The Sharp Form of the Strong Szegő Theorem

Barry Simon

In memoriam, Robert Brooks (1952–2002)

ABSTRACT. Let f be a function on the unit circle and $D_n(f)$ be the determinant of the $(n+1) \times (n+1)$ matrix with elements $\{c_{j-i}\}_{0 \leq i,j \leq n}$ where $c_m = \hat{f}_m \equiv \int e^{-im\theta} f(\theta) \frac{d\theta}{2\pi}$. The sharp form of the strong Szegő theorem says that for any real-valued L on the unit circle with L, e^L in $L^1(\frac{d\theta}{2\pi})$, we have

$$\lim_{n \to \infty} D_n(e^L) e^{-(n+1)\hat{L}_0} = \exp\left(\sum_{k=1}^{\infty} k|\hat{L}_k|^2 \right)$$

where the right side may be finite or infinite. We focus on two issues here: a new proof when $e^{i\theta} \to L(\theta)$ is analytic and known simple arguments that go from the analytic case to the general case. We add background material to make this article self-contained.

1. Introduction

Let $\{c_m\}_{m=-\infty}^{\infty}$ be a two-sided sequence of complex numbers. A *Toeplitz matrix* is a finite matrix constant along diagonals:

$$(1.1) \qquad T_{n+1} = \begin{pmatrix} c_0 & c_1 & c_2 & \cdots & c_n \\ c_{-1} & c_0 & c_1 & \cdots & c_{n-1} \\ \vdots & & & & \vdots \\ c_{-n} & c_{-n+1} & & \cdots & c_0 \end{pmatrix}$$

It turns out that the natural way to label T is in terms of the Fourier transform of c, that is,

$$(1.2) \qquad f(\theta) = \sum_{m=-\infty}^{\infty} c_m e^{im\theta}$$

on $\partial\mathbb{D}$ ($\mathbb{D} = \{z \mid |z| < 1\}$, $\partial\mathbb{D} = \{z \mid |z| = 1\}$). f is called the *symbol* of the Toeplitz matrix.

One can define a symbol as a distribution so long as $|c_m|$ is polynomially bounded in m, but we will discuss the case where there is a signed measure, $d\mu$, so that

$$(1.3) \qquad c_n = \int e^{-in\theta} \, d\mu(\theta) \equiv \hat{\mu}_n$$

As usual, for $f \in L^1(\partial\mathbb{D}, \frac{d\theta}{2\pi})$, we define $\hat{f}_n$ to be the Fourier coefficients of the measure $f \frac{d\theta}{2\pi}$. We will most often discuss the case where $d\mu$ is absolutely continuous, that is, $d\mu = w(\theta)\frac{d\theta}{2\pi}$ and where $w \geq 0$ or even that $w = e^L$.

$D_n(d\mu)$ is the determinant of T_{n+1}. The strong Szegő theorem says that if $L, e^L \in L^1$ with L real, then

$$(1.4) \qquad \log D_n\left(e^L \frac{d\theta}{2\pi}\right) \sim (n+1)\hat{L}_0 + \sum_{k=1}^{\infty} |k| \, |\hat{L}_k|^2$$

There are a number of remarkable aspects of (1.4). The first term was found in 1915 and the second in 1952. Despite the 37-year break, they were both found by Szegő — the twenty-year old in 1915 [30] and the 57-year old in 1952 [34]! You might wonder about whether (1.4) is the leading term in a systematic $1/n$ series. In fact, if L is real-valued and if $e^{i\theta} \mapsto L(\theta)$ is analytic in the neighborhood of $\partial\mathbb{D}$, then the error in (1.4) is $O(e^{-Bn})$ — there are no more terms in the series (this follows from (2.21) and (5.17) below). Lest you be shocked by this, we note that for many models in statistical mechanics, the free energy has a volume term, a surface term, and then, if the interaction is short-range, exponentially small errors.

A second remarkable aspect is the subtlety. Why should $\log w$ enter at all, and then in both linear and quadratic terms? There is a fascination with this subject among mathematicians who have extended the result both in the context of function algebras [38, 15, 22] and in the context of pseudodifferential operators on manifolds [39, 14] (see [28] for literally dozens of papers on each aspect).

A third aspect is that there are a remarkable number of applications of this result. Szegő returned to find the second term because of a question raised by Onsager who ran into Toeplitz determinants in his work on the Ising model (see [23, 5] for a discussion of this). They enter in the study of some Coulomb gases [20, 21, 10] and in electrical engineering applications [18, 3]. And they have had a surge of interest recently because of their role in random matrix theory [24].

When Szegő [34] proved (1.4), he assumed L was $C^{1+\varepsilon}$. There were many papers on this subject which improved this incrementally until Ibragimov [16], fifteen years later, proved the following sharp form:

THEOREM 1.1 ([16, 12]). *Let L be a real-valued function on $\partial\mathbb{D}$ so that $L, e^L \in L^1(\partial\mathbb{D}, \frac{d\theta}{2\pi})$. Then*

$$(1.5) \qquad \lim_{n\to\infty} D_n\left(e^L \frac{d\theta}{2\pi}\right) e^{-(n+1)\hat{L}_0} = \exp\left(\sum_{k=1}^{\infty} |k| \, |\hat{L}_k|^2 \right)$$

Thus, (1.4) holds whenever the right side makes sense, that is, L in $H^{1/2}$, the Sobolev space of order $\frac{1}{2}$. This should be supplemented with a result of Golinskii-Ibragimov [12]:

THEOREM 1.2. *If $d\mu = e^L \frac{d\theta}{2\pi} + d\mu_s$ with $d\mu_s$ singular is a positive measure on $\partial\mathbb{D}$ and $\lim_{n\to\infty} D_n(d\mu)e^{-(n+1)\hat{L}_0} < \infty$, then $d\mu_s = 0$.*

The combination of these two theorems has a spectral theory consequence that links it up to the theme of the conference and to Bob Brooks' interests. As we will see in Section 2, probability measures on $\partial\mathbb{D}$ have associated parameters $\{\alpha_n(d\mu)\}_{n=0}^\infty$ called Verblunsky coefficients. It can be shown using Theorems 1.1, 1.2, and 2.4 that

THEOREM 1.3. *Let $d\mu$ be a probability measure on $\partial\mathbb{D}$ and $\{\alpha_n(d\mu)\}_{n=0}^\infty$ its Verblunsky coefficients. Then the following are equivalent:*

(i) $\sum_{n=0}^\infty n|\alpha_n|^2 < \infty$

(ii) $d\mu_s = 0$ and $d\mu = e^L \frac{d\theta}{2\pi}$ where $\sum_{k=1}^\infty k|\hat{L}_k|^2 < \infty$.

This is one of those gems of spectral theory that give necessary and sufficient conditions relating properties of a measure and its inverse spectral parameters.

We have two main themes in this article. First, we wish to note that despite it taking fifteen years to go from $L \in C^{1+\varepsilon}$ to $L \in H^{1/2}$, there is an elegant and simple argument to do this jump. This combines arguments of Golinskii-Ibragimov [12] and Johansson [17] whose proofs have "easy" halves that handle opposite sides of the extension. It does not appear to be widely appreciated that their arguments can be combined in this way. In fact, these results reduce (1.4) to the case where the $\hat{L}_k$ decay exponentially, that is, $e^{i\theta} \mapsto L(\theta)$ is real analytic on $\partial\mathbb{D}$.

Second, we have a new proof of (1.4) in this real analytic case that is perhaps less mysterious than the elaborate calculation in Szegő [34]. From our point of view, the two terms in (1.4) come from two terms in the Christoffel-Darboux formula for $z = e^{i\theta}$.

While the arguments of [12, 17] are simple, they depend on considerable general machinery relating Toeplitz determinants to orthogonal polynomials on the unit circle (OPUC) on the one hand, and to the statistical mechanics of Coulomb gases on the other (both themes go back to Szegő's early work: [30] has the Coulomb gas representation and the main point of [31, 32] is to discuss the connection to OPUC), so this article, in attempting to be self-contained, provides this background.

Sections 2 and 3 discuss the basics of OPUC. In Section 4, we get the leading term in (1.4), not only for its own sake, but to define in Section 5 the Szegő function which will play a critical role in many aspects of the remainder. These preliminaries allow us to present the Golinskii-Ibragimov half of the extension in Section 6. After proving the Coloumb gas representation in Section 7, we can prove the Johansson half of the extension in Section 8. The final three sections provide the proof of (1.4) in the analytic case: Section 9 has a preliminary proving the Christoffel-Darboux formula, and the last two sections finish the proof.

I have written a comprehensive book on OPUC [28, 29] and everything in this paper appears there, but it seemed sensible, given the fact that the material is spread through a long book, to pull out exactly what is needed to prove Ibragimov's theorem.

While (1.4) is sharp in one sense, it is not the end of the story by any means. First, there is a simple argument of Johansson [17] that drops the requirement of reality from L: if $e^L, L \in L^1$ and $L \in H^{1/2}$, then an extension (1.4) holds in the sense that

$$e^{-(n+1)\hat{L}_0} D_n\left(e^L \frac{d\theta}{2\pi}\right) \to \exp\left(\sum_{k=1}^\infty |k|\hat{L}_k\hat{L}_{-k}\right)$$

There are also subtleties in extending (1.4) to allow matrix-valued symbols, to allow complex w's with nonzero winding number, and to determine the leading behavior when $L \notin L^1$. The reader can consult [**28**] for references on all these issues.

Over the course of studying asymptotics of Toeplitz determinants, I have learned a lot in discussions with Percy Deift, Rowan Killip, and Irina Nenciu, and I would like to thank them for their insights.

Bob Brooks was a substantial mathematician and wonderful person. We lost him too soon. I'm glad to dedicate this article to his memory.

