\nu_\infty = \rho_e$. □

The analogue of the almost converse of this last theorem has an extra subtlety relative to the scalar case:

Theorem 5.3. *Let $d\mu$ be a non-trivial $l \times l$ matrix-valued measure on $\mathbb{R}$ with $E = \operatorname{supp}(d\mu)$ compact. If $d\nu_n \to d\rho_E$, then either μ is regular or else, with $d\mu = M(x)\,d\mu_{\mathrm{tr}}(x)$, there is a set S of capacity zero, so $\det(M(x)) = 0$ for $d\mu_{\mathrm{tr}}$-a.e. $x \notin S$.*

Remark. By taking direct sums of regular point measures, it is easy to find regular measures where $\det(M(x)) = 0$ for $d\mu_{\mathrm{tr}}$-a.e. x.

Proof. For a.e. x with $\det(M(x)) \neq 0$, we have (see Lemma 5.7 below)

$$p_n^R(x) \leq C(n+1)\mathbf{1}.$$

The theorem then follows from the proof of Theorem 2.5 of [171]. □

5.3 General Asymptotics

The following generalizes Theorem 1.10 of [171] from OPRL to MOPRL—it is the matrix analogue of a basic result of Stahl–Totik [180]. Its proof is essentially the same as in [171]. By $\sigma_{\rm ess}(\mu)$, we mean the essential spectrum of the block Jacobi matrix associated to μ.

Theorem 5.4. *Let $E \subset \mathbb{R}$ be compact and let μ be an $l \times l$ matrix-valued measure of compact support with $\sigma_{\rm ess}(\mu) = E$. Then the following are equivalent:*
(i) *μ is regular, that is,* $\lim_{n\to\infty} |\det(A_1 \cdots A_n)|^{1/n} = C(E)^l$.
(ii) *For all z in $\mathbb{C}$, uniformly on compacts,*

$$\limsup |\det(p_n^R(z))|^{1/n} \le e^{G_E(z)}. \tag{5.6}$$

(iii) *For q.e. z in E, we have*

$$\limsup |\det(p_n^R(z))|^{1/n} \le 1. \tag{5.7}$$

Moreover, if (i)–(iii) *hold, then*
(iv) *For every $z \in \mathbb{C} \setminus \mathrm{cvh}(\mathrm{supp}(d\mu))$, we have*

$$\lim_{n\to\infty} |\det(p_n^R(z))|^{1/n} = e^{G_E(z)}. \tag{5.8}$$

(v) *For q.e. $z \in \partial\Omega$,*

$$\limsup_{n\to\infty} |\det(p_n^R(z))|^{1/n} = 1. \tag{5.9}$$

Remarks. 1. G_E, the potential theorists' Green's function for E, is defined by $G_E(z) = -\log(C(E)) - \Phi_{\rho_E}(z)$.

2. There is missing here one condition from Theorem 1.10 of [171] involving general polynomials. Since the determinant of a sum can be much larger than the sum of the determinants, it is not obvious how to extend this result.

5.4 Weak Convergence of the CD Kernel and Consequences

The results of Simon in [174] extend to the matrix case. The basic result is:

Theorem 5.5. *The measures $d\nu_n$ and $\frac{1}{(n+1)l}\,\mathrm{Tr}(K_n(x,x))\,d\mu(x)$ have the same weak limits. In particular, if $d\mu$ is regular,*

$$\frac{1}{(n+1)l}\ \mathrm{Tr}(K_n(x,x)d\mu(x)) \xrightarrow{w} d\rho_E. \tag{5.10}$$

As in [174], $(\pi_n M_x \pi_n)^j$ and $(\pi_n M_x^j \pi_n)$ have a difference of traces which is bounded as $n \to \infty$, and this implies the result. Once one has this, combining it with Theorem 2.20 leads to:

Theorem 5.6. *Let $I = [a,b] \subset E \subset \mathbb{R}$ with E compact. Let $\sigma_{\rm ess}(d\mu) = E$ for an $l\times l$ matrix-valued measure, and suppose $d\mu$ is regular for E and*

$$d\mu = W(x)\,dx + d\mu_{\rm s} \tag{5.11}$$

where $d\mu_{\rm s}$ is singular and $\det(W(x)) > 0$ *for a.e. $x \in I$. Then,*

$$\text{(1)} \qquad \lim_{n\to\infty} \int_I p_n^R(x)^\dagger\, d\mu_{\rm s}(x) p_n^R(x) \to \mathbf{0}, \tag{5.12}$$

$$\text{(2)} \qquad \int_I \left\| \frac{1}{n+1} \sum_{j=0}^n p_j^R(x)^\dagger W(x) p_j(x) - \rho_E(x)\mathbf{1} \right\| dx \to 0. \tag{5.13}$$

5.5 Widom's Theorem

Lemma 5.7. *Let $d\mu$ be an $l \times l$ matrix-valued measure supported on a compact $E \subset \mathbb{R}$ and let $d\eta$ be a scalar measure on $\mathbb{R}$ so*

$$d\mu(x) = W(x)\, d\eta(x) + d\mu_{\mathrm{s}}(x) \tag{5.14}$$

where $d\mu_{\mathrm{s}}$ is $d\eta$ singular. Suppose for $d\eta$-a.e. x,

$$\det(W(x)) > 0. \tag{5.15}$$

Then for $d\eta$-a.e. x, there is a positive real function $C(x)$ so that

$$\|p_n^R(x)\| \le C(x)(n+1)\mathbf{1}. \tag{5.16}$$

In particular,

$$|\det(p_n^R(x))| \le C(x)^l (n+1)^l. \tag{5.17}$$

Proof. Since $\|p_n^R\|_R^2 = 1$, we have

$$\sum_{n=0}^{\infty} (n+1)^{-2} \|p_n^R\|_R^2 < \infty \tag{5.18}$$

so

$$\sum_{n=0}^{\infty} (n+1)^{-2} \operatorname{Tr}(p_n^R(x)^\dagger W(x) p_n^R(x)) < \infty$$

for $d\eta$-a.e. x. Since (5.15) holds, for a.e. x,

$$W(x) \ge b(x)\mathbf{1}$$

for some scalar function $b(x)$. Thus, for a.e. x,

$$\sum_{n=0}^{\infty} (n+1)^{-2} \operatorname{Tr}(p_n^R(x)^\dagger p_n^R(x)) \le C(x)^2.$$

