FINITE GAP JACOBI MATRICES, II. THE SZEGŐ CLASS

JACOB S. CHRISTIANSEN[1], BARRY SIMON[2,3], AND MAXIM ZINCHENKO[2]

ABSTRACT. Let $\mathfrak{e} \subset \mathbb{R}$ be a finite union of disjoint closed intervals. We study measures whose essential support is $\mathfrak{e}$ and whose discrete eigenvalues obey a 1/2-power condition. We show that a Szegő condition is equivalent to

$$\limsup \frac{a_1 \cdots a_n}{\mathrm{cap}(\mathfrak{e})^n} > 0$$

(this includes prior results of Widom and Peherstorfer–Yuditskii). Using Remling's extension of the Denisov–Rakhmanov theorem and an analysis of Jost functions, we provide a new proof of Szegő asymptotics, including L^2 asymptotics on the spectrum. We use heavily the covering map formalism of Sodin–Yuditskii as presented in our first paper in this series.

1. Introduction

In this paper, we study Jacobi matrices, J, and asymptotics of the associated orthogonal polynomials (OPRL), where $\sigma_{\mathrm{ess}}(J)$ is a finite gap set, $\mathfrak{e}$. By this we mean that $\mathfrak{e}$ is a finite union of disjoint closed intervals,

$$\mathfrak{e} = \bigcup_{j=1}^{\ell+1} [\alpha_j, \beta_j] \qquad \alpha_1 < \beta_1 < \alpha_2 < \cdots < \beta_{\ell+1} \tag{1.1}$$

ℓ counts the number of gaps, that is, bounded open intervals in $\mathbb{R} \setminus \mathfrak{e}$.

Date: November 16, 2009.

2000 *Mathematics Subject Classification.* 42C05, 58J53, 14H30.

Key words and phrases. Isospectral torus, Szegő asymptotics, orthogonal polynomials.

[1] Department of Mathematical Sciences, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen, Denmark. E-mail: stordal@math.ku.dk.
Supported in part by a Steno Research Grant from FNU, the Danish Research Council.

[2] Mathematics 253-37, California Institute of Technology, Pasadena, CA 91125. E-mail: bsimon@caltech.edu; maxim@caltech.edu.

[3] Supported in part by NSF grant DMS-0652919.

We recall that a Jacobi matrix is a tridiagonal matrix which we label

$$J = \begin{pmatrix} b_1 & a_1 & 0 & \cdots \\ a_1 & b_2 & a_2 & \cdots \\ 0 & a_2 & b_3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{1.2}$$

The Jacobi parameters $\{a_n, b_n\}_{n=1}^\infty$ have $a_n > 0$ and $b_n \in \mathbb{R}$. There is a one-one correspondence between probability measures, $d\mu$, of compact support on $\mathbb{R}$ and bounded Jacobi matrices where $d\mu$ is the spectral measure for J and the vector $(1, 0, \dots)^t$. Moreover, $d\mu$ determines J via recursion relations for the orthonormal polynomials, $p_n(x)$, which are ($a_0 \equiv 0$)

$$xp_n(x) = a_{n+1}p_{n+1}(x) + b_{n+1}p_n(x) + a_n p_{n-1}(x) \tag{1.3}$$

See [32, 9, 23, 26] for background on OPRL.

This paper is the second in a series—the first, [2], henceforth called paper I, studied the isospectral torus, an ℓ-dimensional family of two-sided almost periodic Jacobi matrices with essential spectrum, $\mathfrak{e}$, about which we'll say more later in this introduction. We note for now that these matrices have periodic coefficients if and only if the harmonic measure of the intervals $[\alpha_j, \beta_j]$ are all rational (i.e., if $d\rho_\mathfrak{e}$ is the potential theoretic equilibrium measure for $\mathfrak{e}$, then each $\rho_\mathfrak{e}([\alpha_j, \beta_j])$ is rational; for background on potential theory in spectral analysis, see [29, 25]). We'll call this the periodic case.

In the current paper, we want to study Szegő's theorem for the general finite gap case. Of course, the phrase "Szegő's theorem" can be ambiguous since Szegő was so prolific, but by this we mean a set of results concerned with leading asymptotics in the theory of orthogonal polynomials on the unit circle (OPUC). Even here, there is ambiguity since some of the results can be interpreted in terms of Toeplitz determinants and there are several related objects. Indeed, we'll distinguish between what we call Szegő's theorem and Szegő asymptotics.

In the OPUC case, the recursion parameters $\{\alpha_n\}_{n=0}^\infty$ lie in $\mathbb{D} = \{z \mid |z| < 1\}$ and are called Verblunsky coefficients. We use $\varphi_n(z)$ for the orthonormal polynomials and write the measure $d\mu$ as

$$d\mu(\theta) = w(\theta)\,\frac{d\theta}{2\pi} + d\mu_\text{s}(\theta) \tag{1.4}$$

where $d\mu_\text{s}$ is $d\theta/2\pi$-singular. One also defines $\rho_n = (1 - |\alpha_n|^2)^{1/2}$ (see [32, 9, 23, 24, 22] for background on OPUC).

Then what we'll call Szegő's theorem for OPUC says that

$$\lim_{N\to\infty}\prod_{n=0}^{N}\rho_n = \exp\biggl(\int_0^{2\pi}\log(w(\theta))\frac{d\theta}{2\pi}\biggr) \tag{1.5}$$

Notice that since $\rho_n \leq 1$, the limit on the left always exists, although it may be 0. By Jensen's inequality, the integral on the right is nonpositive, but may diverge to $-\infty$, in which case we interpret the exponential as 0. It is easy to see that the left side is nonzero if and only if $\sum_{n=0}^\infty |\alpha_n|^2 < \infty$. Thus, (1.5) implies

$$\sum_{n=0}^\infty |\alpha_n|^2 < \infty \quad\Leftrightarrow\quad \int \log(w(\theta))\,\frac{d\theta}{2\pi} > -\infty \tag{1.6}$$

By Szegő asymptotics, we mean the fact that when both conditions in (1.6) hold, there is an explicit nonvanishing function, G, on $\mathbb{C}\setminus\overline{\mathbb{D}}$ so that for z in that set,

$$\lim_{n\to\infty} z^{-n}\varphi_n(z) = G(z) \tag{1.7}$$

In terms of the conventional Szegő function,

$$D(z) = \exp\biggl(\int \frac{e^{i\theta}+z}{e^{i\theta}-z}\log(w(\theta))\frac{d\theta}{2\pi}\biggr), \quad z\in\mathbb{D} \tag{1.8}$$

we have $G(z) = \overline{D(1/\bar z)}^{\,-1}$.

Analogs of Szegő's theorem for OPRL, where $\mathfrak{e}$ is a single interval (typically $\mathfrak{e} = [-1,1]$ or $[-2,2]$), were found initially by Szegő [31], with important developments by Shohat [20] and Nevai [13]. These works suppose no or finitely many eigenvalues outside $\mathfrak{e}$. The natural condition on eigenvalues (see (1.10) and (1.13) below) was found by Killip–Simon [11] and Peherstorfer–Yuditskii [15]. The best form of Szegő's theorem (with a Szegő condition; see below) is

Theorem 1.1 (Simon–Zlatoš)**.** *Let J be a Jacobi matrix with essential spectrum $[-2,2]$, $\{a_n,b_n\}_{n=1}^\infty$ its Jacobi parameters, $\{x_k\}$ a listing of its eigenvalues outside $[-2,2]$, and*

$$d\mu(x) = w(x)\,dx + d\mu_{\rm s}(x) \tag{1.9}$$

its spectral measure. Define

$$\mathcal{E}(J) = \sum_k (|x_k|-2)^{1/2} \tag{1.10}$$

and

$$A_n = a_1\cdots a_n \qquad \bar A = \limsup A_n \qquad \underline{A} = \liminf A_n \tag{1.11}$$

Consider the three conditions:

(i) *Szegő condition*

$$\int_{-2}^{2} \log(w(x))(4 - |x|^2)^{-1/2}\, dx > -\infty \tag{1.12}$$

(ii) *Blaschke condition*

$$\mathcal{E}(J) < \infty \tag{1.13}$$

(iii) *Widom condition*

$$0 < \underline{A} \le \bar{A} < \infty \tag{1.14}$$

Then any two of (i)–(iii) *imply the third, and if they hold, the following have limits as $N \to \infty$:*

$$A_N, \quad \sum_{n=1}^{N} b_n, \quad \sum_{n=1}^{N} (a_n - 1) \tag{1.15}$$

and

$$\sum_{n=1}^{\infty} |a_n - 1|^2 + |b_n|^2 < \infty \tag{1.16}$$

Before leaving our summary of the case $\mathfrak{e} = [-2, 2]$, we note that Damanik–Simon [5] have proven Szegő asymptotics in some cases where the Szegő condition fails. This will not concern us here, but will be studied in the finite gap case in paper III [3].

In Section 4, we prove a precise analog of the statement "any two of (i)–(iii) imply the third" for general finite gap sets, $\mathfrak{e}$. We note that for the periodic case, this is a prior result of Damanik–Killip–Simon [4]. There are also prior results for the general finite gap case in Widom [33], Aptekarev [1], and Peherstorfer–Yuditskii [16, 17]; see Section 4 for more details.

The limit results, (1.15) and (1.16), need modification, however. First, even in the general one-interval case, one needs $a_1 \cdots a_n / C^n$ for a suitable constant C. The theory of regular measures [29, 25] says the right value of C must be $\operatorname{cap}(\mathfrak{e})$, the logarithmic capacity of $\mathfrak{e}$—a result that, in this context, goes back at least to Widom [33] who also discovered that $a_1 \cdots a_n / \operatorname{cap}(\mathfrak{e})^n$ doesn't have a limit but is only asymptotically almost periodic.

These limit results are expressed most naturally in terms of the isospectral torus associated to $\mathfrak{e}$. For any Jacobi matrix obeying the analogs of (i)–(iii), there is an element $\{\tilde{a}_n, \tilde{b}_n\}_{n=1}^{\infty}$ of the isospectral torus so that

$$\lim_{n \to \infty} |a_n - \tilde{a}_n| + |b_n - \tilde{b}_n| = 0 \tag{1.17}$$

This result, which goes back to Aptekarev [1] and Peherstorfer–Yuditskii [16, 17] using variational methods, will be proven with our techniques in Section 6, where we'll also prove that $\lim(a_1 \cdots a_n / \tilde{a}_1 \cdots \tilde{a}_n)$ exists and is nonzero. (In paper I, we proved that in the isospectral torus, $\tilde{a}_1 \cdots \tilde{a}_n / \operatorname{cap}(\mathfrak{e})^n$ is almost periodic in n.)

An interesting open question concerns the analog of (1.16):

Open Question 1. Is $\sum_{n=1}^\infty |a_n - \tilde{a}_n|^2 + |b_n - \tilde{b}_n|^2 < \infty$ when the analogs of (i)–(iii) hold?

In Section 7, we'll prove an analog of Szegő asymptotics, namely, away from the interval $[\alpha_1, \beta_{\ell+1}]$, the ratio $p_n(z)/\tilde{p}_n(z)$ has a nonzero limit where $\tilde{p}_n$ are the OPRL for $\{\tilde{a}_n, \tilde{b}_n\}_{n=1}^\infty$.

Let us next summarize some of the techniques we'll use below, in part to standardize some notation. Coefficient stripping plays an important role in the analysis: if J has Jacobi parameters $\{a_k, b_k\}_{k=1}^\infty$, then the n-times stripped Jacobi matrix, $J^{(n)}$, is the one with parameters $\{a_{n+k}, b_{n+k}\}_{k=1}^\infty$, that is, with

$$a_k(J^{(n)}) = a_{k+n}(J) \qquad b_k(J^{(n)}) = b_{n+k}(J) \tag{1.18}$$

If the m-function of J is defined on $\mathbb{C}_+ = \{z \mid \operatorname{Im} z > 0\}$ by

$$m(z, J) = \langle \delta_1, (J - z)^{-1} \delta_1 \rangle = \int \frac{d\mu(x)}{x - z} \tag{1.19}$$

then we have the coefficient stripping relation that goes back to Jacobi and Stieltjes,

$$m(z, J)^{-1} = -z + b_1 - a_1^2 m(z, J^{(1)}) \tag{1.20}$$

We'll make heavy use of the covering space formalism introduced in spectral theory by Sodin–Yuditskii [28] and presented with our notation in paper I. $\mathbf{x}(z)$ is the unique meromorphic map of $\mathbb{D}$ to $\mathbb{C} \cup \{\infty\} \setminus \mathfrak{e}$ which is locally one-one with

$$\mathbf{x}(z) = \frac{x_\infty}{z} + O(1) \tag{1.21}$$

near $z = 0$ and $x_\infty > 0$.

There is a (Fuchsian) group, Γ, of Möbius transformations of $\mathbb{D}$ onto itself so that

$$\mathbf{x}(z) = \mathbf{x}(w) \quad \Leftrightarrow \quad \exists \gamma \in \Gamma \text{ so that } \gamma(z) = w \tag{1.22}$$

A natural fundamental set, $\mathcal{F}$, is defined as follows:

$$\mathcal{F}^{\text{int}} = \{z \mid |z| < |\gamma(z)|, \text{ all } \gamma \neq 1, \gamma \in \Gamma\} \tag{1.23}$$

$\partial\mathcal{F}^{\text{int}} \cap \mathbb{D}$ is then 2ℓ orthocircles, ℓ in each half-plane. $\mathcal{F}$ is $\mathcal{F}^{\text{int}}$ union the ℓ orthocircles in $\mathbb{C}_+$. $\mathbf{x}$ is then one-one and onto from $\mathcal{F}$ to $\mathbb{C} \cup \{\infty\} \setminus \mathfrak{e}$.

$\mathcal{L}$, the set of limit points of Γ, is defined as $\overline{\{\gamma(0) \mid \gamma \in \Gamma\}} \cap \partial\mathbb{D}$. $\mathbf{x}$ can be meromorphically extended from $\mathbb{D}$ to all of $\mathbb{C} \cup \{\infty\} \setminus \mathcal{L}$, or alternatively, there is a map $\mathbf{x}^\sharp\colon \mathbb{C}\cup\{\infty\}\setminus\mathcal{L}$ to $\mathcal{S}$, the two-sheeted Riemann surface of $[\prod_{j=1}^{\ell+1}(z-\alpha_j)(z-\beta_j)]^{1/2}$. All this is described in more detail in paper I of this series.

That paper also discusses Blaschke products, $B(z,w)$, of the Blaschke factors at $\{\gamma(w)\}_{\gamma\in\Gamma}$. $B(z)\equiv B(z,0)$ is related to the potential theoretic Green's function, $G_{\mathfrak{e}}(x)$, for $\mathfrak{e}$ by

$$|B(z)| = e^{-G_{\mathfrak{e}}(\mathbf{x}(z))} \tag{1.24}$$

which, in particular, implies that near $z=0$,

$$B(z) = \frac{\text{cap}(\mathfrak{e})}{x_\infty}\, z + O(z^2) \tag{1.25}$$

Finally, we use heavily the pullback of m to $\mathbb{D}$ via

$$M(z) = -m(\mathbf{x}(z)) \tag{1.26}$$

We end this introduction with a sketch of the contents of this paper. Our approach to Szegő's theorem is a synthesis of the covering map method and the approach of Killip–Simon [11], Simon–Zlatoš [27], and Simon [21] used for $\mathfrak{e}=[-2,2]$. As such, step-by-step sum rules are critical. These are found in Section 2. One disappointment is that we have thus far not succeeded in finding an analog of what has come to be called the Killip–Simon theorem (from [11]). This result gives necessary and sufficient conditions for the case $\mathfrak{e}=[-2,2]$ that $\sum_{n=1}^\infty (a_n-1)^2 + b_n^2 < \infty$.