2. Verblunsky Coefficients and Toeplitz Matrices

If c_n are the moments of a measure, μ, it is natural to form the monic orthogonal polynomials, $\Phi_n(z; d\mu)$, defined by

$$(2.1) \qquad \Phi_n(z) = z^n + \text{ lower order}$$

$$(2.2) \qquad \langle z^j, \Phi_n \rangle_\mu = 0 \qquad \text{if } j = 0, 1, \ldots, n-1$$

where

$$(2.3) \qquad \langle f, g \rangle = \int \overline{f(e^{i\theta})}\, g(e^{i\theta})\, d\mu(\theta)$$

is the $L^2(\partial\mathbb{D}, d\mu)$ inner product. In order to form Φ_n for all n, we need the measures $d\mu$ to be nontrivial, that is, not supported on a finite set of points.

The matrix elements $c_{k-\ell} = \int e^{i(\ell-k)\theta}\, d\mu = \langle z^k, z^\ell \rangle_\mu$, so $D_n(d\mu)$ is a Gram determinant. Such determinants allow change of basis, that is, if $P_k(z) = z^k + \text{ lower order}$, $\det(\langle P_k, P_\ell \rangle)_{0 \le k, \ell \le n} = D_n(d\mu)$. We can take $P_k = \Phi_k$, in which case, $\langle P_k, P_\ell \rangle$ is a diagonal matrix (!), and so,

THEOREM 2.1.

$$(2.4) \qquad D_n(d\mu) = \prod_{j=0}^{n} \|\Phi_j\|_{d\mu}^2$$

Φ_j is the orthogonal projection in L^2 of z^j onto the orthogonal complement of $[1, \ldots, z^{j-1}]$, the span of $\{1, \ldots, z^{j-1}\}$. Since multiplication by z is unitary, $z\Phi_j$ is the projection of z^{j+1} onto $[z, \ldots, z^j]^\perp$ while Φ_{j+1} is the projection of z^{j+1} onto $[1, \ldots, z^j]^\perp$, so

$$(2.5) \qquad \|\Phi_{j+1}\| \le \|z\Phi_j\| = \|\Phi_j\|$$

Thus, $\|\Phi_j\|$ is decreasing in j. It follows that

THEOREM 2.2.

$$(2.6) \qquad (a) \qquad \lim D_n(d\mu)^{1/n+1} = \lim \frac{D_{n+1}(d\mu)}{D_n(d\mu)} = \lim_{n\to\infty} \|\Phi_n\|^2$$

$$(2.7) \qquad (b) \qquad \frac{D_{n+1}}{D_n} \le \frac{D_n}{D_{n-1}}$$

These ideas all go back to Szegő [**31**, **32**]. The next idea, which is a set of recursion relations of the Φ_n, was first written down by Szegő [**33**], but the basic parameters occurred in a related context in Verblunsky [**36**, **37**]. To state them, we need to define the reversed polynomials

$$(2.8) \qquad \Phi_n^*(z) = z^n\, \overline{\Phi_n(1/\bar{z})}$$

THEOREM 2.3. *For any nontrivial measure, $d\mu$, there exists a sequence of numbers $\{\alpha_n\}_{n=0}^{\infty}$ so that*

$$(2.9) \qquad \Phi_{n+1}(z) = z\Phi_n(z) - \bar{\alpha}_n \Phi_n^*(z)$$

Moreover, $\alpha_n \in \mathbb{D}$ and

$$(2.10) \qquad \|\Phi_{n+1}\|^2 = (1 - |\alpha_n|^2)\|\Phi_n\|^2$$

and

$$(2.11) \qquad \|\Phi_n\|^2 = c_0(d\mu) \prod_{j=0}^{n-1} (1 - |\alpha_j|^2)$$

REMARKS.

(1) In the next section, we will see that $\mu \mapsto \{\alpha_n\}_{n=0}^{\infty}$ is one-one for μ's which are normalized.
(2) There is a converse going back to Verblunsky [36] (see [28] for many other proofs) that the map from probability measures to $\times_{n=0}^{\infty} \mathbb{D}$ by $\mu \mapsto \{\alpha_n\}_{n=0}^{\infty}$ is onto.
(3) The α_n are called the *Verblunsky coefficients* for $d\mu$.
(4) Applying $Q^*(z) = z^{n+1} \overline{Q(1/\bar{z})}$ to (2.9) yields

$$(2.12) \qquad \Phi_{n+1}^*(z) = \Phi_n^*(z) - \alpha_n z \Phi_n(z)$$

(5) The proof below is a variant of one of Atkinson [1]. Szegő's proof first proves the Christoffel-Darboux formula (see Section 9) and uses that to prove (2.9).

PROOF. Let $V_n g = z^n \bar{g}$ on $L^2(d\mu)$. V_n is anti-unitary, maps $\mathcal{P}_n$, the polynomials of degree n, to themselves, and maps Φ_n to Φ_n^*. Since Φ_n is the unique element (up to constants) of $\mathcal{P}_n$ orthogonal to $\{1, z, \ldots, z^{n-1}\}$ and V_n is anti-unitary, Φ_n^* is the unique element of $\mathcal{P}_n$ orthogonal to $\{V_n 1, \ldots, V_n z^{n-1}\} = \{z^n, z^{n-1}, \ldots, z\}$.

Now, for $j = 1, \ldots, n$,

$$\langle z^j, z\Phi_n \rangle = \langle z^{j-1}, \Phi_n \rangle = 0$$

and clearly, $\langle z^j, \Phi_{n+1} \rangle = 0$. Thus,

$$\langle z^j, \Phi_{n+1} - z\Phi_n \rangle = 0$$

for $j = 1, \ldots, n$. Since Φ_n and Φ_{n+1} are monic, $\Phi_{n+1} - z\Phi_n \in \mathcal{P}_n$. So, by the first part of the proof, (2.9) holds for a suitable constant α_n. Thus

$$(2.13) \qquad \alpha_n = -\overline{\Phi_{n+1}(0)}$$

Since $\Phi_n^* \perp \Phi_{n+1}$, we have

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

which implies (2.10). (2.10) in turn implies $|\alpha_n| < 1$. (2.11) follows by induction and $\|\Phi_0\|^2 = \|1\|^2 = c_0(d\mu)$. □

This leads to

THEOREM 2.4. *Suppose* $\int d\mu = 1$. *We have*

$$(2.14) \qquad F(d\mu) = \lim_{n\to\infty} \frac{D_{n+1}(d\mu)}{D_n(d\mu)} = \prod_{j=0}^{\infty} (1 - |\alpha_j|^2)$$

where the product always converges although the limit may be zero.

If $F(d\mu) > 0$, then

$$(2.15) \qquad G_n = \frac{D_n}{F^{n+1}}$$

obeys

$$(2.16) \qquad G_{n+1} \geq G_n$$

The limit always exists (but may be infinite) and is given by

$$(2.17) \qquad G(d\mu) = \lim_{n\to\infty} G_n(d\mu) = \prod_{j=0}^{\infty} (1 - |\alpha_j|^2)^{-j-1}$$

REMARKS.

(1) In particular,

$$(2.18) \qquad F > 0 \Leftrightarrow \sum_{n=0}^{\infty} |\alpha_n|^2 < \infty$$

$$(2.19) \qquad G < \infty \Leftrightarrow \sum_{n=0}^{\infty} (n+1)|\alpha_n|^2 < \infty$$

(2) (2.14) in a sense goes back to Verblunsky [37]. (2.17) seems to have only been noted by Baxter [4] many years later.

(3) If $c_0 \neq 1$, $F(d\mu) = c_0 \prod_{j=0}^{\infty}(1 - |\alpha_j|^2)$ while $G(d\mu)$ is still given by (2.17). Indeed, $G_n(d\mu) = G_n(d\mu / \int d\mu)$.

(4) (2.7) says $\log D_n$ is concave in n. The monotonicity of G_n is a standard fact about concave functions with finite asymptotics.

PROOF. By (2.4) and then (2.11),

$$(2.20) \qquad \frac{D_{n+1}}{D_n} = \|\Phi_{n+1}\| = \prod_{j=0}^{n} (1 - |\alpha_j|^2)$$

(if $\|\Phi_1\|^2 = c_0 = 1$). From this and $|\alpha_j| < 1$, (2.14) is immediate. If F is nonzero,

$$\frac{D_{n+1}}{D_n} \frac{1}{F} = \prod_{j=n+1}^{\infty} (1 - |\alpha_j|^2)^{-1}$$

so that

$$G_n = \frac{c_0}{F} \prod_{k=1}^{n} \left[\frac{D_k}{D_{k-1}} \frac{1}{F} \right]$$

$$(2.21) \qquad = \prod_{j=0}^{\infty} (1 - |\alpha_j|^2)^{-\min(n,j)-1}$$

from which $G_{n+1} \geq G_n$ and (2.17) are immediate. $\square$

3. Bernstein-Szegő Approximations

Given a nontrivial probability measure, $d\mu$, with Verblunsky coefficients $\{\alpha_n\}_{n=0}^{\infty}$, we will identify the measures, $d\mu^{(N)}$, with

$$(3.1) \qquad \alpha_j(d\mu^{(N)}) = \begin{cases} \alpha_j(d\mu) & j = 0, 1, \ldots, N-1 \\ 0 & j \geq N \end{cases}$$

and see $d\mu^{(N)} \to d\mu$ weakly. In many ways, the general proof of the strong Szegő theorem will play off this approximation and the distinct approximation obtained by truncating the Fourier series for L in $e^L \frac{d\theta}{2\pi}$. As a preliminary, we need

THEOREM 3.1. $\Phi_n(z)$ has all its zeros in $\mathbb{D}$. $\Phi_n^*(z)$ has all its zeros in $\mathbb{C}\backslash\bar{\mathbb{D}}$.

