FINE STRUCTURE OF THE ZEROS OF ORTHOGONAL POLYNOMIALS, IV. A PRIORI BOUNDS AND CLOCK BEHAVIOR

YORAM LAST[1,3] AND BARRY SIMON[2,3]

Dedicated to Percy Deift on the occasion of his sixtieth birthday

Abstract. We prove locally uniform spacing for the zeros of orthogonal polynomials on the real line under weak conditions (Jacobi parameters approach the free ones and are of bounded variation). We prove that for ergodic discrete Schrödinger operators, Poisson behavior implies positive Lyapunov exponent. Both results depend on a priori bounds on eigenvalue spacings for which we provide several proofs.

1. Introduction

Our primary goal in this paper concerns the fine structure of the zeros of orthogonal polynomials on the real line (OPRL), although we will say something about zeros of paraorthogonal polynomials on the unit circle (POPUC) (see Section 10). Specifically, $d\mu$ will be a measure of compact support which is nontrivial (i.e., not supported on a finite set), usually a probability measure. $P_n(x)$ or $P_n(x;d\mu)$ will be the monic orthogonal polynomials and $p_n(x) = P_n/\|P_n\|$ the orthonormal polynomials. The Jacobi parameters, $\{a_n, b_n\}_{n=1}^\infty$, are defined by

$$xP_n(x) = P_{n+1}(x) + b_{n+1}P_n(x) + a_n^2 P_{n-1}(x) \tag{1.1}$$

$n = 0, 1, 2, \dots$, where $P_{-1}(x) \equiv 0$. It follows that (when $\mu(\mathbb{R}) = 1$)

$$\|P_n\| = a_1 \dots a_n \tag{1.2}$$

so

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

Date: May 16, 2006.

[1] Institute of Mathematics, The Hebrew University, 91904 Jerusalem, Israel. E-mail: ylast@math.huji.ac.il. Supported in part by The Israel Science Foundation (Grant No. 188/02).

[2] Mathematics 253-37, California Institute of Technology, Pasadena, CA 91125. E-mail: bsimon@caltech.edu. Supported in part by NSF grant DMS-0140592.

[3] Research supported in part by Grant No. 2002068 from the United States-Israel Binational Science Foundation (BSF), Jerusalem, Israel.

and x has a matrix representation in the orthonormal basis $\{p_n\}_{n=0}^\infty$,

$$J = \begin{pmatrix} b_1 & a_1 & 0 & \cdots \\ a_1 & b_2 & a_2 & \cdots \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix} \tag{1.4}$$

called the Jacobi matrix. The finite Jacobi matrix, $J_{n;F}$, is the $n \times n$ submatrix of J in the upper left corner. It is easy to see that (see (b) of the Appendix)

$$P_n(x) = \det(x\mathbf{1} - J_{n;F}) \tag{1.5}$$

Let $\{x_j^{(n)}\}_{j=1}^n$ be the zeros of $P_n(x)$, so (1.5) says that the $x_j^{(n)}$ are eigenvalues of $J_{n;F}$. Let $d\nu_n$ be the pure point probability measure that gives weight $1/n$ to each $x_j^{(n)}$. We say the *density of states* exists if $d\nu_n$ has a weak limit $d\nu_\infty$. By (1.5), one sees that

$$\int x^k \, d\nu_n(x) = \frac{1}{n}\operatorname{Tr}(J_{n;F}^k) \tag{1.6}$$

which is often useful.

The existence of the limit for a large class of regular measures on $[-2,2]$ goes back to Erdös-Turán [9]. Nevai [34] realized all that was used is

$$\lim_{n\to\infty} |b_n| + |a_n - 1| = 0 \tag{1.7}$$

(following the convention in the Schrödinger operator community, we use $a_n \equiv 1$ as a "free" case, while the OP community uses $a_n \equiv \frac{1}{2}$). Indeed,

Theorem 1.1 (known)**.** *If $d\mu$ is a measure where* (1.7) *holds, the density of states exists and is given by*

$$d\nu = \pi^{-1}\big(\sqrt{4-x^2}\,\big)^{-1}\chi_{[-2,2]}\, dx \tag{1.8}$$

The modern proof is not hard. For $a_n \equiv 1$, $b_n \equiv 0$, the OPRL are explicitly given by

$$P_n(2\cos\theta) = \frac{\sin((n+1)\theta)}{\sin\theta}$$

from which one computes $d\nu$ exactly for this case. If $J_{0,n;F}$ is the corresponding cutoff J_0, then (1.7) implies

$$\frac{1}{n}\operatorname{Tr}(J_{n;F}^k - J_{0,n;F}^k) \to 0 \tag{1.9}$$

for each $k = 0,1,2,\dots$ which, by (1.6), implies $\int x^k\, d\nu_n$ has the same limits as for the free case.

Another case where it is known that $d\nu$ exists is when a_n, b_n are samples of an ergodic family, that is, $a_n^{(\omega)} = f(T^n\omega)$, $b_n^{(\omega)} = g(T^n\omega)$

with $T\colon \Omega \to \Omega$ an ergodic transformation on $(\Omega, d\rho)$, a probability measure space. In that case, it is known (going back to the physics literature and proven rigorously by Pastur [38], Avron-Simon [1], and Kirsch-Martinelli [28]):

Theorem 1.2 (known)**.** *For ergodic Jacobi matrices, $d\nu_{n,\omega}$ has a limit $d\nu$ for a.e. ω and $d\nu$ is a.e. ω-independent.*

Again, the proof uses (1.6) plus in this case that, by ergodicity, $\frac{1}{n}\operatorname{Tr}(J_{n;F}^k)$ has a limit a.e. by the Birkhoff ergodic theorem.

The most important examples of the ergodic case are periodic, almost periodic, and random.

One easily combines the two ideas to see that $d\nu$ exists (and does not depend on $\delta a_n, \delta b_n$) if

$$a_n = a_n^{(\omega)} + \delta a_n \qquad b_n = b_n^{(\omega)} + \delta b_n$$

with $a_n^{(\omega)}, b_n^{(\omega)}$ ergodic and $|\delta a_n| + |\delta b_n| \to 0$.

These results describe the bulk features of the zeros. Here we are interested in the fine structure, on the level of individual eigenvalues; specifically, the focus in [41, 42, 43] and a main focus in this paper is what we call clock behavior, that the spacing locally is equal spacing. The term clock comes from the case of orthogonal polynomials on the unit circle (OPUC) where $d\nu$ is typically Lebesgue measure on a circle and the equal space means the zeros look like the numbers on a clock.