Open Question 2. Is there a Killip–Simon theorem for the general finite gap Jacobi matrix?

We note that Damanik–Killip–Simon [4] have found an analog for the case where each band has harmonic measure exactly $(\ell+1)^{-1}$.

Section 3 provides a technical interlude on eigenvalue limit theorems needed in the later sections. Section 4 proves a Szegő-type theorem for general finite gap $\mathfrak{e}$. Section 5 defines Jost functions and Jost solutions. Section 6 proves the existence of the claimed $\{\tilde a_n, \tilde b_n\}_{n=1}^\infty$ in the isospectral torus and asymptotics of Jost solutions. Section 7 proves asymptotic formulae for the orthogonal polynomials away from the convex hull of $\mathfrak{e}$ (i.e., the interval $[\alpha_1,\beta_{\ell+1}]$), and Section 8 L^2 asymptotics on $\mathfrak{e}$.

The idea that we use in Sections 6 and 7 of first proving Jost asymptotics and using that to get Szegő asymptotics is borrowed from an analog for $\mathfrak{e}=[-2,2]$ of Damanik–Simon [5]. But Section 7 has a simplification of their equivalence argument that is an improvement even for $\mathfrak{e}=[-2,2]$. Most of the results in Sections 6–8 are explicit or implicit in Peherstorfer–Yuditskii [16, 17]. We claim two novelties here.

First, the underlying mechanism of our proof of asymptotics is different from their variational approach. Instead, we use a recent theorem of Remling [18] about approach to the isospectral torus, together with an analysis of automorphic characters of Jost functions. Second, by using ideas in a different paper of Peherstorfer–Yuditskii [15], we can clarify the L^2-convergence result of Section 8.

We would like to thank F. Peherstorfer and P. Yuditskii for the private communication [14]. J.S.C. would like to thank M. Flach and A. Lange for the hospitality of Caltech where this work was completed.

2. Step-by-Step Sum Rules

As noted in the introduction, a key to the approach to Szegő-type theorems for $\mathfrak{e} = [-2,2]$ that we'll follow is step-by-step sum rules. Our goal in this section is to prove those for a general finite gap $\mathfrak{e}$. In Theorem 7.5 of paper I, we proved such results for measures in the isospectral torus, and our discussion here will closely follow the proof there. The major change is that there, with finitely many eigenvalues in $\mathbb{R} \setminus \mathfrak{e}$, we could use finite Blaschke products. Here, because we do not wish to suppose a priori a 1/2-power condition on the eigenvalues, we'll need the alternating Blaschke products of Theorem 4.9 of paper I. Here is the result:

Theorem 2.1 (Nonlocal step-by-step sum rule)**.** *Let J be a Jacobi matrix with $\sigma_{\text{ess}}(J) = \mathfrak{e}$. Let $J^{(1)}$ be the once-stripped Jacobi matrix and let $\{p_j\}_{j=1}^\infty$ be the points in $\mathcal{F}$ that are mapped by the covering map, $\mathbf{x}$, to the eigenvalues of J and $\{z_j\}_{j=1}^\infty$ the corresponding points for the eigenvalues of $J^{(1)}$. Let B_∞ be the alternating Blaschke product with poles at $\{\gamma(p_j)\}_{j=1;\gamma\in\Gamma}^\infty$ and zeros at $\{\gamma(z_j)\}_{j=1;\gamma\in\Gamma}^\infty$. Let $B(z)$ be the Blaschke product with zeros at $\{\gamma(0)\}_{\gamma\in\Gamma}$. Let $M(z)$ be the m-function,* (1.26)*, for J, and $M^{(1)}(z)$ the one for $J^{(1)}$. Then*

(a) $\lim_{r\uparrow 1} M(re^{i\theta}) \equiv M(e^{i\theta})$ *and* $\lim_{r\uparrow 1} M^{(1)}(re^{i\theta}) \equiv M^{(1)}(e^{i\theta})$ *exist for $d\theta/2\pi$-a.e. θ.*

(b) *Up to sets of $d\theta/2\pi$ measure zero,*

$$\{\theta \mid \operatorname{Im} M(e^{i\theta}) \neq 0\} = \{\theta \mid \operatorname{Im} M^{(1)}(e^{i\theta}) \neq 0\} \tag{2.1}$$

(c)

$$\log\left(\frac{\operatorname{Im} M(e^{i\theta})}{\operatorname{Im} M^{(1)}(e^{i\theta})}\right) \in \bigcap_{p<\infty} L^p\left(\partial\mathbb{D}, \frac{d\theta}{2\pi}\right) \tag{2.2}$$

(d) *We have*

$$a_1 M(z) = B(z)B_\infty(z) \exp\biggl(\int \frac{e^{i\theta}+z}{e^{i\theta}-z} \log\biggl(\frac{\operatorname{Im} M(e^{i\theta})}{\operatorname{Im} M^{(1)}(e^{i\theta})}\biggr)\frac{d\theta}{4\pi}\biggr) \tag{2.3}$$

Remarks. 1. We've labeled the p's and z's to be infinite in number, although there may be only finitely many. Moreover, we need to group them into not one sequence but potentially $2\ell+2$ if each of the points in $\{\alpha_j, \beta_j\}_{j=1}^{\ell+1}$ is a limit point of eigenvalues in $\mathbb{R}\setminus\mathfrak{e}$. Once this is done, one forms an alternating Blaschke product for each sequence (the p's and z's in each sequence alternate along a boundary arc of $\mathcal{F}$ or on $(0,1)$ or $(-1,0)$), and then takes the product of these $2\ell+2$ alternating Blaschke products.

2. $\operatorname{Im} M$ and $\operatorname{Im} M^{(1)}$ have the same sign at each point of $\partial\mathbb{D}$, positive or negative, depending on whether $\mathbf{x}$ maps to an upper or lower lip of $\mathfrak{e}$.

3. We've written (c) and (d) assuming that the set in (2.1) is all of $\partial\mathbb{D}$ (up to sets of Lebesgue measure zero). A more proper version is that $\lim_{r\uparrow 1}|M(re^{i\theta})|^2$ has a limit as $r\uparrow 1$ which, when multiplied by a_1^2, is the ratio $\operatorname{Im} M/\operatorname{Im} M^{(1)}$ at points in the set in (2.1). It is that boundary value that enters in (2.2) and (2.3).

Proof. We follow the arguments used for Theorem 7.5 of paper I. For $z\in\mathbb{D}$, not a pole or zero of M, let

$$h(z) = \frac{a_1 M(z)}{B(z)B_\infty(z)} \tag{2.4}$$

At the poles and zeros of M, $h(z)$ has removable singularities and no zero values, so h is nonvanishing and analytic in all of $\mathbb{D}$.

All of M, B, and B_∞ are positive on $(0,\varepsilon)$ for ε small, so one can choose a branch of $\log(h(z))$ which has $\operatorname{Im}(\log(h(z))) = 0$ on $(0,\varepsilon)$. Since $\operatorname{Im} M>0$ on $\mathbb{C}_+\cap\mathcal{F}$ and $\operatorname{Im} M<0$ on $\mathbb{C}_-\cap\mathcal{F}$, with this choice,

$$|\arg(M(z))| \le \pi \text{ on } \mathcal{F} \tag{2.5}$$

By eqn. (4.84) in Theorem 4.9 of paper I, there is a constant C so that

$$|\arg(B_\infty(z)B(z))| \le C \text{ on } \mathcal{F} \tag{2.6}$$

As in the proof of Theorem 7.5 of paper I, this plus the fact that $h(z)$ is character automorphic implies that

$$\sup_{0<r<1} \int |\operatorname{Im}(\log(h(re^{i\theta})))|^p \frac{d\theta}{2\pi} < \infty \tag{2.7}$$

Thus, by the M. Riesz theorem,

$$\log(h) \in \bigcap_{p<\infty} H^p(\mathbb{D}) \tag{2.8}$$

This implies that $\log(h)$, and so M, has boundary values and

$$\log|M(e^{i\theta})| \in \bigcap_{p<\infty} L^p\left(\partial\mathbb{D}, \frac{d\theta}{2\pi}\right) \tag{2.9}$$

Taking boundary values in (see (1.20))

$$M(z)^{-1} = \mathbf{x}(z) - b_1 - a_1^2 M^{(1)}(z) \tag{2.10}$$

shows that (2.1) holds, and on the set where $\operatorname{Im} M \neq 0$,

$$|a_1 M(e^{i\theta})|^2 = \frac{\operatorname{Im} M(e^{i\theta})}{\operatorname{Im} M^{(1)}(e^{i\theta})} \tag{2.11}$$

This and (2.9) imply (2.2), and (2.3) is just the Poisson representation for $\log(h(z))$. □

The main use we'll make of (2.3) is to divide by $B(z)$ and take $z \to 0$ using (1.21) and (1.25). The result is:

Theorem 2.2 (Step-by-step C_0 sum rule)**.**

$$\frac{a_1}{\operatorname{cap}(\mathfrak{e})} = B_\infty(0) \exp\left(\int_0^{2\pi} \log\left(\frac{\operatorname{Im} M(e^{i\theta})}{\operatorname{Im} M^{(1)}(e^{i\theta})}\right) \frac{d\theta}{4\pi}\right) \tag{2.12}$$

3. Fun and Games with Eigenvalues

Sum rules include eigenvalue sums—it appears somewhat hidden in (2.12) as $B_\infty(0)$. Since, in exploiting sum rules, we'll be looking at the behavior of sums over families, often with infinitely many elements, we'll need control on such sums. This was true already in the single interval case as studied by [11, 27], but there the main tool needed was a simple variational principle. Eigenvalues above or below the essential spectrum are given by a linear variational principle. This is not true for eigenvalues in gaps, and so we'll need some extra techniques, which we put in the current section. We note that there are still limitations on what can be done in gaps. For example, for perturbations of elements of the finite gap isospectral torus, there is a 1/2 critical Lieb–Thirring bound at the external edges [7] but not yet one known for internal gap edges [10].

We begin with two results about the relation of eigenvalues of J and $J^{(n)}$, the n-times stripped Jacobi matrix of (1.18).

Theorem 3.1. *Let J be a Jacobi matrix with $\sigma_{\text{ess}}(J) = \mathfrak{e}$. Let $c \in (\beta_j, \alpha_{j+1})$, one of the gaps of $\mathbb{R} \setminus \mathfrak{e}$. Suppose f is defined, positive, and monotone on (β_j, c) with $\lim_{x \downarrow \beta_j} f(x) = 0$. Let $c > x_1(J) > x_2(J) > \dots > \beta_j$ be the eigenvalues of J in (β_j, c). Then the eigenvalues of J and $J^{(1)}$ strictly interlace, that is, either*

$$x_1(J) > x_1(J^{(1)}) > x_2(J) > x_2(J^{(1)}) > \dots \tag{3.1}$$

or

$$x_1(J^{(1)}) > x_1(J) > x_2(J^{(1)}) > x_2(J) > \dots \tag{3.2}$$

In particular, $\sum_{k=1}^\infty [f(x_k(J)) - f(x_k(J^{(1)}))]$ is always conditionally convergent.

Remarks. 1. For simplicity of notation, we stated this and the following theorem for (β_j, c), but a similar result holds for (c, α_{j+1}) and also for $(-\infty, \alpha_1)$ and $(\beta_{\ell+1}, \infty)$.

2. By iteration, we also get convergence of $\sum_{k=1}^\infty [f(x_k(J)) - f(x_k(J^{(n)}))]$ for each n.

Proof. By the fact that $x_k(J)$ are the poles of $m(z)$ in (β_j, c) and $x_k(J^{(1)})$ the zeros, and since $\frac{d}{dz} m(z) = \int \frac{d\mu(x)}{(x-z)^2} > 0$ for $z \in (\beta_j, c)$, we see the interlacing, which implies (3.1) (if $m(c) \le 0$) or (3.2) (if $m(c) > 0$). The conditional convergence of the sum is standard for alternating sums. ☐

Theorem 3.2. *Under the hypotheses of Theorem 3.1, if*

$$S \equiv \sup_n \biggl| \sum_{k=1}^\infty f(x_k(J)) - f(x_k(J^{(n)})) \biggr| < \infty \tag{3.3}$$

then

$$\sum_{k=1}^\infty f(x_k(J)) < \infty \tag{3.4}$$

Proof. We will need the fact proven below (in Theorem 3.4) that for each $j \in \{1, \dots, \ell\}$ and $\varepsilon > 0$, there is an N so for $n \ge N$, $J^{(n)}$ has either 0 or 1 eigenvalue in $(\beta_j + \varepsilon, \alpha_{j+1} - \varepsilon)$.

So for $n \ge N$ we may have $x_1(J^{(n)}) > \beta_j + \varepsilon$, but $x_k(J^{(n)}) \le \beta_j + \varepsilon$ for all $k \ge 2$. Hence, for $n \ge N$,

$$\begin{aligned} &\sum_{\{k \mid \beta_j + \varepsilon < x_k(J) < c\}} [f(x_k(J)) - f(\beta_j + \varepsilon)] \\ &\qquad \le f(c) + \sum_{\{k \mid \beta_j + \varepsilon < x_k(J) < c\}} [f(x_k(J)) - f(x_k(J^{(n)}))] \end{aligned} \tag{3.5}$$

Recall now that $J^{(n)}$ can be obtained by decoupling J with a rank 2 perturbation (which is the sum of a positive and a negative rank 1 perturbation) and removing the finite block. Therefore, if we pick $\varepsilon > 0$ so small that $x_3(J) > \beta_j + \varepsilon$, it follows that $x_k(J) > x_k(J^{(n)})$ for all $k \geq 2$ (when $n \geq N$). This implies that

$$\sum_{\{k \mid \beta_j + \varepsilon < x_k(J) < c\}} [f(x_k(J)) - f(x_k(J^{(n)}))] \leq S \tag{3.6}$$

So, for sufficiently small ε_0 and $\varepsilon_1 < \varepsilon_0$,

$$\sum_{\{k \mid \beta_j + \varepsilon_0 < x_k(J) < c\}} [f(x_k(J)) - f(\beta_j + \varepsilon_1)] \leq f(c) + S \tag{3.7}$$

Taking $\varepsilon_1 \downarrow 0$ and then $\varepsilon_0 \downarrow 0$ yields (3.4). □

The following lemma is well known, used for example in Denisov [6]:

Lemma 3.3. *Let A be a bounded operator with*

$$\gamma = \inf(\sigma_{\text{ess}}(A)) \tag{3.8}$$

Let P_n be a family of orthogonal projections with

$$\text{s-lim}\, P_n = 0 \tag{3.9}$$

Then for any ε, we can find N so that for $n \geq N$,

$$\sigma(P_n A P_n \restriction \text{ran}(P_n)) \subset [\gamma - \varepsilon, \infty) \tag{3.10}$$

Proof. Since (3.8) holds, for any ε, we can write

$$A = A_\varepsilon + B_\varepsilon \tag{3.11}$$

where $A_\varepsilon \geq \gamma - \varepsilon/2$ and B_ε is finite rank, and so compact.