REMARK. This proof is due to Landau [19]. See [28] for many other proofs of this fact.

PROOF. Let $\Phi_n(z_0) = 0$. Then $\pi_{n-1} = \Phi_n(z)/(z-z_0)$ is a polynomial of degree $n-1$, and so $\langle \pi_{n-1}, \Phi_n \rangle = 0$. Since $(z-z_0)\pi_{n-1} = \Phi_n$,

$$\|\pi_{n-1}\|^2 = \|z\pi_{n-1}\|^2 = \|z_0\pi_{n-1} + \Phi_n\|^2$$
$$= |z_0|^2 \|\pi_{n-1}\|^2 + \|\Phi_n\|^2$$

or

$$(3.2) \qquad (1-|z_0|^2)\|\pi_{n-1}\|^2 = \|\Phi_n\|^2$$

from which we conclude $|z_0| < 1$, that is, the zeros of Φ_n lie in $\mathbb{D}$. Since $\Phi_n^*(z_0) = 0$ if and only if $\Phi_n(1/\bar{z}_0) = 0$, the zeros of Φ_n^* lie in $\mathbb{C}\backslash\bar{\mathbb{D}}$. □

Given μ, define a measure, $\tilde{\mu}^{(N)}$, by

$$(3.3) \qquad d\tilde{\mu}^{(N)} = \frac{d\theta}{2\pi|\Phi_N(e^{i\theta}; d\mu)|^2}$$

which can be defined since $\Phi_N(e^{i\theta}) \neq 0$ for all θ by Theorem 3.1. We have

LEMMA 3.2. For $j \geq 0$,

$$(3.4) \qquad \Phi_{N+j}(z; d\tilde{\mu}^{(N)}) = z^j \Phi_N(z; d\mu)$$

Moreover,

$$(3.5) \qquad \alpha_\ell(d\tilde{\mu}^{(N)}) = 0$$

for $\ell \geq N$.

PROOF. Since

$$(3.6) \qquad |\Phi_N(e^{i\theta})|^2 = e^{-iN\theta}\Phi_N(e^{i\theta})\Phi_N^*(e^{i\theta})$$

we have for $k \in \mathbb{Z}$, $k \leq N-1$ (includes $k < 0$),

$$\int_0^{2\pi} e^{-ik\theta} \frac{\Phi_N(e^{i\theta})}{|\Phi_N(e^{i\theta})|^2} \frac{d\theta}{2\pi} = \frac{1}{2\pi i} \oint_{|z|=1} z^{N-k} \frac{dz}{z\Phi_N^*(z)}$$
$$= 0$$

since $N-k-1 \geq 0$ and $\Phi_N^*(z)^{-1}$ is analytic in a neighborhood of $\bar{\mathbb{D}}$ by Theorem 3.1. This says

$$(3.7) \qquad \langle z^\ell, z^j \Phi_N \rangle_{\tilde{\mu}^{(N)}} = 0$$

for $\ell = 0, 1, \ldots, N + j - 1$. Since $z^j \Phi_N$ is monic, (3.4) holds.

By (2.13), if $\Phi_{k+1}(0) = 0$, then $\alpha_k = 0$, so (3.4) implies $\Phi_{N+j}(0; d\tilde{\mu}^{(N)}) = 0$, which in turn implies (3.5). $\qquad\square$

THEOREM 3.3 (Geronimus [11]). *Let $d\mu, d\nu$ be two nontrivial measures on $\partial\mathbb{D}$. Suppose that for some fixed N,*

$$(3.8) \qquad \Phi_N(z; d\mu) = \Phi_N(z; d\nu)$$

Then

$$(3.9) \qquad \Phi_j(z; d\mu) = \Phi_j(z; d\nu) \qquad\qquad j = 0, 1, \ldots, N-1$$

$$(3.10) \qquad \alpha_j(d\mu) = \alpha_j(d\nu) \qquad\qquad j = 0, 1, \ldots, N-1$$

$$(3.11) \qquad \frac{c_j(d\mu)}{c_0(d\mu)} = \frac{c_j(d\nu)}{c_0(d\nu)} \qquad\qquad j = 0, 1, \ldots, N$$

$$(3.12) \qquad \|\Phi_j\|^2_{d\mu} = c_0(d\mu)c_0(d\nu)^{-1}\|\Phi_j\|^2_{d\nu} \qquad j = 0, 1, \ldots, N$$

PROOF. (2.9) and (2.12) can be written in matrix form

$$(3.13) \qquad \begin{pmatrix} \Phi_{j+1}(z) \\ \Phi^*_{j+1}(z) \end{pmatrix} = \begin{pmatrix} z & -\bar{\alpha}_j \\ -z\alpha_j & 1 \end{pmatrix} \begin{pmatrix} \Phi_j(z) \\ \Phi^*_j(z) \end{pmatrix}$$

The 2×2 matrix in (3.13) has an inverse $z^{-1}\rho_j^{-2} \left(\begin{smallmatrix} 1 & \bar{\alpha}_j \\ \alpha_j z & z \end{smallmatrix} \right)$ where

$$(3.14) \qquad \rho_j = (1 - |\alpha_j|^2)^{1/2}$$

Thus (3.14) implies the inverse Szegő recursions:

$$(3.15) \qquad \Phi_j(z) = \rho_j^{-2} \frac{[\Phi_{j+1}(z) + \bar{\alpha}_j \Phi^*_{j+1}]}{z}$$

$$(3.16) \qquad \Phi^*_j(z) = \rho_j^{-2}[\Phi^*_{j+1}(z) + \alpha_j \Phi_{j+1}(z)]$$

(3.8) implies

$$\alpha_N(d\mu) = \overline{-\Phi_N(0; d\mu)} = \overline{-\Phi_N(0; d\nu)} = \alpha_N(d\nu)$$

and thus, by (3.15) with $j = N - 1$, we have (3.9) for $j = N - 1$. By iterating this argument, we conclude that (3.9) and (3.10) hold.

We need only prove (3.11) and (3.12), assuming $c_0(d\mu) = c_0(d\nu) = 1$ since with $d\tilde{\mu} = d\mu/c_0(d\mu)$, we have $c_n(d\mu) = c_0(d\mu)c_n(d\tilde{\mu})$ and $\|\Phi_j\|^2_{d\mu} = c_0(d\mu)\|\Phi_j\|^2_{d\tilde{\mu}}$. (3.12) is immediate from (2.11).

We prove (3.11) when $c_0(d\mu) = c_0(d\nu)$ by induction and noting $\langle 1, \Phi_k \rangle = 0$ yields a formula for c_k in terms of the coefficients of Φ_k and $c_0, c_1, \ldots, c_{k-1}$. $\qquad\square$

REMARK. The last paragraph of the proof shows that the α's determine the c's and proves that the map of μ to α is one-one.

To succinctly state this section's final result, we introduce

$$(3.17) \qquad \varphi_N(z; d\mu) = \frac{\Phi_N(z; d\mu)}{\|\Phi_N\|}$$

the orthonormal polynomials. By (2.11),

$$(3.18) \qquad \varphi_N(z; d\mu) = \prod_{j=0}^{N-1} \rho_j^{-1} c_0(d\mu)^{-1/2} \Phi_N(z; d\mu)$$

We define

$$\text{(3.19)} \qquad \kappa_n = \left(\prod_{j=0}^{n-1} \rho_j^{-1} \right) (c_0(d\mu))^{-1/2} = \|\Phi_n\|^{-1}$$

and

$$\text{(3.20)} \qquad \kappa_\infty = \lim_{n\to\infty} \|\Phi_n\|^{-1}$$

which exists (but it may be $+\infty$) by $\rho_j < 1$. The infinite product in (3.19), and so κ_∞, is finite if and only if

$$\text{(3.21)} \qquad \kappa_\infty < \infty \Leftrightarrow \sum_{j=0}^{\infty} |\alpha_j|^2 < \infty$$

We also note the translation of (2.9)/(2.12) from Φ to φ:

$$\text{(3.22)} \qquad \varphi_{n+1}(z) = \rho_n^{-1}(z\varphi_n(z) - \bar{\alpha}_n \varphi_n^*(z))$$

$$\text{(3.23)} \qquad \varphi_{n+1}^*(z) = \rho_n^{-1}(\varphi_n^*(z) - \alpha_n z\varphi_n^*(z))$$

THEOREM 3.4. *Let $d\mu$ be a nontrivial probability measure on $\partial\mathbb{D}$. Define*

$$\text{(3.24)} \qquad d\mu^{(N)} = \frac{d\theta}{2\pi|\varphi_N(e^{i\theta})|^2}$$

Then $d\mu^{(N)}$ is a probability measure on $\partial\mathbb{D}$ for which (3.1) holds. As $N \to \infty$, $d\mu^{(N)} \to d\mu$ weakly.

REMARK. We call measures of the form (3.24) *BS measures* and $d\mu^{(N)}$ the *BS approximation.*

PROOF. Since $d\mu^{(N)}$ is a multiple of $d\tilde{\mu}^{(N)}$, we have the bottom half of (3.1) by (3.5). Since (3.4) holds for $j = 0$, Theorem 3.3 and (3.10) imply the top half of (3.1).

Since $\Phi_N = \|\Phi_N\|_\mu \varphi_N$, we clearly have

$$\|\Phi_N\|_{\mu^{(N)}}^2 = \|\Phi_N\|_\mu^2$$

so, by (3.13), $c_0(d\mu^{(N)}) = c_0(d\mu) = 1$, that is, $d\mu_N$ is a probability measure.