In order for the density of zeros to be $d\nu$, the equal spacing must be $1/(d\nu/dE)$. The symmetric derivative $d\nu/dE$ exists for a.e. E and, of course, $(d\nu/dE)\,dE$ is the a.c. part of $d\nu$. It is known (see, e.g., Avron-Simon [1]) that $d\nu$ has no pure points and, in many cases, it is known that $d\nu/dE$ is a continuous function, or even C^∞. To be formal, we define first

Definition. Let $E_0 \in \operatorname{supp}(d\nu)$. We let $z_n^{(j)}(E_0)$ be the zeros nearest E_0 so that

$$z_n^{(-2)}(E_0) < z_n^{(-1)}(E_0) \le E_0 < z_n^{(1)}(E_0) < z_n^{(2)}(E_0) < \cdots$$

if such zeros exist. If $E_0 \in [\operatorname{supp}(d\nu)]^{\text{int}}$, then $z_n^j(E_0)$ exists for j fixed and n large.

Definition. Let $E_0 \in \operatorname{supp}(d\nu)$. We say there is weak clock behavior at E_0 if $d\nu/dE_0$ exists and

$$\lim_{n\to\infty} n[z_n^{(1)}(E_0) - z_n^{(-1)}(E_0)]\,\frac{d\nu}{dE}\,(E_0) = 1 \tag{1.10}$$

We say there is strong clock behavior at E_0 if $d\nu/dE_0$ exists, (1.10) holds, and for $j = 1, \pm 2, \pm 3, \dots$ fixed,

$$\lim_{n\to\infty} n[z_n^{(j)}(E_0) - z_n^{(j+1)}(E_0)]\,\frac{d\nu}{dE}\,(E_0) = 1 \tag{1.11}$$

Definition. We say there is uniform clock behavior on $[\alpha, \beta]$ if $d\nu/dE$ is continuous and nonvanishing on $[\alpha, \beta]$ and

$$\lim_n \left[\sup\left\{\left|n[E-E'] - \left(\frac{d\nu}{dE}\right)^{-1}\right| \;\middle|\; E, E' \text{ are successive zeros of } p_n \text{ in } [\alpha,\beta]\right\}\right] = 0 \tag{1.12}$$

It is obvious that uniform clock behavior implies strong clock behavior at each interior point.

In the earlier papers in this series that discussed clock behavior for OPRL [41, 43], there was a technical issue that severely limited the results in general situations. In all cases, a Jost function-type analysis was used to show that a suitably rescaled p_n converged, that is,

$$c_n p_n \left(E_0 + n(x - E_0)\,\frac{d\nu}{dE}\,(E_0) + \xi_n \right) \to f_\infty(x) \tag{1.13}$$

where $|\xi_n| \le c/n$ and f_∞ has zeros at $0, \pm 1, \pm 2, \dots$. This would naively seem to show that p_n has clock-spaced zeros and, indeed, it does imply at least one zero near $E_0 + \xi_n + \frac{j}{n}(\frac{d\nu}{dE_0})^{-1}$ consistent with clock spacing.

The snag involves uniqueness, for the function

$$\sin(\pi x) - \frac{1}{n}\sin(n^2 \pi x) \tag{1.14}$$

has a limit like f_∞ but has more and more zeros near $x = 0$. That is, one needs to prove uniqueness of the zeros near $E_0 + \xi_n + \frac{j}{n}(\frac{d\nu}{dE_0})^{-1}$.

In previous papers in this series, two methods were used to solve this uniqueness problem. One relied on some version of the argument principle, essentially Rouché's theorem. This requires analyticity which, typically, severely restricts what recursion coefficients are allowed. In the case of OPUC where one needs to control zeros in the complex plane, some kind of analyticity argument seems to be necessary. The second method relies on the fact that if (1.13) also holds for derivatives and $f'_\infty(j) \ne 0$, then there is a unique zero. This argument also requires extra restrictions on the recursion coefficients, albeit not as severe as the analyticity requirement. For example, in [41], one only needed $\sum_{n=1}^\infty |a_n - 1| + |b_n| < \infty$ to get (1.13) for $E \in [-2+\varepsilon, 2-\varepsilon]$. But to control derivatives, we needed $\sum_{n=1}^\infty n(|a_n - 1| + |b_n|) < \infty$. In fact, the general argument failed to capture Jacobi polynomials (whose clock estimates were earlier obtained by Vértesi and Szabados [49, 52, 53, 54]) where separate arguments are needed. In [43], it was decided not to deal with asymptotically periodic OPRL since the derivative arguments looked to be tedious.

Note. The papers [41, 42, 44] discussed asymptotics of OPUC when the Verblunsky coefficients decay exponentially. We have discovered a paper of Pan [37] whose results overlap those in [41, 42, 44].

The key realization of this paper is that there is a more efficient way to eliminate pathologies like those in (1.14). Namely, we will seek a priori lower bounds on eigenvalue spacings. If we find any $O(1/n)$ lower bound, that implies the rescaled p_n of (1.13) has at most one zero near any zero of f_∞. Such lower bounds are not new in suitable situations. Erdös-Turán [9] already have such bounds if the measure is purely absolutely continuous in an interval with a.c. weights bounded away from zero and infinity. These ideas were developed by Nevai [34] and Golinskii [14]. Under suitable hypotheses on the transfer matrix, lower bounds are known in the Schrödinger operator community; see, for example, Jitomirskaya et al. [21]. (*Note*: They use "equal spacing" for $O(1/n)$ lower bounds and do not mean clock behavior by this term.)

While we could have used these existing bounds in the proofs of Theorems 1.3 and 1.4 below, we have found a new approach which allows us to also prove Theorem 1.5 below, and this approach is discussed in Section 2. With these lower bound ideas, we can prove:

Theorem 1.3 ($\equiv$ Theorem 3.2)**.** *Suppose that*

$$\sum_{n=1}^{\infty} (|a_n - 1| + |b_n|) < \infty \tag{1.15}$$

Then there is uniform clock behavior on each interval $[-2+\varepsilon, 2-\varepsilon]$ *for any* $\varepsilon > 0$.

Theorem 1.4 ($\equiv$ Theorem 3.4)**.** *Suppose that*

$$\lim_{n\to\infty} a_n = 1 \qquad \lim_{n\to\infty} b_n = 0 \tag{1.16}$$

$$\sum_{n=1}^{\infty} (|a_{n+1} - a_n| + |b_{n+1} - b_n|) < \infty \tag{1.17}$$

Then there is uniform clock behavior on each interval $[-2+\varepsilon, 2-\varepsilon]$ *for any* $\varepsilon > 0$.