By (3.9), $P_n B_\varepsilon P_n \to 0$ in $\|\cdot\|$, so we can find N so that, for $n \geq N$, $\|P_n B_\varepsilon P_n\| \leq \varepsilon/2$. Then for each $n \geq N$,

$$P_n A P_n \geq P_n \left(\gamma - \frac{\varepsilon}{2} - \frac{\varepsilon}{2}\right) P_n \geq (\gamma - \varepsilon) P_n \tag{3.12}$$

proving (3.10). □

Theorem 3.4. *Let J be a bounded Jacobi matrix with $(\alpha, \beta) \cap \sigma_{\text{ess}}(J) = \emptyset$. Let $J^{(n)}$ be the n-times stripped Jacobi matrix. Then for any ε, we can find N so that, for $n \geq N$, $J^{(n)}$ has at most one eigenvalue in $(\alpha + \varepsilon, \beta - \varepsilon)$.*

Proof. Let P_n be the projection onto $\text{span}\{\delta_j\}_{j=n+1}^{\infty}$, so

$$J^{(n)} = P_n J P_n \restriction \text{ran}(P_n) \tag{3.13}$$

Let $\gamma = \frac{1}{2}(\alpha+\beta)$ and $A = (J-\gamma)^2$, $A^{(n)} = P_n A P_n \restriction \operatorname{ran}(P_n)$. By the spectral mapping theorem,

$$\inf(\sigma_{\text{ess}}(A)) \geq [\tfrac{1}{2}\,(\beta-\alpha)]^2 \tag{3.14}$$

so, by the lemma, for any ε', there is N so for $n \geq N$,

$$\inf \sigma(A^{(n)}) \geq [\tfrac{1}{2}\,(\beta-\alpha)]^2 - \varepsilon' = [\tfrac{1}{2}\,(\beta-\alpha) - \varepsilon]^2 \tag{3.15}$$

where ε' is chosen so that (3.15) holds.

Since

$$A^{(n)} - (J^{(n)} - \gamma)^2 = P_n(J-\gamma)(1-P_n)(J-\gamma)P_n \tag{3.16}$$

is rank one, $(J^{(n)}-\gamma)^2$ has at most one eigenvalue (which is simple) below $[\frac{1}{2}(\beta-\alpha)-\varepsilon]^2$, which proves the claimed result by the spectral mapping theorem. □

Next, we turn to estimating eigenvalue sums like

$$\mathcal{E}(J) = \sum_{x \in \sigma(J)\setminus\mathfrak{e}} \operatorname{dist}(x,\mathfrak{e})^{1/2} \tag{3.17}$$

with a goal of showing, for example, that if $\mathcal{E}(J)$ is finite, then so is $\sup_n \mathcal{E}(J^{(n)})$.

Definition. Let A be a bounded selfadjoint operator with $(a,b) \cap \sigma_{\text{ess}}(A) = \emptyset$. We set

$$\Sigma_{(a,b)}(A) = \sum_{x \in \sigma(A)\cap(a,b)} \operatorname{dist}(x, \mathbb{R}\setminus(a,b))^{1/2} \tag{3.18}$$

where the sum includes x as many times as the multiplicity of that eigenvalue.

Theorem 3.5. *Let A be a bounded selfadjoint operator with $(a,b) \cap \sigma_{\text{ess}}(A) = \emptyset$ and $\Sigma_{(a,b)}(A) < \infty$. Then*

(i) *If B is another bounded selfadjoint operator with* $\operatorname{rank}(B-A) = r < \infty$, *then*

$$\Sigma_{(a,b)}(B) \leq \Sigma_{(a,b)}(A) + r\left(\frac{b-a}{2}\right)^{1/2} \tag{3.19}$$

(ii) *If P is an orthogonal projection so that* $\operatorname{rank}(PA(1-P)) = r < \infty$ *and* $B = PAP \restriction \operatorname{ran}(P)$, *then* (3.19) *holds.*

Proof. For simplicity of notation, we can suppose A has both a and b as limit points of eigenvalues (from above and below, respectively). It is easy to modify the arguments if there are only finitely many eigenvalues.

(i) By induction, it suffices to prove this for $r = 1$. Label the eigenvalues of A in (a, b), counting multiplicity, by

$$a < \cdots \leq x_{-2}(A) \leq x_{-1}(A) < \tfrac{1}{2}\,(a+b) \leq x_0(A) \leq x_1(A) \leq \cdots < b \tag{3.20}$$

For A's with a cyclic vector φ, and $B = A + \lambda(\varphi,\,\cdot\,)\varphi$, it is well known that eigenvalues of A and B strictly interlace. By writing A as a direct sum of its restriction to the cyclic subspace for φ and the restriction to the orthogonal complement, we can label all the eigenvalues of B in such a way that

$$x_k(A) \leq x_{k+1}(B) \leq x_{k+1}(A) \tag{3.21}$$

With that labeling,

$$\sum_{k=1}^{\infty} \text{dist}(x_k(B), \mathbb{R} \setminus (a,b))^{1/2} \leq \sum_{k=0}^{\infty} \text{dist}(x_k(A), \mathbb{R} \setminus (a,b))^{1/2} \tag{3.22}$$

$$\sum_{k=1}^{\infty} \text{dist}(x_{-k}(B), \mathbb{R} \setminus (a,b))^{1/2} \leq \sum_{k=1}^{\infty} \text{dist}(x_{-k}(A), \mathbb{R} \setminus (a,b))^{1/2} \tag{3.23}$$

so that

$$\Sigma_{(a,b)}(B) \leq \text{dist}(x_0(B), \mathbb{R} \setminus (a,b))^{1/2} + \Sigma_{(a,b)}(A) \tag{3.24}$$

which implies (3.19) for $r = 1$.

(ii) By scaling and adding a constant to A, we can suppose $b = -a = 1$. For $C \geq 0$ with $\sigma_{\text{ess}}(C) \subset [1, \|C\|]$, let

$$\widetilde{\Sigma}(C) = \sum_{x \in \sigma(C) \cap [0,1)} \bigl(1 - \sqrt{x}\bigr)^{1/2} \tag{3.25}$$

so that

$$\Sigma_{(-1,1)}(A) = \widetilde{\Sigma}(A^2) \tag{3.26}$$

By mimicking the proof of (i), we see

$$\text{rank}(D - C) = r,\ D \geq 0 \;\Rightarrow\; \widetilde{\Sigma}(D) \leq \widetilde{\Sigma}(C) + r \tag{3.27}$$

Notice, next, that by the min-max principle, $x_k(PCP \restriction \text{ran}(P)) \geq x_k(C)$ so that

$$\widetilde{\Sigma}(PCP \restriction \text{ran}(P)) \leq \widetilde{\Sigma}(C) \tag{3.28}$$

Notice also that

$$PA^2P - (PAP)^2 = PA(1-P)AP \tag{3.29}$$

is at most rank r. Thus,

$$\begin{aligned} \Sigma_{(-1,1)}(PAP \restriction \text{ran}(P)) &= \widetilde{\Sigma}((PAP \restriction \text{ran}(P))^2) && \text{(by (3.26))} \\ &\leq r + \widetilde{\Sigma}(PA^2P \restriction \text{ran}(P)) && \text{(by (3.27))} \end{aligned}$$

$$\begin{aligned} &\leq r + \widetilde{\Sigma}(A^2) && \text{(by (3.28))} \\ &= r + \Sigma_{(-1,1)}(A) && \text{(by (3.26))} \end{aligned}$$

□

We also want to know that one can make the eigenvalue sum small, uniformly in B, by summing only over eigenvalues sufficiently near a or b. Thus, we prove (for simplicity, we state the result for a; a similar result holds for b):

Theorem 3.6. *Let* $(a,b) \cap \sigma_{\text{ess}}(A) = \emptyset$, $\Sigma_{(a,b)}(A) < \infty$, *and suppose* B *is related to* A *as in either* (i) *or* (ii) *of Theorem 3.5. Then for any* $\delta < \frac{1}{4}(b-a)$,

$$\sum_{x_k(B)\in(a,a+\delta)} (x_k(B)-a)^{1/2} \leq r\delta^{1/2} + \sum_{x_k(A)\in(a,a+2\delta)} (x_k(A)-a)^{1/2} \tag{3.30}$$

Proof. We have

$$\begin{aligned} \text{LHS of (3.30)} &\leq \Sigma_{(a,a+2\delta)}(B) \\ &\leq \Sigma_{(a,a+2\delta)}(A) + r\delta^{1/2} && \text{(by Theorem 3.5)} \\ &= \text{RHS of (3.30)} \end{aligned}$$

□

As a corollary, we have (since $J^{(n)} = P_n J P_n \restriction \text{ran}(P_n)$ with $\text{rank}((1-P_n)JP_n) = 1$):

Theorem 3.7. *Let* J *be a Jacobi matrix with* $\sigma_{\text{ess}}(J) = \mathfrak{e}$. *Given* (3.17), *let* $\mathcal{E}(J)$ *be finite and let* $J^{(n)}$ *be the* n*-times stripped Jacobi matrix. Then*

(i)

$$\mathcal{E}(J^{(n)}) \leq \mathcal{E}(J) + \ell \max_{j=1,\dots,\ell} \bigl(\tfrac{1}{2}\,|\alpha_{j+1}-\beta_j|\bigr)^{1/2} \tag{3.31}$$

(ii) *For any* $j \in \{1,\dots,\ell+1\}$ *and* $\varepsilon > 0$, *there is a* $\delta > 0$ *so that for all* n,

$$\sum_{x_k(J^{(n)})\in(\beta_j,\beta_j+\delta)} (x_k(J^{(n)})-\beta_j)^{1/2} \leq \tfrac{1}{2}\,\varepsilon \tag{3.32}$$

$$\sum_{x_k(J^{(n)})\in(\alpha_j-\delta,\alpha_j)} (\alpha_j - x_k(J^{(n)}))^{1/2} \leq \tfrac{1}{2}\,\varepsilon \tag{3.33}$$

Proof. (i) By the min-max principle for eigenvalues above and below the essential spectrum, the sums for eigenvalues below α_1 or above $\beta_{\ell+1}$ get smaller. In each gap, we use Theorem 3.5 (ii). This yields (3.31) as $r = 1$.

(ii) We prove (3.32); the proof of (3.33) is similar. Take $\delta_0 < \frac{1}{4}(\alpha_{j+1} - \beta_j)$ so that

$$\sum_{x_k(J)\in(\beta_j,\beta_j+2\delta_0)} (x_k(J)-\beta_j)^{1/2} < \tfrac{1}{4}\,\varepsilon \tag{3.34}$$

Then pick $\delta < \delta_0$ so that $\delta^{1/2} < \frac{1}{4}\varepsilon$. (3.30) implies (3.32). $\square$

Theorem 3.8. *Let J, $\tilde J$ be two Jacobi matrices with $\sigma_{\text{ess}}(J) = \sigma_{\text{ess}}(\tilde J) = \mathfrak{e}$ and $\mathcal{E}(J)$, $\mathcal{E}(\tilde J) < \infty$. For $m, q \geq 0$, let $J_{m,q}$ be the Jacobi matrix with*

$$a_n(J_{m,q}) = \begin{cases} a_n(J) & n = 1, \dots, m \\ a_{n-m+q}(\tilde J) & n = m+1, \dots \end{cases} \tag{3.35}$$

$$b_n(J_{m,q}) = \begin{cases} b_n(J) & n = 1, \dots, m \\ b_{n-m+q}(\tilde J) & n = m+1, \dots \end{cases} \tag{3.36}$$

Then for a constant, K, independent of m and q,

$$\mathcal{E}(J_{m,q}) \leq \mathcal{E}(J) + \mathcal{E}(\tilde J) + K \tag{3.37}$$

and for any $j \in \{1, \dots, \ell+1\}$ and $\varepsilon > 0$, there is a $\delta > 0$ so that for all m, q,

$$\sum_{x_k(J_{m,q})\in(\beta_j,\beta_j+\delta)} (x_k(J_{m,q})-\beta_j)^{1/2} < \tfrac{1}{2}\,\varepsilon \tag{3.38}$$

A similar result holds near α_j.

Proof. Let Q_m be the projection onto $\operatorname{span}\{\delta_j\}_{j=1}^m$ and $P_m = 1 - Q_m$. Then $J_{m,q} - Q_m J Q_m - P_m \tilde J^{(q)} P_m$ is rank two. Thus, for $j = 1, \dots, \ell$ and $\gamma = \max_{j=1,\dots,\ell}(\frac{1}{2}|\alpha_{j+1} - \beta_j|)^{1/2}$,

$$\begin{aligned}
\Sigma_{(\beta_j,\alpha_{j+1})}(J_{m,q}) &\leq 2\gamma + \Sigma_{(\beta_j,\alpha_{j+1})}(Q_m J Q_m) + \Sigma_{(\beta_j,\alpha_{j+1})}(P_m \tilde J^{(q)} P_m) \\
&\leq 4\gamma + \Sigma_{(\beta_j,\alpha_{j+1})}(J) + \Sigma_{(\beta_j,\alpha_{j+1})}(\tilde J^{(q)}) \\
&\leq 5\gamma + \Sigma_{(\beta_j,\alpha_{j+1})}(J) + \Sigma_{(\beta_j,\alpha_{j+1})}(\tilde J)
\end{aligned}$$

For eigenvalues below α_1 (or above $\beta_{\ell+1}$), we use the fact that $|a_n(J_{m,q})| \leq \|J\|$ to see that $\|J_{m,q}\| \leq 2\|J\| + \|\tilde J\|$ (a crude over-estimate). Hence we can do a similar bound on some $\Sigma_{(\kappa,\alpha_1)}(J_{m,q})$ with κ independent of m and q.

The passage from the proof of (3.37) to the proof of (3.38) is similar to the argument in the proof of Theorem 3.7. $\square$

It is a well-known phenomenon that, under strong limits, spectrum can get lost (e.g., if J_n is a Jacobi matrix which is the free J_0, except that for $m \in (n^2 - n, n^2 + n)$, $b_m = -2$, then $J_n \xrightarrow{s} J_0$ but J_n has more and more eigenvalues in $(-4,-2)$). We are going to be interested in situations where this doesn't happen, which is the last subject we consider in this section.

Theorem 3.9. *Let J be a Jacobi matrix with $\sigma_{\text{ess}}(J)=\mathfrak{e}$. Suppose that $J^{(n_k)} \to \tilde{J}$ in the sense that for each $m \geq 1$,*

$$a_{n_k+m} \to \tilde{a}_m \qquad b_{n_k+m} \to \tilde{b}_m \tag{3.39}$$

Then $\tilde{J}$ has at most one eigenvalue in (β_j,α_{j+1}), and for each δ small and n_k large, $J^{(n_k)}$ has the same number of eigenvalues in $(\beta_j+\delta,\alpha_{j+1}-\delta)$ as $\tilde{J}$. In fact, if $\tilde{J}$ has an eigenvalue $\tilde{\lambda}$ there, the eigenvalue of $J^{(n_k)}$ in that interval converges to $\tilde{\lambda}$.

Proof. If $\tilde{\lambda}$ is an eigenvalue of $\tilde{J}$ in (β_j,α_{j+1}) with $\tilde{J}\tilde{u}=\tilde{\lambda}\tilde{u}$ (and $\|\tilde{u}\|=1$), then $\varepsilon_{n_k} \equiv \|(J^{(n_k)}-\tilde{\lambda})\tilde{u}\| \to 0$. Thus, $(\tilde{\lambda}-\varepsilon_{n_k},\tilde{\lambda}+\varepsilon_{n_k})\cap\sigma(J^{(n_k)}) \neq \emptyset$. Since the interval for small enough ε_{n_k} is disjoint from $\sigma_{\text{ess}}(J^{(n_k)})$, we conclude that there is at least one eigenvalue λ_{n_k} in the interval, and clearly, $\lambda_{n_k} \to \tilde{\lambda}$.

This fact plus Theorem 3.4 implies that $\tilde{J}$ has at most one eigenvalue in (β_j,α_{j+1}).

