BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 42, Number 4, Pages 431–460
S 0273-0979(05)01075-X
Article electronically published on June 23, 2005

OPUC ON ONE FOOT

BARRY SIMON

Abstract. We present an expository introduction to orthogonal polynomials on the unit circle (OPUC).

1. Introduction

Orthogonal polynomials are the Rodney Dangerfield [112] of analysis. Because of the impact of Stieltjes' great 1895 paper on F. Riesz, Nevanlinna, and Hilbert's school, the moment problem and the closely related subject of orthogonal polynomials on the real line (OPRL) were central in the revolution in analysis from 1900 to 1920 and provided critical precursors to the Hahn-Banach theorem, the Riesz-Markov theorem, the spectral theorem, and the theory of selfadjoint extensions. But in recent years, too often the subject is dismissed as "classical" and not worthy of further study.

With developments in random matrix theory and combinatorics (e.g., [4, 5, 6, 7, 12, 46, 65, 82]), it is clear that orthogonal polynomials still have a lot to contribute. From one point of view, what makes them relevant is that they are the simplest of inverse spectral problems — indeed, Gel'fand-Levitan [28] explicitly note that their approach to inverse theory for Schrödinger operators is motivated by OPRL. Recently, OPUC ideas have provided a matrix realization of Lax pairs for the (defocusing) Ablowitz-Ladik equation [67].

What is true for OPRL is even more true for orthogonal polynomials on the unit circle (OPUC). While the closely related area of positive harmonic functions on the open unit disk, $\mathbb{D} = \{z \in \mathbb{C} \mid |z| < 1\}$, drew the attention of Carathéodory, Fejér, Herglotz, F. Riesz, Schur, and Toeplitz in the second decade of the 1900's, the subject was invented by Szegő only about 1920, especially in his deep 1920–1921 paper [99]. So OPUC never had its era of centrality but has had a steady but small following over the years. Traditionally, the book references for the subject were Szegő's book [101], which has only one full and several partial chapters on OPUC, Geronimus' book [36] and review [34], and a chapter in Freud [27], all of which are very dated. With a major development published only in 2003 (the CMV matrix of Section 5 below), it is hard not to be dated. Motivated by this dearth of review literature and by the opportunity to use Schrödinger operator techniques in a new setting, I published two volumes [91, 92] on the subject. Many friends asked if

Received by the editors February 2, 2005, and, in revised form, April 19, 2005.
2000 *Mathematics Subject Classification.* Primary 42C05, 30E05, 42A70.
Key words and phrases. Orthogonal polynomials, Verblunsky coefficients, Szegő's theorem.
Supported in part by NSF grant DMS-0140592.

there wasn't some way to learn about the subject in less than 1,100 pages, and this expository note is the result.

Throughout, we use $\mathbb{D}$ for the open unit disk in $\mathbb{C}$, and $\partial\mathbb{D}$ for the unit circle. Our inner products, $\langle f, g\rangle$, are linear in g and antilinear in f. Significant missing material involves some explicit examples — these are discussed in Section 1.6 of [91]: my favorite is the Rogers-Szegő polynomials (Example 1.6.5). This article undergoes a kind of phase transition in the middle of Section 5 in that before this section, most results have proofs or at least sketches given, and afterwards there aren't many proofs. This is because the earlier material is more central and also because the later proofs are lengthier.

To put OPUC in context, recall some basics of OPRL. Since the fascinating issues of indeterminate moment problems (see [1, 88]) are irrelevant to OPUC, we will assume all measures have compact support. A measure is called *trivial* if it is supported on a finite set of points and *nontrivial* if the support is infinite.

(1) If μ is a nontrivial probability measure on $\mathbb{C}$ (i.e., positive with $\mu(\mathbb{C}) = 1$) with compact support and $X_n(z)$ are the monic orthogonal polynomials (i.e., $X_n(z) = z^n +$ lower order, $X_n \perp z^\ell$, $\ell = 0, \dots, n-1$), then

$$zX_n(z) = X_{n+1}(z) + \sum_{j=0}^{n} a_j^{(n)} X_j(z) \tag{1.1}$$

$$a_j^{(n)} = \frac{\langle X_j, zX_n\rangle}{\|X_j\|^2}. \tag{1.2}$$

What makes OPRL special is that multiplication by x is selfadjoint, so if we use P_n in place of X_n for OPRL and ρ for μ,

$$\langle P_j, xP_n\rangle = \langle xP_j, P_n\rangle = 0 \qquad j = 0, \dots, n-2$$

and thus (1.2) becomes

$$xP_n(x) = P_{n+1}(x) + b_{n+1}P_n(x) + a_n^2 P_{n-1}(x) \tag{1.3}$$

for *Jacobi parameters*, a_n, b_n; $n = 1, 2, \dots$ (a more common convention is to start the numbering at $n = 0$). If $p_n = P_n/\|P_n\|$ are the orthonormal OPRL, the matrix elements of multiplication by x in p_n basis have the form:

$$J = \begin{pmatrix} b_1 & a_1 & 0 & 0 & \dots \\ a_1 & b_2 & a_2 & 0 & \dots \\ 0 & a_2 & b_3 & a_3 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}. \tag{1.4}$$

(2) There is a one-one correspondence between bounded J's (i.e., $\sup_n(|a_n| + |b_n|) < \infty$) and measures, ρ, on $\mathbb{R}$ with compact but infinite support. This is sometimes called Favard's theorem.

(3) If A is a bounded selfadjoint operator on a separable Hilbert space, $\mathcal{H}$, and φ is a cyclic unit vector (i.e., $\{A^n\varphi\}_{n=0}^\infty$ span $\mathcal{H}$), one can use the spectral theorem to find a measure $d\rho$ on $[-\|A\|, \|A\|]$ with $\int x^n \, d\rho = \langle\varphi, A^n\varphi\rangle$ and then the OPRL for this measure to find a semi-infinite Jacobi matrix unitarily equivalent to A with φ mapped to $(1\,0\,0\,\dots)^t$. This realization is unique; that is, the a_n's and b_n's are intrinsic to the pair (A, φ). It was Stone who emphasized this point of view that the study of Jacobi matrices was the same as the study of selfadjoint operators with a distinguished cyclic vector.

(4) A key role is played by the Stieltjes transform of ρ, that is, the function, m, on $\mathbb{C}\backslash\operatorname{supp}(d\rho)$ given by

$$m(z) = \int \frac{d\rho(x)}{x-z}. \tag{1.5}$$

(5) The Jacobi parameters can also be captured from $m(z)$ via a continued fraction expansion (of Stieltjes) at ∞:

$$m(z) = \cfrac{1}{-z + b_1 - \cfrac{a_1^2}{-z + b_2 - a_2^2 \dots}}. \tag{1.6}$$

We will not discuss applications of OPUC in detail but note its important applications to linear prediction and filtering theory. The basics are due to Wiener [111], Kolmogorov [54], Krein [55, 56], and Levinson [60]. The ideas have been especially developed by Kailath [48, 49, 50].

The title of this article is based on an incident reported in the Talmud [103] that someone asked the famous first-century rabbi Hillel to describe Judaism to him while he stood on one foot. Hillel's answer was: "Do not do unto others that which is hateful to you. The rest is commentary. Go forth and study." This article is OPUC on one foot. [91, 92] are commentary.

It is a pleasure to thank M. Aizenman for pushing me to write such an article. I would like to thank S. Denisov, F. Gesztesy, L. Golinskii, R. Killip, D. Lubinsky, F. Marcellán, P. Nevai, and G. Stolz for useful input. This paper was started while I was a visitor at the Courant Institute and completed during my stay as a Lady Davis Visiting Professor at Hebrew University, Jerusalem. I would like to thank P. Deift and C. Newman for the hospitality of Courant and H. Farkas and Y. Last for the hospitality of the Mathematics Institute at Hebrew University.

2. The Szegő recursion

OPUC is the study of probability measures on $\partial\mathbb{D}$, that is, positive measures, μ, with

$$\mu(\partial\mathbb{D}) = 1. \tag{2.1}$$

The *Carathéodory function* (after [15]) of μ is defined on $\mathbb{D}$ by

$$F(z) = \int \frac{e^{i\theta}+z}{e^{i\theta}-z}\, d\mu(\theta). \tag{2.2}$$

This analog of (1.5) is an analytic function on $\mathbb{D}$ which obeys

$$F(0) = 1 \qquad z \in \mathbb{D} \Rightarrow \operatorname{Re} F(z) > 0. \tag{2.3}$$

The *Schur function* (after [87]) is then defined by

$$F(z) = \frac{1+zf(z)}{1-zf(z)} \tag{2.4}$$

and is an analytic function mapping $\mathbb{D}$ to $\overline{\mathbb{D}}$; that is,

$$\sup_{z\in\mathbb{D}} |f(z)| \leq 1 \tag{2.5}$$

($f(z) \equiv e^{i\theta_0}$ is included and produced by μ, a point mass at $-\theta_0$).

(2.2) sets up a one-one correspondence between probability measures μ and analytic functions obeying (2.3) — this is essentially a form of the Herglotz representation (see [86, pp. 247]) and can be realized via (we use w-lim for the limit in the vague or weak *-topology on measures)

$$d\mu = \operatorname*{w\text{-}lim}_{r\uparrow 1} \operatorname{Re} F(re^{i\theta}) \, \frac{d\theta}{2\pi} \tag{2.6}$$

or by

$$F(z) = 1 + 2\sum_{n=1}^{\infty} c_n z^n \tag{2.7}$$

where c_n are the *moments* of μ given by

$$c_n = \int e^{-in\theta} \, d\mu(\theta). \tag{2.8}$$

(2.4) sets up a bijection between f's obeying (2.5) and F's obeying (2.3).

Recall that we call a measure *trivial* if it is supported on a finite set and *nontrivial* otherwise. We will mainly be interested in nontrivial measures. μ is trivial if and only if its Schur function is a finite Blaschke product

$$f(z) = e^{i\theta_0} \prod_{j=1}^{n-1} \frac{z - z_j}{1 - \bar{z}_j z} \tag{2.9}$$

with $z_1, \dots, z_{n-1} \in \mathbb{D}$. Here n is the number of points in the support of $d\mu$. Later (see the remark after Theorem 7.1) we will interpret (2.9) in terms of OPUC.

If μ is a nontrivial probability measure on $\partial\mathbb{D}$, we define the monic orthogonal polynomials $\Phi_n(z; d\mu)$ (or $\Phi_n(z)$ if $d\mu$ is understood) by

(2.10)

$$\Phi_n(z) = z^n + \text{ lower order terms} \qquad \int e^{-ij\theta} \Phi_n(e^{i\theta}) \, d\mu(\theta) = 0 \quad j = 0, 1, 2, \dots, n-1;$$

so in $L^2(\partial\mathbb{D}, d\mu)$, $\langle \Phi_n, \Phi_m \rangle = 0$ if $n \neq m$. The orthonormal polynomials φ_n are defined by

$$\varphi_n(z) = \frac{\Phi_n(z)}{\|\Phi_n\|} \tag{2.11}$$

where $\|\cdot\|$ is the L^2-norm. $\{\varphi_n\}_{n=0}^{\infty}$ is an orthonormal set in L^2. It may not be a basis (e.g., $d\mu(\theta) = d\theta/2\pi$ where $\varphi_n(z) = z^n$ and $\overline{z^j}$, $j = 1, \dots$, are orthogonal to all φ_n). We will discuss this further below (see Theorem 2.2).

If $d\mu$ is trivial, say $\operatorname{supp}(d\mu) = \{z_j\}_{j=1}^{k}$, we can still define Φ_n, φ_n for $n = 0, 1, \dots, k-1$. We can even define Φ_k (but not φ_k) as the unique monic polynomial of degree k with $\|\Phi_k\| = 0$, that is,

$$\Phi_k(z) = \prod_{j=1}^{k} (z - z_j) \qquad (\mu \text{ trivial}). \tag{2.12}$$

Clearly, (2.10) and the fact that the polynomials of degree at most n have dimension $n+1$ implies.

$$\deg(P) \leq n, \quad P \perp z^j, \quad j = 0, \dots, n-1 \Rightarrow P = c\Phi_n. \tag{2.13}$$

On $L^2(\partial\mathbb{D}, d\mu)$, define the anti-unitary map, ^{*,n}, by

$$f^{*,n}(e^{i\theta}) = e^{in\theta}\,\overline{f(e^{i\theta})}. \tag{2.14}$$

One mainly considers ^{*,n} on the set of polynomials of degree n which is left invariant

$$P(z) = \sum_{j=0}^{n} c_j z^j \Rightarrow P^{*,n}(z) = \sum_{j=0}^{n} \bar{c}_j z^{n-j} = z^n\,\overline{P(1/\bar{z})}. \tag{2.15}$$

Henceforth, following a standard, but unfortunate, convention, we drop the "$,n$" and just use P^*, hoping the n is implicit. Note that $1^* = z^n$, depending on n!

Since * is anti-unitary, (2.13) implies

$$\deg(P) \leq n, \quad P \perp z^j, \quad j = 1, \dots, n \Rightarrow P = c\Phi_n^*. \tag{2.16}$$

Since $\langle f, zg\rangle = \langle z^{-1}f, g\rangle$, it is easy to see that $\Phi_{n+1} - z\Phi_n \perp z^j$ for $j = 1, 2, \dots, n$. Since Φ is monic, this difference is of degree n, so (2.16) implies

$$\Phi_{n+1}(z) = z\Phi_n(z) - \bar{\alpha}_n \Phi_n^*(z) \tag{2.17}$$

for some complex numbers α_n, called the *Verblunsky coefficients* (in the older literature, also called reflection, Schur, Szegő, or Geronimus coefficients). (2.17) is called *Szegő recursion* after its first occurrence in Szegő's book [101]. In the engineering literature, it is called the Levinson algorithm after its rediscovery in linear prediction theory [60]. The choice of minus and $\bar{\alpha}_n$ rather than α_n will be made clear by Geronimus' theorem (see Theorem 3.1). Since Φ_n is monic, (2.15) implies $\Phi_n^*(0) = 1$, so (2.17) at $z = 0$ implies

$$\alpha_n = -\overline{\Phi_{n+1}(0)}. \tag{2.18}$$

Theorem 2.1. *We have*

$$\|\Phi_{n+1}\|^2 = (1 - |\alpha_n|^2)\|\Phi_n\|^2 \tag{2.19}$$

$$\|\Phi_n\| = \prod_{j=0}^{n-1} (1 - |\alpha_j|^2)^{1/2}. \tag{2.20}$$

For any nontrivial μ, we have $\alpha_j(d\mu) \in \mathbb{D}$ for all j. If μ is trivial with n points in its support, then $\alpha_0(d\mu), \dots, \alpha_{n-2}(d\mu) \in \mathbb{D}$ and $\alpha_{n-1}(d\mu) \in \partial\mathbb{D}$.