Remarks. 1. Theorem 1.4 implies Theorem 1.3, but we state them as separate theorems since the proof of the first is easier.

2. We will also prove results of this genre for perturbations of periodic recursion coefficients and for where $n+1$ in (1.17) is replaced by $n+p$ for some p.

3. We also obtain results (see Theorem 3.8) near ± 2 if

$$\sum_{n=1}^{\infty} n(|a_n - 1| + |b_n|) < \infty \tag{1.18}$$

Using our strong lower bound, we will also prove the following:

Theorem 1.5. *Let $a_n^{(\omega)}, b_n^{(\omega)}$ be ergodic Jacobi parameters. Let E_0 be such that*

(i) *The Lyapunov exponent $\gamma(E_0) = 0$.*

(ii) *The symmetric derivative of ν exists at E_0 and is finite and nonzero.*

Then there exists $C > 0$ so that with probability 1,

$$\liminf_{n\to\infty}\, n[z_n^{(1)}(E_0) - z_n^{(-1)}(E_0)] > C \tag{1.19}$$

This result is especially interesting because it implies that the zeros cannot have Poisson behavior. It implies that Poisson behavior *and ergodicity* require $\gamma(E_0) > 0$. We will say more about these issues in Sections 4 and 11.

While our initial interest in a priori lower bounds came from clock theorems and we could have finished the paper with Section 4 if our sole purpose was to prove Theorems 1.3–1.5, it seemed natural to also consider upper bounds. Moreover, in looking over the upper/lower bound results in the OP literature, we realized one could get more from these methods, so we discuss that also.

Broadly speaking, we have two sets of results and methods. The methods rely on either transfer matrices with hypotheses on recursion coefficients or on OP methods with hypotheses on the measure. We believe that the results are of interest to both the Schrödinger operator and OP communities. Because the OP methods are unfamiliar to many Schrödinger operator theorists and are easy to prove (albeit very powerful), we have included an appendix with some major OP methods.

The detailed plan of the paper is as follows: In Section 2, we prove our a priori lower bounds involving the transfer matrix (or, more precisely, the growth of subordinate and nonsubordinate solutions). In Section 3, we prove a variety of clock theorems, including Theorems 1.3 and 1.4. In Section 4, using ideas of Deift-Simon [8], we prove Theorem 1.5. In Section 5, we obtain upper bounds on eigenvalue spacing using the transfer matrix. Section 6 discusses using suitable Prüfer angles to control spacing of zeros.

In Section 7, we begin our discussion of OP methods with a technical result on L^p bounds on w^{-1} and bounds of the Christoffel function. These bounds, which we will need for examples later are local versions of some bounds of Geronimus [11] with a rather different method of proof. In Section 8, we discuss upper bounds on eigenvalue spacing using OP methods and, in particular, find a remarkable lower bound on the density of states that is a kind of microlocal version of some bounds of Deift-Simon [8]. In Section 9, we discuss lower bounds on eigenvalue spacing. The methods in Sections 8 and 9 are borrowed from Erdös-Turán [9, 51], Nevai [34], and Golinskii [14], but we show how to localize them and squeeze out stronger results. In Section 10, we briefly discuss the analogs of our results for zeros of POPUC, and in Section 11 discuss a number of examples, counterexamples, conjectures, and questions.

It is a pleasure to thank L. Golinskii, S. Jitomirskaya, R. Killip, P. Nevai, and M. Stoiciu for useful discussions. This research was begun during B. S.'s stay as a Lady Davis Visiting Professor at The Hebrew University of Jerusalem. He would like to thank H. Farkas for the hospitality of the Einstein Institute of Mathematics at The Hebrew University, and E. B. Davies and A. N. Pressley for the hospitality of King's College, London where this was work was continued. Y. L. would like to thank G. A. Lorden and T. A. Tombrello for the hospitality of Caltech, where this work was completed.

Percy Deift has long been a player in spectral theory and more recently, a champion for orthogonal polynomials. In particular, this paper exploits the work of Deift-Simon [8]. It is a pleasure to dedicate this paper to Percy.

2. Variation of Parameters and Lower Bounds via Transfer Matrices

Our goal here is to use variation of parameters to study eigenvalue spacing. Variation of parameters has an ancient history going back to Lagrange [15] and it was extensively used to study variation of solution with change in potential, for example, to study asymptotics in tunnelling problems [16]. The usefulness of the method as a tool in spectral theory goes back at least to the work of Gilbert-Pearson [13] (also see [12, 23]) with significant later contributions by Jitomirskaya-Last [18, 19, 20] and Killip-Kiselev-Last [24]. It is essentially their equation we will use although, interestingly enough, the earlier applications are to E, E' with $E \in \mathbb{R}$ and $E' = E + i\varepsilon$, while our application is to E, E' both in $\mathbb{R}$.

Given $E \in \mathbb{C}$, we consider solutions of

$$a_n u_{n+1} + (b_n - E)u_n + a_{n-1}u_{n-1} = 0 \tag{2.1}$$

for $n = 1, 2, \dots$. Here $\{a_n, b_n\}_{n=1}^\infty$ are the Jacobi parameters of the measure we are considering and

$$a_0 \equiv 1 \tag{2.2}$$

For $\theta \in [0, \pi)$, we denote by $u_n(E, \theta)$ the solution of (2.1) with

$$u_0(E,\theta) = \sin(\theta) \qquad u_1(E,\theta) = \cos(\theta) \tag{2.3}$$

In particular,

$$u_n(E, \theta = 0) = p_{n-1}(E) \qquad n = 1, 2, \dots \tag{2.4}$$

The transfer matrix is defined by

$$T(n,E) = \begin{pmatrix} u_{n+1}(E,\theta=0) & u_{n+1}(E,\theta=\pi/2) \\ u_n(E,\theta=0) & u_n(E,\theta=\pi/2) \end{pmatrix} \tag{2.5}$$

so for any solution of (2.1),

$$\begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix} = T(n,E) \begin{pmatrix} u_1 \\ u_0 \end{pmatrix} \tag{2.6}$$

Let $K(n, m; E)$ be the kernel

$$K(n,m;E) = u_n(E,0)u_m(E,\pi/2) - u_n(E,\pi/2)u_m(E,0) \tag{2.7}$$

Define the operator $A_L(E)$ on $\mathbb{C}^L = \{v. = \{v_k\}_{k=1}^L \mid v_k \in \mathbb{C}\}$ by

$$(A_L(E)v)_n = \sum_{m=1}^n K(n,m;E)v_m \tag{2.8}$$

(note that $K(n, n) = 0$, so the sum also goes to $n - 1$). The following summarizes results from [18, 19, 20, 24]:

Theorem 2.1. *Let $w_n(E), w_n(E')$ solve (2.1) for E, E', and suppose*

$$w_0(E) = w_0(E') \qquad w_1(E) = w_1(E') \tag{2.9}$$

Then

$$w_n(E') = w_n(E) + (E' - E) \sum_{m=1}^n K(n,m;E)w_m(E') \tag{2.10}$$

that is,

$$w.(E') = w.(E) + (E' - E)(A_L(E)w(E')). \tag{2.11}$$

Moreover, with $\|\cdot\|_{\mathrm{HS}}$ the Hilbert-Schmidt norm on $\mathbb{C}^L$ and $\|\cdot\|_L$ defined by

$$\|v.\|_L = \left(\sum_{n=1}^{L} |v_j|^2\right)^{1/2} \tag{2.12}$$

and $\langle\,\cdot\,,\,\cdot\,\rangle_L$ the associated inner product, we have that

$$\|A_L(E)\|_{\mathrm{HS}}^2 = \|u(E,0)\|_L^2\|u(E,\pi/2/)\|_L^2 - |\langle u(E,0),u(E,\pi/2)\rangle_L|^2 \tag{2.13}$$

$$= \max_\theta \|u(E,\theta)\|_L^2 \min_\theta \|u(E,\theta)\|_L^2 \tag{2.14}$$

In particular,

$$\|A_L(E)\| \le \sum_{j=0}^{L-1} \|T(j,E)\|^2 \tag{2.15}$$

Remarks. 1. That (2.10)/(2.11) hold is either a direct calculation verifying the formula or a calculation obtained by expanding $\binom{w_{n+1}(E')}{w_n(E')}$ in terms of $\binom{u_{n+1}(E,\theta)}{u_n(E,\theta)}$ for $\theta = 0, \pi/2$. (2.13) is a direct calculation from (2.7) and (2.14) is a clever observation [24].

2. Clearly, $\|T(n,E)\binom{1}{0}\|^2 = |u_{n+1}(E)|^2 + |u_n(E)|^2 \ge |u_n(E)|^2$, so $\|u(E,0)\|_L^2 \le$ RHS of (2.15). Similarly for $\|u(E,\pi/2)\|$, so $\|u(E,0)\|_L\|u(E,\pi/2)\| \le$ RHS of (2.15), so (2.13) implies (2.15).

3. $\|T(n)\|$ measures the growth of the fastest growing solution, so the RHS of (2.15) in fact measures ($\max_\theta \|u(E,\theta)\|_L^2$) and thus, by (2.14), one could place $\min_\theta \|u(E,\theta)\|/\max_\theta \|u(E,\theta)\|$ in front of the RHS of (2.15).

Here is the key lower bound:

Theorem 2.2. *Let E', E'' be two distinct zeros of $p_L(x)$. Then*

$$|E'-E_0| + |E''-E_0| \ge \|A_L(E_0)\|^{-1} \tag{2.16}$$

In particular,

$$|z_L^{(+1)}(E_0) - z_L^{(-1)}(E_0)| \ge \|A_L(E_0)\|^{-1} \tag{2.17}$$

Proof. By (2.11), for $E_j = E', E''$,

$$p_{\cdot-1}(E_j) = p_{\cdot-1}(E_0) + (E_j - E_0)A_L(E_0)p_{\cdot-1}(E_j) \tag{2.18}$$

To say $p_L(E') = p_L(E'') = 0$ says E, E'' are eigenvalues of $J_{L;F}$ and $p_{\cdot-1}(\cdot)$ are the eigenvectors. So, by orthogonality of eigenvectors,

$$\langle p_{\cdot-1}(E'), p_{\cdot-1}(E'')\rangle_L = 0 \tag{2.19}$$

By interchanging E' and E'', if necessary, suppose

$$\|p_{\cdot-1}(E')\|_L \geq \|p_{\cdot-1}(E'')\|_L \tag{2.20}$$

Take (2.18) for E' and E'' and take the inner product with $p_{\cdot-1}(E')$ and subtract to get

$$\begin{aligned}\|p_{\cdot-1}(E')\|^2 &\leq |E'-E_0|\,|\langle p_{\cdot-1}(E'), A_L(E_0)p_{\cdot-1}(E')\rangle| \\ &\qquad + |E''-E_0|\,|\langle p_{\cdot-1}(E'), A_L(E_0)p_{\cdot-1}(E'')\rangle| \\ &\leq (|E'-E_0| + |E''-E_0|)\|A_L(E_0)\|\,\|p_{\cdot-1}(E')\|^2\end{aligned}$$

on account of (2.20). (2.16) is immediate.

(2.13) follows from (2.16) and $z^{(-1)} \leq E_0 \leq z^{(+1)}$. □

While our main applications are to clock theorems and Poisson statistics, there is a universal tunnelling bound.

Theorem 2.3. *Let $J_{L;F}$ be a finite Jacobi matrix with $\alpha_- = \inf a_n$, $\alpha_+ = \max a_n$, $\beta = \max b_n - \min b_n$. Let*

$$\gamma = \alpha_-^{-1}|(\beta+2\alpha_+)^2 + \alpha_+^2 + 1|^{1/2} \tag{2.21}$$

Then any pair of eigenvalues, E, E', of $J_{L;F}$ obeys

$$|E - E'| \geq \frac{\gamma^2-1}{\gamma^{2L}-1} \tag{2.22}$$

Remark. This bound is exponential, $\sim \gamma^{-2L}$, for L large.

Proof. Adding a constant to b_n does not change eigenvalue differences, so we can suppose that

$$\max b_n = -\min b_n = \frac{\beta}{2} \tag{2.23}$$

Then any E_0 in the convex hull of $\operatorname{spec}(J_{L;F})$ obeys $|E_0| \leq \frac{\beta}{2} + 2\alpha_+$, so $|E_0 - b_n| \leq \beta + 2\alpha_+$. Thus, γ is an upper bound on the Hilbert-Schmidt norm of

$$\frac{1}{a_{n+1}}\begin{pmatrix} E - b_{n+1} & -a_n \\ 1 & 0 \end{pmatrix}$$

and so on the norm.

Pick E between E' and E''. It follows that

$$\|T(j,E)\| \leq \gamma^j$$

so (2.21) follows from (2.17) and (2.15). □

One can also use our proof to see that one cannot have too many zeros near E_0.