Suppose next that $J^{(n_k)}u_{n_k}=\lambda_{n_k}u_{n_k}$ with $\|u_{n_k}\|=1$ and $\lambda_{n_k} \to \tilde{\lambda} \in (\beta_j,\alpha_{j+1})$. Given $v \in \ell^2(\mathbb{N})$ and n_k, define

$$(v^{(n_k)})_m = \begin{cases} 0 & m \leq n_k \\ v_{m-n_k} & m > n_k \end{cases} \tag{3.40}$$

Then

$$\left[Jv^{(n_k)} - (J^{(n_k)}v)^{(n_k)} \right]_m = \begin{cases} 0 & m \neq n_k \\ a_{n_k}v_1 & m = n_k \end{cases} \tag{3.41}$$

We conclude that

$$\|(J-\lambda_{n_k})u_{n_k}^{(n_k)}\| = a_{n_k}|(u_{n_k})_1| \tag{3.42}$$

If $(u_{n_k})_1 \to 0$, this implies $\tilde{\lambda} \in \sigma_{\text{ess}}(J)$ since $u_{n_k}^{(n_k)} \xrightarrow{w} 0$. But that is impossible, so $(u_{n_k})_1 \nrightarrow 0$. By compactness of the unit ball in the weak topology, we conclude u_{n_k} has a weak limit point $\tilde{u}$ with $(\tilde{u})_1 \neq 0$, so $\tilde{u} \not\equiv 0$. But $(\tilde{J}-\tilde{\lambda})\tilde{u}=0$, so $\tilde{\lambda} \in \sigma(\tilde{J})$.

We have thus proven the final sentence in the theorem, given Theorem 3.4, which says $J^{(n_k)}$ for k large has at most one eigenvalue in $(\beta_j+\delta,\alpha_{j+1}-\delta)$. □

The final theorem of the section deals with a specialized situation that we'll need later.

Theorem 3.10. *Let J be a Jacobi matrix with $\sigma_{\rm ess}(J)=\mathfrak{e}$. Suppose that, as $n_k\to\infty$, (3.39) holds for some two-sided $\tilde J$ and all $m\in\mathbb{Z}$. Let J_k be defined by*

$$a_n(J_k)=\begin{cases} a_m & m\le n_k \\ \tilde a_{m-n_k} & m>n_k \end{cases} \tag{3.43}$$

$$b_m(J_k)=\begin{cases} b_m & m\le n_k \\ \tilde b_{m-n_k} & m>n_k \end{cases} \tag{3.44}$$

Then for any $\delta>0$, with $\{\beta_j+\delta,\alpha_{j+1}-\delta\}\notin\sigma(J)$, all the eigenvalues of J_k in $(\beta_j+\delta,\alpha_{j+1}-\delta)$ for k large are near eigenvalues of J in that interval, and these eigenvalues converge to those for J. Moreover, there is exactly one eigenvalue of J_k near a single eigenvalue of J in that interval.

Proof. We follow the first part of the proof of the last theorem until the analysis of $J_k u_k=\lambda_k u_k$ with $\lambda_k\to\lambda_\infty\in(\beta_j+\delta,\alpha_j-\delta)$. If we prove that $\lambda_\infty\in\sigma(J)$ and u_k converges in norm to the corresponding eigenvector, we are done. For we immediately get existence of eigenvalues near λ_∞, and uniqueness follows from the orthogonality of eigenvectors and the norm convergence.

Define $\tilde u_k\in\ell^2(\mathbb{Z})$ by

$$(\tilde u_k)_m=\begin{cases} (u_k)_{m+n_k} & m>-n_k \\ 0 & m\le -n_k \end{cases} \tag{3.45}$$

and suppose $\tilde u_k$ has a nonzero weak limit $\tilde u_\infty$. Then $(\tilde J-\lambda_\infty)\tilde u_\infty=0$, so $\lambda_\infty\in\sigma(\tilde J)$. As $\sigma(\tilde J)\subset\sigma_{\rm ess}(J)=\mathfrak{e}$ by approximate eigenvector arguments (see, e.g., [12]), we arrive at a contradiction. Thus, $\tilde u_k$ converges weakly to zero. This implies that its projection $P\tilde u_k$ onto $\ell^2(\mathbb{N})$ converges to zero in norm since otherwise $\|(\tilde J-\lambda_\infty)P\tilde u_k\|\to 0$ which is again impossible because $\lambda_\infty\notin\sigma(\tilde J)$.

Therefore, we conclude that $\|(J-\lambda_\infty)u_k\|\to 0$. Since λ_∞ is a simple discrete point of $\sigma(J)$, this can only happen if λ_∞ is an eigenvalue of J and $\|(1-P')u_k\|\to 0$, where P' is the projection onto the eigenvector of λ_∞; that is, u_k converges to that eigenvector in norm. □

4. Szegő's theorem

Our goal in this section is the following. Let $\mathfrak{e}$ be a finite gap set, J a bounded Jacobi matrix with $\sigma_{\rm ess}(J)=\mathfrak{e}$, and $\{a_n,b_n\}_{n=1}^\infty$ its Jacobi parameters. Let $\{x_k\}$ be the eigenvalues of J outside $\mathfrak{e}$, and write

$$d\mu(x)=w(x)\,dx+d\mu_{\rm s}(x) \tag{4.1}$$

where $d\mu$ is the spectral measure for J.

Next, define

$$A_n = \frac{a_1 \cdots a_n}{\text{cap}(\mathfrak{e})^n} \qquad \bar{A} = \limsup A_n \qquad \underline{A} = \liminf A_n \tag{4.2}$$

Consider the three conditions:

(i) Szegő condition

$$\int_{\mathfrak{e}} \log(w(x)) \text{dist}(x, \mathbb{R} \setminus \mathfrak{e})^{-1/2}\, dx > -\infty \tag{4.3}$$

(ii) Blaschke condition

$$\mathcal{E}(J) = \sum_k \text{dist}(x_k, \mathfrak{e})^{1/2} < \infty \tag{4.4}$$

(iii) Widom condition

$$0 < \underline{A} \leq \bar{A} < \infty \tag{4.5}$$

Theorem 4.1. *Any two of* (i)–(iii) *imply the third.*

Remarks. 1. We'll eventually prove more; for example, if (ii) holds, then (i) $\Leftrightarrow$ $\bar{A} > 0$; and if either holds, then (iii) holds.

2. This is a precise analog of a result for $\mathfrak{e} = [-2, 2]$ of Simon–Zlatoš [27] (cf. Theorem 1.1) who relied in part on Killip–Simon [11] and Simon [21].

3. For $\mathfrak{e} = [-2, 2]$, the relevance of (4.4) to Szegő-type theorems is a discovery of Killip–Simon [11] and Peherstorfer–Yuditskii [15].

4. When there are no eigenvalues, the implication (i) $\Rightarrow$ (iii) is a result of Widom [33]; see also Aptekarev [1]. Peherstorfer–Yuditskii [16] allowed infinitely many bound states, and in [17], they proved (i) $\Rightarrow$ (iii) if (ii) holds. The other parts of Theorem 4.1 are new, although as noted to us by Peherstorfer and Yuditskii [14], there is an argument to go from [16, 17] to (iii) $\Rightarrow$ (i) if (ii) holds (see Remark 3 following Theorem 4.5 below).

Recall that, given any pair of Baire measures, $d\mu$, $d\nu$, on a compact Hausdorff space, we define their relative entropy by

$$S(\mu \mid \nu) = \begin{cases} -\infty & \text{if } d\mu \text{ is not } d\nu\text{-a.c.} \\ -\int \log(\frac{d\mu}{d\nu})\, d\mu & \text{if } d\mu \text{ is } d\nu\text{-a.c.} \end{cases} \tag{4.6}$$

It is a fundamental fact (see, e.g., [23, Thm. 2.3.4]) that $S(\mu \mid \nu)$ is jointly concave and jointly weakly upper semicontinuous in $d\mu$ and $d\nu$, and that

$$\mu(X) = \nu(X) = 1 \Rightarrow S(\mu \mid \nu) \leq 0 \tag{4.7}$$

S is relevant because we define

$$Z(J) = -\tfrac{1}{2}\, S(\rho_{\mathfrak{e}} \mid \mu_J) \tag{4.8}$$

with $d\mu_J$ the spectral measure of J and $d\rho_{\mathfrak{e}}$ the potential theoretic equilibrium measure for $\mathfrak{e}$. Then, by (4.7),

$$Z(J) \geq 0 \tag{4.9}$$

More importantly,

$$(4.3) \;\Leftrightarrow\; Z(J) < \infty \tag{4.10}$$

We have (4.10) because (see eqn. (4.31) and Theorem 4.4 of paper I) $d\rho_{\mathfrak{e}}$ is $dx \restriction \mathfrak{e}$ a.c. and

$$C_1 \operatorname{dist}(x, \mathbb{R} \setminus \mathfrak{e})^{-1/2} \leq \frac{d\rho_{\mathfrak{e}}}{dx} \leq C_2 \operatorname{dist}(x, \mathbb{R} \setminus \mathfrak{e})^{-1/2} \tag{4.11}$$

for $0 < C_1 < C_2 < \infty$.

Given the connection (1.24) between Blaschke products and $G_{\mathfrak{e}}$, the potential theoretic Green's function for $\mathfrak{e}$, and the symmetry of Blaschke products (eqn. (4.19) of paper I), one can rewrite the step-by-step C_0 sum rule, Theorem 2.2, as

Theorem 4.2. *For each n, $Z(J) < \infty \Leftrightarrow Z(J^{(n)}) < \infty$, and in that case,*

$$\frac{a_1 \cdots a_n}{\operatorname{cap}(\mathfrak{e})^n} = K_n \exp[Z(J^{(n)}) - Z(J)] \tag{4.12}$$

where

$$K_n = \exp\biggl(\sum_k [G_{\mathfrak{e}}(x_k(J)) - G_{\mathfrak{e}}(x_k(J^{(n)}))]\biggr) \tag{4.13}$$

Remark. By Theorem 3.1, and the monotonicity of $G_{\mathfrak{e}}$ near gap edges (eqns. (4.45) and (4.46) of paper I), the sum in (4.13) is always conditionally convergent if ordered properly.

Proof. By iterating, it suffices to prove the result for $n = 1$. As noted, K_1 is always finite and the remarks before the statement of the theorem show that for $n = 1$, $K_1 = B_\infty(0)$. Thus, the step-by-step C_0 sum rule says

$$\frac{a_1}{\operatorname{cap}(\mathfrak{e})} = K_1 \exp\biggl(\frac{1}{2}\int_0^{2\pi} \log\biggl(\frac{\operatorname{Im} M(e^{i\theta})}{\operatorname{Im} M^{(1)}(e^{i\theta})}\biggr)\frac{d\theta}{2\pi}\biggr) \tag{4.14}$$

Since M and so $\operatorname{Im} M$ is automorphic, Corollary 4.6 of paper I implies

$$\int_0^{2\pi} \log\biggl(\frac{\operatorname{Im} M(e^{i\theta})}{\operatorname{Im} M^{(1)}(e^{i\theta})}\biggr)\frac{d\theta}{2\pi} = \int_{\mathfrak{e}} \log\biggl(\frac{w(x;J)}{w(x;J^{(1)})}\biggr) d\rho_{\mathfrak{e}}(x) \tag{4.15}$$

where we use

$$w(x;J)=\frac{1}{\pi}\,\operatorname{Im} m(x+i0,J) \tag{4.16}$$

Thus,

$$\int_{\mathfrak{e}}\log(w(x;J^{(1)}))\,d\rho_{\mathfrak{e}}(x)>-\infty\;\Leftrightarrow\;\int_{\mathfrak{e}}\log(w(x;J))\,d\rho_{\mathfrak{e}}(x)>-\infty \tag{4.17}$$

showing $Z(J^{(1)})<\infty\Leftrightarrow Z(J)<\infty$. Moreover, if both are finite,

$$\text{RHS of (4.15)}=2Z(J^{(1)})-2Z(J) \tag{4.18}$$

(4.14)–(4.18) imply (4.12). □

Proposition 4.3. *We have that*

$$K_n\le A_n e^{Z(J)} \tag{4.19}$$

In particular, for some constant C_1*,*

$$\underline{A}(J)\ge e^{-Z(J)}\liminf[\exp(-C_1\mathcal{E}(J^{(n)}))] \tag{4.20}$$

and

$$\limsup K_n\le \bar{A}(J)e^{Z(J)} \tag{4.21}$$

Proof. (4.19) is immediate from (4.12) if we note that $Z(J^{(n)})\ge 0$ so that $\exp(-Z(J^{(n)}))\le 1$. (4.20) follows from noting that $K_n\ge \exp(-\sum_k G_{\mathfrak{e}}(x_k(J^{(n)})))$ since $G_{\mathfrak{e}}(x_k(J))\ge 0$ and then, that for some C_1 (depending only on $\mathfrak{e}$),

$$G_{\mathfrak{e}}(x)\le C_1\operatorname{dist}(x,\mathfrak{e})^{1/2} \tag{4.22}$$

by Theorem 4.4 of paper I. Finally, (4.21) is immediate by taking $\limsup$ in (4.19). □

Proposition 4.4. *Let* $J_{\mathfrak{e}}$ *be the Jacobi matrix with spectral measure* $d\rho_{\mathfrak{e}}$ *and let* $\{a_n^{(\mathfrak{e})},b_n^{(\mathfrak{e})}\}_{n=1}^\infty$ *be its Jacobi parameters. Let* J_n *be the Jacobi matrix with parameters*

$$a_m(J_n)=\begin{cases}a_m & m=1,\dots,n\\ a_{m-n}^{(\mathfrak{e})} & m>n\end{cases} \tag{4.23}$$

$$b_m(J_n)=\begin{cases}b_m & m=1,\dots,n\\ b_{m-n}^{(\mathfrak{e})} & m>n\end{cases} \tag{4.24}$$

Then

$$A_n(J)=\exp\biggl(\sum_k G_{\mathfrak{e}}(x_k(J_n))\biggr)\exp(-Z(J_n)) \tag{4.25}$$

In particular, for some C_1 *(depending only on* $\mathfrak{e}$*),*

$$A_n(J)\le\exp(C_1\mathcal{E}(J_n)-Z(J_n)) \tag{4.26}$$

Proof. J_n is defined so that

$$(J_n)^{(n)} = J_{\mathfrak{e}} \tag{4.27}$$

and

$$A_n(J_n) = A_n(J) \tag{4.28}$$

Thus, since $Z(J_{\mathfrak{e}}) = 0$ and $J_{\mathfrak{e}}$ has no eigenvalues outside $\mathfrak{e}$, (4.12) for J_n is (4.25). (4.26) is then immediate from (4.22). □

Theorem 4.5. *If* $\mathcal{E}(J) < \infty$, *then*

$$\bar{A}(J) > 0 \Leftrightarrow Z(J) < \infty \tag{4.29}$$

and if these are true, the Widom condition holds:

$$0 < \underline{A}(J) \le \bar{A}(J) < \infty \tag{4.30}$$

Proof. By (4.20) and Theorem 3.7,

$$\mathcal{E}(J), Z(J) < \infty \Rightarrow \underline{A}(J) > 0 \Rightarrow \bar{A}(J) > 0 \tag{4.31}$$

By (4.26) and Theorem 3.8, going through a subsequence with $A_{n_j}(J) \to \bar{A}(J)$, we see that

$$\mathcal{E}(J) < \infty,\ \bar{A}(J) > 0 \Rightarrow \limsup[\exp(-Z(J_{n_j}))] > 0 \tag{4.32}$$

Thus, for some subsequence,

$$\liminf Z(J_{n_j}) < \infty \tag{4.33}$$

Since $J_{n_j} \xrightarrow{s} J$, the spectral measures converge weakly. Since S is upper semicontinuous, $Z = -\frac{1}{2}S$ is lower semicontinuous, and thus,

$$Z(J) \le \liminf Z(J_{n_j}) \tag{4.34}$$

so (4.33) implies $Z(J) < \infty$. That is, we have proven

$$\mathcal{E}(J) < \infty,\ \bar{A}(J) > 0 \Rightarrow Z(J) < \infty \tag{4.35}$$

If we have $Z(J) < \infty$ and $\mathcal{E}(J) < \infty$, we get $\underline{A}(J) > 0$ by (4.31), and since $Z(J_n) \ge 0$, (4.26) implies

$$\bar{A}(J) \le \limsup[\exp(C_1\mathcal{E}(J_n))] < \infty \tag{4.36}$$

by Theorem 3.8. □

Remarks. 1. The above proof shows that even without $Z(J) < \infty$, we have $\mathcal{E}(J) < \infty \Rightarrow \bar{A}(J) < \infty$.