Proof. (2.17), unitarity of multiplication by z, and $\Phi_n^* \perp \Phi_{n+1}$ imply

$$\|\Phi_n\|^2 = \|z\Phi_n\|^2 = \|\Phi_{n+1} + \bar{\alpha}_n \Phi_n^*\|^2 = \|\Phi_{n+1}\|^2 + |\alpha_n|^2 \|\Phi_n\|^2,$$

which implies (2.19). Induction and $\Phi_0 = 1$ implies (2.20). By (2.19), $|\alpha_j| < 1$ in the nontrivial case and for $j = 0, \dots, n-2$ in the trivial case. Since $\|\Phi_n\| = 0 \neq \|\Phi_{n-1}\|$ in the trivial case, (2.19) implies $|\alpha_{n-1}| = 1$. □

Since it arises often, we define

$$\rho_j = (1 - |\alpha_j|^2)^{1/2} \qquad |\alpha_j|^2 + \rho_j^2 = 1. \tag{2.21}$$

One can use (2.20) to relate completeness of $\{\varphi_n\}_{n=0}^\infty$ to the Verblunsky coefficients:

Theorem 2.2. *For any nontrivial measure, the following are equivalent:*
(a) $\lim_{n\to\infty} \|\Phi_n\| = 0$,
(b) $\sum_{j=0}^\infty |\alpha_j|^2 = \infty$,
(c) $\{\varphi_n\}_{n=0}^\infty$ *are a basis for* $L^2(\partial\mathbb{D}, d\mu)$.

Remark. We will see later that there is an additional equivalence via Szegő's theorem (see (8.9)). The equivalence of a Szegő condition to completeness is due to Kolmogorov [54] and Krein [55, 56].

Sketch. By (2.20), (a) $\Leftrightarrow$ (b). If

$$P_{[k,\ell]} = \text{projection in } L^2(\partial\mathbb{D}, d\mu) \text{ onto } \operatorname{span}\{z^m\}_{m=k}^{\ell}, \tag{2.22}$$

we have that

$$\|\Phi_n\| = \|(1 - P_{[0,n-1]})z^n\| \tag{2.23}$$

$$= \|(1 - P_{[1,n]})1\| \tag{2.24}$$

$$= \|(1 - P_{[0,n-1]})z^{-1}\| \tag{2.25}$$

where (2.23) follows from the definition of Φ_n, (2.24) by applying ^{*,n} to z^n and $P_{[0,n-1]}$, and (2.25) by using the fact that multiplication by z^{-1} is unitary. It follows that

$$\|(1 - P_{[0,\infty)})z^{-1}\| = \lim_{n\to\infty} \|\Phi_n\|, \tag{2.26}$$

so (a) $\Leftrightarrow$ $z^{-1} \in \operatorname{span}\{\varphi_n\}_{n=0}^{\infty}$. If $z^{-1} \notin \operatorname{span}\{\varphi_n\}_{n=0}^{\infty}$, clearly they are not complete. If $z^{-1} \in \operatorname{span}\{\varphi_j\}_{j=0}^{\infty}$, an argument (see the proof of Theorem 1.5.7 in [91]) taking powers of z^{-1} shows $z^{-\ell} \in \operatorname{span}\{\varphi_n\}_{n=0}^{\infty}$ for all ℓ, so $\{\varphi_n\}_{n=0}^{\infty}$ are complete. $\square$

Let $\mathbb{D}^{\infty,c}$ denote the set of complex sequences $\{\alpha_j\}_{j=0}^{N}$ where either $N = \infty$ and $|\alpha_j| < 1$ for all j, or else $N < \infty$ and $\alpha_0, \dots, \alpha_{N-1} \in \mathbb{D}$ while $\alpha_N \in \partial\mathbb{D}$. In the topology of componentwise convergence, $\mathbb{D}^{\infty,c}$ is compact (and is a compactification of $\mathbb{D}^\infty$). The map, $\mathcal{S}$, from $\mu \mapsto \{\alpha_j(d\mu)\}_{j=0}^{N}$ is a well-defined map from $\mathcal{M}_{+,1}(\partial\mathbb{D})$, the probability measures on $\partial\mathbb{D}$, to $\mathbb{D}^{\infty,c}$. By (2.17), the α's determine the Φ_n's. Since $\int \Phi_n(z)\, d\mu = \delta_{n0}$, the Φ_n's determine the moments inductively, and so $d\mu$, since $\{z^\ell\}_{\ell=-\infty}^{\infty}$ span a dense set of $C(\partial\mathbb{D})$. Thus $\mathcal{S}$ is one-one. Moreover,

Theorem 2.3 (Verblunsky's Theorem [106])**.** *$\mathcal{S}$ is onto.*

[91] has four proofs of this theorem (Theorems 1.7.11, 3.1.3, 4.1.5, and 4.2.8); see Section 3 below. Given that $\mathcal{S}$ is a bijection, it is easy to see that it is a homeomorphism if $\mathcal{M}_{+,1}(\partial\mathbb{D})$ is given the vague (i.e., $C(\partial\mathbb{D})$-weak *) topology.

Applying * (actually, ${}^{*,n+1}$) to (2.17) yields

$$\Phi_{n+1}^*(z) = \Phi_n^*(z) - \alpha_n z \Phi_n(z). \tag{2.27}$$

Using (2.19) and (2.11), we get the recursion relations for φ_n written in matrix form

$$\begin{pmatrix} \varphi_{n+1}(z) \\ \varphi_{n+1}^*(z) \end{pmatrix} = A(z, \alpha_n) \begin{pmatrix} \varphi_n(z) \\ \varphi_n^*(z) \end{pmatrix} \tag{2.28}$$

where

$$A(z, \alpha) = \rho^{-1} \begin{pmatrix} z & -\bar{\alpha} \\ -z\alpha & 1 \end{pmatrix}. \tag{2.29}$$

Notice that $\det A = z$, so by inverting A, we get inverse recursion relations. We note the one for Φ_{n-1}

$$\Phi_{n-1}(z) = \frac{\rho_{n-1}^{-2}[\Phi_n + \bar{\alpha}_{n-1}\Phi_n^*]}{z}. \tag{2.30}$$

Note that by (2.18), $[\Phi_n + \bar{\alpha}_{n-1}\Phi_n^*]$ vanishes at zero, so the right side of (2.30) is a polynomial of degree $n-1$. This implies

Theorem 2.4 (Geronimus [33]). *Let μ and ν be two probability measures on $\partial\mathbb{D}$ so that for some N_0, $\Phi_{N_0}(z;d\mu) = \Phi_{N_0}(z;d\nu)$. Then $\Phi_j(z;d\mu) = \Phi_j(z;d\nu)$ for $j = 0,1,\dots,N_0-1$, $\alpha_j(d\mu) = \alpha_j(d\nu)$ for $j = 0,1,\dots,N_0-1$, and $\varphi_j(z;d\mu) = \varphi_j(z;d\nu)$ for $j = 0,1,\dots,N_0$.*

Remark. As noted in a footnote in Geronimus [33] and rediscovered by Wendroff [109], the result for OPRL requires equality for P_{N_0} and P_{N_0-1} and, in particular, it often happens that $P_{N_0}(x,d\gamma) = P_{N_0}(x,d\rho)$, but no other P_j's are equal.

Proof. By (2.18), Φ_{N_0} at 0 determines α_{N_0-1}, and so ρ_{N_0-1}, and thus Φ_{N_0-1} by (2.30). By induction, all α_j, $j \le N_0-1$, and Φ_j, $j \le N_0$, are equal and so, by (2.20), $\|\Phi_j\|$, and so φ_j. $\square$

As a final aspect of Szegő recursion, we turn to the Christoffel-Darboux (CD) formula (proven by Szegő [101] for OPUC; Christoffel [16] and Darboux [19] had a similar formula for OPRL), which is an analog of an iterated Wronskian formula for ODE's. With A given by (2.29), one finds, by matrix multiplication, that

$$A(\zeta,\alpha_n)^* \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} A(z,\alpha_n) = \begin{pmatrix} -z\bar{\zeta} & 0 \\ 0 & 1 \end{pmatrix}, \tag{2.31}$$

so

$$\begin{aligned} &\overline{\varphi_{n+1}^*(\zeta)}\,\varphi_{n+1}^*(z) - \overline{\varphi_{n+1}(\zeta)}\,\varphi_{n+1}(z) \\ &\qquad = \overline{\varphi_n^*(\zeta)}\,\varphi_n^*(z) - z\bar{\zeta}\,\overline{\varphi_n(\zeta)}\,\varphi_n(z) \\ &\qquad = (1-z\bar{\zeta})\,\overline{\varphi_n(\zeta)}\,\varphi_n(z) + [\,\overline{\varphi_n^*(\zeta)}\,\varphi_n^*(z) - \overline{\varphi_n(\zeta)}\,\varphi_n(z)], \end{aligned}$$

which, iterated to $n=0$ (where the term in brackets $[\dots]=0$ since $\varphi_0 = 1$), yields

Theorem 2.5 (Szegő [101]; CD Formula for OPUC).

$$(1-z\bar{\zeta})\sum_{j=0}^{n} \overline{\varphi_n(\zeta)}\,\varphi_n(z) = \overline{\varphi_{n+1}^*(\zeta)}\,\varphi_{n+1}^*(z) - \overline{\varphi_{n+1}(\zeta)}\,\varphi_{n+1}(z). \tag{2.32}$$

If $z=\zeta$ and lie in $\mathbb{D}$, we have various positivity facts that imply (the first since $\varphi_0(z)=1$):

$$|\varphi_n^*(z_0)| \ge (1-|z_0|^2)^{1/2} \qquad \text{for } z_0 \in \mathbb{D} \tag{2.33}$$

$$\lim |\varphi_{n+1}^*(z_0)| = \infty \Leftrightarrow \sum_{j=0}^{\infty} |\varphi_j(z_0)|^2 = \infty \qquad \text{for } z_0 \in \mathbb{D}. \tag{2.34}$$

3. Verblunsky's and Geronimus' theorems

In this section, we will prove Verblunsky's theorem (Theorem 2.3) and also a celebrated theorem of Geronimus. Our approach follows Section 3.1 of [91], which claims a new proof of Geronimus' theorem assuming Verblunsky's theorem. But in preparing this article, we realized the argument can be slightly modified to also prove Verblunsky's theorem.

To state Geronimus' theorem, we need to describe the Schur algorithm [87]. Given a Schur function f, define

$$\gamma_0(f) = f(0) \qquad f(z) = \frac{\gamma_0 + z f_1(z)}{1 + \bar{\gamma}_0 z f_1(z)}. \tag{3.1}$$

If $\gamma_0 \in \partial\mathbb{D}$ (i.e., $f(z) \equiv \gamma_0$), we do not define f_1. Otherwise, f_1 defined by (3.1) is also a Schur function since $w \to (\gamma_0 + w)/(1 + \bar{\gamma}_0 w)$ is a biholomorphic bijection of $\mathbb{D}$ to $\mathbb{D}$ if $|\gamma_0| < 1$, and g a Schur function with $g(0) = 0$ implies $g(z)/z$ is a Schur function (the Schwarz lemma).

(3.1) is called the Schur algorithm. It can be iterated; that is, we define $\gamma_n(f)$, the *Schur parameters*, and f_{n+1}, the *Schur iterates*, inductively by

$$\gamma_n(f) = f_n(0) \qquad f_n(z) = \frac{\gamma_n + z f_{n+1}(z)}{1 + \bar{\gamma}_n z f_{n+1}(z)}. \tag{3.2}$$

If, for some n, $f_n(z) = e^{i\theta_0}$, we set $\gamma_n = e^{i\theta_0}$ and stop. In this way, we map any Schur function f to a sequence in $\mathbb{D}^{\infty,c}$, the set defined after Theorem 2.2. We can now state Geronimus' theorem:

Theorem 3.1 (Geronimus' Theorem)**.** *Let μ be a probability measure on $\partial\mathbb{D}$, f its Schur function, and $\gamma_n(d\mu) \equiv \gamma_n(f)$ the Schur parameters of f. Then*

$$\gamma_n(d\mu) = \alpha_n(d\mu). \tag{3.3}$$

This gives a continued fraction expansion of F whose coefficients are α_n, and so is an analog of (1.6). This formula explains why we took a minus and conjugate in (2.17). The procedure of dropping a Verblunsky coefficient from the start can be understood by using the recursion relations and the relation of F to the OPUC (see Theorem 4.4 below). This approach to proving Theorem 3.1, due to Peherstorfer [72], is discussed in Section 3.3 of [91].

(3.1)/(3.2) can be rewritten and then iterated following Schur [87]:

$$f(z) = \gamma_0 + (1 - \bar{\gamma}_0 f) z f_1 \tag{3.4}$$

$$= \gamma_0 + \sum_{j=1}^{n-1} \biggl[\prod_{k=0}^{j-1} (1 - \bar{\gamma}_k f_k) \biggr] z^j \gamma_j + \prod_{k=0}^{n-1} (1 - \bar{\gamma}_k f_k) z^n f_n, \tag{3.5}$$

which implies that if $f(z) = \sum_{n=0}^{\infty} a_n(f) z^n$, then

$$a_n(f) = \gamma_n \prod_{j=0}^{n-1} (1 - |\gamma_j|^2) + \text{ polynomial in } (\gamma_0, \bar{\gamma}_0, \dots, \gamma_{n-1}, \bar{\gamma}_{n-1}). \tag{3.6}$$

Plugging this into (2.4) and using (2.7) implies

$$c_n(d\mu) = \gamma_{n-1} \prod_{j=0}^{n-2} (1 - |\gamma_j|^2) + \text{ polynomial in } (\gamma_0, \bar{\gamma}_0, \dots, \gamma_{n-2}, \bar{\gamma}_{n-2}) \tag{3.7}$$

(the polynomials are different, but the leading terms are the same up to a shift of index).