Since $\|A\|^2 \le \operatorname{Tr}(A^\dagger A)$, we find (5.16), which in turn implies (5.17). □

This lemma replaces Lemma 4.1 of [171] and then the proof there of Theorem 1.12 extends to give (a matrix version of the theorem of Widom [195]):

Theorem 5.8. *Let $d\mu$ be an $l \times l$ matrix-valued measure with $\sigma_{\mathrm{ess}}(d\mu) = E \subset \mathbb{R}$ compact. Suppose*

$$d\mu(x) = W(x)\, d\rho_E(x) + d\mu_{\mathrm{s}}(x) \tag{5.19}$$

with $d\mu_{\mathrm{s}}$ singular with respect to $d\rho_E$. Suppose for $d\rho_E$-a.e. x, $\det(W(x)) > 0$. Then μ is regular.

5.6 A Conjecture

We end our discussion of regular MOPRL with a conjecture—an analog of a theorem of Stahl–Totik [180]; see also Theorem 1.13 of [171] for a proof and references. We expect the key will be some kind of matrix Remez inequality. For direct sums, this conjecture follows from Theorem 1.13 of [171].

Conjecture 5.9. Let E be a finite union of disjoint closed intervals in $\mathbb{R}$. Suppose μ is an $l \times l$ matrix-valued measure on $\mathbb{R}$ with $\sigma_{\text{ess}}(d\mu) = E$. For each $\eta > 0$ and $m = 1, 2, \dots,$ define

$$S_{m,\eta} = \{x : \mu([x - \tfrac{1}{m}, x + \tfrac{1}{m}]) \geq e^{-\eta m}\mathbf{1}\}. \tag{5.20}$$

Suppose that for each η (with $|\cdot|$ = Lebesgue measure)

$$\lim_{m\to\infty} |E \setminus S_{m,\eta}| = 0.$$

Then μ is regular.

Acknowledgments. It is a pleasure to thank Alexander Aptekarev, Christian Berg, Antonio Durán, Jeff Geronimo, Fritz Gesztesy, Alberto Grünbaum, Paco Marcellán, Ken McLaughlin, Hermann Schulz-Baldes, and Walter Van Assche for useful correspondence. D.D. was supported in part by NSF grants DMS–0500910 and DMS-0653720. B.S. was supported in part by NSF grants DMS–0140592 and DMS-0652919 and U.S.–Israel Binational Science Foundation (BSF) Grant No. 2002068.

References

[1] R. Ackner, H. Lev-Ari, and T. Kailath, *The Schur algorithm for matrix-valued meromorphic functions*, SIAM J. Matrix Anal. Appl. **15** (1994), 140–150.

[2] N. Akhiezer and I. Glazman, *Theory of Linear Operators in Hilbert Space*, Dover Publications, New York (1993).

[3] D. Alpay, *The Schur Algorithm, Reproducing Kernel Spaces and System Theory*, SMF/AMS Texts and Monographs **5**, American Mathematical Society, Providence, R.I.; Société Mathématique de France, Paris (2001).

[4] A. Aptekarev and E. Nikishin, *The scattering problem for a discrete Sturm–Liouville operator*, Mat. Sb. **121(163)** (1983), 327–358.

[5] F. V. Atkinson, *Discrete and Continuous Boundary Problems*, Mathematics in Science and Engineering **8**, Academic Press, New York-London (1964).

[6] M. Bakonyi and T. Constantinescu, *Schur's Algorithm and Several Applications*, Pitman Research Notes in Math. **261**, Longman, Essex, U.K. (1992).

[7] S. Basu and N. K. Bose, *Matrix Stieltjes series and network models*, SIAM J. Math. Anal. **14** (1983), 209–222.

[8] E. D. Belokolos, F. Gesztesy, K. A. Makarov, and L. A. Sakhnovich, *Matrix-valued generalizations of the theorems of Borg and Hochstadt*, Evolution Equations (G. Ruiz Goldstein et al., eds.), pp. 1–34, Lecture Notes in Pure and Applied Mathematics **234**, Marcel Dekker, New York (2003).

[9] Ju. M. Berezans'ki, *Expansions in Eigenfunctions of Selfadjoint Operators*, Translations of Mathematical Monographs **17**, American Mathematical Society, R.I. (1968).

[10] C. Berg, *The matrix moment problem*, to appear in Coimbra Lecture Notes on Orthogonal Polynomials (A. Foulquié Moreno and Amilcar Branquinho, eds.).

[11] C. Berg and A. J. Durán, *Measures with finite index of determinacy or a mathetmatical model for Dr. Jekyll and Mr. Hyde*, Proc. Amer. Math. Soc. **125** (1997), 523–530.

[12] C. Berg and A. J. Durán, *Orthogonal polynomials and analytic functions associated to positive definite matrices*, J. Math. Anal. Appl. **315** (2006), 54–67.

[13] R. Bhatia, *Matrix Analysis*, Graduate Texts in Mathematics **169**, Springer-Verlag, New York (1997).

[14] M. J. Cantero, M. P. Ferrer, L. Moral, and L. Velázquez, *A connection between orthogonal polynomials on the unit circle and matrix orthogonal polynomials on the real line*, J. Comput. Appl. Math. **154** (2003), 247–272.

[15] M. J. Cantero, L. Moral, and L. Velázquez, *Measures and para-orthogonal polynomials on the unit circle*, East J. Approx. **8** (2002), 447–464.