Theorem 2.4. *Define $\delta_{L,n}(E_0)$ to be the smallest number so that*

$$\#\{z_j^{(n)} \mid |z_j^{(n)} - E_0| < \delta\} \leq n-1 \tag{2.24}$$

Then, for $n \geq 2$,

$$\delta_{L,n}(E_0) \geq \tfrac{1}{2} \, \|A_L(E_0)\|_{\text{HS}}^{-1} \sqrt{n} \tag{2.25}$$

Remarks. 1. If one has strong clock behavior, $\delta_{L,n} \sim cn/L$ for n fixed and L large, so $\sqrt{n}$ is worse than one expects in nice cases.

2. Our proof shows that (2.17) can be "improved," if $\|\cdot\|_{\text{HS}}$ is used, to

$$|z_L^{(+1)}(E_0) - z_L^{(-1)}(E_0)| \geq \frac{1}{\sqrt{2}} \, \|A_L(E_0)\|_{\text{HS}}^{-1} \tag{2.26}$$

Proof. There are at least n zeros, $z_1, \dots, z_n$, in $\{z \mid |z-E_0| \leq \delta_{L,n}\}$. Order them so that if $\varphi_{j,m} = p_m(z_j)$, then $\|\varphi_j\|_L \geq \|\varphi_2\|_L \geq \cdots \geq \|\varphi_n\|_L$. Let $\tilde{\varphi}_j = \varphi_j / \|\varphi_j\|_L$. Then the argument in the proof of Theorem 2.2 says that, for $j \leq k$,

$$\begin{aligned} 1 &\leq [\delta_{L,n}(|A_{jj}| + |A_{jk}|)]^2 \\ &\leq 2\delta_{L,n}^2(|A_{jj}|^2 + |A_{jk}|^2) \end{aligned} \tag{2.27}$$

where

$$A_{jk} = \langle \varphi_j, A_L(E_0)\varphi_k \rangle \tag{2.28}$$

and (2.27) comes from $(a+b)^2 \leq 2(a^2+b^2)$. Summing over all pairs and noting that $|A_{11}|^2$ occurs $n-1$ times, we find that

$$\frac{n(n-1)}{2} \geq 2\delta_{L,n}^2 (n-1) \|A\|_{\text{HS}}^2$$

which is (2.25). $\square$

3. Clock Theorems for Bounded Variation Perturbations of Free and Periodic OPRL

The basic result from which we will derive all our clock theorems presupposes the existence of a complex solution to the difference equation (2.1) for which we have precise information on the phase. The model is the Jost solution which is asymptotic to $e^{i\theta(E)n}$, where $E = 2\cos(\theta)$, $0 \leq \theta \leq \pi$.

Theorem 3.1. *Let $\{a_n, b_n\}_{n=1}^\infty$ be a set of Jacobi parameters and Δ a closed interval in $\mathbb{R}$. Suppose there exists a solution $u_n(E)$ of (2.1) for $E \in \Delta$ which obeys*

(i)

$$u_0(E) > 0 \qquad \operatorname{Im} u_1(E) > 0 \tag{3.1}$$

(ii)

$$u_n(E) = |u_n(E)| \exp(i[n\alpha(E) + \beta_n(E)]) \tag{3.2}$$

where α is C^1 with

$$-\frac{d\alpha}{dE} > 0 \qquad \text{all } E \in \Delta \tag{3.3}$$

each β_n continuous on Δ, and

$$\lim_{n\to\infty} \Bigg[\sup_{\substack{E',E\in\Delta \\ |E-E'|\le 1/n}} |\beta_n(E') - \beta_n(E)| \Bigg] = 0 \tag{3.4}$$

(iii) *The transfer matrix $T_n(E)$ of* (2.5) *obeys*

$$\tau \equiv \sup_{n,E\in\Delta} \|T_n(E)\| < \infty \tag{3.5}$$

Then the density of states exists on Δ,

$$d\nu(E) = -\frac{1}{\pi}\frac{d\alpha}{dE}\, dE \tag{3.6}$$

and there is clock behavior uniformly in Δ.

Remarks. 1. α is, of course, a rotation number and (3.6) an expression of the connection between the density of states and rotation numbers; see Johnson-Moser [22].

2. (3.4) implies β_n is irrelevant for eigenvalue spacing comparable to $1/n$. To control possible spacings with ΔE small compared to $1/n$, one needs some Lipschitz control of β_n, that is,

$$\sup_{E\neq E'} \frac{|\beta_n(E') - \beta_n(E)|}{n|E-E'|} \to 0$$

which is where differentiability of β_n and so moment conditions on $\{a_n, b_n\}$ came into [41]. We avoid this by using (3.5) to get a priori bounds.

3. (3.4) implies the same if $1/n$ is replaced by A/n for any fixed A. Define $\zeta_n(A)$ by

$$\zeta_n(A) = \sup_{\substack{E,E'\in\Delta \\ |E-E'|\le A/n}} |\beta_n(E') - \beta_n(E)| \tag{3.7}$$

4. (3.4) is implied by an equicontinuity assumption, for example, uniform convergence of β_n to a continuous limit.

Proof. By (3.1), u_n and $\bar u_n$ are independent solutions of (2.1) and so cannot vanish at any points. Moreover,

$$p_{n-1}(E) = A(E)u_n(E) + \overline{A(E)}\;\overline{u_n(E)} \tag{3.8}$$

where

$$A(E) = \frac{W(p_{\cdot-1}, \bar{u})}{W(u, \bar{u})} \tag{3.9}$$

Here, W is the Wronskian. Given sequences, f_n, g_n,

$$W_n(f, g) = a_n(f_{n+1}g_n - f_n g_{n+1}) \tag{3.10}$$

is constant (call it $W(f, g)$) if f, g both solve (2.1). Since $p_{-1} = 0$, $p_0 = 1$, $a_0 = 1$, we have $W(p_{\cdot-1}, \bar{u}) = u_0(E)$, and clearly, $W(u, \bar{u}) = 2i \operatorname{Im}(u_1 \bar{u}_0) = 2iu_0 \operatorname{Im} u_1$, so

$$A(E) = -\frac{i}{2} \frac{1}{\operatorname{Im} u_1} \tag{3.11}$$

is pure imaginary. Thus, (3.8) says $p_{n-1}(E)$ vanishes if and only if u_n is real, that is, by (3.2),

$$p_{n-1}(E) = 0 \Leftrightarrow n\alpha(E) + \beta_n(E) = k\pi \qquad (k \in \mathbb{Z}) \tag{3.12}$$