(3.6) also shows that if $\gamma_j(f) = \gamma_j(g)$ for $j = 0, \dots, n-1$, then the Schur function $\frac{1}{2}(f - g) = O(z^n)$, so, by the Schwarz lemma,

$$\gamma_j(f) = \gamma_j(g), \quad j = 0, \dots, n-1 \Rightarrow |f(z) - g(z)| \le 2|z|^n. \tag{3.8}$$

Lemma 3.2. *The map from Schur functions to $\mathbb{D}^{\infty,c}$ is one-one and onto.*

Proof. (3.8) shows that if $\gamma_j(f) = \gamma_j(g)$ for all j, then $f = g$ on $\mathbb{D}$. Given a sequence in $\mathbb{D}^\infty$, define the *Schur approximates*, $f^{[n]}$, by setting $f^{[n]}_{n+1}$ to 0 in (3.2) and using $\{\gamma_j\}_{j=0}^n$ to define $f^{[n]}_n, f^{[n]}_{n-1}, \dots, f^{[n]}$. By construction,

$$\gamma_j(f^{[n]}) = \begin{cases} \gamma_j & j \le n \\ 0 & j > n \end{cases}. \tag{3.9}$$

Since $\gamma_j(f^{[n]}) = \gamma_j(f^{[m]})$ for $j \le \min(n, m)$, we have, by (3.8), that $f^{[n]}$ converge uniformly on compacts and the limit clearly has the prescribed set of γ's. Given a sequence in $\mathbb{D}^{\infty,c} \backslash \mathbb{D}^\infty$, suppose $\gamma_{n+1} = e^{i\theta_0} \in \partial\mathbb{D}$, set $f_{n+1} \equiv e^{i\theta_0}$ and use (3.2) to define f with the prescribed γ's. □

Proof of Theorems 2.3 *and* 3.1. $(z^n - \Phi_n) \perp \Phi_n \Rightarrow \|\Phi_n\|^2 = \langle z^n, \Phi_n \rangle$, so applying ^{*,n},

$$\langle \Phi^*_n, 1 \rangle = \|\Phi_n\|^2 = \prod_{j=0}^{n-1} (1 - |\alpha_j|^2). \tag{3.10}$$

Taking the inner product of (2.17) with the function, 1, and using $\langle \Phi_{n+1}, 1 \rangle = 0$, we see

$$\langle z\Phi_n, 1 \rangle = \alpha_n \prod_{j=0}^{n-1} (1 - |\alpha_j|^2). \tag{3.11}$$

By (2.17) and induction, the coefficients of Φ_j are polynomials $\alpha_0, \bar\alpha_0, \dots, \alpha_{j-1}, \bar\alpha_{j-1}$, and so, by induction, the moments c_{j+1} are polynomials in the same α's. Then (3.11) becomes (a formula of Verblunsky):

$$c_{n+1} = \alpha_n \prod_{j=0}^{n-1} (1 - |\alpha_j|^2) + \text{ polynomial in } (\alpha_0, \bar\alpha_0, \dots, \alpha_{n-1}, \bar\alpha_{n-1}). \tag{3.12}$$

We will now prove Theorem 3.1 by induction, and then Theorem 2.3 follows from Lemma 3.2. For $n = 0$, we have, by (3.12) and (3.7), that

$$c_1 = \alpha_0 = \gamma_0. \tag{3.13}$$

Suppose we know $\alpha_j = \gamma_j$ for $j = 0, 1, \dots, n-1$. We fix those n values in $\mathbb{D}$ and ask what values of c_{n+1} can occur. By (3.7), it is a solid disk in $\mathbb{C}$ of radius $\prod_{j=0}^{n-1}(1 - |\alpha_j|^2)$ since γ_n can run through $\overline{\mathbb{D}}$. The center of the disk is some fixed point (given fixed $\{\gamma_j\}_{j=0}^{n-1}$).

By (3.12), it is also a subset of the disk of radius $\prod_{j=0}^{n-1}(1 - |\alpha_j|^2)$ with possibly another center. But since the sets are the same, the centers must be the same, and all α_j must occur. Once we know the centers and radii are the same, the equality of the formulae for c_{n+1} implies $\alpha_n = \gamma_n$. □

4. Zeros, the Bernstein-Szegő approximation, and boundary conditions

Our first goal in this section is to prove that the zeros of OPUC lie in $\mathbb{D}$. There are six proofs of this in [91]. We pick the one that is shortest, using the same argument that led to (2.19).

Theorem 4.1. *Φ_n has all its zeros in $\mathbb{D}$ and Φ^*_n has all its zeros in $\mathbb{C} \backslash \overline{\mathbb{D}}$.*

Proof. (Landau [59]) Let $\Phi_n(z_0) = 0$ and define $P(z) = \Phi_n(z)/(z - z_0)$. Since $\deg P = n - 1$, $P \perp \Phi_n$. Thus

$$\|P\|^2 = \|zP\|^2 = \|z_0 P + \Phi_n\|^2 = |z_0|^2 \|P\|^2 + \|\Phi_n\|^2, \tag{4.1}$$

so $\|\Phi_n\|^2 = (1 - |z_0|^2)\|P\|^2$, implying $|z_0| < 1$. Since $\Phi_n^*(z_0) = 0 \Leftrightarrow \Phi_n(1/\bar{z}) = 0$, the result for Φ_n implies the result for Φ_n^*. □

Next, we will identify measures with $\alpha_j(d\mu) = 0$ for $j \geq n_0$. The key is a calculation that goes back at least to Erdélyi et al. [24].

Proposition 4.2. *Let P_n be a polynomial of degree n with all zeros in $\mathbb{D}$. Let*

$$d\mu = \frac{c\, d\theta}{2\pi |P_n(e^{i\theta})|^2} \tag{4.2}$$

where c is picked to make $d\mu$ a probability measure. Then for all integral $j < n$ (including $j < 0$),

$$\langle z^j, P \rangle_{L^2(\partial\mathbb{D}, d\mu)} = 0. \tag{4.3}$$

Proof.

$$\begin{aligned} \langle z^j, P \rangle_{L^2(\partial\mathbb{D}, d\mu)} &= \int e^{-ij\theta} P(e^{i\theta}) \, \frac{d\theta}{2\pi z^{-n} P^*(z) P(z)|_{z=e^{i\theta}}} \\ &= \frac{1}{2\pi i} \oint z^{n-j-1} \, \frac{dz}{P^*(z)} \end{aligned}$$

is zero for $n - j - 1 \geq 0$ since $P^*(z)$ is nonvanishing on $\overline{\mathbb{D}}$. □

Theorem 4.3. *Let $d\mu$ be a nontrivial probability measure on $\partial\mathbb{D}$. Let*

$$d\mu_n = \frac{d\theta}{2\pi |\varphi_n(e^{i\theta}, d\mu)|^2}. \tag{4.4}$$

Then $d\mu_n$ is a probability measure with

$$\alpha_j(d\mu_n) = \begin{cases} \alpha_j(d\mu) & j \leq n-1 \\ 0 & j \geq n \end{cases}. \tag{4.5}$$

Proof. Let $d\nu = c\, d\mu_n$ where c is picked so that $\int d\nu = 1$ (eventually, we will prove $c = 1$). By Proposition 4.2, $\langle z^j, \Phi_n(\,\cdot\,; d\mu) \rangle_{L^2(\partial\mathbb{D}, d\nu)} = 0$ for $j = 0, 1, \dots, n-1$, so $\Phi_n(z; d\nu) = \Phi_n(z; d\mu)$. It follows from Theorem 2.3 that $\alpha_j(d\nu) = \alpha_j(d\mu)$ for $j = 0, \dots, n-1$ and $\varphi_n(z; d\mu) = \varphi_n(z; d\nu)$. Therefore, $1 = \int |\varphi_n|^2 \, d\nu = c$, so $d\nu = d\mu_n$.

By Proposition 4.3, for any $k \geq 0$,

$$\langle z^j, z^k \Phi_n \rangle_{L^2(\partial\mathbb{D}, d\mu_n)} = 0 \qquad j = 0, \dots, n+k-1. \tag{4.6}$$

It follows that $\Phi_{n+k}(z; d\mu_n) = z^k \Phi_n(z; d\mu_n)$, and thus $\Phi_{n+k}(0) = 0$ for $k \geq 1$. Therefore, by (2.18), $\alpha_j(d\mu_n) = 0$ for $j \geq n$. □

Even though Theorem 4.3 was proven by Verblunsky [107] and rediscovered by Geronimus [33] (to whom it is often credited), $d\mu_n$ are called Bernstein-Szegő approximations, since Szegő [98] first considered measures of this form (3.2) and Bernstein [11] their OPRL analog. Since, for each fixed j, $\alpha_j(d\mu_n) \to \alpha_j(d\mu)$ (indeed, they are equal for $n > j$), $d\mu_n \to d\mu$ weakly since $\mathcal{S}$ is a homeomorphism.

Some thought about the form of $d\mu_n$ suggests its Carathéodory function should be a rational function whose denominator is φ_n^*. We will prove this by identifying the numerator. The *second kind polynomials*, ψ_n, are the OPUC for the measure $d\mu_{-1}$ with $\alpha_j(d\mu_{-1}) = -\alpha_j(d\mu)$. Notice that in terms of the matrix A of (2.29),

$$\begin{pmatrix} \psi_{n+1} \\ -\psi_{n+1}^* \end{pmatrix} = A(z, \alpha_n(d\mu)) \begin{pmatrix} \psi_n \\ -\psi_n^* \end{pmatrix} \tag{4.7}$$

(note $\alpha_n(d\mu)$, not $\alpha_n(d\mu_{-1})$). Thus

$$\begin{pmatrix} \psi_n & \varphi_n \\ -\psi_n^* & \varphi_n^* \end{pmatrix} = A(z, \alpha_{n-1}) \dots A(z, \alpha_0) \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}. \tag{4.8}$$

Taking determinants, using $\det(A) = z$,

$$\varphi_n^* \psi_n + \varphi_n \psi_n^* = 2z^n. \tag{4.9}$$

Theorem 4.4 (Verblunsky [107])**.** *Let $d\mu_n$ be given by* (4.4)*. Then*

$$F(z, d\mu_n) = \frac{\psi_n^*(z; d\mu)}{\varphi_n^*(z; d\mu)}. \tag{4.10}$$

Proof. For $z = e^{i\theta}$, (4.9) can be rewritten as $\mathrm{Re}(\overline{\varphi_n(e^{i\theta})}\, \psi_n(e^{i\theta})) = 1$. Thus, if $G(z)$ is the right side of (4.10),

$$\mathrm{Re}(G(e^{i\theta})) = \frac{1}{|\varphi_n(e^{i\theta})|^2}. \tag{4.11}$$

Since G is analytic in a neighborhood of $\overline{\mathbb{D}}$, $\mathrm{Re}\, G > 0$ on $\mathbb{D}$. Since $G(0) = 1$, the complex Poisson representation (see Rudin [86, pg. 235]) and (4.10) imply that $G(z)$ is the Carathéodory function of $d\mu_n$. □

It is useful to think of $d\mu$ and $d\mu_{-1}$ as embedded in a family $d\mu_\lambda$ for $\lambda \in \partial\mathbb{D}$. The *Aleksandrov family* associated to $d\mu$ is defined by

$$\alpha_j(d\mu_\lambda) = \lambda \alpha_j(d\mu). \tag{4.12}$$

Given Geronimus' theorem (Theorem 3.1), it is easy to see that

$$f(z, d\mu_\lambda) = \lambda f(z, d\mu) \tag{4.13}$$

(for $\gamma_0(\lambda f) = \lambda \gamma_0(f)$ and $(\lambda f)_1 = \lambda(f_1)$). So, by (2.4) and its inverse, $zf(z) = (F(z) - 1)/(F(z) + 1)$,

$$F(z, d\mu_\lambda) = \frac{(1 - \lambda) + (1 + \lambda) F(z, d\mu)}{(1 + \lambda) + (1 - \lambda) F(z, d\mu)}, \tag{4.14}$$

which is the original definition of Aleksandrov [2]; it is Golinskii-Nevai [42] who realized its relevance to OPUC and boundary conditions. If $\varphi_n^{(\lambda)}(z) = \varphi_n(z; d\mu_\lambda)$, then

$$\begin{pmatrix} \varphi_{n+1}^{(\lambda)} \\ \bar\lambda (\varphi_{n+1}^{(\lambda)})^* \end{pmatrix} = A(z, \alpha_n) \begin{pmatrix} \varphi_n^{(\lambda)} \\ \bar\lambda (\varphi_n^{(\lambda)})^* \end{pmatrix}; \tag{4.15}$$

so φ_n and $\varphi_n^{(\lambda)}$ obey the same difference equation, but the $n = 0$ boundary values change from $\binom{1}{1}$ to $\binom{1}{\bar\lambda}$. The Aleksandrov family is the analog of variation of boundary conditions in second-order ODE's.

A direct calculation (via contour integrals) shows that if $\mathrm{Re}(a) > 0$, then

$$\int_0^{2\pi} \frac{(1 - e^{i\theta}) + (1 + e^{i\theta})a}{(1 + e^{i\theta}) + (1 - e^{i\theta})a} \, \frac{d\theta}{2\pi} = 1. \tag{4.16}$$

Since 1 is the Carathéodory function of $d\theta/2\pi$, (4.16) and (4.14) imply

Theorem 4.5 (Aleksandrov [2], Golinskii [41]). *For the Aleksandrov family, we have*

$$\int_\theta [d\mu_{e^{i\theta}}(\varphi)]\,\frac{d\theta}{2\pi} = \frac{d\varphi}{2\pi}. \tag{4.17}$$

This is the OPUC analog of the Javrjan [45]-Wegner [108] averaging for Schrödinger operators, which is the basis of the localization proof of Simon-Wolff [97]. It can be used [92] to prove localization for suitable random OPUC.

5. The CMV matrix

Perturbation theory involves looking at similarities of measures when their Verblunsky coefficients are close in some suitable sense. In the analogous OPRL situation, the Jacobi matrices, (1.4), are an invaluable tool. If one defines the essential support of a measure to be the support with isolated points removed, and if ρ and γ are measures on $[c,d]\subset\mathbb{R}$ with Jacobi parameters a_n, b_n and $\tilde a_n, \tilde b_n$, then ρ and γ have the same essential support if $|a_n-\tilde a_n|+|b_n-\tilde b_n|\to 0$. This can be seen by noting that the difference of the Jacobi matrices is compact and then appealing to Weyl's theorem on the invariance of essential spectrum.

In this section, we discuss a suitable matrix representation for multiplication by z in $L^2(\partial\mathbb{D}, d\mu)$. There is an obvious choice, namely, $\langle\varphi_n, z\varphi_m\rangle$, but this is not the "right" one. It has two problems. If $\sum|\alpha_j|^2<\infty$, $\{\varphi_n\}_{n=0}^\infty$ is not a basis, and so this matrix is not unitary. Even worse, although this matrix has finite columns ($\langle\varphi_n, z\varphi_m\rangle = 0$ if $n>m+1$), in general, it does not have finite rows.

The right basis, as discovered by Cantero, Moral, and Velázquez [14], is the one $\chi_0,\chi_1,\chi_2,\dots$, obtained by orthonormalizing $1, z, z^{-1}, z^2, z^{-2},\dots$. We will also want to consider the basis $x_0, x_1, x_2, \dots$ obtained by orthonormalizing $1, z^{-1}, z, z^{-2},\dots$. Remarkably, the χ's can be expressed in terms of φ's and φ^*'s, and the matrix elements in terms of α's and ρ's.