[16] M. J. Cantero, L. Moral, and L. Velázquez, *Differential properties of matrix orthogonal polynomials*, J. Concr. Appl. Math. **3** (2005), 313–334.

[17] M. J. Cantero, L. Moral, and L. Velázquez, *Matrix orthogonal polynomials whose derivatives are also orthogonal*, J. Approx. Theory **146** (2007), 174–211.

[18] E. Cartan, *Sur les domaines bornés homogènes de l'espace de n variables complexes*, Abh. Math. Sem. Univ. Hamburg **11** (1935), 116–162.

[19] M. Castro and F. A. Grünbaum, *Orthogonal matrix polynomials satisfying first order differential equations: a collection of instructive examples*, J. Nonlinear Math. Physics **12** (2005), 63–76.

[20] M. Castro and F. A. Grünbaum, *The algebra of differential operators associated to a family of matrix valued orthogonal polynomials: five instructive examples*, Int. Math. Res. Not. **2006:7** (2006), 1–33.

[21] M. Castro and F. A. Grünbaum, *The noncommutative bispectral problem for operators of order one*, Constr. Approx. **27** (2008), 329–347.

[22] G. Chen and Y. Hu, *The truncated Hamburger matrix moment problems in the nondegenerate and degenerate cases, and matrix continued fractions*, Linear Algebra Appl. **277** (1998), 199–236.

[23] T. S. Chihara, *An Introduction to Orthogonal Polynomials*, Mathematics and Its Applications **13**, Gordon and Breach, New York-London-Paris (1978).

[24] S. Clark and F. Gesztesy, *Weyl–Titchmarsh M-function asymptotics for matrix-valued Schrödinger operators*, Proc. London Math. Soc. **82** (2001), 701–724.

[25] S. Clark and F. Gesztesy, *Weyl–Titchmarsh M-function asymptotics, local uniqueness results, trace formulas, and Borg-type theorems for Dirac operators*, Trans. Amer. Math. Soc. **354** (2002), 3475–3534.

[26] S. Clark and F. Gesztesy, *On Povzner–Wienholtz-type self-adjointness results for matrix-valued Sturm–Liouville operators*, Proc. Math. Soc. Edinburgh **133A** (2003), 747–758.

[27] S. Clark and F. Gesztesy, *On Weyl–Titchmarsh theory for singular finite difference Hamiltonian systems*, J. Comput. Appl. Math. **171** (2004), 151–184.

[28] S. Clark, F. Gesztesy, H. Holden, and B. M. Levitan, *Borg-type theorems for matrix-valued Schrödinger operators*, J. Diff. Eqs. **167** (2000), 181–210.

[29] S. Clark, F. Gesztesy, and W. Renger, *Trace formulas and Borg-type theorems for matrix-valued Jacobi and Dirac finite difference operators*, J. Diff. Eqs. **219** (2005), 144–182.

[30] S. Clark, F. Gesztesy, and M. Zinchenko, *Weyl–Titchmarsh theory and Borg–Marchenko-type uniqueness results for CMV operators with matrix-valued Verblunsky coefficients*, Operators and Matrices **1** (2007), 535–592.

[31] W. Craig and B. Simon, *Log Hölder continuity of the integrated density of states for stochastic Jacobi matrices*, Comm. Math. Phys. **90** (1983), 207–218.

[32] D. Damanik and R. Killip, *Half-line Schrödinger operators with no bound states*, Acta Math. **193** (2004), 31–72.

[33] D. Damanik, R. Killip, and B. Simon, *Perturbations of orthogonal polynomials with periodic recursion coefficients*, preprint.

[34] P. Dean and J. L. Martin, *A method for determining the frequency spectra of disordered lattices in two-dimensions*, Proc. Roy. Soc. (London) A **259** (1960), 409–418.

[35] A. M. Delgado, J. S. Geronimo, P. Iliev, and F. Marcellán, *Two variable orthogonal polynomials and structured matrices*, SIAM J. Matrix Anal. Appl. **28** (2006), 118–147 [electronic].

[36] P. Delsarte and Y. V. Genin, *On a generalization of the Szegő–Levinson recurrence and its application in lossless inverse scattering*, IEEE Trans. Inform. Theory **38** (1992), 104–110.

[37] P. Delsarte, Y. V. Genin, and Y. G. Kamp, *Orthogonal polynomial matrices on the unit circle*, IEEE Trans. Circuits and Systems **CAS-25** (1978), 149–160.

[38] P. Delsarte, Y. V. Genin, and Y. G. Kamp, *Planar least squares inverse polynomials, I. Algebraic properties*, IEEE Trans. Circuits and Systems **CAS-26** (1979), 59–66.

[39] P. Delsarte, Y. V. Genin, and Y. G. Kamp, *The Nevanlinna–Pick problem for matrix-valued functions*, SIAM J. Appl. Math. **36** (1979), 47–61.

[40] P. Delsarte, Y. V. Genin, and Y. G. Kamp, *Schur parametrization of positive definite block-Toeplitz systems*, SIAM J. Appl. Math. **36** (1979), 34–46.

[41] P. Delsarte, Y. V. Genin, and Y. G. Kamp, *Generalized Schur representation of matrix-valued functions*, SIAM J. Algebraic Discrete Methods **2** (1981), 94–107.

[42] S. Denisov, *On Rakhmanov's theorem for Jacobi matrices*, Proc. Amer. Math. Soc. **132** (2004), 847–852.

[43] H. Dette and W. J. Studden, *Matrix measures, moment spaces and Favard's theorem for the interval* $[0,1]$ *and* $[0,\infty)$, Linear Algebra Appl. **345** (2002), 169–193.

[44] V. Dubovoj, B. Fritzsche, and B. Kirstein, *Matricial Version of the Classical Schur Problem*, Teubner Texts in Mathematics **129**, Teubner Verlag, Stuttgart (1992).

[45] A. J. Durán, *A generalization of Favard's theorem for polynomials satisfying a recurrence relation*, J. Approx. Theory **74** (1993), 83–109.