2. The proof borrows heavily from ideas of Killip–Simon [11] and Simon–Zlatoš [27].

3. As noted, $\mathcal{E}(J), Z(J) < \infty \Rightarrow$ (4.30) is a prior result (using variational methods) of Peherstorfer–Yuditskii [16, 17]. Peherstorfer and Yuditskii [14] have pointed out that their results can be used to

prove $\mathcal{E}(J) < \infty$, $\bar{A}(J) > 0 \Rightarrow Z(J) < \infty$ by the following argument: While it is not explicitly stated, [16, 17] prove that for any K, there is a constant C so that for all measures with $Z(J) < \infty$ and $\mathcal{E}(J) \le K$,

$$\limsup_{n\to\infty} \frac{a_1 \cdots a_n}{\text{cap}(\mathfrak{e})^n} \le C e^{-Z(J)} \tag{4.37}$$

Given $d\mu$ with $Z(J) = \infty$ and $\mathcal{E}(J) \le K$, let $d\tilde{\mu}_\varepsilon$ be the measure $d\mu + \varepsilon\, dx \restriction \mathfrak{e}$. Then with $d\mu_\varepsilon$ the normalized measure and $a_n(\varepsilon)$ the corresponding a's, (4.37) implies (since $Z(J_\varepsilon) < \infty$)

$$\limsup \frac{a_1(\varepsilon) \cdots a_n(\varepsilon)}{\text{cap}(\mathfrak{e})^n} \le C e^{-Z(J_\varepsilon)} \tag{4.38}$$

By the variational principle for $a_1 \cdots a_n = \|P_n\|$, we have

$$a_1 \cdots a_n \le [a_1(\varepsilon) \cdots a_n(\varepsilon)](1 + \varepsilon|\mathfrak{e}|)^{1/2} \tag{4.39}$$

Since $Z(J_\varepsilon) - \frac{1}{2}\log(1+\varepsilon|\mathfrak{e}|) \uparrow Z(J)$, (4.38)–(4.39) imply that $\bar{A}(J) = 0$ if $Z(J) = \infty$. This argument for the classical Szegő case is in Garnett [8].

Theorem 4.6. $\bar{A}(J), Z(J) < \infty \Rightarrow \mathcal{E}(J) < \infty$

Proof. This is immediate from (4.21) and Theorem 3.2. □

Remark. This argument follows ideas of Simon–Zlatoš [27].

Theorems 4.5 and 4.6 imply Theorem 4.1.

5. Jost Functions and Jost Solutions

In Section 8 of paper I, we defined the Szegő class for $\mathfrak{e}$, which we'll denote $\text{Sz}(\mathfrak{e})$, to be the set of probability measures, $d\mu$, of the form (4.1) that obey (4.3) and (4.4). As usual, we associate $d\mu$ with its Jacobi matrix and Jacobi parameters $\{a_n, b_n\}_{n=1}^\infty$, which we will write as $\{a_n(\mu), b_n(\mu)\}_{n=1}^\infty$ if we need to be explicit about the measures. Of course, the a's obey the Widom condition (4.5) for all measures in the Szegő class.

In this section, we want to recall the definitions of Jost function and Jost solution from Sections 8 and 9 of paper I, extend some results on Jost solutions to the full Szegő class, and state the main theorem that we'll prove in the next section about their asymptotics.

Jost functions require a reference measure, and we'll use the one from paper I. Let $\tilde{\zeta}_j \in \widetilde{C}_j^+$, the full orthocircle, be the point farthest from 0 on $\widetilde{C}_j^+$ and let $w_j \in \mathcal{S}$, the Riemann surface for $\mathfrak{e}$, be given by $w_j = \mathbf{x}^\sharp(\tilde{\zeta}_j)$. Each w_j lies in $G_j = \pi^{-1}([\beta_j, \alpha_{j+1}])$, so $\vec{w} = (w_1, \dots, w_\ell) \in \mathbb{G} = G_1 \times \cdots \times G_\ell$, which can be associated with the isospectral torus.

Our reference measure is the measure in $\mathcal{T}_\mathfrak{e}$ associated to $\vec{w}$. We denote it by

$$d\nu_\mathfrak{e}(x) = v_\mathfrak{e}(x)\,dx \tag{5.1}$$

We point out that while our choice of the reference measure is convenient, one can take any other measure in the Szegő class to be the reference measure.

Given $d\mu \in \mathrm{Sz}(\mathfrak{e})$, let $\{x_k\}$ be the eigenvalues of J in $\mathbb{R}\setminus\mathfrak{e}$ and define $z_k \in \mathcal{F}$ by

$$\mathbf{x}(z_k) = x_k \tag{5.2}$$

The Jost function is then defined on $\mathbb{D}$ by

$$u(z;\mu) = \prod_k B(z,z_k)\exp\biggl(\frac{1}{4\pi}\int_0^{2\pi} \frac{e^{i\theta}+z}{e^{i\theta}-z}\,\log\biggl(\frac{v_\mathfrak{e}(\mathbf{x}(e^{i\theta}))}{w(\mathbf{x}(e^{i\theta}))}\biggr)d\theta\biggr) \tag{5.3}$$

Since (4.16) implies

$$\frac{v_\mathfrak{e}(\mathbf{x}(e^{i\theta}))}{w(\mathbf{x}(e^{i\theta}))} = \frac{\operatorname{Im} M_{\nu_\mathfrak{e}}(e^{i\theta})}{\operatorname{Im} M_\mu(e^{i\theta})} \tag{5.4}$$

we could use that ratio instead. By the Blaschke condition and Proposition 4.8 of paper I, the product in (5.3) (which we'll call the Blaschke part) converges. By eqn. (4.54) of paper I and the Szegő condition for $d\mu$ and $d\nu_\mathfrak{e}$, the log in (5.3) is in $L^1(\partial\mathbb{D}, d\theta/2\pi)$. We call the exponential in (5.3) the Szegő part. As proven in Theorem 8.2 of paper I, u is a character automorphic function on $\mathbb{D}$.

For any Jacobi matrix, J, with $\sigma_{\mathrm{ess}}(J) = \mathfrak{e}$, we let $M^{(n)}$ be the m-function (1.26) of the n-times stripped Jacobi matrix, $J^{(n)}$, and define the *Weyl solution* by

$$W_n(z) = M(z)(a_1 M^{(1)}(z))\cdots(a_{n-1}M^{(n-1)}(z)) \tag{5.5}$$

$M^{(k)}$ has poles at the inverse images of eigenvalues of $J^{(k)}$ and zeros at the inverse images of eigenvalues of $J^{(k+1)}$, so there is a cancellation, and W_n can be defined as meromorphic on $\mathbb{D}$ with poles exactly at the points ζ with $\mathbf{x}(\zeta)$ an eigenvalue of J.

The name, Weyl solution, comes from the fact that because m is a ratio of solutions L^2 at $n = +\infty$, W_n obeys

$$W_n(z) = -\langle \delta_n, (J-\mathbf{x}(z))^{-1}\delta_1\rangle \tag{5.6}$$

so that for $k \geq 2$,

$$[(J-\mathbf{x}(z))W.(z)]_k = 0 \tag{5.7}$$

where $W.(z)$ is the vector $(W_1(z), W_2(z), \dots)$. That is,

$$a_n W_n(z) + b_{n+1}W_{n+1}(z) + a_{n+1}W_{n+2}(z) = \mathbf{x}(z)W_{n+1}(z) \tag{5.8}$$

for $n = 1, 2, \dots$.

The Jost solution is defined by

$$u_n(z;\mu) = u(z;\mu)W_n(z) \tag{5.9}$$

Since $u(z;\mu)$ is n-independent, (5.8) holds for u_n also. Since u has zeros at the points where M, and so W_n, has poles, u_n is analytic on $\mathbb{D}$.

Theorem 5.1.

$$a_n M^{(n-1)}(z) = B(z)\,\frac{u(z;\mu_n)}{u(z;\mu_{n-1})} \tag{5.10}$$

where $M^{(0)} = M$, $d\mu_0 = d\mu$, and $d\mu_n$, $M^{(n)}$ are associated to $J^{(n)}$, the n-times stripped Jacobi matrix.

Proof. This is a rewrite of (2.3) for $J^{(n-1)}$. □

Theorem 5.2. *Let $d\mu \in \mathrm{Sz}(\mathfrak{e})$. Then*

$$u_n(z;\mu) = a_n^{-1} B(z)^n u(z;\mu_n) \tag{5.11}$$

where $d\mu_n$ is the spectral measure for $J^{(n)}$, the n-times stripped Jacobi matrix.

Proof. By (5.10) and (5.5),

$$a_n W_n(z) = B(z)^n\,\frac{u(z;\mu_n)}{u(z;\mu)} \tag{5.12}$$

which by (5.9) implies (5.11). □

The key asymptotic result of the next section is the following:

Theorem 5.3. *Suppose $d\mu \in \mathrm{Sz}(\mathfrak{e})$ and that for some subsequence $n_j \to \infty$ and all $m \in \mathbb{Z}$,*

$$a_{n_j+m}(J_\mu) \to a_m^\sharp \qquad b_{n_j+m}(J_\mu) \to b_m^\sharp \tag{5.13}$$

for some point $\{a_n^\sharp, b_n^\sharp\}_{n=-\infty}^\infty$ in the isospectral torus. If $d\mu^\sharp$ is the spectral measure for the Jacobi matrix with parameters $\{a_n^\sharp, b_n^\sharp\}_{n=1}^\infty$, then

$$u(z;\mu_{n_j}) \to u(z;\mu^\sharp) \tag{5.14}$$

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

We note, as will be explained in the next section, that there is no loss in supposing that the limit $J^\sharp$ is in the isospectral torus. We'll also show that Theorem 5.3 allows the proof of (1.17) for a point $\tilde{J}$ in the isospectral torus.

6. Jost Asymptotics

In this section, we'll prove Theorem 5.3, use this result to prove that for $d\mu \in \mathrm{Sz}(\mathfrak{e})$, the Jacobi parameters a_n, b_n are asymptotic to a fixed element of $\mathcal{T}_{\mathfrak{e}}$, and prove an asymptotic formula for the Jost solution.

The key to our proof of the existence of an $\{\tilde a_n, \tilde b_n\}_{n=1}^\infty$ obeying (1.17) is the Denisov–Rakhmanov–Remling theorem for $\mathfrak{e}$ ([18]) which implies that any right limit of J lies in the isospectral torus. Tracking the characters of the Jost functions will determine exactly which right limits occur. This leads to a proof quite different from the variational approach of [33, 1, 16].

We write

$$u(z;\mu) = \beta(z;\mu)\varepsilon(z;\mu) \tag{6.1}$$

where β is the Blaschke part and ε the Szegő part. We'll prove (5.14) by proving separately the convergence of the two parts.

Theorem 6.1. *Under the hypotheses of Theorem 5.3, uniformly on compact subsets of* $\mathbb{D}$*,*

$$\beta(z;\mu_{n_j}) \to \beta(z;\mu^\sharp) \tag{6.2}$$

Proof. By Theorem 3.7 of this paper and Proposition 4.8 of paper I (and its proof), given a compact set $K \subset \mathbb{D}$ and $\varepsilon > 0$, we can find $\delta > 0$ so that the product of the contributions to β from x's with $\mathrm{dist}(x,\mathfrak{e}) < \delta$ are within ε of 1 for all $z \in K$. Thus, it suffices to prove convergence of individual x's for μ_{n_j} to those for $\mu^\sharp$, and this follows from Theorem 3.9. □

To control the Szegő part, we first need the following lemma of Simon–Zlatoš [27]:

Theorem 6.2 ([27])**.** *Let X be a compact Hausdorff measure space, $d\rho$, $d\mu_n$, $d\mu_\infty$ probability measures with $d\mu_n \to d\mu_\infty$ weakly, and*

$$d\mu_n = f_n\, d\rho + d\mu_{n;\mathrm{s}} \tag{6.3}$$

Suppose that

$$S(\rho \mid \mu_n) \to S(\rho \mid \mu_\infty) \tag{6.4}$$

with all relative entropies finite. Then

$$\log(f_n)\, d\rho \xrightarrow{w} \log(f_\infty)\, d\rho \tag{6.5}$$

Proof. If h is continuous and strictly positive, by upper semicontinuity,

$$\limsup S(h\rho \mid \mu_n) \le S(h\rho \mid \mu_\infty) \tag{6.6}$$

or

$$\limsup \int \log(f_n h^{-1}) h\, d\rho \le \int \log(f_\infty h^{-1}) h\, d\rho \tag{6.7}$$

so that

$$\limsup \int \log(f_n) h\, d\rho \leq \int \log(f_\infty) h\, d\rho \tag{6.8}$$

For arbitrary continuous real-valued g, let $h = 2\|g\|_\infty \pm g$ to get

$$\lim \int \log(f_n) g\, d\rho = \int \log(f_\infty) g\, d\rho \tag{6.9}$$

□

Proposition 6.3. *To get*

$$\varepsilon(z; \mu_{n_j}) \to \varepsilon(z; \mu^\sharp) \tag{6.10}$$

uniformly for z in compact subsets of $\mathbb{D}$, it suffices to prove that

$$\lim_{j\to\infty} S(\rho_\mathfrak{e} \mid \mu_{n_j}) = S(\rho_\mathfrak{e} \mid \mu^\sharp) \tag{6.11}$$

Proof. By definition of ε, it suffices that as measures on $\partial\mathbb{D}$,

$$\log\Big(\frac{1}{\pi}|\mathrm{Im}\, M_{\mu_{n_j}}(e^{i\theta})|\Big)\frac{d\theta}{2\pi} \xrightarrow{w} \log\Big(\frac{1}{\pi}|\mathrm{Im}\, M_{\mu^\sharp}(e^{i\theta})|\Big)\frac{d\theta}{2\pi}$$

Given $g \in C(\partial\mathbb{D})$, define

$$\tilde g(e^{i\theta}) = \frac{\sum_{\gamma\in\Gamma} g(\gamma(e^{i\theta}))|\gamma'(e^{i\theta})|}{\sum_{\gamma\in\Gamma}|\gamma'(e^{i\theta})|} \tag{6.12}$$

and h on $\mathfrak{e}$ by

$$h(\mathbf{x}(e^{i\theta})) = \tfrac{1}{2}\,[\tilde g(e^{i\theta}) + \tilde g(e^{-i\theta})] \tag{6.13}$$

Note that h is continuous on $\mathfrak{e}$ since $\tilde g$ is continuous on $\partial\mathcal{F} \cap \partial\mathbb{D}$ by eqn. (3.4) of paper I.