Proposition 5.1.

$$\text{(a)}\qquad \chi_{2n}(z) = z^{-n}\varphi_{2n}^*(z) \qquad \chi_{2n-1}(z) = z^{-n+1}\varphi_{2n-1}(z) \tag{5.1}$$

$$\text{(b)}\qquad x_{2n}(z) = z^{-n}\varphi_{2n}(z) \qquad x_{2n-1}(z) = z^{-n}\varphi_{2n-1}^*(z). \tag{5.2}$$

Proof. In terms of the projections $P_{[k,\ell]}$ of (2.22), we have

$$\varphi_m = \frac{(1-P_{[0,m-1]})z^m}{\|\dots\|} \qquad \varphi_m^* = \frac{(1-P_{[1,m]})1}{\|\dots\|} \tag{5.3}$$

where $\|\dots\|$ is the norm of the numerator. Since multiplication by z^ℓ is unitary,

$$z^{-n}\varphi_{2n} = \frac{(1-P_{[-n,n-1]})z^n}{\|\dots\|} = x_{2n},$$

proving the first half of (5.2). The others are similar. $\square$

We define four matrices ($\mathcal{C}$ = *CMV matrix*) by

$$\mathcal{C}_{k\ell} = \langle\chi_k, z\chi_\ell\rangle \quad \tilde{\mathcal{C}}_{k\ell} = \langle x_k, z x_\ell\rangle \quad \mathcal{L}_{k\ell} = \langle\chi_k, z x_\ell\rangle \quad \mathcal{M}_{k\ell} = \langle x_k, \chi_\ell\rangle. \tag{5.4}$$

Clearly,

$$\mathcal{C} = \mathcal{L}\mathcal{M} \qquad \tilde{\mathcal{C}} = \mathcal{M}\mathcal{L} \qquad \tilde{\mathcal{C}} = \mathcal{C}^t, \tag{5.5}$$

where the last comes from the fact that the explicit formulae below show $\mathcal{L}$ and $\mathcal{M}$ are (complex) symmetric. Define, for $\alpha \in \overline{\mathbb{D}}$, the 2×2 symmetric matrix

$$\Theta(\alpha) = \begin{pmatrix} \bar\alpha & \rho \\ \rho & -\alpha \end{pmatrix}. \tag{5.6}$$

Theorem 5.2. *Let* $\mathbf{1}$ *be the* 1×1 *unit matrix. Then*

$$\mathcal{M} = \mathbf{1} \oplus \Theta(\alpha_1) \oplus \Theta(\alpha_3) \oplus \cdots \qquad \mathcal{L} = \Theta(\alpha_0) \oplus \Theta(\alpha_2) \oplus \Theta(\alpha_4) \oplus \cdots . \tag{5.7}$$

Proof. This is an expression of the Szegő recursion formula. For example, the $2n$ row (labelling rows $0,1,2,\dots$) of $\mathcal{L}$ says that $zx_{2n} = \bar\alpha_{2n}\chi_{2n} + \rho_{2n}\chi_{2n+1}$, which, by Proposition 5.1, is equivalent to $z\varphi_{2n} = \bar\alpha_{2n}\varphi_{2n}^* + \rho_{2n}\varphi_{2n+1}$, which is the top row of (2.28). □

While $\mathcal{L}$ and $\mathcal{M}$ have direct sum structures, in general (i.e., if all $|\alpha_j| < 1$), $\mathcal{C}$ does not. Indeed, by (5.5) and (5.7),

$$\mathcal{C} = \begin{pmatrix} \bar\alpha_0 & \bar\alpha_1\rho_0 & \rho_1\rho_0 & 0 & 0 & \dots \\ \rho_0 & -\bar\alpha_1\alpha_0 & -\rho_1\alpha_0 & 0 & 0 & \dots \\ 0 & \bar\alpha_2\rho_1 & -\bar\alpha_2\alpha_1 & \bar\alpha_3\rho_2 & \rho_3\rho_2 & \dots \\ 0 & \rho_2\rho_1 & -\rho_2\alpha_1 & -\bar\alpha_3\alpha_2 & -\rho_3\alpha_2 & \dots \\ 0 & 0 & 0 & \bar\alpha_4\rho_3 & -\bar\alpha_4\alpha_3 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}. \tag{5.8}$$

Thus $\mathcal{C}$ has a 4×2 block structure and is generally five-diagonal. It is the simplest unitary matrix with a cyclic vector; for example [13], any tridiagonal semi-infinite unitary is a direct sum of 1×1 and 2×2 matrices.

Theorem 5.3. *If* $\mathcal{C}^{(N)}$ *is the top left* $N\times N$ *block of* $\mathcal{C}$*, then*

$$\det(z_0\mathbf{1} - \mathcal{C}^{(N)}) = \Phi_N(z_0). \tag{5.9}$$

Sketch of proof. If ζ is the operator of multiplication by z in $L^2(\partial\mathbb{D}, d\mu)$, then $\mathcal{C}^{(N)} = P_{[-\ell,N-\ell-1]}\zeta P_{[-\ell,N-\ell-1]}$ restricted to $\operatorname{ran}P_{[-\ell,N-\ell-1]}$ where $P_{[j,k]}$ is given by (2.22) and ℓ is either $(N-1)/2$ or $N/2$. Since multiplication by z^ℓ is unitary, $\mathcal{C}^{(N)}$ is unitarily equivalent to $P_{[0,N-1]}\zeta P_{[0,N-1]}$ on $\operatorname{ran}P_{[0,N-1]}$.

z_0 is an eigenvalue of $P_{[0,N-1]}\zeta P_{[0,N-1]}$ if and only if there is Q of degree $N-1$ so $(z-z_0)Q = \Phi_N(z)$, that is, if and only if $\Phi_N(z_0) = 0$. This proves (5.9) if Φ_n has distinct zeros. By a limiting argument (see Theorem 1.7.18 of [91]), (5.9) holds in general. □

Remark. This theorem sheds light on a result of Fejér [25] that for OP's of general measures on $\mathbb{C}$, their zeros lie in the convex hull of $\operatorname{supp}(d\mu)$, for (5.9) implies the zeros are in the numerical range of $\mathcal{C}^{(N)}$, so in the numerical range of $\mathcal{C}$, which is the convex hull of $\operatorname{supp}(d\mu)$ by the spectral theorem. In particular, Fejér's theorem implies in the OPUC case that if $\zeta \in \partial\mathbb{D}$ with $d = \operatorname{dist}(\zeta, \operatorname{supp}(d\mu)) > 0$ and $\Phi_n(z_0) = 0$, then $|z_0 - \zeta| \geq \frac{1}{2}d^2$.

Notice that if $|\alpha_j| = 1$, $\Theta(\alpha_j) = \left(\begin{smallmatrix} \bar\alpha_j & 0 \\ 0 & -\alpha_j \end{smallmatrix}\right)$ is a direct sum, and so $\mathcal{C} = \mathcal{L}\mathcal{M}$ has a $(j+1)\times(j+1)$ unitary block in the upper corner. This implies that if $\beta \in \partial\mathbb{D}$, then $\Phi_N^{(\beta)} \equiv z\Phi_{N-1} - \beta\Phi_{N-1}^*$ has all its zeros on $\partial\mathbb{D}$ since they are eigenvalues of a unitary matrix. The $\Phi_N^{(\beta)}$ are called *paraorthogonal polynomials* and are studied in [33, 47, 40].

Dombrowski [22] proved that a Jacobi matrix with $\liminf a_n = 0$ has no a.c. spectrum by picking a subsequence with $\sum_{j=0}^{\infty} a_{n(j)} < \infty$ and trace class perturbing to a decoupled direct sum of finite rank matrices. Unaware of this work, Simon-Spencer [96] proved a similar result if $\limsup |b_n| = \infty$. As noted by Golinskii-Simon [43], the same idea and CMV matrices prove the following, originally proven by other means [84].

Theorem 5.4 (Rakhmanov's Lemma [84])**.** *If μ is a probability measure on $\partial\mathbb{D}$ so $\limsup |\alpha_n(d\mu)| = 1$, then μ is singular with respect to $d\theta/2\pi$.*

Golinskii-Simon also use perturbations of CMV matrices to prove

Theorem 5.5 ([43])**.** *If μ,ν are two probability measures on $\partial\mathbb{D}$ so $|\alpha_n(d\mu) - \alpha_n(d\nu)| \to 0$, then $\operatorname{ess\,sup}(d\mu) = \operatorname{ess\,sup}(d\nu)$. If $\sum_n |\alpha_n(d\mu) - \alpha_n(d\nu)| < \infty$, then the absolutely continuous parts of μ and ν are mutually absolutely continuous.*

Aleksandrov families fit into CMV matrices with a twist. $\mathcal{C}(\{\lambda\alpha_n\})$ and $\mathcal{C}(\{\alpha_n\})$ do not differ by a rank one perturbation — rather they do up to a unitary equivalence. Specifically:

Theorem 5.6. *Let $\lambda \in \partial\mathbb{D}$ and $\{\alpha_n\} \in \mathbb{D}^\infty$. Let D be the diagonal matrix with elements $1, \lambda^{-1}, 1, \lambda^{-1}, \dots$. Then $D\mathcal{C}(\{\lambda\alpha_n\})D^{-1} = \mathcal{L}(\{\alpha_n\})\mathcal{M}_\lambda(\{\alpha_n\})$ where $\mathcal{M}_\lambda$ differs from $\mathcal{M}$ by having λ in the $(1,1)$ position instead of 1.*

This is a restatement of Theorem 4.2.9 of [91]. A generalization to rank one perturbation in the n-th diagonal can be found in Simon [93].

CMV matrices have been generalized in two directions. First, OPUC can be thought of as an analog of half-line ODE. The whole-line analog is an extended CMV matrix, $\mathcal{E}$, defined on $\ell_2(-\infty,\infty)$ by a two-sided sequence $\{\alpha_n\}_{n=-\infty}^{\infty}$ as a product of $\cdots \oplus \Theta(\alpha_{-2}) \oplus \Theta(\alpha_0) \oplus \Theta(\alpha_2) \oplus \cdots$ and $\cdots \oplus \Theta(\alpha_{-1}) \oplus \Theta(\alpha_1) \oplus \cdots$ where $\Theta(\alpha_j)$ acts on the span of δ_j and δ_{j+1}. This is discussed in Sections 4.5, 10.5, and 10.16 of [91, 92]. It is useful in the study of ergodic (Section 6) and periodic (Section 10) OPUC. Gesztesy-Zinchenko [37, 38] have further results on $\mathcal{E}$.

Second, if U is an $n\times n$ unitary matrix and φ is cyclic in that $\{U^j\varphi\}_{j=0}^{n-1}$ is a basis, then the spectral measure for φ has n points. One can then define polynomials $\Phi_0,\dots,\Phi_n$ and Verblunsky coefficients $\alpha_0,\dots,\alpha_{n-2} \in \mathbb{D}$ and $\alpha_{n-1} \in \partial\mathbb{D}$. U is unitarily equivalent to a finite CMV matrix, the upper block of an infinite matrix where α_{n-1} is taken in $\partial\mathbb{D}$.

Just as the theory of selfadjoint matrices with cyclic vector is identical to the theory of Jacobi matrices, the theory of unitary matrices with cyclic vector (i.e., $\{U^j\varphi\}_{j=-\infty}^{\infty}$ spanning) is identical to the theory of CMV matrices. The Verblunsky coefficients are a complete set of unitary invariants.

In this regard, there is a natural question answered by Killip-Nenciu [53]. Let $\mathbb{U}(n)$ be the group of $n\times n$ unitary matrices and consider Haar measure on $\mathbb{U}(n)$. For a.e. U, $(1\,0\dots0)^t$ is cyclic, so there is induced a measure on Verblunsky coefficients $\alpha_0,\dots,\alpha_{n-2} \in \mathbb{D}$ and $\alpha_{n-1} \in \partial\mathbb{D}$. The measure is the same if $(1\,0\dots0)^t$ is replaced by any other vector or by a random choice (say, uniform distribution on the unit sphere in $\mathbb{C}^n$).

Theorem 5.7 ([53])**.** *Under the measure induced by Haar measure on $\mathbb{U}(n)$, the α_j's are independent (i.e., the induced measure is a product measure), α_{n-1} is* *uniformly distributed on $\partial\mathbb{D}$, and α_j, $j = 1, \dots, n-2$, is distributed via*

$$\frac{(n-j-1)}{\pi}\,(1-|\alpha|^2)^{(n-j-2)}\,d^2\alpha. \tag{5.10}$$

6. Transfer matrices, Weyl solutions, and Lyapunov exponents

In this section we present a potpourri of results connected with solutions of Szegő recursion (2.28) where the two components are freed of $u_2^* = u_1$. Indeed, we look at solutions for a fixed z. Thus, solutions have the form

$$u(z;n) = T_n(z)u(z;0) \qquad T_n(z) = A(z,\alpha_{n-1})\dots A(z,\alpha_0) \tag{6.1}$$

with A given by (2.29). T_n is called the *transfer matrix*. By (4.8), we have

$$\begin{aligned} T_n(z) &= \tfrac{1}{2}\begin{pmatrix} \varphi_n(z)+\psi_n(z) & \varphi_n(z)-\psi_n(z) \\ \varphi_n^*(z)-\psi_n^*(z) & \varphi_n^*(z)+\psi_n^*(z)\end{pmatrix} \\ &= \biggl(\prod_{j=0}^{n-1}\rho_j^{-1/2}\biggr)\begin{pmatrix} zB_{n-1}^*(z) & -A_{n-1}^*(z) \\ -zA_{n-1}(z) & B_{n-1}(z)\end{pmatrix} \end{aligned} \tag{6.2}$$

where A_{n-1} and B_{n-1} are polynomials of degree $n-1$ and the * term is $^{*,n-1}$. The degree count uses $\varphi_n(0) = -\psi_n(0)$, $\varphi_n^*(0) = \psi_n^*(0)$. A_n and B_n are the *Wall polynomials* which are related to the Schur approximants, $f^{[n]}$, of (3.9) by $f^{[n]}(z) = A_n(z)/B_n(z)$.

For $z \in \partial\mathbb{D}$, T_n lies in the group $\mathbb{U}(1,1)$ of matrices obeying $M^*\left(\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right)M = \left(\begin{smallmatrix}1&0\\0&-1\end{smallmatrix}\right)$. Features of this group play a role in advanced aspects of the theory; see [92], especially Section 10.4.

The solutions $u_\varphi = (\varphi_n, \varphi_n^*)$ and $u_\psi = (\psi_n, -\psi_n^*)$ of (6.1) can be combined into an ℓ_2 solution for $|z| < 1$:

Theorem 6.1 (Geronimo [29]; Golinskii-Nevai [42])**.** *Fix $z \in \mathbb{D}$. Then $u_\psi(z;n) + ru_\varphi(z;n) \to 0$ as $n \to \infty$ for fixed $r \in \mathbb{C}$ if and only if $r = F(z)$. Moreover, $u_\psi + F(z)u_\varphi$ is in ℓ^2.*

Remark. In analogy to ODE theory, $u_\psi + F(z)u_\varphi$ is called the *Weyl solution*.