By the above and (3.11), we have

$$\text{(3.25)} \qquad c_j(d\mu^{(N)}) = c_j(d\mu) \qquad j = 0, 1, \ldots, N$$

This and its complex conjugate implies that for any Laurent polynomial, f (polynomial in z and z^{-1}),

$$\text{(3.26)} \qquad \lim_{N\to\infty} \int f(e^{i\theta}) \, d\mu^{(N)} = \int f(e^{i\theta}) \, d\mu$$

since the left side is equal to the right for N large. Since Laurent polynomials are dense in $C(\partial\mathbb{D})$, (3.26) holds for all f, that is, we have weak convergence. $\square$

We note we have proven that

$$\text{(3.27)} \qquad \alpha_j(d\mu^{(N)}) = \begin{cases} \alpha_j & j \le N - 1 \\ 0 & j \ge N \end{cases}$$

The ideas of this section go back to Geronimus [11] and were rediscovered in [9] and [7]. In particular, the use of inverse recursion to prove Geronimus' theorem is taken from [7].

4. Szegő's Theorem

In this section, our main goal is to prove

THEOREM 4.1. *For any* $w \in L^1(\frac{d\theta}{2\pi})$,

$$(4.1) \qquad \lim_{N\to\infty} \frac{1}{N} \log D_N(w) = \int \log w(\theta) \frac{d\theta}{2\pi}$$

REMARKS.
(1) Since $\log(x) \leq x - 1$ and $w(\theta) \in L^1$, $\int \max(0, \log w(x)) \frac{d\theta}{2\pi} < \infty$, so the integral on the right side of (4.1) is either convergent or diverges to $-\infty$, in which case (4.1) says $D_n(w)^{1/n} \to 0$.
(2) This was conjectured by Pólya [26] and proven by the twenty-year old Szegő in 1915 [30]. Our proof here is essentially the one Szegő presented in [31, 32].
(3) The same result is true for the symbol $d\mu = w(\theta)\frac{d\theta}{2\pi} + d\mu_s$, that is, the limit is independent of $d\mu_s$. This extension was first proven by Verblunsky [37]. We will not prove this more general result here ([28] has four different proofs in Chapter 2) since it is peripheral to our main result.

The first half of the theorem is a simple use of Jensen's inequality:

PROPOSITION 4.2. *Let* $w = e^L$ *with* $w, L \in L^1$. *Then*

$$(4.2) \qquad \|\Phi_n\|^2_{w\frac{d\theta}{2\pi}} \geq \exp\left(\int L(\theta) \frac{d\theta}{2\pi}\right)$$

In particular,

$$(4.3) \qquad D_n(w) \geq \exp\left(\int (n+1)L(\theta) \frac{d\theta}{2\pi}\right)$$

PROOF. (4.3) follows from (4.2) and (2.4). To prove (4.2), we write

$$\|\Phi_n\|^2 = \|\Phi_n^*\|^2 = \int \exp(2\log|\Phi_n^*(e^{i\theta})| + L(e^{i\theta})) \frac{d\theta}{2\pi}$$

$$(4.4) \qquad \geq \exp\left(\int [2\log|\Phi_n^*(e^{i\theta})| + L(e^{i\theta})] \right)\frac{d\theta}{2\pi}$$

by Jensen's inequality.

By Theorem 3.1, $\log(\Phi_n^*(z))$ is analytic in $\mathbb{D}$, so

$$\int \log|\Phi_n^*(e^{i\theta})| \frac{d\theta}{2\pi} = \text{Re} \int \log(\Phi_n^*(e^{i\theta})) \frac{d\theta}{2\pi}$$

$$= \log|\Phi_n^*(0)| = 0$$

since Φ_n monic implies $\Phi_n^*(0) = 1$. $\qquad\square$

The other half of the theorem depends on a variational principle noted by Szegő:

PROPOSITION 4.3. *We have for any* $w \in L^1(\partial\mathbb{D}, \frac{d\theta}{2\pi})$,

$$(4.5) \qquad \lim_{n\to\infty} [D_n(w)]^{1/n} = \inf\left\{ \int |f(e^{i\theta})|^2 w(\theta) \frac{d\theta}{2\pi} \,\middle|\, f \in H^\infty(\mathbb{D}); f(0) = 1 \right\}$$

REMARK. $H^\infty(\mathbb{D})$ is the Hardy space of bounded analytic functions on $\mathbb{D}$. By general principles [8, 27], for $\frac{d\theta}{2\pi}$ a.e. $e^{i\theta} \in \partial\mathbb{D}$, $\lim_{r\uparrow 1} f(re^{i\theta})$ exists, and that is what we mean by $f(e^{i\theta})$ in (4.5).

PROOF. Since $\|\Phi_n^*\| = \|\Phi_n\|$, by (2.6),

$$\text{LHS of (4.5)} = \lim_{n \to \infty} \|\Phi_n^*\|^2$$

(4.6)
$$= \inf_n \|\Phi_n^*\|^2$$

by (2.5). By the argument at the start of the proof of Theorem 2.3,

$$\Phi_n^* = \pi_n 1$$

where π_n is the projection in the space of polynomials, $\mathcal{P}_n$, of degree n onto the orthogonal component of the span of $z, z^2, z^3, \ldots, z^n$. This span is $\{P \in \mathcal{P}_n \mid P(0) = 0\}$, so by standard geometry,

(4.7)
$$\|\Phi_n^*\|^2 = \inf\{\|P\|^2 \mid P \in \mathcal{P}_n; P(0) = 1\}$$

proving again that $\|\Phi_n^*\|^2$ is decreasing in n and proving (4.5) if H^∞ is replaced by the set of all polynomials.

To complete the proof, we need only show that for any $f \in H^\infty$ with $f(0) = 1$, there are polynomials $P_\ell(z)$ so that $P_\ell(0) = 1$ and

(4.8)
$$\int |P_\ell(e^{i\theta})|^2 w(\theta) \, \frac{d\theta}{2\pi} \to \int |f(e^{i\theta})|^2 w(\theta) \, \frac{d\theta}{2\pi}$$

If f is analytic in a neighborhood of $\mathbb{D}$, the Taylor approximations converge uniformly on $\bar{\mathbb{D}}$, so (4.8) holds. For general f, by the dominated convergence theorem,

$$\lim_{r \uparrow 1} \int |f(re^{i\theta})|^2 w(\theta) \, \frac{d\theta}{2\pi} = \int |f(e^{i\theta})|^2 w(\theta) \, \frac{d\theta}{2\pi}$$

so, by a two-step approximation, we find P_ℓ's so (4.8) holds. $\qquad \square$

PROOF OF THEOREM 4.1. We will prove that for any $\varepsilon > 0$, there is an f in H^∞ with $f(0) = 1$ and

(4.9)
$$\int |f(e^{i\theta})|^2 w(\theta) \, \frac{d\theta}{2\pi} \leq \exp\left(\int \log(w(\theta) + \varepsilon) \, \frac{d\theta}{2\pi}\right)$$

so taking $\varepsilon \downarrow 0$ yields the opposite inequality to (4.3).

Define

(4.10)
$$g(z) = \exp\left(-\int \log(w(\theta) + \varepsilon) \left(\frac{e^{i\theta} + z}{e^{i\theta} - z}\right) \frac{d\theta}{4\pi}\right)$$

and $f(z) = g(z)/g(0)$, so $f(0) = 1$. Moreover, $|g(z)| \leq \varepsilon^{-1/2}$ by the fact that the Poisson kernel

$$P_r(\theta, \varphi) = \text{Re}\left(\frac{e^{i\theta} + re^{i\varphi}}{e^{i\theta} - re^{i\varphi}}\right)$$

is nonnegative and $\int \frac{d\theta}{2\pi} P_r(\theta, \varphi) = 1$. By standard maximal function arguments (see [27]), $|g(e^{i\theta})| = |w(\theta) + \varepsilon|^{-1/2}$, so

$$\int |f(e^{i\theta})|^2 w(\theta) \, \frac{d\theta}{2\pi} \leq g(0)^{-2} = \text{RHS of (4.9)}$$

$\qquad \square$

5. The Szegő Function

When $w(\theta) = e^{L(\theta)}$ with $L \in L^1$, Szegő [31, 32] introduced a natural function, $D(z)$ on $\mathbb{D}$, which will play a critical role in several places below:

$$(5.1) \qquad D(z) = \exp\left(\int \left(\frac{e^{i\theta} + z}{e^{i\theta} - z} \right) L(\theta) \, \frac{d\theta}{4\pi} \right)$$

Do not confuse D_n and $D(z)$. Both symbols are standard, but the objects are very different.

THEOREM 5.1.

(a) *If* (2.18) *holds, then for* $|z| < 1$,

$$(5.2) \qquad D(z) = D(0) \exp\left(\sum_{k=1}^{\infty} \hat{L}_k z^k \right)$$

$$(5.3) \qquad D(0) = \left[c_0\left(w \, \frac{d\theta}{2\pi} \right) \right]^{1/2} \prod_{n=0}^{\infty} (1 - |\alpha_n|^2)^{1/2}$$

(b) $D(z)$ *lies in* $H^2(\mathbb{D})$.
(c) $\lim_{r \uparrow 1} D(re^{i\theta}) \equiv D(e^{i\theta})$ *exist for a.e.* θ *and*

$$(5.4) \qquad\qquad |D(e^{i\theta})|^2 = w(\theta)$$

(d) D *is nonvanishing on* $\mathbb{D}$.

PROOF.

(a) We get (5.2) and

$$(5.5) \qquad\qquad D(0) = \exp(\tfrac{1}{2} L_0)$$

from (5.1) and

$$\frac{e^{i\theta} + z}{e^{i\theta} - z} = 1 + 2 \sum_{j=0}^{\infty} (e^{-i\theta} z)^n$$

uniformly in $e^{i\theta} \in \partial\mathbb{D}$ and $z \in \{|z| < r\}$. (5.3) then follows from (2.14), (4.1), and (5.5).