By Corollary 4.6 of paper I,

$$\int_0^{2\pi} g(e^{i\theta}) \log\bigg(\frac{1}{\pi}\,|\mathrm{Im}\, M_\mu(e^{i\theta})|\bigg)\frac{d\theta}{2\pi} = \int_\mathfrak{e} h(x)\log(w_\mu(x))\, d\rho_\mathfrak{e}(x) \tag{6.14}$$

so the necessary weak convergence on $\partial\mathbb{D}$ is implied by weak convergence of $\log(f_{n_j})\, d\rho_\mathfrak{e}$ to $\log(f_\infty)\, d\rho_\mathfrak{e}$. This in turn follows from (6.11) and Theorem 6.2. □

Theorem 6.4. *Under the hypotheses of Theorem 5.3, uniformly on compact subsets of $\mathbb{D}$,*

$$\varepsilon(z; \mu_{n_j}) \to \varepsilon(z; \mu^\sharp) \tag{6.15}$$

Proof. By Proposition 6.3, it suffices to prove (6.11). Since $\mu_{n_j} \xrightarrow{w} \mu^\sharp$, upper semicontinuity of S implies

$$\limsup S(\rho_\mathfrak{e} \mid \mu_{n_j}) \leq S(\rho_\mathfrak{e} \mid \mu^\sharp) \tag{6.16}$$

So it suffices to prove that

$$\underline{S} \equiv \liminf S(\rho_{\mathfrak{e}} \mid \mu_{n_j}) \geq S(\rho_{\mathfrak{e}} \mid \mu^\sharp) \tag{6.17}$$

Pick a subsequence (that we'll still denote by n_j) so that $S(\rho_{\mathfrak{e}} \mid \mu_{n_j}) \to \underline{S}$ and so that $\tau_j \to \tau_\infty$ for some $\tau_\infty > 0$, where

$$\tau_j = \frac{a_1 \cdots a_{n_j}}{\operatorname{cap}(\mathfrak{e})^{n_j}} \tag{6.18}$$

Note that by Theorem 4.1 and $d\mu \in \operatorname{Sz}(\mathfrak{e})$, the original τ_j's are bounded, so we can pick such a convergent subsequence.

For $k < \ell$, let $J_{k,\ell}$ be the Jacobi matrix obtained by starting with $J^{(n_k)}$ and then putting $J^\sharp$ at sites beyond n_ℓ, that is,

$$a_m(J_{k,\ell}) = \begin{cases} a_{n_k+m} & 1 \leq m \leq n_\ell - n_k \\ a^\sharp_{m-n_\ell+n_k} & m > n_\ell - n_k \end{cases} \tag{6.19}$$

$$b_m(J_{k,\ell}) = \begin{cases} b_{n_k+m} & 1 \leq m \leq n_\ell - n_k \\ b^\sharp_{m-n_\ell+n_k} & m > n_\ell - n_k \end{cases} \tag{6.20}$$

Thus, $(J_{k,\ell})^{(n_\ell - n_k)} = J^\sharp$, so the iterated step-by-step C_0 sum rule says that

$$\frac{\tau_\ell}{\tau_k} = \frac{\beta(0;\mu^\sharp)}{\beta(0;\mu_{k,\ell})} \exp\bigl[\tfrac{1}{2}\, S(\rho_{\mathfrak{e}} \mid \mu_{k,\ell}) - \tfrac{1}{2}\, S(\rho_{\mathfrak{e}} \mid \mu^\sharp)\bigr] \tag{6.21}$$

We claim that

$$\lim_{\ell\to\infty} \beta(0;\mu_{k,\ell}) = \beta(0;\mu_{n_k}) \tag{6.22}$$

Accepting this for now, we take $\ell \to \infty$ in (6.21), using the upper semicontinuity of $S(\rho_{\mathfrak{e}} \mid \mu)$ in μ to get

$$\exp\bigl[\tfrac{1}{2}\, S(\rho_{\mathfrak{e}} \mid \mu_{n_k}) - \tfrac{1}{2}\, S(\rho_{\mathfrak{e}} \mid \mu^\sharp)\bigr] \geq \frac{\tau_\infty}{\tau_k}\, \frac{\beta(0;\mu_{n_k})}{\beta(0;\mu^\sharp)} \tag{6.23}$$

Now take $k \to \infty$ using the assumption that $S(\rho_{\mathfrak{e}} \mid \mu_{n_k}) \to \underline{S}$. Since $\tau_\infty/\tau_k \to 1$ and, by (6.2),

$$\frac{\beta(0;\mu_{n_k})}{\beta(0;\mu^\sharp)} \to 1$$

we get (6.17).

Thus, we need only prove (6.22), which follows the proof of Theorem 6.1, but using Theorems 3.8 and 3.10. □

Proof of Theorem 5.3. By (6.1), this follows from Theorems 6.1 and 6.4. □

We can now prove (1.17).

Theorem 6.5. *Let $d\mu \in \mathrm{Sz}(\mathfrak{e})$. Take $d\tilde{\mu}$ to be the unique element in $\mathcal{T}_{\mathfrak{e}}$ so that $u(z;\mu)$ and $u(z;\tilde{\mu})$ have the same automorphic character. Then*

$$\lim_{n\to\infty} |a_n - \tilde{a}_n| + |b_n - \tilde{b}_n| = 0 \tag{6.24}$$

Remark. The existence and uniqueness of $d\tilde{\mu} \in \mathcal{T}_{\mathfrak{e}}$ follows from Theorem 7.3 of paper I.

Proof. If not, by compactness, there is a right limit $J^\sharp$ so that

$$a_{m+n_j} \to a_m^\sharp \qquad b_{m+n_k} \to b_m^\sharp \tag{6.25}$$

and so that

$$\tilde{a}_{m+n_j} \to a_m^{(\infty)} \qquad \tilde{b}_{m+n_j} \to b_m^{(\infty)} \tag{6.26}$$

with

$$J^\sharp \neq J^{(\infty)} \tag{6.27}$$

By the Denisov–Rakhmanov–Remling theorem [18], $J^\sharp$ and $J^{(\infty)}$ lie in the isospectral torus. Let $\chi_B(\gamma)$ be the automorphic character of $B(z)$. Then with $\chi_J(\gamma)$ the character of the Jost function for J, (5.10) and the fact that $M^{(n-1)}$ is automorphic implies that

$$\chi_{J^{(n)}} = \chi_J \chi_B^{-n} \qquad \chi_{\tilde{J}^{(n)}} = \chi_{\tilde{J}} \chi_B^{-n} \tag{6.28}$$

Since the definition of $\tilde{J}$ is $\chi_{\tilde{J}} = \chi_J$, we see that

$$\chi_{J^{(n)}} = \chi_{\tilde{J}^{(n)}} \tag{6.29}$$

By Theorem 5.3 and the fact that uniform convergence of character automorphic functions implies convergence of their characters, we get

$$\chi_{J^\sharp} = \chi_{J^{(\infty)}} \tag{6.30}$$

But $J^\sharp$ and $J^{(\infty)}$ lie in the isospectral torus, so by Theorem 7.3 of paper I,

$$J^\sharp = J^{(\infty)} \tag{6.31}$$

This contradiction to (6.27) implies that (6.24) holds. □

As a corollary, we get convergence of Jost solutions.

Theorem 6.6. *Uniformly on compact subsets of $\mathbb{D}$,*

$$\frac{u_n(z;\mu) - u_n(z;\tilde{\mu})}{B(z)^n} \to 0 \tag{6.32}$$

Moreover,

$$\frac{u_n(z;\mu)}{u_n(z;\tilde{\mu})} \to 1 \tag{6.33}$$

uniformly on compact subsets of $\mathcal{F}^{\mathrm{int}}$.

Remark. At each point in $\{\gamma(0) \mid \gamma \in \Gamma\}$, u_n and B^n have zeros of order n, so $u_n B^{-n}$ has removable singularities at those points.

Proof. Since $J^{(n)}$ and $\tilde{J}^{(n)}$ (by Theorem 6.5) have the same right limits, by Theorem 5.3,

$$|u(z;\mu_n) - u(z;\tilde{\mu}_n)| \to 0 \tag{6.34}$$

uniformly on $\mathbb{D}$. Since $a_n/\tilde{a}_n \to 1$, (5.11) implies (6.32).

As $u_n(z;\tilde{\mu})$ is bounded away from zero (uniformly in n) on compact subsets of $\mathcal{F}^{\text{int}}$, (6.34) implies (6.33). □

Corollary 6.7. *Let $d\mu \in \text{Sz}(\mathfrak{e})$ and let $d\tilde{\mu} \in \mathcal{T}_\mathfrak{e}$ be the measure for which* (6.24) *holds. Then, as $n \to \infty$,*

$$\frac{a_1 \cdots a_n}{\tilde{a}_1 \cdots \tilde{a}_n} \to \frac{u(0;\tilde{\mu})}{u(0;\mu)} \tag{6.35}$$

In particular, $a_1 \cdots a_n / \operatorname{cap}(\mathfrak{e})^n$ is asymptotically almost periodic.

Proof. The final sentence follows from (6.35) and Corollary 7.4 of paper I. To obtain (6.35), note that (5.10) at $z = 0$ and (2.12) implies

$$\frac{u(0;\mu_n)}{u(0;\mu)} = \frac{a_1 \cdots a_n}{\operatorname{cap}(\mathfrak{e})^n} \tag{6.36}$$

Thus,

$$\frac{a_1 \cdots a_n}{\tilde{a}_1 \cdots \tilde{a}_n} = \frac{u(0;\tilde{\mu})}{u(0;\mu)} \, \frac{u(0;\mu_n)}{u(0;\tilde{\mu}_n)} \tag{6.37}$$

Since $u(0;\nu)$ is bounded away from 0 as $d\nu$ runs through the isospectral torus, (6.34) implies that

$$\frac{u(0;\mu_n)}{u(0;\tilde{\mu}_n)} \to 1$$

proving (6.35). □

7. Szegő Asymptotics

In Section 6, we proved that if u_n is the Jost solution of a J_μ with $d\mu \in \text{Sz}(\mathfrak{e})$ and $\tilde{u}_n$ is the Jost solution for the element of the isospectral torus to which J_μ is asymptotic (in the sense of (1.17)), then, as $n \to \infty$, $u_n(z)/\tilde{u}_n(z) \to 1$ uniformly on compact subsets of $\mathcal{F}^{\text{int}}$. Our goal in this section is to prove that if p_n and $\tilde{p}_n$ are the corresponding orthonormal polynomials, then also on $\mathcal{F}^{\text{int}}$, $p_n(z)/\tilde{p}_n(z)$ has a limit (which will not be identically 1 and which we'll write explicitly in terms of Jost functions).

The passage from Jost asymptotics to Szegő asymptotics in the case $\mathfrak{e} = [-2, 2]$ was studied by Damanik–Simon [5] using constancy of the

Wronskian. Our first approach for general $\mathfrak{e}$ mimicked that of [5] but was awkward because certain objects which were constant in the case $\mathfrak{e} = [-2, 2]$ were instead almost periodic. To overcome this, we found a new approach which, even for $\mathfrak{e} = [-2, 2]$, is somewhat simpler than the approach in [5].

The idea is to exploit the formula for the diagonal Green's function for $x \in \mathbb{C}_+$,

$$G_{nn}(x) = \langle \delta_n, (J - x)^{-1} \delta_n \rangle \tag{7.1}$$

namely (see, e.g., [26]),

$$G_{nn}(x) = \frac{p_{n-1}(x) U_n(x)}{\mathrm{Wr}(x)} \tag{7.2}$$

where $U_n(x)$ is defined by

$$U_n(x) = u_n(\zeta) \qquad \mathbf{x}(\zeta) = x \qquad \zeta \in \mathcal{F}^{\mathrm{int}} \tag{7.3}$$

and $\mathrm{Wr}(x)$ is defined by

$$\mathrm{Wr}(x) = a_m \big(U_{m+1}(x) p_{m-1}(x) - U_m(x) p_m(x) \big) \tag{7.4}$$

for $m \geq 1$. The right-hand side is independent of m. The funny indices in (7.4) compared to Wronskians come from the fact that U_m and $V_m = p_{m-1}$ obey the same difference equation, and RHS of (7.4) is nothing but $a_m(U_{m+1} V_m - U_m V_{m+1})$.

In (7.4), we can also take $m = 0$ if we set $a_0 = 1$, $p_{-1}(x) = 0$, and

$$U_0(x) = u(\zeta; \mu) \tag{7.5}$$

With this choice of p_{-1}, U_0, and a_0, U_m obeys $a_0 U_0 + b_1 U_1 + a_1 U_2 = x U_1$, and similarly for V_m. Since $p_{-1} = 0$ and $p_0 = 1$, (7.4) for $m = 0$ says

$$\mathrm{Wr}(x) = -u(\zeta; \mu) \tag{7.6}$$

Here is the key to going from Jost to Szegő asymptotics:

Theorem 7.1. *Suppose* $\{a_n, b_n\}_{n=1}^\infty$ *obey* (1.17) *for some* $\{\tilde{a}_n, \tilde{b}_n\}_{n=1}^\infty$ *in* $\mathcal{T}_{\mathfrak{e}}$. *Then, uniformly for* z *in compact subsets of* $\mathbb{C} \setminus ([\alpha_1, \beta_{\ell+1}] \cup \sigma(J))$,

$$\lim_{n \to \infty} [G_{nn}(z) - \widetilde{G}_{nn}(z)] = 0 \tag{7.7}$$

where $\widetilde{G}_{nn}$ *is given by* (7.1) *with* J *replaced by* $\tilde{J}$.

Proof. By the resolvent formula,

$$G_{nn}(z) - \widetilde{G}_{nn}(z) = \sum_{m,k} G_{nm}(z) (\tilde{J} - J)_{mk} \, \widetilde{G}_{kn}(z) \tag{7.8}$$

On compact subsets of $\mathbb{C}_+$,

$$|G_{kn}(z)| + |\widetilde{G}_{kn}(z)| \leq C e^{-D|k-n|} \tag{7.9}$$

for suitable $C, D > 0$. Since $(\tilde{J} - J)_{mk} \to 0$ as $m, k \to \infty$, we get (7.7) from (7.8) and (7.9). Using the maximum principle, one extends the result to compact subsets of $\mathbb{C} \setminus ([\alpha_1, \beta_{\ell+1}] \cup \sigma(J))$. $\square$

Theorem 7.2. *Under the hypotheses of Theorem 7.1, uniformly on the same compact subsets of* $\mathbb{C}$*, we have that*

$$\lim_{n\to\infty} \frac{G_{nn}(z)}{\widetilde{G}_{nn}(z)} = 1 \tag{7.10}$$

Proof. For each fixed n, $\widetilde{G}_{nn}(z)$ is nonvanishing on the compact subsets under discussion since neither $\tilde{u}_n$ nor $\tilde{p}_{n-1}$ have zeros there. Since shifting n is equivalent to moving on the torus, $\widetilde{G}_{nn}$ is uniformly bounded away from zero as n varies (cf. (7.13) below). Therefore, (7.7) implies (7.10). $\square$

As a final preliminary on Szegő asymptotics, we look at the isospectral torus. If $d\nu \in \mathcal{T}_{\mathfrak{e}}$, then reflection of the Jacobi parameters about $n = 0$,

$$b_n^{(r)} = b_{-n}, \quad a_n^{(r)} = a_{1-n}, \quad n \in \mathbb{Z} \tag{7.11}$$

gives an almost periodic Jacobi matrix in the isospectral torus, so a point we will call $d\nu^{(r)} \in \mathcal{T}_{\mathfrak{e}}$.

For $n \in \mathbb{Z}$, we denote by $d\nu_n \in \mathcal{T}_{\mathfrak{e}}$ the spectral measure of the two-sided Jacobi matrix $\tilde{J}_\nu$ when restricted to $\ell^2(\{n+1, n+2, \dots\})$. In particular, $d\nu_0 = d\nu$.

Following paper I, for $x \in \mathbb{C} \cup \{\infty\} \setminus \mathfrak{e}$, we define $\mathbf{z}(x) \in \mathcal{F}$ to be the unique point with $\mathbf{x}(\mathbf{z}(x)) = x$, and for $x \in \mathfrak{e}$, we set $\mathbf{z}(x) = \mathbf{z}(x - i0)$.