[46] A. J. Durán, *On orthogonal polynomials with respect to a positive definite matrix of measures*, Canad. J. Math. **47** (1995), 88–112.

[47] A. J. Durán, *Markov's theorem for orthogonal matrix polynomials*, Canad. J. Math. **48** (1996), 1180–1195.

[48] A. J. Durán, *Matrix inner product having a matrix symmetric second order differential operator*, Rocky Mountain J. Math. **27** (1997), 585–600.

[49] A. J. Durán, *Ratio asymptotics for orthogonal matrix polynomials*, J. Approx. Theory **100** (1999), 304–344.

[50] A. J. Durán, *How to find weight matrices having symmetric second order differential operators with matrix leading coefficient*, preprint.

[51] A. J. Durán, *Generating weight matrices having symmetric second order differential operators from a trio of upper triangular matrices*, preprint.

[52] A. J. Durán and E. Daneri, *Ratio asymptotics for orthogonal matrix polynomials with unbounded recurrence coefficients*, J. Approx. Theory **110** (2001), 1–17.

[53] A. J. Durán and E. Daneri, *Weak convergence for orthogonal matrix polynomials*, Indag. Math. (N.S.) **13** (2002), 47–62.

[54] A. J. Durań and E. Defez, *Orthogonal matrix polynomials and quadrature formulas*, Linear Algebra Appl. **345** (2002), 71–84.

[55] A. J. Durán and F. A. Grünbaum, *Orthogonal matrix polynomials satisfying second order differential equations*, Int. Math. Res. Not. **2004:10** (2004), 461–484.

[56] A. J. Durán and F. A. Grünbaum, *A survey on orthogonal matrix polynomials satisfying second order differential equations*, J. Comput. Appl. Math. **178** (2005), 169–190.

[57] A. J. Durán and F. A. Grünbaum, *Structural formulas for orthogonal matrix polynomials satysfying second order differential equations, I*, Constr. Approx. **22** (2005), 255–271.

[58] A. J. Durán and F. A. Grünbaum, *Orthogonal matrix polynomials, scalar type Rodrigues' formulas and Pearson equations*, J. Approx. Theory. **134** (2005), 267–280.

[59] A. J. Durán and F. A. Grünbaum, *A characterization for a class of weight matrices with orthogonal matrix polynomials satisfying second order differential equations*, Int. Math. Res. Not. **23** (2005), 1371–1390.

[60] A. J. Durán and F. A. Grünbaum, *P. A. M. Dirac meets M. G. Krein: matrix orthogonal polynomials and Diracs equation*, J. Phys. A: Math. Gen. **39** (2006), 3655–3662.

[61] A. J. Durán and F. A. Grünbaum, *Matrix orthogonal polynomials satisfying second-order differential equations: Coping without help from group representation theory*, J. Approx. Theory **148** (2007), 35–48.

[62] A. J. Durán and F. A. Grünbaum, *Matrix differential equations and scalar polynomials satisfying higher order recursions*, preprint.

[63] A. Duran, F. A. Grünbaum, I. Pacharoni, and J. A. Tirao, *Matrix orthogonal polynomials and differential operators, with examples and applications*, in preparation.

[64] A. J. Durán and M. D. de la Iglesia, *Some examples of orthogonal matrix polynomials satisfying odd order differential equations*, to appear in J. Approx. Theory.

[65] A. J. Durán and M. D. de la Iglesia, *Second order differential operators having several families of orthogonal matrix polynomials as eigenfunctions*, preprint.

[66] A. J. Durán and P. López-Rodríguez, *Orthogonal matrix polynomials: zeros and Blumenthal's theorem*, J. Approx. Theory **84** (1996), 96–118.

[67] A. J. Durán and P. López-Rodríguez, *The L^p space of a positive definite matrix of measures and density of matrix polynomials in L^1*, J. Approx. Theory **90** (1997), 299–318.

[68] A. J. Durán and P. López-Rodríguez, *Density questions for the truncated matrix moment problem*, Canad. J. Math. **49** (1997), 708–721.

[69] A. J. Durán and P. López-Rodríguez, *N-extremal matrices of measures for an indeterminate matrix moment problem*, J. Funct. Anal. **174** (2000), 301–321.

[70] A. J. Durán and P. López-Rodríguez, *The matrix moment problem*, Margarita Mathematica en memoria de José Javier Guadalupe (L. Español and J. L. Varona, eds.), pp. 333–348, Universidad de La Rioja, Logroño (2001).

[71] A. J. Durán and P. López-Rodríguez, *Orthogonal matrix polynomials*, Laredo's SIAG Lecture Notes, (R. Álvarez-Nodarse et al., eds.), pp. 11–43, Advances in the Theory of Special Functions and Orthogonal Polynomials, Nova Science Publishers **1** (2003).

[72] A. J. Durán and P. López-Rodríguez, *Structural formulas for orthogonal matrix polynomials satisfying second order differential equations, I*, Constr. Approx. **26** (2007), 29–47.

[73] A. J. Durán and P. López-Rodríguez, *Structural formulas for orthogonal matrix polynomials satisfying second order differential equations, II*, Constr. Approx. **26** (2007) 29–47.

[74] A. J. Durán, P. López-Rodríguez, and E. B. Saff, *Zero asymptotic behaviour for orthogonal matrix polynomials*, J. d'Analyse Math. **78** (1999), 37–60.

[75] A. J. Durán and B. Polo, *Gaussian quadrature formulae for matrix weights*, Linear Algebra Appl. **355** (2002), 119–146.

[76] A. J. Durán and B. Polo, *Matrix Christoffel functions*, Constr. Approx. **20** (2004), 353–376.

[77] A. J. Durán and W. Van Assche, *Orthogonal matrix polynomials and higher-order recurrence relations*, Linear Algebra Appl. **219** (1995), 261–280.