Let

$$\eta = \min_{E \in \Delta} \left[-\frac{d\alpha}{dE} \right] \tag{3.13}$$

Pick N_0 so that $n > N_0$ implies

$$\frac{\eta}{\tau^2} \geq 4\zeta_n\left(\frac{1}{\tau^2}\right) \tag{3.14}$$

This can be done since $\zeta_n(A) \to 0$ as $n \to \infty$ by hypothesis. Since $\zeta_n(A)$ is increasing in A and $\zeta_n(A + B) \leq \zeta_n(A) + \zeta_n(B)$, we have $\zeta_n(x\beta) \leq \zeta_n([x]\beta) + \zeta_n(x\beta - [x]\beta) \leq ([x]+1)\zeta_n(\beta) \leq 2x\zeta_n(\beta)$, if $x \geq 1$, (3.14) implies

$$q \geq \frac{1}{\tau^2} \Rightarrow \eta q \geq 2\zeta_n(q) \tag{3.15}$$

This in turn implies

$$E - E' \geq \frac{1}{n\tau^2} \Rightarrow \big[n[\alpha(E') - \alpha(E)] + [\beta_n(E') - \beta_n(E)]\big] \geq \frac{n}{2}\,\eta|E' - E| \tag{3.16}$$

By (2.17) and (2.15), any two successive zeros obey

$$|E' - E| \geq \frac{1}{n\tau^2} \tag{3.17}$$

Thus, (3.17) implies that for $n > N_0$, any two solutions of (3.12) have distinct values of k. We also see from (3.16) and continuity that if E is a solution of (3.12), there is another solution in $(E - \frac{2\pi}{n\eta}, E)$ and it has the next larger value of k (i.e., $k + 1$).

Subtracting (3.12) for two successive values of (3.12) and using $\zeta_n(2\pi/\eta) \to 0$, we see that

$$\sup_{\substack{E'<E \text{ successive} \\ \text{eigenvalues in } \Delta}} |n(\alpha(E') - \alpha(E)) - \pi| \to 0 \tag{3.18}$$

Given the uniformity of convergence of the difference quotient to the derivative, (3.18) implies that

$$\sup_{\substack{E'<E \text{ successive} \\ \text{eigenvalues in } \Delta}} \left| n(E' - E)\, \frac{d\alpha}{dE} - \pi \right| \to 0 \tag{3.19}$$

This implies the density of states exists and is given by (3.6) and that one has uniform clock behavior. □

Theorem 3.2 (≡ Theorem 1.3). *Let $\{a_n, b_n\}_{n=0}^{\infty}$ be a set of Jacobi parameters obeying*

$$\sum_{n=1}^{\infty} |b_n| + |a_n - 1| < \infty \tag{3.20}$$

so $\operatorname{ess\,supp}(d\mu) = [-2, 2]$. *For any $\varepsilon > 0$, we have uniform clock behavior on $[-2 + \varepsilon, 2 - \varepsilon]$.*

Remarks. 1. This includes Jacobi polynomials (rescaled to $[-2, 2]$) for which

$$|b_n| + |a_n - 1| = O(n^{-2})$$

2. Of course, the density of states is the free one.

Proof. It is well-known (see, e.g., [26, 6]) that when (3.20) holds, there exists, for all $x \in (-2, 2)$, a solution $\tilde{u}_n(x)$ so that if

$$z + z^{-1} = x \tag{3.21}$$

(i.e., $z = e^{i\theta}$ with $x = 2\cos\theta$ and $\theta \in (0, \pi)$), then

$$z^{-n}\tilde{u}_n(x) \to 1 \tag{3.22}$$

uniformly on compact subsets of $(-2, 2)$. Moreover, $\tilde{u}_n(x)$ is continuous on $(-2, 2)$ for each fixed n. By evaluating the Wronskian near $n = \infty$, we see

$$W(\tilde{u}_n, \bar{\tilde{u}}_n) = z - z^{-1} \tag{3.23}$$

Thus, if

$$u_n(E) = \frac{\tilde{u}_n(E)}{\tilde{u}_0(E)} \tag{3.24}$$

then (3.1) and (3.2) hold. If

$$u_0(E) = |u_0(E)| e^{i\beta_\infty(E)} \tag{3.25}$$

then we have that

$$\beta_n(E) \to \beta_\infty(E) \tag{3.26}$$

uniformly on $[-2+\varepsilon, 2-\varepsilon]$. By continuity, (3.4) holds.

α is given by

$$\alpha(E) = \arccos\biggl(\frac{E}{2}\biggr) \tag{3.27}$$

so

$$-\frac{1}{\pi}\,\frac{d\alpha}{dE} = \frac{1}{\pi}\,\frac{1}{\sqrt{4-E^2}} \tag{3.28}$$

and (3.3) holds.

Finally, standard variation of parameters about z^n, z^{-n} shows that (3.5) holds for each $\Delta = [2+\varepsilon, 2-\varepsilon]$.

Thus, Theorem 3.1 applies, and we have clock behavior. $\square$

In the above, we used the fact that $\beta_n \to \beta_\infty$ uniformly to obtain (3.4). In the bounded variation case, we will instead use:

Lemma 3.3. *If $\beta_n = \beta_n^{(1)} + \beta_n^{(2)}$ where $\beta_n^{(1)}$ is C^1 and*

$$\frac{1}{n}\sup_{E\in\Delta}\biggl|\frac{\partial\beta_n^{(1)}}{\partial E}\biggr| \to 0 \tag{3.29}$$

and $\beta_n^{(2)} \to \beta_\infty^{(2)}$ uniformly, then (3.4) holds.

Proof. Immediate since $E, E' \in \Delta$ implies

$$|\beta_n(E) - \beta_n(E')| \le |E-E'| \sup_{E\in\Delta}\biggl|\frac{\partial\beta_n^{(1)}}{\partial x}\biggr| + |\beta_n^{(2)}(E) - \beta_n^{(2)}(E')| \quad \square$$

Theorem 3.4 ($\equiv$ Theorem 1.4)**.** *Let $\{a_n, b_n\}_{n=0}^\infty$ be a set of Jacobi parameters with*

$$a_n \to 1 \qquad b_n \to 0 \tag{3.30}$$

and

$$\sum_{n=1}^\infty (|a_{n+1}-a_n| + |b_{n+1}-b_n|) < \infty \tag{3.31}$$

then for any $\varepsilon > 0$, we have uniform clock behavior in $[-2+\varepsilon, 2+\varepsilon]$.

Remark. Again, the density of states is the free one by (3.30).