Sketch of proof ([42]) . Looking at the second component, we see that if $u_\psi + ru_\varphi \to 0$, then $-\psi_n^* + r\varphi_n^* \to 0$. By (2.33), $r - \psi_n^*/\varphi_n^* \to 0$, so by (4.10), $r = F(z)$. The ℓ^2 proof below implies that the first components go to zero for $r = F(z)$.

By using the CD formula (2.32) for φ and ψ plus a mixed CD formula obtained from (2.31) by using $\binom{1}{1}$ and $\binom{1}{-1}$, one finds that

$$(1-|z|^2)\sum_{j=0}^{n-1}|\psi_j(z)+r\varphi_j(z)|^2 = 4\operatorname{Re}(r)+|\psi_n^*(z)-r\varphi_n^*|^2-|\psi_n(z)+r\varphi_n(z)|^2. \tag{6.3}$$

Taking $r = \psi_n^*(z)/\varphi_n^*(z)$, one finds

$$k \le n-1 \Rightarrow \sum_{j=0}^{k}\left|\psi_j + \frac{\psi_n^*}{\varphi_n^*}\varphi_j\right|^2 \le 4\operatorname{Re}\biggl(\frac{\psi_n^*}{\varphi_n^*}\biggr). \tag{6.4}$$

Taking $n \to \infty$ and then $k \to \infty$ shows

$$\sum_{j=0}^{\infty}|\psi_j + F\varphi_j|^2 \le 4\operatorname{Re}(F). \tag{6.5}$$

The inequality in (6.5) plus the equality in (6.3) imply that $|\psi_j^* - F\varphi_j^*| \leq |\psi_j + F\varphi_j|$, so (6.5) implies $u_\psi + Fu_\varphi \in \ell^2$. $\square$

Another way of proving the ℓ^2 result, from [29], is illuminating. It starts from a formula which was Geronimus' original definition of the second kind polynomials,

$$\psi_n(z) = \int \frac{e^{i\theta}+z}{e^{i\theta}-z}\,[\varphi_n(e^{i\theta}) - \varphi_n(z)]\,d\mu(\theta). \tag{6.6}$$

This and its image under the map * imply

$$\begin{aligned} F(z)\varphi_n(z) + \psi_n(z) &= \int \frac{e^{i\theta}+z}{e^{i\theta}-z}\,\varphi_n(e^{i\theta})\,d\mu(\theta) \\ F(z)\varphi_n^*(z) - \psi_n^*(z) &= z^n \int \frac{e^{i\theta}+z}{e^{i\theta}-z}\,\overline{\varphi_n(e^{i\theta})}\,d\mu(\theta). \end{aligned} \tag{6.7}$$

Using $\int |\varphi_n|\,d\mu \leq 1$ and $(e^{i\theta}+z)(e^{i\theta}-z)^{-1} = 1 + \sum_{n=1}^\infty 2(e^{-i\theta}z)^n$, we see the Taylor coefficients of each expression in (6.7) are bounded by 2. Since $\int e^{-ik\theta}\varphi_n(e^{i\theta})\,d\mu(\theta) = 0$ for $k = 0,\dots,n-1$, we see $|F\varphi_n + \psi_n| \leq 2|z|^n(1-|z|)^{-1}$, while $|F\varphi_n^* - \psi_n^*| \leq 2|z|^{n+1}(1-|z|)^{-1}$. This proves not only an ℓ^2 property but exponential decay uniformly on compact subsets of $\mathbb{D}$.

The next issue we want to discuss is Lyapunov exponents. To understand them, it pays to also discuss the *density of zeros*, an object of independent interest. Given $d\mu$, a nontrivial probability measure in $\partial\mathbb{D}$, define the measure $d\nu_n$ on $\overline{\mathbb{D}}$ to be the point measure which gives weight k/n to a zero of Φ_n of multiplicity k. On account of (5.9) for $\ell = 0,1,2,\dots$,

$$\int z^\ell\,d\nu_n(z) = \frac{1}{n}\,\mathrm{Tr}([\mathcal{C}^{(n)}]^\ell), \tag{6.8}$$

which can help show that $d\nu_n$ sometimes has a weak limit. If it does, we say the limit is the *density of zeros*. The limit may not exist; there even exist examples (see Example 1.7.17 of [91]) where the set of limit points of $d\nu_n$ is all measures on $\overline{\mathbb{D}}$! Here is how (6.8) can be used:

Theorem 6.2 (Mhaskar-Saff [66]). *If*

$$\lim_{n\to\infty} |\alpha_n|^{1/n} = r \qquad \textit{and} \qquad \frac{1}{n}\sum_{j=0}^{n-1} |\alpha_j| \to 0 \tag{6.9}$$

(*automatic if* $r < 1$), *then* $d\nu_n$ *converges weakly to the uniform measure on the circle of radius* r.

Sketch of proof. (See Sections 8.1 and 8.2 of [91] for details.) An argument that exploits Theorem 9.1 below and the fact that $|\alpha_{n-1}| = |\Phi_n(0)|$ is the product of zeros shows that when the first equation in (6.9) holds, then all limits of $d\nu_n$ are concentrated on the circle of radius r. The second equation and (6.8) show that any such limit, $d\nu$, has $\int z^\ell\,d\nu = \delta_{\ell 0}$ for $\ell \geq 0$. $\square$

The other case where we know ν_n has a limit is ergodic families of Verblunsky coefficients. Let $(\Omega, d\beta)$ be a probability measure space, $T : \Omega \to \Omega$, an invertible ergodic transformation, and $V : \Omega \to \mathbb{D}$. For each $\omega \in \Omega$, define a measure $d\mu_\omega$ by

$$\alpha_j(d\mu_\omega) = V(T^j\omega). \tag{6.10}$$

An argument using the ergodic theorem, (6.8), and control of $\lim|\alpha_n(d\mu_\omega)|^{1/n}$ show that so long as $\int[-\log V(\omega)]\,d\beta(\omega) < \infty$, then $d\mu_\omega$ has for a.e. ω a limit supported on $\partial\mathbb{D}$ and ω-independent. The most important examples of ergodic families are random, periodic, almost periodic, and subshifts (see Chapters 10–12 of [92]).

Before leaving the subject of zeros, we note:

Theorem 6.3 (Widom [110]). *If* $\operatorname{supp}(d\mu)$ *is not all of* $\partial\mathbb{D}$*, then for any* $r < 1$, $\sup_n(\#$ *of zeros of* Φ_n *in* $|z| < r) < \infty$*. In particular, any limit of* $d\nu_n$ *is supported on* $\partial\mathbb{D}$*.*

Theorem 6.4 (see Theorem 8.1.11 of [91]). *If* z_0 *in* $\partial\mathbb{D}$ *is an isolated point of* $\operatorname{supp}(d\mu)$*, there is precisely a single zero of* Φ_n *near* z_0 *for n large, and it approaches* z_0 *exponentially fast.*

Finally, we discuss the Lyapunov exponent and Thouless formula.

Theorem 6.5 (see Theorem 10.5.8 of [92]). *If the density of zeros measure,* $d\nu$*, exists and is supported on* $\partial\mathbb{D}$*, and if*

$$\rho_\infty = \lim_{n\to\infty}\left(\prod_{j=0}^{n-1}\rho_j\right)^{1/n} \tag{6.11}$$

exists, then for $z \notin \partial\mathbb{D}$*, the following limit exists and is given by*

$$\gamma(z) \equiv \lim_{n\to\infty}\frac{1}{n}\log\|T_n(z)\| = -\log\rho_\infty - \int \log|e^{i\theta}-z|^{-1}\,d\nu(\theta). \tag{6.12}$$

γ is called the *Lyapunov exponent.* (6.12) is called the *Thouless formula.* For $|z| > 1$, $|\varphi_n| > |\varphi_n^*|$ and $|\psi_n| > |\psi_n^*|$, we need only control the growth of $|\varphi_n|$ and $|\psi_n|$. By (6.8), φ_n and ψ_n have the same density of zeros. Writing $\varphi_n = \prod_{j=0}^{n-1}\rho_j^{-1}\prod_{\text{zeros}}(z-z_\ell)$ easily yields (6.12).

See [92] for discussion of when (6.12) holds on $\partial\mathbb{D}$ and for further study of ergodic OPUC.

7. Khrushchev's formula, CMV resolvents, and Rakhmanov's theorem

In two remarkable papers [51, 52], Khrushchev found deep connections between Schur iterates and the structure of OPUC. A key input for the theory is:

Theorem 7.1 (Khrushchev's Formula). *The Schur function for the measure* $|\varphi_n(e^{i\theta}, d\mu)|^2\,d\mu(\theta)$ *is given by* $b_n(z)f_n(z)$*, where* f_n *is the n-th Schur iterate (by Geronimus' theorem, this is the Schur function of the measure with Verblunsky coefficients* $\{\alpha_{n+j}\}_{j=0}^\infty$*) and* b_n *is the Blaschke product,*

$$b_n(z;d\mu) = \frac{\varphi_n(z;d\mu)}{\varphi_n^*(z;d\mu)}. \tag{7.1}$$

Remark. Khrushchev's formula illuminates (2.9). In this trivial measure case, $\{z_j\}_{j=1}^{n-1}$ are the zeros of Φ_{n-1} and $e^{i\theta_0}$ is the Schur parameter, γ_{n-1}.

In terms of the CMV matrix, this gives a formula for $\langle\delta_n, (\mathcal{C}+z)(\mathcal{C}-z)^{-1}\delta_n\rangle$, and so, when $n = m$, for

$$G_{nm}(z) = \langle\delta_n, (\mathcal{C}-z)^{-1}\delta_m\rangle, \tag{7.2}$$

the analog of the Green's function in ODE's [17]. Half-line Green's functions for ODE's have the form $f_-(\min(x,y))f_+(\max(x,y))$ where f_- (resp. f_+) obeys boundary conditions at $x=0$ (resp. $x=\infty$). There is an analogous formula, due to Simon (even if $n\neq m$), for G_{nm} in terms of the OPUC and Weyl solutions. It can be found in Section 4.4 of [91] and generalizes Theorem 7.1. Other proofs of Theorem 7.1 appear in Theorem 4.5.10 of [91] and Theorem 9.2.4 of [92]. The most important consequence of Khrushchev's formula is

Theorem 7.2 (Khrushchev [51])**.** *The essential support of the a.c. part of $d\mu$ is all of $\partial\mathbb{D}$ if and only if*

$$\lim_{n\to\infty}\int_0^{2\pi}|f_n(e^{i\theta},d\mu)|^2\,\frac{d\theta}{2\pi}=0. \tag{7.3}$$

Since

$$\int_0^{2\pi} f_n(e^{i\theta},d\mu)\,\frac{d\theta}{2\pi}=f_n(0)=\alpha_n, \tag{7.4}$$

an immediate corollary is

Theorem 7.3 (Rakhmanov's Theorem)**.** *If the essential support of the a.c. part of $d\mu$ is all of $\partial\mathbb{D}$, then*

$$\lim_{n\to\infty}\alpha_n=0. \tag{7.5}$$

This result is originally due to Rakhmanov [83, 84, 85] with important further developments by Máté-Nevai-Totik [62, 63, 68, 69]. Bello-López [10] extended this result to arcs, and Denisov [21] to OPRL. Here are some other important results of Khrushchev's theory:

Theorem 7.4.

$$\text{w-lim}|\varphi_n(e^{i\theta})|^2\,d\mu=\frac{d\theta}{2\pi}\Leftrightarrow(\forall j\neq 0)\ \lim_{n\to\infty}\alpha_{n+j}\alpha_n=0.$$

Remark. (6.8) can be reinterpreted as saying weak Cesàro limits of $|\varphi_n|^2\,d\mu$ are the density of zeros when the latter is supported on $\partial\mathbb{D}$; see Section 8.2 of [91].

Theorem 7.5. *Let $f^{[n]}$ be the Schur approximates* (*given by* (3.9))*. Then*

$$\int|f^{[n]}(e^{i\theta})-f(e^{i\theta})|^2\,\frac{d\theta}{2\pi}\to 0 \tag{7.6}$$

if and only if either

(i) *$d\mu_{\rm ac}=0$, that is, μ is purely singular; or*

(ii) *$\alpha_n(d\mu)\to 0$.*

Moreover, if $\text{w-lim}|\varphi_n(e^{i\theta})|^2\,d\mu=d\theta/2\pi$*, then* (7.6) *holds.*

As a consequence of these theorems, we get a result for sparse α's.

Corollary 7.6. *If $\lim_{n\to\infty}\alpha_{n+j}\alpha_n=0$ for all $j\neq 0$, but $\limsup_n|\alpha_n|\neq 0$, then μ is purely singular continuous.*

Theorem 7.7. *Suppose that uniformly on compacts of $\partial\mathbb{D}$,*

$$\lim_{n\to\infty}\frac{\Phi_{n+1}^*(z)}{\Phi_n^*(z)}=G(z). \tag{7.7}$$

Then either $G(z)\equiv 1$ or else for some $a\in(0,1]$ and $\lambda\in\partial\mathbb{D}$,

$$G(z)=\tfrac12\left[(1+\lambda z)+\sqrt{(1-\lambda z)^2+4a^2\lambda z}\,\right]. \tag{7.8}$$

Note we have that $G \equiv 1$ if and only if $\lim_{n\to\infty} \alpha_{n+j}\alpha_n = 0$ for all $j \neq 0$ and that Barrios-López have proven that (7.8) holds if and only if $\lim_{n\to\infty}|\alpha_n| = a$ and $\lim_{n\to\infty} \alpha_{n+1}\alpha_n^{-1} = \lambda$.

Khrushchev has also described all possible $d\nu$'s that can occur as w-lim$|\varphi_n|^2\, d\mu$ (i.e., for which the limit exists) and when they can occur (essentially, asymptotically periodic with period 1 or 2). The analogs of these w-limit and ratio asymptotic results for OPRL were found by Simon [90].

8. Szegő's and Baxter's theorems

Szegő's theorems may well be the most celebrated in OPUC. While they have expressions purely in terms of OPUC objects, for historical reasons, one should state them in terms of Toeplitz determinants, $D_n(d\mu)$. This is defined as the determinant of the $(n+1)\times(n+1)$ matrix $\{c_{k-\ell}\}_{0\leq k,\ell\leq n}$ with c given by (2.8). D_n is the Gram determinant of $\{z^k\}_{k=0}^n$ since $\langle z^k, z^\ell\rangle_{L^2(d\mu)} = c_{k-\ell}$. The invariance of such determinants under triangular change of basis implies (using also (2.20))

$$D_n(d\mu) = \prod_{j=0}^{n} \|\Phi_j\|^2 = \prod_{j=0}^{n-1}(1-|\alpha_j|^2)^{n-j}, \tag{8.1}$$

which immediately implies

$$F(d\mu) \equiv \lim_{n\to\infty} D_n(d\mu)^{1/n} = \prod_{j=0}^{\infty}(1-|\alpha_j|^2) = \lim_{n\to\infty} \|\Phi_n\|^2 \tag{8.2}$$

$$G(d\mu) \equiv \lim_{n\to\infty} \frac{D_n(d\mu)}{F(d\mu)^{n+1}} = \prod_{j=0}^{\infty}(1-|\alpha_j|^2)^{-j-1}. \tag{8.3}$$

F is always defined, although it may be 0. G is defined so long as $F > 0$, that is, so long as $\sum_{j=0}^\infty|\alpha_j|^2 < \infty$. G may be infinite and is finite if and only if $\sum_{j=0}^\infty j|\alpha_j|^2 < \infty$.