[78] H. Dym, *J Contractive Matrix Functions, Reproducing Kernel Hilbert Spaces and Interpolation*, CBMS Regional Conference Series in Math. **71**, American Mathematical Society, Providence, R.I. (1989).

[79] Yu. M. Dyukarev, *Deficiency numbers of symmetric operators generated by block Jacobi matrices*, Sb. Math. **197** (2006), 1177–1203.

[80] T. Erdélyi, P. Nevai, J. Zhang, and J. Geronimo, *A simple proof of "Favard's theorem" on the unit circle*, Atti Sem. Mat. Fis. Univ. Modena **39** (1991), 551–556.

[81] L. Faybusovich and M. Gekhtman, *On Schur flows*, J. Phys. A **32** (1999), 4671–4680.

[82] G. Freud, *Orthogonal Polynomials*, Pergamon Press, Oxford-New York (1971).

[83] P. A. Fuhrmann, *Orthogonal matrix polynomials and system theory*, Rend. Sem. Mat. Univ. Politec. Torino **1987**, Special Issue, 68–124.

[84] J. S. Geronimo, *Matrix orthogonal polynomials on the unit circle*, J. Math. Phys. **22** (1981), 1359–1365.

[85] J. S. Geronimo, *Scattering theory and matrix orthogonal polynomials on the real line*, Circuits Systems Signal Process **1** (1982), 471–495.

[86] J. S. Geronimo, *Polynomials orthogonal on the unit circle with random recurrence coefficients*, Methods of Approximation Theory in Complex Analysis and Mathematical Physics (Leningrad, 1991), pp. 43–61, Lecture Notes in Math. **1550**, Springer, Berlin (1993).

[87] J. S. Geronimo and K. M. Case, *Scattering theory and polynomials orthogonal on the unit circle*, J. Math. Phys. **20** (1979), 299–310.

[88] J. S. Geronimo and H. Woerdeman, *Two variable orthogonal polynomials on the bicircle and structured matrices*, SIAM J. Matrix Anal. Appl. **29** (2007), 796–825.

[89] Ya. L. Geronimus, *On polynomials orthogonal on the circle, on trigonometric moment problem, and on allied Carathéodory and Schur functions*, Mat. Sb. **15** (1944), 99–130 [Russian].

[90] Ya. L. Geronimus, *On the trigonometric moment problem*, Ann. of Math. (2) **47** (1946), 742–761.

[91] Ya. L. Geronimus, *Polynomials Orthogonal on a Circle and Their Applications*, Amer. Math. Soc. Translation **1954** (1954), no. 104, 79 pp.

[92] Ya. L. Geronimus, *Orthogonal Polynomials: Estimates, Asymptotic Formulas, and Series of Polynomials Orthogonal on the Unit Circle and on an Interval*, Consultants Bureau, New York (1961).

[93] Ya. L. Geronimus, *Orthogonal polynomials*, Engl. translation of the appendix to the Russian translation of Szegő's book [184], in "Two Papers on Special Functions," Amer. Math. Soc. Transl., Ser. 2, Vol. 108, pp. 37–130, American Mathematical Society, Providence, R.I. (1977).

[94] F. Gesztesy, N. J. Kalton, K. A. Makarov, and E. Tsekanovskii, *Some applications of operator-valued Herglotz functions*, Operator Theory, System Theory and Related Topics (Beer-Sheva/Rehovot, 1997), pp. 271–321, Oper. Theory Adv. Appl. **123**, Birkhäuser, Basel (2001).

[95] F. Gesztesy, A. Kiselev, and K. Makarov, *Uniqueness results for matrix-valued Schrödinger, Jacobi, and Dirac-type operators*, Math. Nachr. **239/240** (2002), 103–145.

[96] F. Gesztesy and L. A. Sakhnovich, *A class of matrix-valued Schrödinger operators with prescribed finite-band spectra*, Reproducing Kernel Hilbert Spaces, Positivity, System Theory and Related Topics (D. Alpay, ed.), pp. 213–253, Operator Theory: Advances and Applications **143**, Birkhäuser, Basel (2003).

[97] F. Gesztesy and B. Simon, *On local Borg–Marchenko uniqueness results*, Comm. Math. Phys. **211** (2000), 273–287.

[98] F. Gesztesy and E. Tsekanovskii, *On matrix-valued Herglotz functions*, Math. Nachr. **218** (2000), 61–138.

[99] I. Gohberg and M. Krein, *Introduction to the Theory of Linear Nonselfadjoint Operators*, Translations of Mathematical Monographs **18**, American Mathematical Society, Providence, R.I. (1969).

[100] L. Golinskii, *Quadrature formula and zeros of para-orthogonal polynomials on the unit circle*, Acta Math. Hungar. **96** (2002), 169–186.

[101] L. Golinskii and S. Khrushchev, *Cesàro asymptotics for orthogonal polynomials on the unit circle and classes of measures*, J. Approx. Theory **115** (2002), 187–237.

[102] F. A. Grünbaum, *Matrix valued Jacobi polynomials*, Bull. Sciences Math. **127** (2003), 207–214.

[103] F. A. Grünbaum, *Random walks and orthogonal polynomials: some challenges*, preprint.

[104] F. A. Grünbaum and M. D. de la Iglesia, *Matrix valued orthogonal polynomials related to* $SU(N+1)$*, their algebras of differential operators and the corresponding curves*, Exp. Math. **16** (2007), 189–207.

[105] F. A. Grünbaum and M. D. de la Iglesia, *Matrix valued orthogonal polynomials arising from group representation theory and a family of quasi-birth-and-death processes*, preprint.

[106] F. A. Grünbaum and P. Iliev, *A noncommutative version of the bispectral problem*, J. Comput. Appl. Math. **161** (2003), 99–118.

[107] F. A. Grünbaum, I. Pacharoni, and J. A. Tirao, *A matrix valued solution to Bochner's problem*, J. Phys. A: Math. Gen. **34** (2001), 10647–10656.