(b) Let $D^{(M)}(z)$ be given by (5.1) with $L(\theta)$ replaced by $\min(L(\theta), M)$. Then $D^{(M)} \in H^\infty$ and $|D^{(M)}(e^{i\theta})|^2 = \min(w(\theta), e^{2M})$,

$$\sup_{0 < r < 1} \int_0^{2\pi} |D^{(M)}(re^{i\theta})|^2 \, \frac{d\theta}{2\pi} = \int_0^{2\pi} \min(w(\theta), e^{2M}) \, \frac{d\theta}{2\pi}$$

$$\leq \int_0^{2\pi} w(\theta) \, \frac{d\theta}{2\pi}$$

Thus, taking $M \to \infty$, we see $D \in H^2$.

(c) is immediate from properties of boundary values of the Poisson integral in (5.1).

(d) is trivial from (5.1).

□

The following simple but powerful L^2 calculation goes back to Szegő [31, 32] (it has a version when $d\mu_s \neq 0$; see [28]):

THEOREM 5.2. *Let $d\mu = w\frac{d\theta}{2\pi}$ where $w = e^L$, $L \in L^1$. Then, as $n \to \infty$,*

(i)

$$\int |D\varphi_n^*(e^{i\theta}) - 1|^2 \frac{d\theta}{2\pi} \to 0 \tag{5.6}$$

(ii)

$$\int |\varphi_n^*(e^{i\theta}) - D^{-1}(e^{i\theta})|^2 \, d\mu \to 0 \tag{5.7}$$

(iii)

$$\varphi_n^*(z) \to D(z)^{-1} \tag{5.8}$$

uniformly on compact subsets of $\mathbb{D}$.

PROOF.

(i) By (5.4), $\int |D\varphi_n^*|^2 \frac{d\theta}{2\pi} = \int |\varphi_n^*|^2 \, d\mu = 1$, so (5.6) is equivalent to

$$\mathrm{Re} \int \varphi_n^*(e^{i\theta}) D(e^{i\theta}) \frac{d\theta}{2\pi} \to 1 \tag{5.9}$$

$D\varphi_n^*$ is in H^2, so the Cauchy integral formula applies, that is,

$$\text{LHS of } (5.9) = \varphi_n^*(0)D(0)$$

$$= \kappa_n \kappa_\infty^{-1} \tag{5.10}$$

since $\Phi_n^*(0) = 1$, $\varphi_n = \kappa_n \Phi_n$ and (5.3) and (3.19) imply $D(0) = \kappa_\infty^{-1}$. (5.10) implies (5.9).

(ii) is immediate from (5.6) and $d\mu = w\frac{d\theta}{2\pi} = |D|^2 \frac{d\theta}{2\pi}$ by (5.4).

(iii) (5.6) says $D\varphi_n^* \to 1$ in $\mathbb{H}^2$ and so, a fortiori, uniformly on compact subsets of $\mathbb{D}$.

$\square$

We need to extend this result to a neighborhood of $\bar{\mathbb{D}}$ when L is real analytic. As a preliminary, we note:

LEMMA 5.3. *Let $d\mu = w\frac{d\theta}{2\pi}$ be a probability measure with $\log w \in L^1(\frac{d\theta}{2\pi})$. Then*

$$\alpha_n = -\kappa_\infty \int \overline{\Phi_{n+1}(e^{i\theta})} \, D(e^{i\theta})^{-1} \, d\mu(\theta) \tag{5.11}$$

PROOF. We will prove for $m \geq n + 1$ that

$$\alpha_n = -\kappa_m \int \overline{\Phi_{n+1}(e^{i\theta})} \, \varphi_m^*(e^{i\theta}) \, d\mu(\theta) \tag{5.12}$$

from which (5.11) follows from (5.7).

φ_m^* is orthogonal to $\{z^\ell\}_{\ell=1}^m$, so if P is any polynomial of degree at most m with $P(0) = 0$,

$$\int \overline{P(e^{i\theta})} \, \varphi_m^*(e^{i\theta}) \, d\mu(\theta) = 0 \tag{5.13}$$

(5.12) follows from (5.13) by taking

$$P(z) = \bar{\alpha}_n \varphi_m^*(z) + \kappa_m \Phi_{n+1}(z)$$

which has $P(0) = \bar{\alpha}_n \kappa_m + \kappa_m(-\bar{\alpha}_n) = 0$.

$\square$

The following is due to Nevai-Totik [25]; it is needed in Section 11:

THEOREM 5.4. *Suppose that $d\mu = e^L \frac{d\theta}{2\pi}$ and $e^{i\theta} \mapsto L(\theta)$ is analytic in a neighborhood of $\partial\mathbb{D}$. Then $\varphi_n^*(z) \to D(z)^{-1}$ uniformly in a neighborhood of $\bar{\mathbb{D}}$. Moreover, the Verblunsky coefficients obey $|\alpha_n| \leq C_2 e^{-An/2}$ for some $A > 0$.*

PROOF. Analyticity says $|\hat{L}_k| \leq Ce^{-A|k|}$ for some $A > 0$. So, by (5.2), $D(z)$ is analytic and nonvanishing in a disk of radius e^A. In particular, if

$$(5.14) \qquad D(z)^{-1} = \sum_{j=0}^{\infty} d_{j,-1} z^j$$

then

$$(5.15) \qquad |d_{j,-1}| \leq C_1 e^{-A|j|/2}$$

Plug (5.14) into (5.11) and note that

$$\int \overline{\Phi_{n+1}(e^{i\theta})} \, e^{ik\theta} \, d\mu(\theta) = 0$$

for $k = 0, 1, \ldots, n$. Thus

$$|\alpha_n| \leq \kappa_\infty \sum_{k=n+1}^{\infty} |d_{k,-1}| \left| \int \overline{\Phi_{n+1}(e^{i\theta})} \, e^{ik\theta} \, d\mu \right|$$

$$\leq \kappa_\infty \|\Phi_{n+1}\| \sum_{k=n+1}^{\infty} |d_{k,-1}|$$

$$(5.16) \qquad \leq \kappa_\infty \sum_{k=n+1}^{\infty} |d_{k,-1}|$$

since $\|\Phi_{n+1}\| \leq 1$. So, by (5.15),

$$(5.17) \qquad |\alpha_n| \leq C_2 e^{-A|n|/2}$$

By (2.12) and $|\Phi_n^*(e^{i\theta})| = |\Phi_n(e^{i\theta})|$, we have

$$\sup_\theta |\Phi_{n+1}^*(e^{i\theta})| \leq (1 + |\alpha_n|) \sup_\theta |\Phi_n^*(e^{i\theta})|$$

$$\leq \prod_{j=0}^{n} (1 + |\alpha_j|)$$

$$(5.18) \qquad \leq \exp\left(\sum_{j=0}^{\infty} |\alpha_j| \right) = C_4 < \infty$$

by iteration. $C_4 < \infty$ follows from (5.17). Since Φ_n^* is analytic, we get

$$(5.19) \qquad \sup_{z \in \bar{\mathbb{D}}} |\Phi_n^*(z)| \leq C_4$$

and thus, since $\Phi_n^*(z) = z^n \overline{\Phi_n(1/\bar{z})}$, we get

$$(5.20) \qquad z \in \mathbb{C} \backslash \mathbb{D} \Rightarrow |\Phi_n(z)| \leq C_4 |z|^n$$

Returning to (2.12),

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

$$\leq C_4 C_2 \sum_{n=0}^{\infty} |ze^{-A/2}|^n$$

showing Φ_n^*, and so φ_n^*, converges uniformly in $\{z \mid |z| \leq e^{A/4}\}$. Since the limit is D^{-1} in $\mathbb{D}$, it is D^{-1} in this larger disk. $\qquad\square$

Finally, in terms of D, we want to rewrite the second term in the Szegő asymptotic formula (1.4):

THEOREM 5.5. *Let $d\mu = e^{L(\theta)} \frac{d\theta}{2\pi}$ with $L \in L^1$. Let $\hat{L}_k$ be given by (1.3). Then*

$$(5.21) \qquad \sum_{k=1}^{\infty} k|\hat{L}_k|^2 = \frac{1}{\pi} \int_{|z| \leq 1} |D(z)|^{-2} \left| \frac{\partial D}{\partial z} \right|^2 d^2 z$$

where both sides can be infinite.

PROOF. (5.21) follows by taking $r \uparrow 1$ in

$$(5.22) \qquad \sum_{k=1}^{\infty} k|\hat{L}_k|^2 r^{2k} = \frac{1}{\pi} \int_{|z| \leq r} |D(z)|^{-2} \left| \frac{\partial D}{\partial z} \right|^2 d^2 z$$

(by using monotone convergence). To prove (5.22), note that by (5.2),

$$\log\left[\frac{D(z)}{D(0)} \right] = \sum_{k=1}^{\infty} \hat{L}_k z^k$$

converges uniformly in $|z| < R$ and that $|D|^{-2}|\frac{\partial D}{\partial z}|^2 = |\frac{\partial}{\partial z} \log D(z)|^2$. Thus (5.22) follows from

$$\frac{1}{\pi} \int_{|z| \leq r} \bar{z}^{k-1} z^{\ell-1} \, d^2 z = k^{-1} \delta_{k\ell} r^{2k}$$

and

$$\frac{\partial}{\partial z} \log D(z) = \sum_{k=1}^{\infty} k \hat{L}_k z^{k-1}$$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

6. Extending the Strong Szegő Theorem, Part I

With the Szegő function and Fatou's lemma, we have the tools for the Golinskii-Ibragimov [12] half of the extension theorem:

THEOREM 6.1 ([12]). *Suppose for any BS measure, $d\mu$, we know that*

$$(6.1) \qquad G(d\mu) = \exp\left(\sum_{k=1}^{\infty} k|\hat{L}_k|^2 \right)$$

Then for any measure $d\mu = e^L \frac{d\theta}{2\pi}$ with $L \in L^1$, we have

$$(6.2) \qquad G(d\mu) \geq \exp\left(\sum_{k=1}^{\infty} k|\hat{L}_k|^2 \right)$$

REMARK. In particular, $G < \infty \Rightarrow \sum k|\hat{L}_k|^2 < \infty$.