In order to prove this theorem, we need the following result:

Theorem 3.5. *Let $B_n(\theta)$ depend continuously on $\theta \in I$, a compact subinterval of $(0,\pi)$, and suppose*

$$\sup_{\theta\in I}\sum_{n=1}^\infty \|B_{n+1}(\theta) - B_n(\theta)\| < \infty \tag{3.32}$$

and

$$B_n(\theta) \to B_\infty(\theta) \tag{3.33}$$

uniformly where $B_\infty(\theta)$ *has eigenvalue* $e^{\pm i\theta}$. *Explicitly for* $V_\infty(\theta)$ *continuous and invertible:*

$$B_\infty(\theta) = V_\infty(\theta) \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix} V_\infty(\theta)^{-1} \tag{3.34}$$

Suppose each $B_n(\theta)$ *has eigenvalues* $e^{\pm i\varphi_n(\theta)}$ *with* $\varphi_n(\theta) \in (0, \pi)$. *Let*

$$\tilde{T}_n(\theta) = B_n(\theta) \dots B_1(\theta) \tag{3.35}$$

Then

(i)

$$\sup_{\theta \in I, n} \|\tilde{T}_n(\theta)\| < \infty \tag{3.36}$$

(ii) *There exists* $S_\infty(\theta)$ *depending continuously on* θ *so that*

$$\|\tilde{T}_n(\theta) - V_\infty(\theta) D_n(\theta) S_\infty(\theta)\| \to 0 \tag{3.37}$$

where

$$D_n(\theta) = \begin{pmatrix} e^{i\eta_n(\theta)} & 0 \\ 0 & e^{-i\eta_n(\theta)} \end{pmatrix} \tag{3.38}$$

with

$$\eta_n(\theta) = \sum_{\eta=1}^{n} \varphi_n(\theta) \tag{3.39}$$

Remark. See the notes to Section 2.1 of [40] for a history of results on bounded variation.

Proof. This is a strong version of Kooman's theorem [29]. In Section 12.1 of [40], (3.36) is proven, and in the notes to that section, it is noted that (3.37) holds. In those notes, there is no V_∞ because the analog of D_n is not diagonal but can be diagonalized in a basis where B_∞ is diagonal. □

Proof of Theorem 3.4. Define

$$B_n(\theta) = \begin{pmatrix} (a_n a_{n+1})^{-1/2}(2\cos(\theta) - b_{n+1}) & -(a_n/a_{n+1})^{1/2} \\ (a_n/a_{n+1})^{-1/2} & 0 \end{pmatrix} \tag{3.40}$$

so the transfer matrix at $E = 2\cos\theta$ is

$$\begin{aligned} T_n(\theta) &= \left(\frac{a_n}{a_{n+1}}\right)^{1/2} B_n(\theta) \left(\frac{a_{n-1}}{a_n}\right)^{1/2} \dots B_1(\theta) \\ &= a_{n+1}^{-1/2} \tilde{T}_n(\theta) \end{aligned} \tag{3.41}$$

Since $\det(B_n) = 1$, $B_n(\theta)$ has eigenvalues $e^{\pm i\varphi_n}$, $\varphi_n \in (0, \pi)$, if and only if

$$(a_n a_{n+1})^{-1/2}(2\cos\theta - b_{n+1}) = 2\cos(\varphi_n) \in (-2, 2) \tag{3.42}$$

and this holds uniformly for $\theta \in (\delta, \pi - \varepsilon)$ and $n > N_0$ for some fixed N_0. Thus, for $n \geq N_0$, we define

$$\eta_n(\theta) = \sum_{j=N_0}^{n} \varphi_j(\theta) \tag{3.43}$$

and define

$$\tilde{D}_n(\theta) = \begin{pmatrix} e^{i\eta_n(\theta)} & 0 \\ 0 & e^{-i\eta_n(\theta)} \end{pmatrix} \tag{3.44}$$

and

$$\tilde{S}_\infty(\theta) = S_\infty(\theta; B_n \dots B_{N_0}) B_{N_0-1} \dots B_1 \tag{3.45}$$

where $S_\infty(\theta; B_n \dots B_{N_0})$ is the S_∞ for the sequence $B_{N_0}, B_{N_0+1}, \dots$. Thus, (3.37) and (3.41) show that

$$\|T_n(\theta) - a_{n+1}^{-1/2} V_\infty(\theta)\tilde{D}_n(\theta)\tilde{S}_\infty(\theta)\| \to 0 \tag{3.46}$$

We now proceed to construct a solution u obeying the hypothesis of Theorem 3.1. Pick a two-vector $x(\theta)$ by $x(\theta) = e^{i\varphi_0(\theta)}\tilde{S}_\infty(\theta)^{-1}\binom{1}{0}$ where φ_0 is chosen below and u_n by

$$T_n(\theta)x(\theta) = \begin{pmatrix} u_{n+1}(\theta) \\ u_n(\theta) \end{pmatrix} \tag{3.47}$$

then (3.46) says that

$$u_n(\theta) = |u_n(\theta)| \exp(i[\eta_n(\theta) + \beta_n^{(2)}(\theta)]) \tag{3.48}$$

where

$$\beta_n^{(2)}(\theta) \to \beta_\infty^{(2)}(\theta) \tag{3.49}$$

uniformly.

Here $\beta_\infty^{(2)}(\theta)$ is $\varphi_0(\theta)$ plus the phase of the 21 element of $V_\infty(\theta)$, and so $i(\eta_n(\theta) + \beta_\infty^{(2)}(\theta))$ is the phase of the lower component of $a_{n+1}^{-1/2} V_\infty(\theta)\tilde{D}_n(\theta)\tilde{S}_\infty(\theta)x(\theta)$. Since u_n is not real, $u_0(\theta) \neq 0$, and so $\varphi_0(\theta)$ can be chosen so that $u_0(\theta) > 0$.

Since $\varphi_j(\theta) \to \theta$, $\eta_{n+1} - \eta_n \to \theta$, and thus the imaginary part of the Wronskian of u and $\bar{u}$ is positive, so u obeys (3.1).

By (3.42),

$$\frac{\partial\varphi_n}{\partial E} = \frac{\sin(\theta)}{\sin(\varphi_n)}\,(a_n a_{n+1})^{-1/2}\,\frac{\partial\theta}{\partial E} \tag{3.50}$$

so

$$\frac{\partial \eta_n}{\partial E} = \frac{\partial \theta}{\partial E}\,(n - N_0) + \sum_{j=N_0}^{n} \biggl(\frac{\sin(\theta)}{\sin(\varphi_n)}\,(a_n a_{n+1})^{-1/2} - 1 \biggr) \tag{3.51}$$

Thus, if

$$\beta_n^{(1)} = \eta_n - n\,\frac{\partial \theta}{\partial E} \tag{3.52}$$

then

$$\frac{1}{n}\,\frac{\partial \beta^{(1)}}{\partial E} = \biggl(-\frac{N_0}{n} + \frac{1}{n} \sum_{j=N_0}^{n} \biggl(\frac{\sin(\theta)}{\sin(\varphi_n)}\,(a_n a_{n+1})^{-1/2} - 1 \biggr) \biggr) \frac{\partial \theta}{\partial E} \tag{3.53}$$

converges uniformly to zero, since

$$\frac{\sin(\theta)}{\sin(\varphi_n)} \to 1 \qquad (a_n a_{n+1})^{-1/2} \to 1 \tag{3.54}$$

uniformly in θ.