[108] F. A. Grünbaum, I. Pacharoni, and J. A. Tirao, *Matrix valued spherical functions associated to the three dimensional hyperbolic space*, Internat. J. of Math. **13** (2002), 727–784.

[109] F. A. Grünbaum, I. Pacharoni, and J. A. Tirao, *Matrix valued spherical functions associated to the complex projective plane*, J. Funct. Anal. **188** (2002), 350–441.

[110] F. A. Grünbaum, I. Pacharoni, and J. A. Tirao, *Matrix valued orthogonal polynomials of the Jacobi type*, Indag. Math. **14** (2003), 353–366.

[111] F. A. Grünbaum, I. Pacharoni, and J. A. Tirao, *An invitation to matrix-valued spherical functions: Linearization of products in the case of complex projective space* $P_2(\mathbb{C})$, Modern Signal Processing, pp. 147–160, Math. Sci. Res. Inst. Publ. **46**, Cambridge University Press, Cambridge (2004).

[112] F. A. Grünbaum, I. Pacharoni, and J. A. Tirao, *Matrix valued orthogonal polynomials of the Jacobi type: The role of group representation theory*, Ann. Inst. Fourier (Grenoble) **55** (2005), 2051–2068.

[113] F. A. Grünbaum and J. A. Tirao, *The algebra of differential operators associated to a weight matrix*, Integr. Equ. Oper. Th. **58** (2007), 449–475.

[114] S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces*, corrected reprint of the 1978 original, Graduate Studies in Mathematics **34**, American Mathematical Society, Providence, R.I. (2001).

[115] L. L. Helms, *Introduction to Potential Theory*, Pure and Applied Mathematics **22**, Wiley-Interscience, New York (1969).

[116] H. Helson and D. Lowdenslager, *Prediction theory and Fourier series in several variables*, Acta Math. **99** (1958), 165–202.

[117] H. Helson and D. Lowdenslager, *Prediction theory and Fourier series in several variables, II*, Acta Math. **106** (1961), 175–213.

[118] M. E. H. Ismail, *Classical and Quantum Orthogonal Polynomials in One Variable*, Encyclopedia of Mathematics and its Applications **98**, Cambridge University Press, Cambridge (2005).

[119] W. B. Jones, O. Njåstad, and W. J. Thron, *Moment theory, orthogonal polynomials, quadrature, and continued fractions associated with the unit circle*, Bull. London Math. Soc. **21** (1989), 113–152.

[120] T. Kailath, *A view of three decades of linear filtering theory*, IEEE Trans. Inform. Theory **IT-20** (1974), 146–181.

[121] T. Kailath, *Signal processing applications of some moment problems*, Moments in Mathematics, (San Antonio, Tex., 1987), Proc. Sympos. Appl. Math. **37**, pp. 71–109, American Mathematical Society, Providence, R.I. (1987).

[122] T. Kailath, *Norbert Wiener and the development of mathematical engineering*, The Legacy of Norbert Wiener: A Centennial Symposium, Proc. Sympos. Pure Math. **60**, pp. 93–116, American Mathematical Society, Providence, R.I. (1997).

[123] T. Kailath, A. Vieira, and M. Morf, *Inverses of Toeplitz operators, innovations, and orthogonal polynomials*, SIAM Rev. **20** (1978), 106–119.

[124] T. Kato, *Perturbation Theory for Linear Operators*, second edition, Grundlehren der Mathematischen Wissenschaften **132**, Springer, Berlin-New York (1976).

[125] S. Khrushchev, *Schur's algorithm, orthogonal polynomials, and convergence of Wall's continued fractions in* $L^2(\mathbb{T})$, J. Approx. Theory **108** (2001), 161–248.

[126] S. Khrushchev, *Classification theorems for general orthogonal polynomials on the unit circle*, J. Approx. Theory **116** (2002), 268–342.

[127] R. Killip and I. Nenciu, *Matrix models for circular ensembles*, Int. Math. Res. Not. **2004:50** (2004), 2665–2701.

[128] R. Killip and B. Simon, *Sum rules for Jacobi matrices and their applications to spectral theory*, Ann. of Math. **158** (2003), 253–321.

[129] A. N. Kolmogorov, *Stationary sequences in Hilbert space*, Bull. Univ. Moscow **2** (1941), 40 pp. [Russian].

[130] S. Kotani and B. Simon, *Stochastic Schrödinger operators and Jacobi matrices on the strip*, Comm. Math. Phys. **119** (1988), 403–429.

[131] M. G. Krein, *On a generalization of some investigations of G. Szegő, W. M. Smirnov, and A. N. Kolmogorov*, Dokl. Akad. Nauk SSSR **46** (1945), 91–94.

[132] M. G. Krein, *On a problem of extrapolation of A. N. Kolmogorov*, Dokl. Akad. Nauk SSSR **46** (1945), 306–309.

[133] M. G. Krein, *Infinite J-matrices and a matrix moment problem*, Dokl. Akad. Nauk SSSR **69** (1949), 125–128 [Russian].

[134] J. Lacroix, *The random Schrödinger operator in a strip*, Probability Measures on Groups, VII (Oberwolfach, 1983), pp. 280–297, Lecture Notes in Math. **1064**, Springer, Berlin (1984).

[135] H. J. Landau, *Maximum entropy and the moment problem*, Bull. Amer. Math. Soc. **16** (1987), 47–77.

[136] N. S. Landkof, *Foundations of Modern Potential Theory*, Springer-Verlag, Berlin-New York (1972).

[137] L. Lerer and A. C. M. Ran, *A new inertia theorem for Stein equations, inertia of invertible Hermitian block Toeplitz matrices and matrix orthogonal polynomials*, Integr. Equ. Oper. Th. **47** (2003), 339–360.

[138] N. Levinson, *The Wiener RMS (root-mean square) error criterion in filter design and prediction*, J. Math. Phys. Mass. Inst. Tech. **25** (1947), 261–278.

[139] P. López-Rodríguez, *Riesz's theorem for orthogonal matrix polynomials*, Constr. Approx. **15** (1999), 135–151.