Szegő's theorems express F and G in terms of the a.c. weight, w, of $d\mu$

$$d\mu = w(\theta)\,\frac{d\theta}{2\pi} + d\mu_{\mathrm{s}} \tag{8.4}$$

where $w \in L^1(\partial\mathbb{D}, \frac{d\theta}{2\pi})$ and $d\mu_{\mathrm{s}}$ is singular with respect to $d\theta/2\pi$.

Theorem 8.1 (Szegő's Theorem)**.**

$$F(d\mu) = \prod_{j=0}^{\infty}(1-|\alpha_j|^2) = \exp\biggl(\int \log(w(\theta))\,\frac{d\theta}{2\pi}\biggr). \tag{8.5}$$

Remark. Szegő proved this when $d\mu_{\mathrm{s}} = 0$ in 1915; the proof below is basically his proof in [99]. The result does not depend on $d\mu_{\mathrm{s}}$ — this was shown first by Verblunsky [107]. [91, 92] have five proofs of Theorem 8.1.

Sketch of proof when $d\mu_{\mathrm{s}} = 0$. Since Φ_n^* is nonvanishing on $\mathbb{D}$ and $\Phi_n^*(0) = 1$, $\int \log|\Phi_n^*(e^{i\theta})|\frac{d\theta}{2\pi} = 1$. Thus, by Jensen's inequality,

$$\|\Phi_n\|^2 \equiv \int |\Phi_n^*(e^{i\theta})|^2 w(\theta)\,\frac{d\theta}{2\pi} \geq \exp\biggl(\int \log(w(\theta))\,\frac{d\theta}{2\pi}\biggr),$$

so $F(d\mu) \geq$ RHS of (8.5). On the other hand, since Φ_n^* is the projection of 1 to the complement of $[z, \dots, z^n]$ in $[1, \dots, z^n]$, we have

$$\|\Phi_n^*\|^2 = \min\{\|P\|^2_{L^2(d\mu)} \mid \deg P \leq n,\ P(0) = 1\}. \tag{8.6}$$

Using (8.2) and a limit argument,

$$F(d\mu) = \min\{\|f\|^2_{L^2(d\mu)} \mid f \in H^\infty,\ f(0) = 1\}. \tag{8.7}$$

Pick the trial functions $f_\varepsilon(z) = g_\varepsilon(z)/g_\varepsilon(0)$ where

$$g_\varepsilon(z) = \exp\biggl(-\int \frac{e^{i\theta}+z}{e^{i\theta}-z} \, \log(w(\theta)+\varepsilon)\frac{d\theta}{4\pi}\biggr)$$

and take $\varepsilon \downarrow 0$ to get $F(d\mu) \leq$ RHS of (8.5). □

Because their singular continuous part is arbitrary, once an ℓ^2 condition is dropped, $d\mu$ can be arbitrarily "bad".

Theorem 8.2. *Let $d\rho$ be a measure on $\partial\mathbb{D}$ with support all of $\partial\mathbb{D}$. Then there exists $d\mu$, a probability measure on $\partial\mathbb{D}$ mutually equivalent to $d\rho$, so that for all $p > 2$,*

$$\sum_{n=0}^\infty |\alpha_n(d\mu)|^p < \infty. \tag{8.8}$$

This is Theorem 2.10.1 of [91], proven using ideas of Totik [104] and the bounds in (8.5).

By (8.5), we get one of the gems of spectral theory: equivalences between some recursion coefficient property and some spectral measure property.

Corollary 8.3.

$$\sum_{j=0}^\infty |\alpha_j|^2 < \infty \Leftrightarrow \int \log(w(\theta))\,\frac{d\theta}{2\pi} > -\infty. \tag{8.9}$$

The equivalent conditions (8.9) are called the *Szegő condition*. When they hold, Szegő defined the *Szegő function* by

$$D(z) = \exp\biggl(\int \frac{e^{i\theta}+z}{e^{i\theta}-z} \, \log(w(\theta))\frac{d\theta}{4\pi}\biggr). \tag{8.10}$$

Standard boundary value theory for the Poisson kernel implies $D(e^{i\theta}) = \lim_{r\uparrow 1} D(re^{i\theta})$ exists for $d\theta/2\pi$-a.e. $\theta \in [0, 2\pi)$ and

$$|D(e^{i\theta})|^2 = w(\theta). \tag{8.11}$$

Theorem 8.4 (Szegő [99])**.** *Suppose either and thus both of the conditions of* (8.9) *hold. Let $D_{\rm ac}(e^{i\theta}) = D(e^{i\theta})$ for a.e. θ and $= 0$ on a supporting set for $d\mu_{\rm s}$. Then*

$$\text{(i)} \qquad \int |\varphi_n^*(e^{i\theta}) - D_{\rm ac}(e^{i\theta})^{-1}|^2 \, d\mu \to 0 \tag{8.12}$$

$$\text{(ii)} \qquad \int |\varphi_n(e^{i\theta})|^2 \, d\mu_{\rm s} \to 0 \tag{8.13}$$

$$\text{(iii)} \qquad \varphi_n^*(z) \to D(z)^{-1} \quad \textit{uniform on compacts in } \mathbb{D}. \tag{8.14}$$

Sketch. A short preliminary argument proves that $D \in \mathbb{H}^2(\mathbb{D})$. Thus Cauchy's formula holds for $\varphi_n^* D$, so

$$\int (\varphi_n^* D)(e^{i\theta})\,\frac{d\theta}{2\pi} = \varphi_n^*(0)D(0) \to 1 \tag{8.15}$$

since, by (8.5), $\varphi_n^*(0)D(0) = \prod_{j=0}^{n-1}(1-|\alpha_j|^2)^{-1/2}\prod_{j=0}^{\infty}(1-|\alpha_j|^2)^{1/2}$. By (8.11),

$$\begin{aligned}\int |\varphi_n^* - D(e^{i\theta})^{-1}|^2 w(\theta)\,\frac{d\theta}{2\pi} + \int |\varphi_n^*(e^{i\theta})|^2\,d\mu_{\rm s}& \\ = \|\varphi_n^*\|^2_{L^2(d\mu)} + 1 - 2\operatorname{Re}(\varphi_n^*(0)D(0)) \to 0&\end{aligned}$$

by (8.15). This implies (i) and (ii). This then implies that $D\varphi_n^* \to 1$ in $L^2(\partial\mathbb{D}, \frac{d\theta}{2\pi})$, so by $\mathbb{H}^2$ theory, (iii) holds. $\square$

(8.14) and the related

$$z^{-n}\varphi_n(z) \to \overline{D(1/\bar{z})} \text{ on } \mathbb{C}\backslash\overline{\mathbb{D}} \tag{8.16}$$

are called *Szegő asymptotics*.

Theorem 8.5 (Sharp Form of the Strong Szegő Theorem [102, 44, 39])**.** *If $d\mu_{\rm s} = 0$, if the Szegő condition holds, and if $\hat{L}_k$ are the Fourier coefficients of $\log w$, then*

$$G(d\mu) = \prod_{j=0}^{\infty}(1-|\alpha_j|^2)^{-j-1} = \exp\biggl(\sum_{n=0}^{\infty} n|\hat{L}_n|^2\biggr). \tag{8.17}$$

Remark. Szegő [102] proved this when $d\mu$ has certain regularity properties. The general result is due to Ibragimov [44]; see also [39].

Seeing when $G(d\mu) < \infty$ leads to a second gem:

Corollary 8.6.

$$\sum_{j=0}^{\infty} j|\alpha_j|^2 < \infty \Leftrightarrow d\mu_{\rm s} = 0 \quad \textit{and} \quad \sum_{n=0}^{\infty} n|\hat{L}_n|^2 < \infty. \tag{8.18}$$

This corollary relies also on a theorem of Golinskii-Ibragimov [39] that the LHS of (8.18) $\Rightarrow d\mu_{\rm s} = 0$. This result plus five distinct proofs of Theorem 8.5 are found in Chapter 6 of [91]. A sixth proof is in Section 9.10 of [92].

A final gem we want to mention is

Theorem 8.7 (Baxter's Theorem [8, 9])**.** *Fix $\ell \geq 0$. The following are equivalent:*

(a) $\displaystyle\sum_{n=0}^{\infty} n^{\ell}|\alpha_n| < \infty$

(b) $\displaystyle d\mu_{\rm s} = 0, \quad \min_{\theta} w(\theta) > 0, \quad \textit{and} \quad \sum_{n=0}^{\infty} n^{\ell}|c_n| < \infty.$

In particular if $d\mu_{\rm s} = 0$ and $\min_\theta w(\theta) > 0$, then w is C^∞ if and only if $\sup_n n^\ell|\alpha_n| < \infty$ for all $\ell \geq 0$.

Remark. When $\sum_{n=0}^{\infty}|c_n| < \infty$, w has a representative which is continuous, and it is that choice we make for the otherwise a.e. defined function.

This is proven in Chapter 5 of [91].

9. Exponential decay

Suppose for some $R > 1$, we have

$$|\alpha_n| \leq CR^{-n}. \tag{9.1}$$

By (2.17) and induction,

$$\sup_{n,\,|z|=1} |\Phi_n(z)| \leq \prod_{j=0}^{\infty}(1 + |\alpha_j|) < \infty, \tag{9.2}$$

so, by the maximum principle and (2.15),

$$\sup_{|z|\geq 1} |z|^{-n}|\Phi_n(z)| = \sup_{|z|\leq 1} |\Phi_n^*(z)| < \infty.$$

Thus, by (2.27), if $|z| < R$,

$$\sum_{n=0}^{\infty} |\Phi_{n+1}^*(z) - \Phi_n^*(z)| \leq \sum_{n=0}^{\infty} |\alpha_n|\,|z|^{n+1} < \infty. \tag{9.3}$$

It follows that $\Phi_n^*(z)$ and so $\varphi_n^*(z)$ converge uniformly on compacts of $\{z \mid |z| < R\}$, and so, by (8.14), $D(z)^{-1}$ has a continuation to this disk. We thus have one-half of

Theorem 9.1 (Nevai-Totik [70])**.** *Fix $R > 1$. The following are equivalent:*
(a) $d\mu_{\mathrm{s}} = 0$*, the Szegő condition holds, and* $D(z)^{-1}$ *has an analytic continuation to* $\{z \mid |z| < R\}$*.*
(b) $\limsup_{n\to\infty} |\alpha_n|^{1/n} = R^{-1}$*.*

The other direction uses the useful formula,

$$d\mu_{\mathrm{s}} = 0 \Rightarrow \alpha_n = -D(0)^{-1} \int \overline{\Phi_{n+1}(e^{i\theta})}\, D(e^{i\theta})^{-1}\, d\mu(\theta). \tag{9.4}$$

Section 7.1 of [91] has a complete proof. One can say more ([91, Section 7.2], [20, 95]). When this holds, $\alpha_n +$ Taylor coefficients of $D(z)^{-1}\,\overline{D(1/\bar{z})}$ decay as $0(R^{-3n+\varepsilon})$. There is also a lot known about asymptotics of the zeros when there is exponential decay (see [91, Sections 8.1 and 8.2], [94, 61, 95] and references therein).

10. Periodic OPUC

The theory of one-dimensional periodic Schrödinger operators (a.k.a. Hill's equation) and of periodic Jacobi matrices has been extensively developed [23, 26, 57, 58, 64, 105]. In the 1940's, Geronimus [32] found the earliest results on OPUC with periodic Verblunsky coefficients; that is, for some integral $p \geq 1$ and $j = 0, 1, 2, \ldots,$

$$\alpha_{j+p} = \alpha_j. \tag{10.1}$$

In particular, the case $\alpha_j \equiv a \in \mathbb{D}\backslash\{0\}$ yields OPUC called *Geronimus polynomials* (see Example 1.6.12 of [91]). Many of the general features for OPUC obeying (10.1) were found by Peherstorfer and his collaborators [71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81]. Geronimo-Johnson [30, 31] have studied almost periodic Verblunsky coefficients. A reworking with some new results is Chapter 11 of [92], which uses methods mimicking the periodic Hill-Jacobi theory.

We suppose henceforth that p is even. A basic object is the *discriminant*,

$$\Delta(z) = \mathrm{Tr}(z^{-p/2}T_p(z)), \tag{10.2}$$

where $T_p(z)$ is the transfer matrix given by (6.2). The $z^{-p/2}$ is included since, by $\det(A) = z$, $\det(z^{-p/2}T_p(z)) = 1$, and so $z^{-p/2}T_p(z)$ has eigenvalues $\frac{\Delta}{2} \pm i\sqrt{1-(\frac{\Delta}{2})^2}$. In particular, these eigenvalues have magnitude 1; that is, $\sup_m \|T_{mp}(z)\| < \infty$ exactly when $\Delta(z) \in [-2,2]$. This is part of

Theorem 10.1. *There exist* $\{x_j\}_{j=1}^{2p}$, $\{y_j\}_{j=1}^{2p}$ *with*

$$x_1 < y_1 \le x_2 < y_2 \le \dots \le x_p < y_p \le x_1 + 2\pi =: x_{p+1}$$

so that the solutions of $\Delta(z) = 2$ (*resp.* -2) *are exactly* $e^{ix_1}, e^{iy_2}, e^{ix_3}, \dots, e^{ix_p}$ (*resp.* $e^{iy_1}, e^{ix_2}, e^{iy_3}, \dots, e^{iy_p}$) *and* $\Delta(z) \in [-2,2]$ *exactly on the* bands

$$B = \bigcup_{j=1}^{p} \{e^{i\theta} \mid x_j \le \theta_j \le y_j\}. \tag{10.3}$$

B is the essential support of $d\mu_{\text{ac}}$, and the only possible singular spectrum is composed of mass points which can occur in open gaps (*i.e., nonempty sets of the form* $\{e^{i\theta} \mid y_j < \theta < x_{j+1}\}$) *with one* (*or zero*) *mass point in each gap.*

Theorem 10.2. *Let $d\rho$ be the equilibrium measure for B* (*i.e., the minimizer for* $\mathcal{E}(\rho) = \int \log|z-w|^{-1}\, d\rho(z)\, d\rho(w)$ *with* $\operatorname{supp}(d\rho) \subset B$ *and* $\rho(B) = 1$). *Let C_B be the logarithmic capacity of B* (*i.e.,* $\exp(-$ *minimizing value of* $\mathcal{E}(p))$) *and $Q(z)$ the logarithmic potential for B* (*i.e.,* $Q(z) = \int \log|z-\omega|^{-1}\, d\rho(\omega)$). *Then*

(i) *$d\rho$ is the density of zeros for $d\mu$,* $-[Q(z) + \log C_B]$ *is the Lyapunov exponent, and* $C_B = \prod_{j=0}^{p-1}(1-|\alpha_j|^2)^{1/2p}$.