PROOF. Let $d\mu^{(N)}$ be the BS approximations to $d\mu$. Let $D^{(N)}(z)$ be the D function for $d\mu^{(N)}$ and let

$$(6.3) \qquad M^{(N)} = \frac{1}{\pi} \int_{|z| \leq 1} |D^{(N)}(z)|^{-2} \left| \frac{\partial D^{(N)}}{\partial z} \right|^2 d^2 z$$

and similarly for M and D.

By (2.17) and (3.27),

$$G(d\mu^{(N)}) = \prod_{j=0}^{N-1} (1 - |\alpha_j|^2)^{-j-1}$$

so $G(d\mu^{(N)})$ is monotone increasing to $G(d\mu)$, that is,

$$(6.4) \qquad G(d\mu) = \lim_{N \to \infty} G(d\mu^{(N)})$$

On the other hand, since $\varphi_{N+j}(z, d\mu^{(N)}) = z^j \varphi_N(z, d\mu^{(N)})$, $\varphi_{N+j}^*(z, d\mu^{(N)}) = \varphi_N^*(z, d\mu^{(N)}) = \varphi_N^*(z, d\mu)$, so by (5.8),

$$(6.5) \qquad D^{(N)}(z) = \varphi_N^*(z, d\mu)^{-1}$$

and thus, by (5.8) again,

$$(6.6) \qquad D^{(N)}(z) \to D(z)$$

uniformly on compacts. By analyticity, the same thing is true for derivatives, so

$$|D(z)|^{-2} \left| \frac{\partial D}{\partial z} \right|^2 = \lim_{N \to \infty} |D^{(N)}(z)|^{-2} \left| \frac{\partial D^{(N)}}{\partial z} \right|^2$$

Thus, by Fatou's lemma,

$$(6.7) \qquad M \leq \liminf M^{(N)}$$

By hypothesis,

$$(6.8) \qquad G(d\mu^{(N)}) = \exp(M^{(N)})$$

(6.8), (6.7), and (6.4) imply that

$$G(d\mu) \geq \exp(M)$$

which is (6.2).   □

7. The Coulomb Gas Representation

In this section, we will provide an integral formula for D_n that will be critical in the next section. This formula appeared in Szegő's first paper [30] on asymptotics of Toeplitz determinants. He used it there for a minor technical purpose — and for us, too, it plays a relatively minor role. That said, it plays a central role in two proofs of the strong Szegő theorem and several applications.

While we will not pursue the Coulomb gas picture, if one uses $|z_j - z_k| = \exp(\log|z_j - z_k|)$, the formula we give for D_n says it is the partition function of a two-dimensional gas, and this point of view is the basis of Johansson's proof [17]. It also explains some interest in Toeplitz matrices in the physics literature [20, 21, 10].

Using Weyl's relation that Haar measure restricted to the classes of $\mathbb{U}(n+1)$, the group of $n \times n$ unitary matrices, is essentially $(2\pi)^{-n-1}[(n+1)!]^{-1} \prod_{k<j} |e^{i\theta_k} - e^{i\theta_j}|^2 \, d\theta_0 \ldots d\theta_n$ with $\{e^{i\theta_j}\}_{j=0}^n$ the eigenvalues of $U \in \mathbb{U}(n+1)$, one can use the formula below to rewrite D_n as an integral over $\mathbb{U}(n+1)$. This is both the starting point of the Bump-Diaconis [6] proof and of the many applications of Toeplitz matrices in the theory of random matrices [24].

It is, of course, well-known that $\prod_{k<j}(z_k - z_j)$ is a Vandermonde determinant. Expanding two such products, we get

$$(7.1) \qquad \left| \prod_{0 \le k < j \le n} (z_k - z_j) \right|^2 = \sum_{\pi,\sigma \in \Sigma_{n+1}} (-1)^\pi (-1)^\sigma \prod_{j=0}^n \bar{z}_j^{\pi(j)} z_j^{\sigma(j)}$$

where Σ_{n+1} is the permutations of $\{0, \ldots, n\}$ to itself.

Setting $z_j = e^{i\theta_j}$ and integrating $d\mu(\theta_0) \ldots d\mu(\theta_n)$, we get that

$$(7.2) \qquad \int |\pi(z_k - z_j)|^2 \, d\mu(\theta_0) \ldots d\mu(\theta_n) = \sum_{\pi,\sigma \in \Sigma_{n+1}} (-1)^\pi (-1)^\sigma \prod_{j=0}^n T_{\pi(j)\sigma(j)}$$

where $T_{k\ell}$ are the matrix elements of the $(n+1) \times (n+1)$ Toeplitz matrix. Now

$$\prod_{j=0}^n T_{\pi(j)\sigma(j)} = \prod_{j=0}^n T_{j(\sigma\pi^{-1})(j)}$$

and $(-1)^\pi (-1)^\sigma = (-1)^{\sigma\pi^{-1}}$. Thus, summing over σ for π fixed and then over π, we see the right side of (7.2) is $(n+1)! D_n(d\mu)$. Specializing to $d\mu = e^L \frac{d\theta}{2\pi}$, we have proven

THEOREM 7.1 (Coulomb Gas Representation for D_n). *Let e^L, $L \in L^1(\partial\mathbb{D}, \frac{d\theta}{2\pi})$. Then with $z_k = e^{i\theta_k}$,*

$$D_n\left(e^L \frac{d\theta}{2\pi} \right) = [(n+1)!]^{-1} \int_{\partial\mathbb{D}^{n+1}} \left| \prod_{0 \le k < j \le n} (z_k - z_j) \right|^2 e^{\sum_{j=0}^n L(\theta_j)} \prod_{j=0}^n \frac{d\theta_j}{2\pi}$$

8. Extending the Strong Szegö Theorem, Part II

The Coulomb representation and Fatou's lemma are the tools for Johansson's half of the extension theorem. A measure of the form $e^L \frac{d\theta}{2\pi}$ where L is a real Laurent polynomial (i.e., $\sum_{k=-n}^n \hat{L}_k e^{ik\theta}$ with $\hat{L}_{-k} = \bar{\hat{L}}_k$) is called a *GI measure* (after its early use in [12]). If $d\mu = e^L \frac{d\theta}{2\pi}$ with $L \in L^1$, we define the *GI approximations*, $d\mu_{(N)}$, by

$$(8.1) \qquad d\mu_{(N)} = \exp\left(\sum_{|k| \le N} \hat{L}_k e^{ik\theta} \right) \frac{d\theta}{2\pi} \equiv \exp(L_{(N)}(\theta)) \frac{d\theta}{2\pi}$$

THEOREM 8.1 ([17]). *Suppose for any GI measure, $d\mu$, we know that*

$$(8.2) \qquad G(d\mu) = \exp\left(\sum_{k=1}^\infty k |\hat{L}_k|^2 \right)$$

Then for any measure $d\mu = e^L \frac{d\theta}{2\pi}$ with $L \in L^1$, we have

$$(8.3) \qquad G(d\mu) \le \exp\left(\sum_{k=1}^\infty k |\hat{L}_k|^2 \right)$$

REMARK. In particular, if $\sum_{k=1}^{\infty} k|\hat{L}_k|^2 < \infty$, then $G(d\mu) < \infty$.

PROOF. Without loss (since we can multiply $d\mu$ by a constant), suppose $\hat{L}_0 = 0$. Given $d\mu = e^L \frac{d\theta}{2\pi}$, let $d\mu_{(N)}$ be its GI approximations and let

$$(8.4) \qquad C_{(N),m} = [(n+1)!]^{-1} \int_{\mathbb{D}^{m+1}} \prod_{0 \le j < k \le m} |z_j - z_k|^2 e^{\sum_{j=0}^m L_{(N)}(\theta_j)} \prod_{j=0}^m \frac{d\theta_j}{2\pi}$$

so, of course, $C_{(N),m} = D_m(d\mu_{(N)})$ by Theorem 7.1. Let C_m be the integral with $L_{(N)}$ replaced by L.

There is nothing to prove in (8.3) if $\sum_{k=1}^{\infty} k|\hat{L}_k|^2 = \infty$, so suppose the sum is finite. Thus, $L_{(N)} \to L$ in L^2, and so, we can find a subsequence $L_{(N_j)} \to L$ pointwise. By Fatou's lemma, for each m,

$$(8.5) \qquad C_m \le \liminf_{j \to \infty} C_{(N_j),m}$$

By Theorem 7.1, (2.16), $\hat{L}_0 = 0$, and the assumption for GI measures,

$$\begin{aligned} C_{(N_j),m} &= G_m(d\mu_{(N_j)}) \\ &\le G(d\mu_{(N_j)}) \\ &= \exp\left(\sum_{k=1}^{N_j} k|\hat{L}_k|^2 \right) \\ &\le \exp\left(\sum_{k=1}^{\infty} k|\hat{L}_k|^2 \right) \end{aligned}$$

By (8.5) and $\hat{L}_0 = 0$,

$$G_m(d\mu) = C_m \le \exp\left(\sum_{k=1}^{\infty} k|\hat{L}_k|^2 \right)$$

Taking $m \to \infty$ yields (8.3). $\square$

Theorems 6.1 and 8.1 reduce the proof of the sharp form of the strong Szegő theorem to proving the result for BS and GI measures. In both cases, $d\mu = e^L \frac{d\theta}{2\pi}$ with $e^{i\theta} \mapsto L(\theta)$ analytic in a neighborhood of $\partial\mathbb{D}$. In the last three sections, we will prove the strong Szegő theorem in that case.