[140] P. López-Rodríguez, *Nevanlinna parametrization for a matrix moment problem*, Math. Scand. **89** (2001), 245–267.

[141] D. S. Lubinsky, *A new approach to universality limits involving orthogonal polynomials*, to appear in Ann. of Math.

[142] F. Marcellán and I. Rodriguez, *A class of matrix orthogonal polynomials on the unit circle*, Linear Algebra Appl. **121**, (1989), 233–241.

[143] F. Marcellán and G. Sansigre, *On a class of matrix orthogonal polynomials on the real line*, Linear Algebra Appl. **181** (1993), 97–109.

[144] F. Marcellán and H. O. Yakhlef, *Recent trends on analytic properties of matrix orthonormal polynomials*, Orthogonal Polynomials, Approximation Theory, and Harmonic Analysis (Inzel, 2000), Electron. Trans. Numer. Anal. **14** (2002), 127–141 [electronic].

[145] F. Marcellán and S. M. Zagorodnyuk, *On the basic set of solutions of a high-order linear difference equation*, J. Diff. Equ. Appl. **12** (2006), 213–228.

[146] A. Máté and P. Nevai, *Bernstein's inequality in L^p for $0 < p < 1$ and $(C, 1)$ bounds for orthogonal polynomials*, Ann. of Math. (2) **111** (1980), 145–154.

[147] A. Máté, P. Nevai, and V. Totik, *Strong and weak convergence of orthogonal polynomials*, Am. J. Math. **109** (1987), 239–281.

[148] L. Miranian, *Matrix valued orthogonal polynomials*, Ph.D. Thesis, University of California at Berkeley (2005).

[149] L. Miranian, *Matrix valued orthogonal polynomials on the real line: Some extensions of the classical theory*, J. Phys. A: Math. Gen. **38** (2005), 5731–5749.

[150] L. Miranian, *On classical orthogonal polynomials and differential operators*, J. Phys. A: Math. Gen. **38** (2005), 6379–6383.

[151] S. Molchanov and B. Vainberg, *Schrödinger operators with matrix potentials. Transition from the absolutely continuous to the singular spectrum*, J. Funct. Anal. **215** (2004), 111–129.

[152] P. Nevai, *Géza Freud, orthogonal polynomials and Christoffel functions. A case study*, J. Approx. Theory **48** (1986), 167 pp.

[153] E. M. Nikishin, *The discrete Sturm–Liouville operator and some problems of function theory*, J. Soviet Math. **35** (1987), 2679–2744; Russian original in Trudy Sem. Petrovsk. **10** (1984), 3–77.

[154] I. Pacharoni and P. Román, *A sequence of matrix valued orthogonal polynomials associated to spherical functions*, Constr. Approx. **28** (2008), 127–147.

[155] I. Pacharoni and J. A. Tirao, *Three term recursion relation for spherical functions associated to the complex projective plane*, Math. Phys. Anal. Geom. **7** (2004), 193–221.

[156] I. Pacharoni and J. A. Tirao, *Matrix valued orthogonal polynomials arising from the complex projective space*, Constr. Approx. **25** (2007), 177–192.

[157] I. Pacharoni and J. A. Tirao, *Matrix valued orthogonal polynomials associated to the group $SU(N)$*, in preparation.

[158] T. Ransford, *Potential Theory in the Complex Plane*, Press Syndicate of the Univesity of Cambridge, New York (1995).

[159] M. Reed and B. Simon, *Methods of Modern Mathematical Physics, IV: Analysis of Operators*, Academic Press, New York (1978).

[160] L. Rodman, *Orthogonal matrix polynomials*, Orthogonal Polynomials (Columbus, OH, 1989), pp. 345–362, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. **294**, Kluwer, Dordrecht (1990).

[161] P. Román and J. A. Tirao, *Spherical functions, the complex hyperbolic plane and the hypergeometric operator*, Internat. J. Math. **17** (2006), 1151–1173.

[162] M. Rosenberg, *The square integrability of matrix valued functions with respect to a non-negative Hermitian measure*, Duke Math. J. **31** (1964), 291–298.

[163] I. Schur, *Über Potenzreihen, die im Innern des Einheitskreises beschr änkt sind, I, II*, J. Reine Angew. Math. **147** (1917), 205–232; **148** (1918), 122–145; English translation in "I. Schur Methods in Operator Theory and Signal Processing" (I. Gohberg, ed.), pp. 31–59, 66–88, Operator Theory: Advances and Applications **18**, Birkhäuser, Basel (1986).

[164] H. Schulz-Baldes, *Perturbation theory for Lyapunov exponents of an Anderson model on a strip*, Geom. Funct. Anal. **14** (2004), 1089–1117.

[165] H. Schulz-Baldes, *Rotation numbers for Jacobi matrices with matrix entries*, Math. Phys. Electron. J. **13** (2007), 40 pp.

[166] B. Schwarz and A. Zaks, *Matrix Möbius transformations*, Comm. Algebra **9** (1981), 1913–1968.

[167] B. Simon, *Orthogonal Polynomials on the Unit Circle. Part 1. Classical Theory*, Colloquium Publications **54.1**, American Mathematical Society, Providence, R.I. (2005).

[168] B. Simon, *Orthogonal Polynomials on the Unit Circle. Part 2. Spectral Theory*, Colloquium Publications **54.2**, American Mathematical Society, Providence, R.I. (2005).

[169] B. Simon, *Trace Ideals and Their Applications*, second edition, Mathematical Surveys and Monographs **120**, American Mathematical Society, Providence, R.I. (2005).

[170] B. Simon, *CMV matrices: Five years after*, J. Comput. Appl. Math. **208** (2007), 120–154.

[171] B. Simon, *Equilibrium measures and capacities in spectral theory*, Inverse Problems and Imaging **1** (2007), 376-382.