(ii) *$d\rho$ can be written in terms of the discriminant as*

$$d\rho(\theta) =: \frac{1}{p}\, \frac{|\partial\Delta(e^{i\theta})/\partial\theta|}{(4-\Delta^2(e^{i\theta}))^{1/2}}\, \frac{d\theta}{2\pi}. \tag{10.4}$$

(iii) *For each* $j = 1, 2, \dots, p$,

$$\rho(\{e^{i\theta} \mid x_j \le \theta \le y_j\}) = \frac{1}{p}. \tag{10.5}$$

The proof of this result (see Section 11.1 of [92]) depends on noting that, by the Thouless formula, $\gamma(z)$ is harmonic on $\mathbb{C}\backslash B$ and $\gamma(z) = 0$ on B. (iii) is related to half of the following result of Peherstorfer motivated by an OPRL result of Aptekarev [3] (based in part on Geronimus [35]).

Theorem 10.3 (Peherstorfer [71])**.** *Let B be a union of ℓ disjoint, closed intervals, $B_1, \dots, B_\ell$, in $\partial\mathbb{D}$. Then B is the set of bands of a period p set of α's if and only if the following conditions are true:*

(1) *If $d\rho$ is the equilibrium measure of B, then $p\rho(B_j) \in \mathbb{Z}$ for $j = 1, \dots, \ell$.*

(2) *Let $z_1, z_2, \dots, z_{2p}$ be defined clockwise around the circle so that z_1 is the lower edge of B_1 and the $2p$ points are the 2ℓ band edges and those interior points in a band with $p\rho(B_j) \ge 2$ that divide B_j into $p\rho(B_j)$ sets with ρ-measure $1/p$, each counted twice. Then* $z_1 z_4 z_5 z_8 z_9 \dots z_{2p} = 1$.

If some $\rho(B_j)$ is irrational, then there is no periodic family of α's with those bands, but there is an almost periodic set, as proven by Geronimo-Johnson [31] (see Section 11.8 of [92]).

Given a measure $d\mu$ on $\partial\mathbb{D}$ so that (10.1) holds, the *Dirichlet data* is defined partly as the p points where $\binom{1}{1}$ is an eigenvector for $T_p(z)$, that is, zeros of $\varphi_p^*(z) - \varphi_p(z)$. There is one such point in each gap, including closed gaps (i.e., e^{iy_j} when $y_j = x_{j+1}$). If the value is at a gap edge, the eigenvalue λ of $z^{-p/2}T_p(z)$ for $\binom{1}{1}$ is ± 1. Otherwise, it is in $\mathbb{R}\backslash\{0,-1,1\}$. In that case, we add $\sigma_j = \pm 1$ to the j-th Dirichlet point with $\sigma_j = +1$ (resp. -1) of the eigenvalue $|\lambda_j| < 1$ (resp. $|\lambda_{-j}| > 1$). The point masses of $d\mu$ are precisely those Dirichlet points inside gaps with $\sigma_j = +1$. The set of allowed Dirichlet data is composed of single points for closed gaps and a circle ($[y_j, z_{j+1}] \times \{-1,1\}$ glued at the ends) for open gaps. Thus, the totality is a torus of dimension $\ell = \#$ of open gaps.

Theorem 10.4. *If Δ has ℓ open gaps, then the subset of $\{\alpha_j\}_{j=0}^{p-1} \in \mathbb{D}^p$ which, when periodized, have discriminant Δ, is a torus of dimension ℓ. The map from these α's to the possible Dirichlet date is a bijection.*

Critical to at least one understanding of this result is that the Carathéodory function F has a minimal degree meromorphic continuation to the genus $\ell - 1$ hyperelliptic Riemann surface associated to $\sqrt{\Delta^2 - 4}$.

There is a natural symplectic form on $\mathbb{D}^p$ so that the real and imaginary parts of the coefficients of Δ comprise the set of integrals of a completely integrable system. This is described in Section 11.11 of [92] and in [67]. The associated flows include the defocusing Ablowitz-Ladik flow.

11. The Szegő mapping and the Geronimus relations

Finally, we discuss a deep connection between OPRL and OPUC found by Szegő [100]. The map $z \mapsto z + z^{-1}$ maps $\mathbb{D}$ biholomorphically to $\mathbb{C} \cup \{\infty\}$ with a cut $[-2,2]$ removed. The map on the boundary $e^{i\theta} \to 2\cos\theta$ is a two-to-one map of $\partial\mathbb{D}$ to $[-2,2]$ that induces a map from $\mathcal{M}_{+,1}([-2,2])$ to those measures on $\partial\mathbb{D}$ which are invariant under complex conjugation. It is easy to see $\mu \in \mathcal{M}_{+,1}(\partial\mathbb{D})$ has such invariance if and only if its Verblunsky coefficients are real. Explicitly, ρ, a probability measure on $[-2,2]$, is associated to $\mu = \mathrm{Sz}(\rho)$, an even probability measure on $\partial\mathbb{D}$, via

$$\int f(x)\, d\rho(x) = \int f(2\cos\theta)\, d\mu(\theta). \tag{11.1}$$

Szegő found the OPRL P_n for ρ in terms of the OPUC Φ_n for μ:

$$P_n\left(z + \frac{1}{z}\right) = [1 - \alpha_{2n-1}(d\mu)]^{-1} z^{-n} [\Phi_{2n}(z) + \Phi_{2n}^*(z)]. \tag{11.2}$$

and used this to convert Szegő asymptotics for OPUC (see (8.16)) to asymptotics for suitable OPRL. This asymptotics is often called *Jost asymptotics* in the discrete Schrödinger literature.

Geronimus [33] found the relation between the Jacobi parameters $\{a_n, b_n\}_{n=1}^\infty$ for ρ and the Verblunsky coefficients $\{\alpha_n\}_{n=0}^\infty$ for μ (with $\alpha_{-1} \equiv 1$):

$$a_{n+1}^2 = (1 - \alpha_{2n-1})(1 - \alpha_{2n}^2)(1 + \alpha_{2n+1}) \tag{11.3}$$

$$b_{n+1} = (1 - \alpha_{2n-1})\alpha_{2n} - (1 + \alpha_{2n-1})\alpha_{2n-2}. \tag{11.4}$$

The map from α to (a,b) is local; that is, changing a single α changes only a finite number of a's and b's. That is not true for the inverse. Scaled Chebyshev polynomials of the first kind have $a_1 = \sqrt{2}$, $a_n = 1$ $(n \geq 2)$, $b_n = 0$, and the corresponding $\alpha_n \equiv 0$. Scaled Chebyshev polynomials of the second kind have $a_n \equiv 1$, $b_n = 0$ (i.e., they differ at a single a_n), but have $\alpha_{2n} = 0$ and $\alpha_{2n-1} = -1/(n+1)$.

Still the inverse can be computed ([33]). Given $\{a_n, b_n\}_{n=1}^\infty$, define $\varphi_n^\pm$ by $\varphi_0 = 0$, $\varphi_1 = 1$, and for $n \geq 1$,

$$\varphi_{n+1}^\pm + a_{n-1}^2 \varphi_{n-1}^\pm + b_n \varphi_n^\pm = \pm 2\varphi_n^\pm. \tag{11.5}$$

By a Sturm oscillation theorem, $\{a_n, b_n\}_{n=1}^\infty$ are the Jacobi parameters of a measure supported on $[-2, 2]$ if and only if $\varphi_n^+ > 0$ and $(-1)^n \varphi_n^- > 0$. The Verblunsky coefficients are given by

$$u_n = \frac{\varphi_{n+2}^+}{\varphi_{n+1}^+} \qquad v_n = -\frac{\varphi_{n+2}^-}{\varphi_{n+1}^-} \tag{11.6}$$

$$\alpha_{2n} = \frac{v_n - u_n}{v_n + u_n} \qquad \alpha_{2n-1} = 1 - \tfrac{1}{2}(u_n + v_n). \tag{11.7}$$

Recently, these mappings have been used by Denisov [21] and Damanik-Killip [18, 89] as a powerful tool in the study of discrete Schrödinger operators and of OPRL. For proofs and references, see Chapter 13 of [92].

References

[1] N. I. Akhiezer, *The Classical Moment Problem and Some Related Questions in Analysis*, Hafner, New York, 1965; Russian original, 1961.

[2] A. B. Aleksandrov, *Multiplicity of boundary values of inner functions*, Izv. Akad. Nauk Arm. SSR **22** (1987), 490–503. MR0931885 (89e:30058)

[3] A. I. Aptekarev, *Asymptotic properties of polynomials orthogonal on a system of contours, and periodic motions of Toda chains*, Math. USSR Sb. **53** (1986), 233–260; Russian original in Mat. Sb. (N.S.) **125(167)** (1984), 231–258. MR0764479 (86g:35166)

[4] J. Baik, P. Deift, and K. Johansson, *On the distribution of the length of the longest increasing subsequence of random permutations*, J. Amer. Math. Soc. **12** (1999), 1119–1178. MR1682248 (2000e:05006)

[5] J. Baik, P. Deift, and K. Johansson, *On the distribution of the length of the second row of a Young diagram under Plancherel measure*, Geom. Funct. Anal. **10** (2000), 702–731. MR1791137 (2001m:05258a)

[6] J. Baik, P. Deift, K. T.-R. McLaughlin, P. Miller, and X. Zhou, *Optimal tail estimates for directed last passage site percolation with geometric random variables*, Adv. Theor. Math. Phys. **5** (2001), 1207–1250. MR1926668 (2003h:60141)

[7] J. Baik and E. Rains, *Algebraic aspects of increasing subsequences*, Duke Math. J. **109** (2001), 1–65. MR1844203 (2002i:05119)

[8] G. Baxter, *A convergence equivalence related to polynomials orthogonal on the unit circle*, Trans. Amer. Math. Soc. **99** (1961), 471–487. MR0126126 (23:A3422)

[9] G. Baxter, *A norm inequality for a "finite-section" Wiener-Hopf equation*, Illinois J. Math. **7** (1963), 97–103. MR0145285 (26:2818)

[10] M. Bello Hernández and G. López Lagomasino, *Ratio and relative asymptotics of polynomials orthogonal on an arc of the unit circle*, J. Approx. Theory **92** (1998), 216–244. MR1604927 (99c:42041)

[11] S. Bernstein, *Sur une classe de polynomes orthogonaux*, Commun. Kharkow **4** (1930), 79–93.

[12] A. Borodin and E. Strahov, *Averages of characteristic polynomials in random matrix theory*, preprint.

[13] O. Bourget, J. S. Howland, and A. Joye, *Spectral analysis of unitary band matrices*, Comm. Math. Phys. **234** (2003), 191–227. MR1962460 (2004c:47063)

[14] M. J. Cantero, L. Moral, and L. Velázquez, *Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle*, Linear Algebra Appl. **362** (2003), 29–56. MR1955452 (2003k:42046)

[15] C. Carathéodory, *Über den Variabilitätsbereich der Koeffizienten von Potenzreihen die gegebene Werte nicht annehmen*, Math. Ann. **64** (1907), 95–115.

[16] E. B. Christoffel, *Über die Gaussische Quadratur und eine Verallgemeinerung derselben*, J. Reine Angew. Math. **55** (1858), 61–82.

[17] E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York-Toronto-London, 1955. MR0069338 (16:1022b)

[18] D. Damanik and R. Killip, *Half-line Schrödinger operators with no bound states*, Acta Math. **193** (2004), 31–72.

[19] G. Darboux, *Mémoire sur l'approximation des fonctions de très-grands nombres, et sur une classe étendue de développements en série*, Liouville J. (3) **4** (1878), 5–56; 377–416.

[20] P. Deift and J. Ostensson, *A Riemann-Hilbert approach to some theorems on Toeplitz operators and orthogonal polynomials*, in preparation.

[21] S. A. Denisov, *On Rakhmanov's theorem for Jacobi matrices*, Proc. Amer. Math. Soc. **132** (2004), 847–852. MR2019964

[22] J. Dombrowski, *Quasitriangular matrices*, Proc. Amer. Math. Soc. **69** (1978), 95–96. MR0467373 (57:7232)

[23] B. A. Dubrovin, V. B. Matveev, and S. P. Novikov, *Nonlinear equations of Korteweg-de Vries type, finite-band linear operators and Abelian varieties*; Uspekhi Mat. Nauk **31** (1976), no. 1(187), 55–136 [Russian]. MR0427869 (55:899)

[24] 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. MR1150798 (93a:42010)

[25] L. Fejér, *Über die Lage der Nullstellen von Polynomen, die aus Minimumforderungen gewisser Art entspringen*, Math. Ann. **85** (1922), 41–48.

[26] H. Flaschka and D. W. McLaughlin, *Canonically conjugate variables for the Korteweg-de Vries equation and the Toda lattice with periodic boundary conditions*, Progr. Theoret. Phys. **55** (1976), 438–456. MR0403368 (53:7179)

[27] G. Freud, *Orthogonal Polynomials*, Pergamon Press, Oxford-New York, 1971.

[28] I. M. Gel'fand and B. M. Levitan, *On the determination of a differential equation from its spectral function*, Amer. Math. Soc. Transl. (2) **1** (1955), 253–304; Russian original in Izvestiya Akad. Nauk SSSR. Ser. Mat. **15** (1951), 309–360. MR0073805 (17:489c); MR0045281 (13:558f)

[29] J. S. Geronimo, *Polynomials orthogonal on the unit circle with random recurrence coefficients*, in "Methods of Approximation Theory in Complex Analysis and Mathematical Physics" (Leningrad, 1991), pp. 43–61, Lecture Notes in Math., 1550, Springer, Berlin, 1993. MR1322290 (95k:42041)

[30] J. S. Geronimo and R. Johnson, *Rotation number associated with difference equations satisfied by polynomials orthogonal on the unit circle*, J. Differential Equations **132** (1996), 140–178. MR1418504 (98a:39002)

[31] J. S. Geronimo and R. Johnson, *An inverse problem associated with polynomials orthogonal on the unit circle*, Comm. Math. Phys. **193** (1998), 125–150. MR1620309 (2000a:39008)

[32] 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]. MR0012715 (7:63e)

[33] Ya. L. Geronimus, *On the trigonometric moment problem*, Ann. of Math. (2) **47** (1946), 742–761. MR0018265 (8:265d)

[34] Ya. L. Geronimus, *Polynomials Orthogonal on a Circle and Their Applications*, Amer. Math. Soc. Translation **1954** (1954), no. 104, 79 pp. MR0061706 (15:869i)

[35] Ya. L. Geronimus, *On some equations in finite differences and the corresponding system of orthogonal polynomials*, Zap. Mat. Otdel Fiz.-Mat. Fak. i Khar'kov. Mat. Obšč **25** (1957), 87-100 [Russian].