Theorem 7.3. *Given $d\nu \in \mathcal{T}_{\mathfrak{e}}$, there exist nonvanishing, continuous functions $\alpha(x;\nu)$ and $\beta(x;\nu)$ for $x \in \mathbb{C} \setminus [\alpha_1, \beta_{\ell+1}]$ so that the orthonormal polynomials are given by*

$$p_{n-1}(x;\nu) = \alpha(x;\nu) \frac{u(\mathbf{z}(x), \nu_{-n}^{(r)})}{a_{-n}^{(r)} B(\mathbf{z}(x))^n} + \beta(x;\nu) \frac{u(\mathbf{z}(x), \nu_n)}{a_n B(\mathbf{z}(x))^{-n}} \tag{7.12}$$

In particular, $p_{n-1}(x;\nu) B(\mathbf{z}(x))^n$ is asymptotically almost periodic. Moreover, on any compact subset, K, of $\mathbb{C} \setminus [\alpha_1, \beta_{\ell+1}]$, there is a constant $C > 1$ so that

$$C^{-1} B(\mathbf{z}(x))^n \le |p_{n-1}(x;\nu)| \le C B(\mathbf{z}(x))^n \tag{7.13}$$

for all $x \in K$ and $d\nu \in \mathcal{T}_{\mathfrak{e}}$.

Proof. Define

$$u_n^+(x;\nu) = u_n(\mathbf{z}(x);\nu) \qquad u_n^-(x;\nu) = u_{-n}^+(x;\nu^{(r)}) \tag{7.14}$$

Then $u_n^\pm$ are two solutions of

$$a_n v_{n+1} + b_n v_n + a_{n-1} v_{n-1} = x v_n \tag{7.15}$$

and they are linearly independent since one is L^2 at $+\infty$ and the other at $-\infty$, and x is not an eigenvalue of $\tilde{J}_\nu$.

Since $p_{n-1}(x;\nu)$ also solves (7.15), we have

$$p_{n-1}(x;\nu) = \alpha(x;\nu) u_n^-(x;\nu) + \beta(x;\nu) u_n^+(x;\nu) \tag{7.16}$$

and Wronskian formulae for α and β show that they are real analytic in $\nu \in \mathcal{T}_{\mathfrak{e}}$ and analytic in $x \in \mathbb{C} \setminus [\alpha_1, \beta_{\ell+1}]$.

(7.12) then follows from Theorem 9.2 of paper I.

Since $|B| < 1$ on $\mathbb{D}$, the second term multiplied by B^n is exponentially small, and the first is almost periodic, so $p_{n-1} B^n$ is almost periodic up to an exponentially small error.

The upper bound in (7.13) is immediate from (7.12), $|B| < 1$, and the almost periodicity of $u(z;\nu_n)$.

Since x is not an eigenvalue of $\tilde{J}_\nu$, α is nonvanishing, which proves that for any K and $n \geq N$, we have a lower bound. Since p_n has no zero in K, a lower bound on $n < N$ is immediate. That proves (7.13). $\square$

Theorem 7.4 (Szegő asymptotics)**.** *Let $d\mu \in \mathrm{Sz}(\mathfrak{e})$ and let $d\tilde{\mu}$ be the measure of the Jacobi matrix in $\mathcal{T}_{\mathfrak{e}}$ for which* (1.17) *holds. Then, uniformly on compact subsets of $\mathbb{C} \setminus [\alpha_1, \beta_{\ell+1}]$,*

$$\frac{p_n(x;\mu)}{p_n(x;\tilde{\mu})} \to \frac{u(\mathbf{z}(x);\mu)}{u(\mathbf{z}(x);\tilde{\mu})} \tag{7.17}$$

In particular, $p_n(x;\mu) B(\mathbf{z}(x))^n$ is asymptotically almost periodic.

Remarks. 1. It is not hard to see that the last statement extends to $\mathbb{C} \setminus \mathfrak{e}$.

2. In the periodic case, one also has Szegő asymptotics in the gaps of $\mathfrak{e}$ except at finitely many points.

3. Since the monic orthogonal polynomials, $P_n(x)$, are related to the orthonormal ones via $P_n(x) = (a_1 \cdots a_n)\, p_n(x)$, Szegő asymptotics for the monic polynomials immediately follows from (6.35) and (7.17),

$$\frac{P_n(x;\mu)}{P_n(x;\tilde{\mu})} \to \frac{u(\mathbf{z}(x);\mu)/u(0;\mu)}{u(\mathbf{z}(x);\tilde{\mu})/u(0;\tilde{\mu})}$$

Proof. It follows from (7.2) and (7.6) that

$$\frac{p_{n-1}(x;\mu)}{p_{n-1}(x;\tilde{\mu})} = \frac{G_{nn}(x)}{\widetilde{G}_{nn}(x)}\, \frac{u_n(\mathbf{z}(x);\tilde{\mu})}{u_n(\mathbf{z}(x);\mu)}\, \frac{u(\mathbf{z}(x);\mu)}{u(\mathbf{z}(x);\tilde{\mu})} \tag{7.18}$$

The result is immediate from (7.10) and (6.33) since we can include points below α_1 and above $\beta_{\ell+1}$ by the maximum principle and the fact that $p_n(x;\tilde\mu)$ is non-vanishing on $\mathbb{R}\setminus[\alpha_1,\beta_{\ell+1}]$. $\square$

8. L^2 Szegő Asymptotics on the Spectrum

By a standard approximation argument going back to Szegő [30], the function

$$\int_0^{2\pi} \frac{e^{i\theta}+z}{e^{i\theta}-z} \log(\operatorname{Im} M(e^{i\theta}))\frac{d\theta}{2\pi}$$

is in $H^2(\mathbb{D})$, so it has nontangential boundary values for a.e. $z\in\partial\mathbb{D}$. Since convergent Blaschke products (with a Blaschke condition) are well known to have boundary values (see [19, pp. 249, 310]), $u(z;\mu)$ has boundary values for a.e. $z\in\partial\mathbb{D}$ and all $d\mu\in\operatorname{Sz}(\mathfrak{e})$, and so does $u_n(z;\mu)$ by (5.11).

Thus, for Lebesgue a.e. $x\in\mathfrak{e}$,

$$u_n^+(x;\mu)\equiv u_n(\mathbf{z}(x-i0);\mu) \tag{8.1}$$

exists. Moreover, since $\operatorname{Im} m(x+i0)\neq 0$ for a.e. $x\in\mathfrak{e}$, we can define a linearly independent solution u_n^- by

$$u_n^-(x;\mu)\equiv\overline{u_n^+(x;\mu)} \tag{8.2}$$

This leads to an expansion:

$$p_n(x) = \frac{\operatorname{Wr}(p_{\cdot-1},u_\cdot^-)u_{n+1}^+(x;\mu)-\operatorname{Wr}(p_{\cdot-1},u_\cdot^+)u_{n+1}^-(x;\mu)}{\operatorname{Wr}(u_\cdot^+,u_\cdot^-)} \tag{8.3}$$

$$= \frac{\overline{u_0^+(x;\mu)}u_{n+1}^+(x;\mu)-u_0^+(x;\mu)\overline{u_{n+1}^+(x;\mu)}}{\operatorname{Wr}(u_\cdot^+,u_\cdot^-)} \tag{8.4}$$

Given the asymptotics of u_n^+ to $\tilde u_n^+$, this explains the expected L^2 asymptotic result we'll prove:

Theorem 8.1. *Let $d\mu\in\operatorname{Sz}(\mathfrak{e})$ have the form* (1.9) *and let $\tilde u_n^+(x)$ be the Jost solution for the asymptotic point in $\mathcal{T}_\mathfrak{e}$ (i.e., the point given by* (1.17)*). Then*

$$\int_\mathfrak{e}\left|p_n(x)-\frac{\operatorname{Im}(\overline{u(\mathbf{z}(x);\mu)}\,\tilde u_{n+1}^+(x))}{\pi v_\mathfrak{e}(x)}\right|^2 w(x)\,dx\to 0 \tag{8.5}$$

and

$$\int |p_n(x)|^2\,d\mu_{\rm s}(x)\to 0 \tag{8.6}$$

where $v_\mathfrak{e}$ is the weight for the reference measure used in (5.3).

Remarks. 1. $\pi v_{\mathfrak{e}}(x)$ enters because of the following calculation:

$$\operatorname{Wr}(\tilde{u}_{\cdot}^{+}, \tilde{u}_{\cdot}^{-}) = \tilde{a}_0(\tilde{u}_1^{+}\, \overline{\tilde{u}_0^{+}} - \overline{\tilde{u}_1^{+}}\, \tilde{u}_0^{+}) \tag{8.7}$$

$$= -(\tilde{a}_0)^2 |\tilde{u}_0^{+}|^2 2i \operatorname{Im} \tilde{m}(x - i0) \tag{8.8}$$

$$= 2i\, \frac{v_{\mathfrak{e}}(x)}{\tilde{w}(x)}\, \pi \tilde{w}(x) \tag{8.9}$$

$$= 2\pi i\, v_{\mathfrak{e}}(x) \tag{8.10}$$

In the above, (8.8) comes from (1.26) and (5.10), and (8.9) comes from (4.16), (5.3) (see Lemma 8.2 below), and (5.11), which says that $\tilde{u}_0^{+} = \tilde{a}_0^{-1} u(\,\cdot\,, \tilde{\mu})$.

2. In case $\mathfrak{e} = [-2, 2]$, (8.5) becomes

$$\int_{-2}^{2} \left| p_n(x) - \frac{\operatorname{Im}(\overline{u(\mathbf{z}(x);\mu)}\, e^{i(n+1)\theta(x)})}{\sin(\theta(x))} \right|^2 w(x)\, dx \to 0$$

where $\theta(x)$ is given by $\mathbf{z}(x) = e^{i\theta(x)}$. This is a result of [15]; see also [5] and [26, Sect. 3.7].

We define

$$k_n^{+}(x) = \frac{\overline{u(\mathbf{z}(x);\mu)}\, \tilde{u}_{n+1}^{+}(x)}{2\pi i v_{\mathfrak{e}}(x)} \tag{8.11}$$

$$k_n^{-}(x) = \overline{k_n^{+}(x)} \tag{8.12}$$

in which case, (8.5)–(8.6) become

$$\|p_n - k_n^{+} - k_n^{-}\|_w^2 + \|p_n\|_{\mathrm{s}}^2 \to 0 \tag{8.13}$$

where $\|\cdot\|_w$ is the $L^2(\mathfrak{e}, w\, dx)$ norm (we use $\langle\ ,\ \rangle_w$ for the inner product) and $\|\cdot\|_{\mathrm{s}}$ is the $L^2(\mathbb{R}, d\mu_{\mathrm{s}})$ norm. Clearly, (8.13) follows from:

$$\|p_n\|_w^2 + \|p_n\|_{\mathrm{s}}^2 = 1 \tag{8.14}$$

$$\|k_n^{\pm}\|_w^2 = \tfrac{1}{2} \tag{8.15}$$

$$\lim_{n\to\infty} \langle k_n^{-}, k_n^{+}\rangle_w = 0 \tag{8.16}$$

$$\lim_{n\to\infty} \operatorname{Re}\langle k_n^{-}, p_n\rangle_w = \tfrac{1}{2} \tag{8.17}$$

(8.14) is the normalization condition on p_n, so we only need to prove (8.15)–(8.17). We'll need some preliminaries:

Lemma 8.2. *For a.e. $z \in \partial\mathbb{D}$, the boundary value of $u(z;\mu)$ obeys*

$$|u(z;\mu)|^2 = \frac{v_{\mathfrak{e}}(\mathbf{x}(z))}{w(\mathbf{x}(z))} \tag{8.18}$$

Proof. In (5.3), $|\prod_k B(z,z_k)|$ has 1 as boundary value, by standard results on Blaschke products. By convergence of the Poisson kernel, for a.e. z in $\partial\mathbb{D}$, the real part of the exponential converges to $\log(\frac{v_{\mathfrak{e}}(\mathbf{x}(z))}{w(\mathbf{x}(z))})$. $\square$

Lemma 8.3. *For any $d\nu \in \mathcal{T}_{\mathfrak{e}}$ with weight w_ν, we have*

$$\int_{\mathfrak{e}} \frac{dx}{w_\nu(x)} = 2\pi^2 a_0(\nu)^2 \tag{8.19}$$

Proof. If $\widetilde{G}_{00}(z;\nu)$ is the Green's function of the whole-line Jacobi matrix $\tilde{J}_\nu$ and $u_n^+(x;\nu) = u_n(\mathbf{z}(x+i0);\nu)$ the boundary value of the Jost solution, then

$$\begin{aligned} \widetilde{G}_{00}(x+i0;\nu) &= \frac{\overline{u_0^+(x;\nu)}\, u_0^+(x;\nu)}{a_0(\nu)[u_1^+(x;\nu)\,\overline{u_0^+(x;\nu)} - \overline{u_1^+(x;\nu)}\, u_0^+(x;\nu)]} && (8.20) \\ &= -\frac{1}{a_0(\nu)^2 2i \operatorname{Im} m(x+i0;\nu)} && (8.21) \\ &= \frac{i}{2\pi a_0(\nu)^2 w_\nu(x)} && (8.22) \end{aligned}$$

so

$$\frac{1}{\pi} \operatorname{Im} \widetilde{G}_{00}(x+i0;\nu) = \frac{1}{2\pi^2 a_0(\nu)^2 w_\nu(x)} \tag{8.23}$$

But the whole-line Jacobi matrix $\tilde{J}_\nu$ has purely a.c. spectrum $\sigma(\tilde{J}_\nu) = \mathfrak{e}$ and the density of the probability spectral measure for $\tilde{J}_\nu$ and δ_0 is $\frac{1}{\pi}\operatorname{Im}\widetilde{G}_{00}(x+i0;\nu)$, so

$$\frac{1}{\pi}\int_{\mathfrak{e}} \operatorname{Im}\widetilde{G}_{00}(x+i0;\nu)\,dx = 1 \tag{8.24}$$

(8.23) and (8.24) imply (8.19). $\square$

Proposition 8.4. (8.15) *holds.*

Proof. By (5.11) and (8.11),

$$|k_n^+(x)|^2 = \frac{|u(\mathbf{z}(x);\mu)|^2 |u(\mathbf{z}(x);\tilde{\mu}_{n+1})|^2}{4\pi^2(\tilde{a}_{n+1})^2 v_{\mathfrak{e}}(x)^2} \tag{8.25}$$

so, by Lemma 8.2,

$$|k_n^+(x)|^2 = \frac{1}{4\pi^2(\tilde{a}_{n+1})^2 w(x)\tilde{w}_{n+1}(x)} \tag{8.26}$$

and so,

$$\int_{\mathfrak{e}} |k_n^+(x)|^2 w(x)\,dx = \frac{1}{4\pi^2(\tilde{a}_{n+1})^2}\int_{\mathfrak{e}}\frac{dx}{\tilde{w}_{n+1}(x)} = \frac{1}{2} \tag{8.27}$$

by Lemma 8.3. Since $|k_n^-| = |k_n^+|$, we get the same result for $\|k_n^-\|_w^2$. $\square$

Lemma 8.5. *Let $f \in L^1(\mathfrak{e}, d\rho_\mathfrak{e})$. Then*

$$\lim_{n\to\infty} \int_\mathfrak{e} B(\mathbf{z}(x))^n f(x)\, d\rho_e(x) = 0 \tag{8.28}$$

Moreover, (8.28) *holds uniformly on norm compact subsets of* $L^1(\mathfrak{e}, d\rho_\mathfrak{e})$.