[172] B. Simon, *Rank one perturbations and the zeros of paraorthogonal polynomials on the unit circle*, J. Math. Anal. Appl. **329** (2007), 376–382.

[173] B. Simon, *Two extensions of Lubinsky's universality theorem*, to appear in J. Anal. Math.

[174] B. Simon, *Weak convergence of CD kernels and applications*, preprint.

[175] B. Simon, *The Christoffel–Darboux kernel*, in preparation, to appear in Maz'ya Birthday volume.

[176] B. Simon, *Szegő's Theorem and Its Descendants: Spectral Theory for L^2 Perturbations of Orthogonal Polynomials*, in preparation; to be published by Princeton University Press.

[177] A. Sinap, *Gaussian quadrature for matrix valued functions on the real line*, J. Comput. Appl. Math. **65** (1995), 369–385.

[178] A. Sinap and W. Van Assche, *Orthogonal matrix polynomials and applications*, J. Comput. Appl. Math. **66** (1996), 27–52.

[179] V. N. Sorokin and J. Van Iseghem, *Matrix continued fractions*, J. Approx. Theory **96** (1999), 237–257.

[180] H. Stahl and V. Totik, *General Orthogonal Polynomials*, Encyclopedia of Mathematics and its Applications **43**, Cambridge University Press, Cambridge (1992).

[181] T. Stieltjes, *Recherches sur les fractions continues*, Ann. Fac. Sci. Univ. Toulouse **8** (1894–1895), J76–J122; ibid. **9**, A5–A47.

[182] G. Szegő, *Über den asymptotischen Ausdruck von Polynomen, die durch eine Orthogonalitätseigenschaft definiert sind*, Math. Ann. **86** (1922), 114–139.

[183] G. Szegő, *Bemerkungen zu einer Arbeit von Herrn M. Fekete: Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten*, Math. Z. **21** (1924), 203–208.

[184] G. Szegő, *Orthogonal Polynomials*, Amer. Math. Soc. Colloq. Publ. **23**, American Mathematical Society, Providence, R.I. (1939); third edition, (1967).

[185] G. Teschl, *Jacobi Operators and Completely Integrable Nonlinear Lattices*, Mathematical Surveys and Monographs **72**, American Mathematical Society, Providence, R.I. (2000).

[186] J. A. Tirao, *The matrix valued hypergeometric equation*, Proc. Nat. Acad. Sci. U.S.A. **100** (2003), 8138–8141.

[187] V. Totik, *Asymptotics for Christoffel functions for general measures on the real line*, J. Anal. Math. **81** (2000), 283–303.

[188] V. Totik, *Polynomial inverse images and polynomial inequalities*, Acta Math. **187** (2001), 139–160.

[189] V. Totik, *Universality and fine zero spacing on general sets*, in preparation.

[190] J. L. Ullman, *On the regular behaviour of orthogonal polynomials*, Proc. London Math. Soc. (3) **24** (1972), 119–148.

[191] W. Van Assche, *Rakhmanov's theorem for orthogonal matrix polynomials on the unit circle*, J. Approx. Theory **146** (2007), 227–242.

[192] S. Verblunsky, *On positive harmonic functions: A contribution to the algebra of Fourier series*, Proc. London Math. Soc. (2) **38** (1935), 125–157.

[193] H. Weyl, *Über gewöhnliche Differentialgleichungen mit Singularitäten und die zugehörigen Entwicklungen willkürlicher Funktionen*, Math. Ann. **68** (1910), 220–269.

[194] P. Whittle, *On the fitting of multivariate autoregressions and the approximate canonical factorization of a spectral density matrix*, Biometrika **50** (1963), 129–134.

[195] H. Widom, *Polynomials associated with measures in the complex plane*, J. Math. Mech. **16** (1967), 997–1013.

[196] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series. With Engineering Applications*, The Technology Press of the Massachusetts Institute of Technology, Cambridge, Mass. (1949).

[197] M.-W. L. Wong, *First and second kind paraorthogonal polynomials and their zeros*, J. Approx. Theory **146** (2007), 282–293.

[198] H. O. Yakhlef, *Relative asymptotics for orthogonal matrix polynomials with unbounded recurrence coefficients*, Integral Transforms Spec. Funct. **18** (2007), 39–57.

[199] H. O. Yakhlef and F. Marcellán, *Orthogonal matrix polynomials, connection between recurrences on the unit circle and on a finite interval*, Approximation, Optimization and Mathematical Economics (Pointe-à-Pitre, 1999), pp. 369–382, Physica, Heidelberg (2001).

[200] H. O. Yakhlef and F. Marcellán, *Relative asymptotics for orthogonal matrix polynomials with respect to a perturbed matrix measure on the unit circle*, Approx. Theory Appl. (N.S.) **18** (2002), 1–19.

[201] H. O. Yakhlef, F. Marcellán, and M. A. Piñar, *Perturbations in the Nevai matrix class of orthogonal matrix polynomials*, Linear Algebra Appl. **336** (2001), 231–254.

[202] H. O. Yakhlef, F. Marcellán, and M. A. Piñar, *Relative asymptotics for orthogonal matrix polynomials with convergent recurrence coefficients*, J. Approx. Theory **111** (2001), 1–30.

[203] D. C. Youla and N. N. Kazanjian, *Bauer-type factorization of positive matrices and the theory of matrix polynomials orthogonal on the unit circle*, IEEE Trans. Circuits and Systems **CAS-25** (1978), 57–69.

[204] M. Zygmunt, *Matrix Chebyshev polynomials and continued fractions*, Linear Algebra Appl. **340** (2002), 155–168.

David Damanik
Department of Mathematics
Rice University
Houston, TX 77005, USA
damanik@rice.edu

Alexander Pushnitski
Department of Mathematics
King's College London
Strand, London WC2R 2LS, UK
alexander.pushnitski@kcl.ac.uk

Barry Simon
Mathematics 253–37
California Institute of Technology
Pasadena, CA 91125, USA
bsimon@caltech.edu