[36] 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. MR0133643 (24:A3469)

[37] F. Gesztesy and M. Zinchenko, *A Borg-type theorem associated with orthogonal polynomials on the unit circle*, preprint, 2004.

[38] F. Gesztesy and M. Zinchenko, *Weyl–Titchmarsh theory for CMV operators associated with orthogonal polynomials on the unit circle*, preprint, 2004.

[39] B. L. Golinskii and I. A. Ibragimov, *On Szegő's limit theorem*, Math. USSR Izv. **5** (1971), 421–444.

[40] L. Golinskii, *Quadrature formula and zeros of para-orthogonal polynomials on the unit circle*, Acta Math. Hungar. **96** (2002), 169–186. MR1919160 (2003e:41048)

[41] L. Golinskii, *Orthogonal polynomials on the unit circle, Szegő difference equations and spectral theory of unitary matrices*, second Doctoral thesis, Kharkov, 2003.

[42] L. Golinskii and P. Nevai, *Szegő difference equations, transfer matrices and orthogonal polynomials on the unit circle*, Comm. Math. Phys. **223** (2001), 223–259. MR1864433 (2002k:42051)

[43] L. Golinskii and B. Simon, Results on spectral theorem using CMV matrices in Section 4.3 of [91].

[44] I. A. Ibragimov, *A theorem of Gabor Szegő*, Mat. Zametki **3** (1968), 693–702 [Russian]. MR0231114 (37:6669)

[45] V. A. Javrjan, *A certain inverse problem for Sturm-Liouville operators*, Izv. Akad. Nauk Armjan. SSR Ser. Mat. **6** (1971), 246–251 [Russian]. MR0301565 (46:723)

[46] K. Johansson, *Shape fluctuations and random matrices*, Comm. Math. Phys. **209** (2000), 437–476. MR1737991 (2001h:60177)

[47] 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. MR0976057 (90e:42027)

[48] T. Kailath, *A view of three decades of linear filtering theory*, IEEE Trans. Inform. Theory **IT-20** (1974), 146–181. MR0465437 (57:5337)

[49] T. Kailath, *Signal processing applications of some moment problems*, in "Moments in Mathematics," (San Antonio, Tex., 1987), pp. 71–109, Proc. Sympos. Appl. Math., 37, American Mathematical Society, Providence, R.I., 1987. MR0921085 (89e:94001)

[50] T. Kailath, *Norbert Wiener and the development of mathematical engineering*, in "The Legacy of Norbert Wiener: A Centennial Symposium," Proc. Sympos. Pure Math., 60, pp. 93–116, American Mathematical Society, Providence, R.I., 1997. MR1460278 (98i:01023)

[51] 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. MR1815919 (2002b:42032)

[52] S. Khrushchev, *Classification theorems for general orthogonal polynomials on the unit circle*, J. Approx. Theory **116** (2002), 268–342. MR1911083 (2003e:42036)

[53] R. Killip and I. Nenciu, *Matrix models for circular ensembles*, Internat. Math. Res. Notices, (2004) no. 50, 2665–2701.

[54] A. N. Kolmogorov, *Stationary sequences in Hilbert space*, Bull. Univ. Moscow **2** (1941), 40 pp. [Russian]. MR0009098 (5:101c)

[55] 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. MR0013457 (7:156b)

[56] M. G. Krein, *On a problem of extrapolation of A.N. Kolmogorov*, Dokl. Akad. Nauk SSSR **46** (1945), 306–309. MR0012700 (7:61e)

[57] I. M. Krichever, *Algebraic curves and nonlinear difference equations*, Uspekhi Mat. Nauk **33** (1978), no. 4(202), 215–216 [Russian]. MR0510681 (80k:58055)

[58] I. M. Krichever, *Appendix to "Theta-functions and nonlinear equations" by B.A. Dubrovin*, Russian Math. Surveys **36** (1981), 11–92 (1982); Russian original in Uspekhi Mat. Nauk **36** (1981), no. 2(218), 11–80. MR0616797 (83i:35149)

[59] H. J. Landau, *Maximum entropy and the moment problem*, Bull. Amer. Math. Soc. **16** (1987), 47–77. MR0866018 (88k:42010)

[60] 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. MR0019257 (8:391e)

[61] A. Martinez-Finkelshtein, K. T.-R. McLaughlin, and E. B. Saff, *Strong asymptotics of Szegő orthogonal polynomials with respect to an analytic weight*, preprint.

[62] A. Máté, P. Nevai, and V. Totik, *Asymptotics for the ratio of leading coefficients of orthonormal polynomials on the unit circle*, Constr. Approx. **1** (1985), 63–69. MR0766095 (85j:42045)

[63] A. Máté, P. Nevai, and V. Totik, *Strong and weak convergence of orthogonal polynomials*, Amer. J. Math. **109** (1987), 239–281. MR0882423 (88d:42040)

[64] H. P. McKean and P. van Moerbeke, *The spectrum of Hill's equation*, Invent. Math. 30 (1975), 217–274. MR0397076 (53:936)

[65] M. Mehta, *Random Matrices*, second ed., Academic Press, Inc., Boston, 1991. MR1083764 (92f:82002)

[66] H. N. Mhaskar and E. B. Saff, *On the distribution of zeros of polynomials orthogonal on the unit circle*, J. Approx. Theory **63** (1990), 30–38. MR1074079 (92f:42030)

[67] I. Nenciu, *Lax pairs for the Ablowitz-Ladik system via orthogonal polynomials on the unit circle*, to appear in Internat. Math. Res. Notices.

[68] P. Nevai, *Characterization of measures associated with orthogonal polynomials on the unit circle*, in "Constructive Function Theory—86 Conference" (Edmonton, Alta., 1986), Rocky Mountain J. Math. **19** (1989), 293–302. MR1016182 (90m:42034a)

[69] P. Nevai, *Weakly convergent sequences of functions and orthogonal polynomials*, J. Approx. Theory **65** (1991), 322–340. MR1109411 (92f:42031)

[70] P. Nevai and V. Totik, *Orthogonal polynomials and their zeros*, Acta Sci. Math. (Szeged) **53** (1989), 99–104. MR1018677 (90i:33021)

[71] F. Peherstorfer, *Orthogonal and extremal polynomials on several intervals*, in "Proc. Seventh Spanish Symposium on Orthogonal Polynomials and Applications (VII SPOA)" (Granada, 1991), J. Comput. Appl. Math. **48** (1993), 187–205. MR1246858 (94m:42058)

[72] F. Peherstorfer, *A special class of polynomials orthogonal on the unit circle including the associated polynomials*, Constr. Approx. **12** (1996), 161–185. MR1393285 (97d:42023)

[73] F. Peherstorfer, *Deformation of minimal polynomials and approximation of several intervals by an inverse polynomial mapping*, J. Approx. Theory **111** (2001), 180–195. MR1849545 (2002g:41009)

[74] F. Peherstorfer, *Inverse images of polynomial mappings and polynomials orthogonal on them*, in "Proc. Sixth International Symposium on Orthogonal Polynomials, Special Functions and their Applications" (Rome, 2001), J. Comput. Appl. Math. **153** (2003), 371–385. MR1985708 (2004e:30009)

[75] F. Peherstorfer and R. Steinbauer, *Perturbation of orthogonal polynomials on the unit circle—a survey*, In "Orthogonal Polynomials on the Unit Circle: Theory and Applications" (Madrid, 1994), pp. 97–119, Univ. Carlos III Madrid, Leganés, 1994. MR1317108 (95j:42022)

[76] F. Peherstorfer and R. Steinbauer, *Orthogonal polynomials on arcs of the unit circle, I*, J. Approx. Theory **85** (1996), 140–184. MR1385813 (97k:42052)

[77] F. Peherstorfer and R. Steinbauer, *Orthogonal polynomials on arcs of the unit circle, II. Orthogonal polynomials with periodic reflection coefficients*, J. Approx. Theory **87** (1996), 60–102. MR1410612 (97k:42053)

[78] F. Peherstorfer and R. Steinbauer, *Asymptotic behaviour of orthogonal polynomials on the unit circle with asymptotically periodic reflection coefficients*, J. Approx. Theory **88** (1997), 316–353. MR1432577 (97m:42019)

[79] F. Peherstorfer and R. Steinbauer, *Asymptotic behaviour of orthogonal polynomials on the unit circle with asymptotically periodic reflection coefficients, II. Weak asymptotics*, J. Approx. Theory **105** (2000), 102–128. MR1768526 (2001i:42042)

[80] F. Peherstorfer and R. Steinbauer, *Orthogonal polynomials on the circumference and arcs of the circumference*, J. Approx. Theory **102** (2000), 96–119. MR1736047 (2001b:42035)

[81] F. Peherstorfer and R. Steinbauer, *Strong asymptotics of orthonormal polynomials with the aid of Green's function*, SIAM J. Math. Anal. **32** (2000), 385–402. MR1781222 (2002k:33010)

[82] M. Praehofer and H. Spohn, *Universal distributions for growth processes in* $1+1$ *dimensions and random matrices*, Phys. Rev. Lett. **84** (2000), 4882–4885.

[83] E. A. Rakhmanov, *On the asymptotics of the ratio of orthogonal polynomials*, Math. USSR Sb. **32** (1977), 199–213.

[84] E. A. Rakhmanov, *On the asymptotics of the ratio of orthogonal polynomials, II*, Math. USSR Sb. **46** (1983), 105–117.

[85] E. A. Rakhmanov, *Asymptotic properties of polynomials orthogonal on the circle with weights not satisfying the Szegő condition*, Math. USSR-Sb. **58** (1987), 149–167; Russian original in Mat. Sb. (N.S.) **130(172)** (1986), 151–169, 284. MR0854969 (88b:42033)

[86] W. Rudin, *Real and Complex Analysis*, 3rd edition, McGraw-Hill, New York, 1987. MR0924157 (88k:00002)

[87] I. Schur, *Über Potenzreihen, die im Innern des Einheitskreises beschränkt 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" (edited by I. Gohberg), pp. 31–59; pp. 66–88, Operator Theory: Advances and Applications, 18, Birkhäuser, Basel, 1986. MR0902600 (88d:00006)

[88] B. Simon, *The classical moment problem as a self-adjoint finite difference operator*, Adv. in Math. **137** (1998), 82–203. MR1627806 (2001e:47020)

[89] B. Simon, *The Golinskii-Ibragimov method and a theorem of Damanik-Killip*, Int. Math. Res. Not. (2003), 1973–1986. MR1991180 (2004d:42045)

[90] B. Simon, *Ratio asymptotics and weak asymptotic measures for orthogonal polynomials on the real line*, J. Approx. Theory. **126** (2004), 198–217. MR2045539 (2005e:42078)

[91] B. Simon, *Orthogonal Polynomials on the Unit Circle, Part 1: Classical Theory*, AMS Colloquium Series, American Mathematical Society, Providence, RI, 2005.

[92] B. Simon, *Orthogonal Polynomials on the Unit Circle, Part 2: Spectral Theory*, AMS Colloquium Series, American Mathematical Society, Providence, RI, 2005. MR2105089

[93] B. Simon, *Aizenman's theorem for orthogonal polynomials on the unit circle*, to appear in Const. Approx.

[94] B. Simon, *Fine structure of the zeros of orthogonal polynomials, II. OPUC with competing exponential decay*, to appear in J. Approx. Theory.

[95] B. Simon, *Meromorphic Szegő functions and asymptotic series for Verblunsky coefficients*, preprint.

[96] B. Simon and T. Spencer, *Trace class perturbations and the absence of absolutely continuous spectra*, Comm. Math. Phys. **125** (1989), 113–125. MR1017742 (91g:81018)

[97] B. Simon and T. Wolff, *Singular continuous spectrum under rank one perturbations and localization for random Hamiltonians*, Comm. Pure Appl. Math. **39** (1986), 75–90. MR0820340 (87k:47032)

[98] G. Szegő, *Über Orthogonalsysteme von Polynomen*, Math. Z. **4** (1919), 139–151.

[99] G. Szegő, *Beiträge zur Theorie der Toeplitzschen Formen, I, II*, Math. Z. **6** (1920), 167–202; **9** (1921), 167–190.

[100] G. Szegő, *Über den asymptotischen Ausdruck von Polynomen, die durch eine Orthogonalitätseigenschaft definiert sind*, Math. Ann. **86** (1922), 114–139.

[101] G. Szegő, *Orthogonal Polynomials*, Amer. Math. Soc. Colloq. Publ., Vol. 23, American Mathematical Society, Providence, R.I., 1939; 3rd edition, 1967.

[102] G. Szegő, *On certain Hermitian forms associated with the Fourier series of a positive function*, Comm. Sém. Math. Univ. Lund **1952** (1952), Tome Supplementaire, 228–238. MR0051961 (14:553d)

[103] *Talmud Bavli*, Tractate Shabbos, 31a; see, for example, Schottenstein Edition, Mesorah Publications, New York, 1996.

[104] V. Totik, *Orthogonal polynomials with ratio asymptotics*, Proc. Amer. Math. Soc. **114** (1992), 491–495. MR1065095 (92e:42017)

[105] P. van Moerbeke, *The spectrum of Jacobi matrices*, Invent. Math. **37** (1976), 45–81. MR0650253 (58:31226)

[106] S. Verblunsky, *On positive harmonic functions: A contribution to the algebra of Fourier series*, Proc. London Math. Soc. (2) **38** (1935), 125–157.

[107] S. Verblunsky, *On positive harmonic functions* (*second paper*), Proc. London Math. Soc. (2) **40** (1936), 290–320.

[108] F. Wegner, *Bounds on the density of states in disordered systems*, Z. Phys. B **44** (1981), 9–15. MR0639135 (83b:82060)

[109] B. Wendroff, *On orthogonal polynomials*, Proc. Amer. Math. Soc. **12** (1961), 554–555. MR0131120 (24:A974)

[110] H. Widom, *Polynomials associated with measures in the complex plane*, J. Math. Mech. **16** (1967), 997–1013. MR0209448 (35:346)

[111] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series. With Engineering Applications*, The Technology Press of the Massachusetts Institute of Technology, Cambridge, MA; John Wiley & Sons, Inc., New York, N. Y.; Chapman & Hall, Ltd., London, 1949. MR0031213 (11:118j)

[112] Wikipedia entry on Rodney Dangerfeld: For those in the international community who don't know of this reference, see `http://en.wikipedia.org/wiki/Rodney_Dangerfield`.

Mathematics 253-37, California Institute of Technology, Pasadena, California 91125
E-mail address: `bsimon@caltech.edu`