9. The CD Formula

If u_1, u_2 solve $-u_j'' + V u_j = \lambda_j u_j$ with $u_j(0) = 0$, then $(\bar{\lambda}_1 - \lambda_2)\bar{u}_1 u_2 = (\bar{u}_1 u_2' - u_2 \bar{u}_1')'$, so

$$(9.1) \qquad (\bar{\lambda}_1 - \lambda_2) \int_0^a \overline{u_1(x)}\, u_2(x)\, dx = \bar{u}_1(a)u_2'(a) - \bar{u}_1'(a)u_2(a)$$

The Christoffel-Darboux formula is just the analog of this Wronskian relation that many undergraduates learn!

THEOREM 9.1 (CD Formula). *Let*

$$(9.2) \qquad K_n(z, \zeta) = \sum_{j=0}^n \overline{\varphi_j(\zeta)}\, \varphi_j(z)$$

Then

$$(9.3) \qquad K_n(z,\zeta) = \frac{\overline{\varphi_{n+1}^*(\zeta)}\,\varphi_{n+1}^*(z) - \overline{\varphi_{n+1}(\zeta)}\,\varphi_{n+1}(z)}{1 - \bar{\zeta}z}$$

$$(9.4) \qquad = \frac{\overline{\varphi_n^*(\zeta)}\,\varphi_n^*(\zeta) - z\bar{\zeta}\,\overline{\varphi_n(\zeta)}\,\varphi_n(z)}{1 - \bar{\zeta}z}$$

PROOF. By using (3.22)/(3.23) and their conjugates to write φ_{n+1} and φ_{n+1}^* in terms of φ_n and φ_n^*, we get

$$(9.5) \qquad \overline{\varphi_{n+1}^*(\zeta)}\,\varphi_{n+1}^*(z) - \overline{\varphi_{n+1}(\zeta)}\,\varphi_{n+1}(z) = \overline{\varphi_n^*(\zeta)}\,\varphi_n^*(z) - z\bar{\zeta}\,\overline{\varphi_n(\zeta)}\,\varphi_n(z)$$

where we used $\rho_n^{-2}(1 - |\alpha_n|^2) = 1$, and that the $\varphi^*\varphi$ cross terms cancel.

First of all, (9.5) says that (9.3) is equivalent to (9.4). In addition, if we let $Q_n(z,\zeta)$ be what we know is the common value of the right sides of (9.3) and (9.4), then subtracting (9.3) for $n-1$ from (9.4) for n, we see that

$$Q_n(z,\zeta) - Q_{n-1}(z,\zeta) = (1 - z\bar{\zeta})\left[\frac{\overline{\varphi_n(\zeta)}\,\varphi_n(z)}{1 - z\bar{\zeta}}\right]$$
$$= K_n(z,\zeta) - K_{n-1}(z,\zeta)$$

so we need only prove $Q_0(z,\zeta) = K_0(z,\zeta)$. Since (9.4) for $n = 0$ is 1 and that is $K_0(z,\zeta)$, (9.3) is proven. $\qquad\square$

COROLLARY 9.2. *For $e^{i\theta} \in \partial\mathbb{D}$,*

$$(9.6) \quad K_n(e^{i\theta}, e^{i\theta}) \equiv \sum_{j=0}^n |\varphi_j(e^{i\theta})|^2 = -\frac{\partial}{\partial r}\,|\varphi_{n+1}^*(re^{i\theta})|^2\bigg|_{r=1} + (n+1)|\varphi_{n+1}^*(e^{i\theta})|^2$$

PROOF. In (9.3), take $z = \zeta = re^{i\theta}$ and take $r \uparrow 1$. If we note $|\varphi_{n+1}(re^{i\theta})|^2 = r^{2n+2}|\varphi_{n+1}^*(r^{-1}e^{i\theta})|^2$, we get two terms: the one from the $-(r^{2n+2} - 1)/(1 - r^2)$ gives the $(n+1)|\varphi_{n+1}^*(e^{i\theta})|^2$ term in (9.6), and the other terms give the derivative in (9.6). $\qquad\square$

The CD formula for OPUC is due to Szegő [33]. Our proof is taken from Golinskii [13].

REMARK. Integrating both sides of (9.6) yields

$$(9.7) \qquad \int \frac{\partial}{\partial r}\,|\varphi_{n+1}^*(re^{i\theta})|^2\bigg|_{r=1} d\mu(\theta) = 1$$

10. The Feynman-Hellman Theorem for Toeplitz Determinants

Here we want to begin with a simple but useful formula for $\frac{\partial}{\partial\lambda}\|\Phi_n\|^2_{d\mu_\lambda}$, where the measure $d\mu_\lambda$ depends smoothly on λ. This appeared in [29], strongly motivated by closely related ideas of Baik et al. [2].

THEOREM 10.1. *Let $d\mu_\lambda$ be a family of positive measures on $\partial\mathbb{D}$ which are C^1 in λ in the sense that $\lambda \mapsto c_n(d\mu_\lambda)$ is C^1 for each n. Then $\|\Phi_n(\,\cdot\,, d\mu_\lambda)\|_{\mu_\lambda}$ is C^1 and*

$$(10.1) \qquad \frac{d}{d\lambda}\log\|\Phi_n\|^2_{\mu_\lambda} = \int |\varphi_n(e^{i\theta}, d\mu_\lambda)|^2\,\frac{d\mu_\lambda(\theta)}{d\lambda}$$

PROOF. Since $\frac{d}{d\lambda} \log \|\Phi_n\|^2 = (\frac{d}{d\lambda} \|\Phi_n\|^2)/\|\Phi_n\|^2$, it suffices to prove that

$$(10.2) \qquad \frac{d}{d\lambda} \|\Phi_n\|_{\mu_\lambda}^2 = \int |\Phi_n(e^{i\theta}, d\mu_\lambda)|^2 \frac{d\mu_\lambda(\theta)}{d\lambda}$$

Since

$$(10.3) \qquad \|\Phi_\lambda\|^2 = \int \overline{\Phi_n(e^{i\theta}, d\mu_\lambda)} \, \Phi_n(e^{i\theta}, d\mu_\lambda) \, d\mu_\lambda$$

its derivative is a sum of three terms, of which one term is the right side of (10.2) and the other two are

$$(10.4) \qquad \int \overline{\frac{\partial}{\partial\lambda} \Phi_n(e^{i\theta}, d\mu_\lambda)} \, \Phi_n(e^{i\theta}, d\mu_\lambda) \, d\mu_\lambda$$

and its conjugate.

But, for all λ, Φ_n is monic, so $\frac{\partial}{\partial\lambda}\Phi_n$ is a polynomial of degree at most $n-1$ and so orthogonal to Φ_n. It follows that the term in (10.4) is zero. $\qquad \square$

REMARK. Those familiar with the Feynman-Hellman theorem [35] will see a metaphysical link to this proof.

We want to do two things with (10.1). First, we sum several logs in

$$\log D_n(d\mu) = \sum_{j=0}^{n} \log \|\Phi_j\|^2$$

(by (2.4)), and second, we make an explicit choice of $d\mu_\lambda$:

THEOREM 10.2. *Let w_t be a family of L^1 functions, C^1 in t so $\log w_t$ is also C^1 in t, and $w_0 = 1$. Then*]

$$(10.5) \qquad \begin{aligned} \log D_n\left(w_1 \frac{d\theta}{2\pi} \right) &= (n+1) \log(\|\Phi_{n+1}\|_{t=1}^2) \\ &\quad - \int_0^1 dt \int \left[\frac{d}{dt} (\log w_t) \right] \left(\frac{\partial}{\partial r} |\varphi_{n+1}^*(re^{i\theta}; d\mu_t)|^2 \Big|_{r=1} \right) d\mu_t(\theta) \end{aligned}$$

PROOF. We have that

$$\frac{d}{dt} w_t \frac{d\theta}{2\pi} = \frac{w_t'}{w_t} d\mu_t = \frac{d}{dt} (\log w_t) \, d\mu_t$$

Using this, (2.4), (10.1), and the definition (9.2) of K, we get

$$(10.6) \qquad \frac{d}{dt} \log D_n\left(w_t \frac{d\theta}{2\pi} \right) = \int \left[\frac{d}{dt} (\log w_t) \right] K(e^{i\theta}, e^{i\theta}; d\mu_t) \, d\mu_t$$

Now use (9.6) to get two terms. If we integrate dt, the first term gives the integral in (10.5). The second term can be integrated using (10.1) to give the first term in (10.5). $\qquad \square$

11. Completion of the Proof

We complete the proof, following [29], by showing:

THEOREM 11.1. *Let $d\mu(\theta) = e^{L(\theta)} \frac{d\theta}{2\pi}$ where $e^{i\theta} \mapsto L(\theta)$ is analytic in a neighborhood of $\partial\mathbb{D}$. Then*

$$(11.1) \qquad \lim_{n\to\infty} [\log D_n(d\mu) - (n+1)\hat{L}_0] = \sum_{k=1}^{\infty} k|\hat{L}_k|^2$$

PROOF. Without loss, we suppose $\int d\mu(\theta) = 1$. We claim first that (under the analyticity assumption)

$$(11.2) \qquad \lim_{n \to \infty} (n+1)[\log \|\Phi_{n+1}\|^2 - \hat{L}_0] = 0$$

For in terms of the Verblunsky coefficients,

$$\|\Phi_{n+1}\|^2 = \prod_{j=0}^{n} (1 - |\alpha_j|^2)$$

(by (2.11)), while (4.1) and (2.14) say that

$$e^{\hat{L}_0} = \prod_{j=0}^{\infty} (1 - |\alpha_j|^2)$$

Thus

$$\|\Phi_{n+1}\|^2 e^{-\hat{L}_0} = \prod_{j=n+1}^{\infty} (1 - |\alpha_j|^2)^{-1}$$
$$= 1 + O(e^{-An/2})$$

by (5.17). This proves (11.2).