Proof. Without loss of generality, assume that f is real-valued. Then by Corollary 4.6 of paper I, we obtain

$$\int_\mathfrak{e} B(\mathbf{z}(x))^n f(x)\, d\rho_e(x) = \int_0^{2\pi} B(e^{i\theta})^n f(\mathbf{x}(e^{i\theta}))\, \frac{d\theta}{2\pi} \tag{8.29}$$

By the Cauchy theorem, $\{B^n\}_{n\in\mathbb{Z}}$ forms an orthonormal system in $L^2(\partial\mathbb{D}, \frac{d\theta}{2\pi})$. Hence it follows from the Bessel inequality that RHS of (8.29) converges to zero for any L^2-function. The general case of L^1-functions and the result on uniform convergence on norm compacts follow by approximation. $\square$

Remark. The above result can be also established via a stationary phase argument.

Proposition 8.6. (8.16) *holds.*

Proof. By the same calculation that was used in the proof of Proposition 8.4,

$$\langle k_n^-, k_n^+\rangle_w = \int_\mathfrak{e} f_n(x) B^{2n+2}(\mathbf{z}(x))\, dx \tag{8.30}$$

where

$$f_n(x) = -\frac{1}{4\pi^2(\tilde{a}_{n+1})^2}\, \frac{u(\mathbf{z}(x);\tilde{\mu}_{n+1})^2}{v_\mathfrak{e}(x)}\, \frac{|u(\mathbf{z}(x);\mu)|^2}{u(\mathbf{z}(x);\mu)^2} \tag{8.31}$$

For $d\nu \in \mathcal{T}_\mathfrak{e}$, let

$$f(x;\nu) = -\frac{1}{4\pi^2 a_0(\nu)^2}\, \frac{u(\mathbf{z}(x);\nu)^2}{v_\mathfrak{e}(x)}\, \frac{|u(\mathbf{z}(x);\mu)|^2}{u(\mathbf{z}(x);\mu)^2} \tag{8.32}$$

By Lemma 8.2, the f's are all in L^1 (with L^1 norm $1/2$ by Lemma 8.3) and f is L^1 continuous in ν. So, since $\mathcal{T}_\mathfrak{e}$ is compact, we see from Lemma 8.5 that the integral in (8.30) goes to zero. $\square$

This leaves (8.17). The argument is somewhat complicated in case there are bound states, especially if there are infinitely many. So let us consider it first when $d\mu$ has no point masses in $\mathbb{R}\setminus\mathfrak{e}$.

Proposition 8.7. *Suppose $d\mu$ has support $\mathfrak{e}$ so that $u(z;\mu)$ is nonvanishing on $\mathbb{D}$. Then* (8.17) *holds.*

Proof. We claim that

$$\begin{aligned}\operatorname{Re}\left[\int_{\mathfrak{e}} \overline{k_n^-(x)}\, p_n(x) w(x)\, dx\right] \\ = \frac{1}{2}\int_{\partial\mathcal{F}\cap\partial\mathbb{D}} \frac{\overline{u(z;\mu)}\,\tilde u_{n+1}(z)}{2\pi i v_{\mathfrak{e}}(\mathbf{x}(z))}\, p_n(\mathbf{x}(z))\, w(\mathbf{x}(z))\mathbf{x}'(z)\, dz\end{aligned} \tag{8.33}$$

where the integral is evaluated counterclockwise. As $\operatorname{Re}\overline{k_n^-} = \frac{1}{2}k_n^+ + \frac{1}{2}k_n^-$ and $\operatorname{Re} p_n(x) = p_n(x)$, the k_n^+ term directly gives the counterclockwise integral over $\mathbb{C}_+ \cap \partial\mathcal{F}\cap\partial\mathbb{D}$ (since $\mathbf{x}'(z)$ is positive there). Since u and $\tilde u_{n+1}^+$ are real on $\mathbb{R}$, and $\mathbf{x}'$ and i flip signs under $e^{i\theta}\to e^{-i\theta}$, the k_n^- term gives the integral over $\partial\mathcal{F}\cap\partial\mathbb{D}\cap\mathbb{C}_-$.

Notice next that, by (8.18),

$$\overline{u(z;\mu)}\,\frac{w(\mathbf{x}(z))}{v_{\mathfrak{e}}(\mathbf{x}(z))} = \frac{1}{u(z;\mu)} \tag{8.34}$$

so

$$\text{LHS of (8.33)} = \frac{1}{4\pi i}\int_{\partial\mathcal{F}\cap\partial\mathbb{D}} \frac{\tilde u_{n+1}(z) p_n(\mathbf{x}(z))}{u(z;\mu)}\,\mathbf{x}'(z)\, dz \tag{8.35}$$

By (6.28), (5.11), and the choice of $d\tilde\mu$, the integrand in (8.35), call it F, is automorphic under Γ. Since F is real on $\mathbb{R}$, we have $F(\bar z) = \overline{F(z)}$. Moreover, there are $\gamma\in\Gamma$ so that for $z\in C_\ell^+$, we have $\overline{\gamma(z)} = z$, so we conclude that F is real on C_ℓ^+ and C_ℓ^-. Thus, orienting the contours counterclockwise about 0, we get

$$\int_{C_\ell^+\cup C_\ell^-} F(z)\, dz = 0$$

since C_ℓ^+ and C_ℓ^- run in opposite directions. It follows that

$$\text{LHS of (8.33)} = \frac{1}{4\pi i}\int_{\partial\mathcal{F}} \frac{\tilde u_{n+1}(z) p_n(\mathbf{x}(z))}{u(z;\mu)}\,\mathbf{x}'(z)\, dz \tag{8.36}$$

Inside $\mathcal{F}$, the integrand is regular except at $z=0$. Since p_n is a polynomial of degree n in $\mathbf{x}(z)$, and $\mathbf{x}(z)$ has a simple pole at $z=0$, $z^n p_n(\mathbf{x}(z))$ is regular at $z=0$. By (5.11), $\tilde u_{n+1}(z)/B(z)^{n+1}$ is regular at $z=0$. Thus, $\tilde u_{n+1}(z)p_n(\mathbf{x}(z))$ has a first-order zero at $z=0$. $u(z)$ is regular there and $\mathbf{x}'(z)$ has a double pole. So the integrand in (8.36) has a simple pole at $z=0$ and we conclude that

$$\text{LHS of (8.33)} = \frac{1}{2}\left[\frac{u_{n+1}(z;\mu)p_n(\mathbf{x}(z))}{zu(z;\mu)}\bigg|_{z=0}\right]\frac{u_{n+1}(0;\tilde\mu)}{u_{n+1}(0;\mu)}\,[z^2\mathbf{x}'(z)|_{z=0}] \tag{8.37}$$

The first factor in (8.37) is $z^{-1}G_{n+1,n+1}(\mathbf{x}(z))|_{z=0}$, which is

$$\lim_{z\to 0} z^{-1}\left(-\frac{1}{\mathbf{x}(z)}+O\left(\frac{1}{\mathbf{x}(z)^2}\right)\right)=-\frac{1}{x_\infty} \tag{8.38}$$

The third factor is

$$\lim_{z\to 0} z^{2}\left(-\frac{x_\infty}{z^2}+O(1)\right)=-x_\infty \tag{8.39}$$

so

$$\text{LHS of (8.33)} = \frac{1}{2}\,\frac{u_{n+1}(0;\tilde\mu)}{u_{n+1}(0;\mu)} \to \frac{1}{2} \tag{8.40}$$

by Theorem 6.6. □

Proposition 8.8. *If $d\mu$ has support $\mathfrak{e}$ plus finitely many mass points in $\mathbb{R}\setminus\mathfrak{e}$, then* (8.17) *holds.*

Proof. We follow the proof of the last proposition until we get to (8.35). However, u now has a pole at each z_k in $\mathcal{F}$ with

$$\mathbf{x}(z_k)=x_k\in\sigma(J) \tag{8.41}$$

Thus, the integrand can have poles (but only finitely many) in $\mathcal{F}^{\text{int}}$ and also on $C_j^\pm$. Interpret (8.36) as taking principal parts at the poles on $C_j^\pm$. Each such pole contributes with half of $2\pi i$ times the residue, so we get $2\pi i$ times the residue if we only count the poles in $\mathcal{F}$ (i.e., in C_j^+ but not in C_j^-).

The residue at z_k is

$$\frac{B(z_k)^{n+1}u(z_k;\tilde\mu_{n+1})p_n(x_k)\mathbf{x}'(z_k)}{2\tilde a_{n+1}u'(z_k;\mu)} \tag{8.42}$$

As $\sum_n |p_n(x_k)|^2 = 1/\mu(\{x_k\})$, $|B(z_k)|<1$ and $\sup_n |u(z_k;\tilde\mu_{n+1})|<\infty$, the quantity in (8.42) goes to zero. Since there are finitely many of these poles, their contribution vanishes in the limit and LHS of (8.33) converges to $1/2$. □

Finally, we turn to the general case. The following completes the proof of Theorem 8.1:

Proposition 8.9. *For any $d\mu\in\mathrm{Sz}(\mathfrak{e})$,* (8.17) *holds.*

Proof. Following Peherstorfer–Yuditskii [15], we'll approximate u by one with a finite number of zeros, but to preserve the fact that we need certain functions to be automorphic, we also modify $\tilde u_n$.

Label all the point masses of $d\mu$ in a single sequence $\{x_k\}_{k=1}^\infty$ with corresponding points $z_k \in \mathcal{F}$ such that $\mathbf{x}(z_k) = x_k$. Let

$$u^{(m)}(z;\mu) = \prod_{k=1}^m B(z,z_k)\varepsilon(z;\mu) \tag{8.43}$$

and denote by $d\tilde\mu^{(m)}$ the measure in the isospectral torus whose Jost function has the same character as $u^{(m)}$. Define

$$k_n^{(m)+}(x) = \frac{\overline{u^{(m)}(\mathbf{z}(x);\mu)}\, u_{n+1}^+(x;\tilde\mu^{(m)})}{2i\pi v_{\mathfrak{e}}(x)} \tag{8.44}$$

Clearly, it suffices to prove that

$$\lim_{m\to\infty} \|k_n^{(m)+} - k_n^+\|_w \to 0 \tag{8.45}$$

uniformly in n, and that

$$\lim_{m\to\infty} \lim_{n\to\infty} |\mathrm{Re}\langle k_n^{(m)+}, p_n\rangle - \tfrac12| = 0 \tag{8.46}$$

Since $\prod_{k=1}^m B(z,z_k) \to \prod_{k=1}^\infty B(z,z_k)$ uniformly on compacts, the characters converge. Moreover, this convergence of B's is pointwise on $\partial\mathbb{D}$. The first implies convergence of $u(\mathbf{z}(x);\tilde\mu_{n+1}^{(m)})$ to $u(\mathbf{z}(x);\tilde\mu_{n+1})$ away from the band edges (uniformly in n and x as $m\to\infty$) with uniform square root bounds. This plus (8.26) yields (8.45).

The proof of (8.46) follows the proof of Proposition 8.8. The fact that we've arranged for the functions to be automorphic allows the cancellation of the C_j^+ and C_j^- integrals, and since there are only finitely many poles away from $z=0$, we get convergence in (8.42) and hence in (8.46). $\square$

References

[1] 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.

[2] J. Christiansen, B. Simon, and M. Zinchenko, *Finite gap Jacobi matrices, I. The isospectral torus*, to appear in Constr. Approx.

[3] J. Christiansen, B. Simon, and M. Zinchenko, *Finite gap Jacobi matrices, III. Beyond the Szegő class*, in preparation.

[4] D. Damanik, R. Killip, and B. Simon, *Perturbations of orthogonal polynomials with periodic recursion coefficients*, to appear in Annals of Math.

[5] D. Damanik and B. Simon, *Jost functions and Jost solutions for Jacobi matrices, I. A necessary and sufficient condition for Szegő asymptotics*, Invent. Math. **165** (2006), 1–50.

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

[7] R. Frank, B. Simon, and T. Weidl *Eigenvalue bounds for perturbations of Schrödinger operators and Jacobi matrices with regular ground states*, Comm. Math. Phys. **282** (2008), 199–208.
[8] J. B. Garnett, *Bounded Analytic Functions*, Pure and Applied Math., **96**, Academic Press, New York-London, 1981.
[9] 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.
[10] D. Hundertmark and B. Simon, *Eigenvalue bounds in the gaps of Schrödinger operators and Jacobi matrices*, J. Math. Anal. Appl. **340** (2008), 892–900.
[11] R. Killip and B. Simon, *Sum rules for Jacobi matrices and their applications to spectral theory*, Annals of Math. **158** (2003), 253–321.
[12] Y. Last and B. Simon, *The essential spectrum of Schrödinger, Jacobi, and CMV operators*, J. Anal. Math. **98** (2006), 183–220.
[13] P. Nevai, *Orthogonal polynomials*, Mem. Amer. Math. Soc. **18** (1979), no. 213, 1–183.
[14] F. Peherstorfer and P. Yuditskii, private communication.
[15] F. Peherstorfer and P. Yuditskii, *Asymptotics of orthonormal polynomials in the presence of a denumerable set of mass points*, Proc. Amer. Math. Soc. **129** (2001), 3213–3220.
[16] F. Peherstorfer and P. Yuditskii, *Asymptotic behavior of polynomials orthonormal on a homogeneous set*, J. Anal. Math. **89** (2003), 113–154.
[17] F. Peherstorfer and P. Yuditskii, *Remark on the paper "Asymptotic behavior of polynomials orthonormal on a homogeneous set"*, arXiv math.SP/0611856.
[18] C. Remling, *The absolutely continuous spectrum of Jacobi matrices*, preprint.
[19] W. Rudin, *Real and Complex Analysis*, 3rd edition, McGraw–Hill, New York, 1987.
[20] J. A. Shohat, *Théorie Générale des Polinomes Orthogonaux de Tchebichef*, Mémorial des Sciences Mathématiques, **66**, pp. 1–69, Paris, 1934.
[21] B. Simon, *A canonical factorization for meromorphic Herglotz functions on the unit disk and sum rules for Jacobi matrices*, J. Funct. Anal. **214** (2004), 396–409.
[22] B. Simon, *OPUC on one foot*, Bull. Amer. Math. Soc. **42** (2005), 431–460.
[23] B. Simon, *Orthogonal Polynomials on the Unit Circle, Part 1: Classical Theory*, AMS Colloquium Publications, **54.1**, American Mathematical Society, Providence, R.I., 2005.
[24] B. Simon, *Orthogonal Polynomials on the Unit Circle, Part 2: Spectral Theory*, AMS Colloquium Publications, **54.2**, American Mathematical Society, Providence, R.I., 2005.
[25] B. Simon, *Equilibrium measures and capacities in spectral theory*, Inverse Problems and Imaging **1** (2007), 713–772.
[26] B. Simon, *Szegő's Theorem and Its Descendants: Spectral Theory for L^2 Perturbations of Orthogonal Polynomials*, in preparation; to be published by Princeton University Press.
[27] B. Simon and A. Zlatoš, *Sum rules and the Szegő condition for orthogonal polynomials on the real line*, Comm. Math. Phys. **242** (2003), 393–423.

[28] M. Sodin and P. Yuditskii, *Almost periodic Jacobi matrices with homogeneous spectrum, infinite-dimensional Jacobi inversion, and Hardy spaces of character-automorphic functions*, J. Geom. Anal. **7** (1997), 387–435.

[29] H. Stahl and V. Totik, *General Orthogonal Polynomials*, in "Encyclopedia of Mathematics and its Applications," **43**, Cambridge University Press, Cambridge, 1992.

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

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

[32] G. Szegő, *Orthogonal Polynomials*, AMS Colloquium Publications, **23**, American Mathematical Society, Providence, R.I., 1939; 3rd edition, 1967.

[33] H. Widom, *Extremal polynomials associated with a system of curves in the complex plane*, Adv. in Math. **3** (1969), 127–232.