Thus, by (10.5), (11.1) is equivalent to there being a choice w_t with

$$(11.3) \qquad -\int_0^1 dt \int \left(\frac{d}{dt} \log w_t \right) \left. \frac{\partial}{\partial r} |\varphi_{n+1}^*(re^{i\theta}; d\mu_t)|^2 \right|_{r=1} d\mu_t(\theta) \to \sum_{k=1}^{\infty} k |\hat{L}_k|^2$$

Define for $0 \le t \le 1$,

$$(11.4) \qquad c_t = \log\left[\int e^{tL(\theta)} \frac{d\theta}{2\pi} \right]$$

and

$$(11.5) \qquad w_t(\theta) = \exp(tL(\theta) - c_t)$$

so $\int d\mu_t = 1$ also. Thus, the Szegő function for μ_t is given by

$$(11.6) \qquad \log D_t(z) = t \log D(z) - \tfrac{1}{2} c_t$$

$$(11.7) \qquad = t \sum_{k=1}^{\infty} z^k \hat{L}_k + \tfrac{1}{2} t \hat{L}_0 - \tfrac{1}{2} c_t$$

It follows that $D_t(z)$ is analytic in a t-independent disk, and then the Verblunsky coefficients obey

$$|\alpha_n(d\mu_t)| \le C e^{-An/2}$$

uniformly in $t \in [0, 1]$, and then, from the proof of Theorem 5.4, that $\varphi_n^*(z, d\mu_t) \to D_t(z)^{-1}$ uniformly in $t \in [0, 1]$ and z in $\{z \mid |z| \le e^{A/4}\}$. Thus, the $\frac{\partial}{\partial r}$ derivative in the integral on the left side of (11.3) converges to a $\frac{\partial}{\partial r}$ derivative of D_t.

We also write

$$(11.8) \qquad \frac{\partial}{\partial r} |\varphi_{n+1}^*|^2 = 2|\varphi_{n+1}^*|^2 \operatorname{Re} \frac{\partial}{\partial r} \log(\varphi_{n+1}^*) \to -2|D_t|^{-2} \operatorname{Re}\left(t \frac{\partial}{\partial r} \log D \right)$$

and we can use $|D_t|^{-2} d\mu_t = \frac{d\theta}{2\pi}$.

By (11.5),

$$(11.9) \qquad \frac{d}{dt} \log w_t = L(\theta) - \frac{d}{dt} c_t$$

Since the $\frac{d}{dt} c_t$ is θ-independent and by using (9.7), we see the $\frac{d}{dt} c_t$ can be integrated to $c_1 - c_0 = 0$. Thus, in (11.3) we can replace $\frac{d}{dt} \log w_t$ by $L(\theta)$. The result is that

$$(11.10) \qquad \text{LHS of } (11.3) \to \int_0^1 dt \int 2t\, L(\theta)\, \text{Re}\left[\frac{\partial}{\partial r} \log D\right] \frac{d\theta}{2\pi}$$

The only t-dependence is the $2t$ and $\int_0^1 2t\, dt = 1$, so

$$\text{RHS of } (11.10) = \int \left[\sum_{k=-\infty}^{\infty} \hat{L}_k e^{ik\theta}\right] \text{Re}\left(\sum_{k=1}^{\infty} k\hat{L}_k e^{ik\theta}\right) \frac{d\theta}{2\pi}$$

$$= \int \left(\sum_{k=-\infty}^{\infty} \hat{L}_k e^{ik\theta}\right)\left[\tfrac{1}{2} \sum_{k=-\infty}^{\infty} |k|\hat{L}_k e^{ik\theta}\right] \frac{d\theta}{2\pi}$$

$$= \tfrac{1}{2} \sum_{k=-\infty}^{\infty} |k|\bar{\hat{L}}_k \hat{L}_k = \sum_{1}^{\infty} k|\hat{L}_k|^2$$

as claimed. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

References

1. F.V. Atkinson, *Discrete and Continuous Boundary Problems*, Academic Press, New York, 1964.

2. J. Baik, P. Deift, K. McLaughlin, and X. Zhou, in preparation.

3. M. Bakonyi and T. Constantinescu, *Schur's Algorithm and Several Applications*, Pitman Research Notes in Math. **261**, Longman, Essex, U.K., 1992.

4. G. Baxter, *A convergence equivalence related to polynomials orthogonal on the unit circle*, Trans. Amer. Math. Soc. **99** (1961), 471–487.

5. A. Böttcher, *The Onsager formula, the Fisher-Hartwig conjecture, and their influence on research in Toeplitz operators*, J. Statist. Phys. **78** (1995), 575–584.

6. D. Bump and P. Diaconis, *Toeplitz minors*, J. Combin. Theory Ser. A **97** (2002), no. 2, 252–271.

7. P. Delsarte, Y.V. Genin, and Y.G. Kamp, *Orthogonal polynomial matrices on the unit circle*, IEEE Trans. Circuits and Systems **CAS-25** (1978), 149–160.

8. P.L. Duren, *Theory of H^p Spaces*, Pure and Applied Math., Vol. 38, Academic Press, New York-London, 1970.

9. 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.

10. M.E. Fisher and R.E. Hartwig, *Toeplitz determinants: Some applications, theorems and conjectures*, Stochastic Processes in Chemical Physics (K.E. Shuler, ed.), Advances in Chemical Physics, Vol. 15, Wiley Interscience, New York, 1969, pp. 333–353.

11. J. Geronimus, *On the trigonometric moment problem*, Ann. of Math. (2) **47** (1946), 742–761.

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

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

14. V. Guillemin and K. Okikiolu, *Spectral asymptotics of Toeplitz operators on Zoll manifolds*, J. Funct. Anal. **146** (1997), no. 2, 496–516.

15. K. Hoffman, *Analytic functions and logmodular Banach algebras*, Acta Math. **108** (1962), 271–317.

16. I.A. Ibragimov, *A theorem of Gabor Szegő*, Mat. Zametki **3** (1968), 693–702. [Russian]

17. K. Johansson, *On Szegő's asymptotic formula for Toeplitz determinants and generalizations*, Bull. Sci. Math. (2) **112** (1988), no. 3, 257–304.

18. T. Kailath, *A view of three decades of linear filtering theory*, IEEE Trans. Inform. Theory **IT-20** (1974), 146–181.

19. H.J. Landau, *Maximum entropy and the moment problem*, Bull. Amer. Math. Soc. **16** (1987), 47–77.

20. A. Lenard, *Momentum distribution in the ground state of the one-dimensional system of impenetrable Bosons*, J. Mathematical Phys. **5** (1964), 930–943.

21. _____, *Some remarks on large Toeplitz determinants*, Pacific J. Math. **42** (1972), 137–145.

22. G. Lumer, *Analytic functions and Dirichlet problem*, Bull. Amer. Math. Soc. **70** (1964), 98–104.

23. B.M. McCoy, *Introductory remarks to Szegő's paper "On certain Hermitian forms associated with the Fourier series of a positive function"*, Gabor Szegő: Collected Papers, Volume 1, 1915–1927, (R. Askey, ed.), Birkhäuser, Boston, 1982, pp. 47–51.

24. M.L. Mehta, *Random Matrices and the Statistical Theory of Energy Levels*, Academic Press, New York, 1967.

25. P. Nevai and V. Totik, *Orthogonal polynomials and their zeros*, Acta Sci. Math. (Szeged) **53** (1989), 99–104.

26. G. Pólya, L'Intermédiaire des Mathématiciens **21** (1914), S. 27.

27. W. Rudin, *Real and Complex Analysis*, 3rd edition, McGraw-Hill, New York, 1987.

28. B. Simon, *Orthogonal Polynomials on the Unit Circle. Part 1. Classical Theory*, Amer. Math. Soc. Colloq. Publ., Vol. 54, Providence, RI, 2005.

29. _____, *Orthogonal Polynomials on the Unit Circle. Part 2. Spectral Theory*, Amer. Math. Soc. Colloq. Publ., Vol. 54, Providence, RI, 2005.

30. G. Szegő, *Ein Grenzwertsatz über die Toeplitzschen Determinanten einer reelen positiven Funktion*, Math. Ann. **76** (1915), 490–503.

31. _____, *Beiträge zur Theorie der Teoplitzschen Formen, I*, Math. Z. **6** (1920), 167–202.

32. _____, *Beiträge zur Theorie der Teoplitzschen Formen, II*, Math. Z. **9** (1921), 167–190.

33. _____, *Orthogonal Polynomials*, Amer. Math. Soc. Colloq. Publ., Vol. 23, Providence, RI, 1939; 3rd edition, 1967.

34. _____, *On certain Hermitian forms associated with the Fourier series of a positive function*, Comm. Sém. Math. Univ. Lund **1952** (1952), Tome Supplementaire, 228–238.

35. W. Thirring, *Quantum Mechanics of Atoms and Molecules*, Springer-Verlag, New York, 1981.

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

37. _____, *On positive harmonic functions (second part)*, Proc. London Math. Soc. **40** (1936), 290–320.

38. J. Wermer, *Dirichlet algebras*, Duke Math. J. **27** (1960), 373–381.

39. H. Widom, *Eigenvalue distribution theorems for certain homogeneous spaces*, J. Funct. Anal. **32** (1979), no. 2, 139–147.

MATHEMATICS 253-37, CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA, CA 91125, U.S.A.
E-mail address: bsimon@caltech.edu