Lemma 3.3 applies, so condition (ii) of Theorem 3.1 holds with $\alpha(E) = \theta(E) = \arccos(\frac{E}{2})$. We thus have clock behavior with density of states the free one, that is, given by (3.28). $\square$

Now consider the periodic case, that is,

$$a_{n+p} = a_n \qquad b_{n+p} = b_p \tag{3.55}$$

The spectrum now has bands (see, e.g., [30]). For any E in the interior of the bands, there is a Floquet solution with $u_{n+p} = e^{i\gamma(E)} u_n$ with $\gamma(E) \in (0,\pi)$ and $\frac{\partial \gamma}{\partial E} < 0$. Thus

$$u_{Lp+r} = |u_r| e^{i(L\gamma(E) + \beta_r^{(\infty)})} \tag{3.56}$$

where $\beta_r^{(\infty)}$ is the phase of u_r. Theorem 3.1 applies with $\alpha = \gamma(E)/p$ and

$$\beta_{Lp+r}(E) = \beta_r^{(\infty)}(E) - \frac{r}{p}\,\gamma(E) \tag{3.57}$$

There are only r such functions so (3.4) holds, and we recover the zero spacing part of Theorem 2.6 of [43].

If $a_n^{(0)}, b_n^{(0)}$ are periodic and $a_n = a_n^{(0)} + \delta a_n$, $b_n = b_n^{(0)} + \delta b_n$ and

$$\sum_{n=1}^{\infty} (|\delta b_n| + |\delta a_n|) < \infty \tag{3.58}$$

then one can construct Jost solutions on the interiors of the bands. All that changes is that (3.57) is replaced by

$$\lim_{L\to\infty} \beta_{Lp+r}(E) = \beta_r^{(\infty)}(E) - \frac{r}{p}\,\gamma(E) \tag{3.59}$$

so Theorem 3.1 still applies. Similarly applying the ideas in the proof of Theorem 3.4, we obtain a bounded variation result. Since it includes the (3.58) result, we summarize in a single theorem:

Theorem 3.6. *Let $a_n^{(0)}, b_n^{(0)}$ obey*

$$a_{n+p}^{(0)} = a_n^{(0)} \qquad b_{n+p}^{(0)} = b_n^{(0)} \tag{3.60}$$

for some p. Let a_n, b_n obey

$$\lim_{n\to\infty} |a_n - a_n^{(0)}| + |b_n - b_n^{(0)}| = 0 \tag{3.61}$$

$$\sum_{n=1}^{\infty} (|a_{n+p} - a_n| + |b_{n+p} - b_n|) < \infty \tag{3.62}$$

Then, for any closed interval which is a subset of the interior of the bands (see Remark), we have uniform clock behavior with density of states $-\frac{1}{p}\frac{d\gamma}{dE}$.

Remark. There are p closed bands, $B_1, \dots, B_p$, generically disjoint but perhaps touching in a single point (closed gap). By the interior of the bands, we mean $\cup_{j=1}^p B_j^{\text{int}}$ which will be smaller than $(\cup_{j=1}^p B_j)^{\text{int}}$ if some gap is closed, that is, we must remove all of the gaps, including those that degenerate to single points.

p need not be the minimal period, so we have that

Corollary 3.7. *Suppose*

$$a_n \to 1 \qquad b_n \to 0$$

and, for some p, (3.62) *holds. Then, in any closed interval in* $\{E = 2\cos(\theta) \mid p\theta \neq 0, \frac{\pi}{p}, \dots, \frac{(p-1)\pi}{p}\}$*, we have uniform clock behavior.*

As a final topic, we want to discuss zeros very near $E = 2$ when

$$\sum_{n=1}^{\infty} n(|a_n - 1| + |b_n|) < \infty \tag{3.63}$$

It should be possible to extend this argument to get uniform clock behavior in $[-2, 2]$, with a suitable modification to take into account the behavior exactly at ± 2. When (3.63) holds, the Jost function, u, can de defined on $[-2, 2]$; see, for example, the appendix to [6]. If $u(2) = 0$, we say there is a resonance at 2, and if $u(2) \neq 0$, we say that 2 is nonresonant.

Theorem 3.8. *Let $\{a_n, b_n\}_{n=1}^\infty$ be a set of Jacobi parameters obeying* (3.63)*. Define $0 \leq \theta_1^{(n)} < \theta_2^{(n)} < \cdots$ so $E_j^{(n)} = 2\cos(\theta_j^{(n)})$ are the zeros of $p_n(x)$ nearest to $x = 2$ and below. Then*

(a) *If* 2 *is a resonance, then*

$$n\theta_j^{(n)} \to (j - \tfrac{1}{2})\pi \tag{3.64}$$

(b) *If* 2 *is nonresonant, then*

$$n\theta_j^{(n)} \to j\pi \tag{3.65}$$

Remarks. 1. The two simplest examples are the nonresonant $a_n \equiv 1$, $b_n \equiv 0$ where

$$p_n(2\cos\theta) = c_n \frac{\sin((n+1)\theta)}{\sin\theta} \tag{3.66}$$

and the resonant $a_n \equiv 1$ $(n \geq 2)$, $a_1 = \sqrt{2}$, and $b_n \equiv 0$, where

$$p_n(2\cos\theta) = d_n \cos(n\theta) \tag{3.67}$$

2. Notice that (for simplicity, consider the nonresonant case)

$$E_j^{(n)} - E_{j+1}^{(n)} \sim \frac{(2j+1)\pi^2}{n^2}$$

the eigenvalue space is $O(n^{-2})$ and not equal in E but has clock spacing in θ.

Remark. The key fact that at a zero energy resonance, the scattering phase is $\pi/2$ (mod π) and otherwise it is 0 (mod π) is well-known in the continuum case, for which there is extensive physics literature; see, for example, Newton [36].

Proof. By the theorems found in the appendix to [6] (which codifies well-known results), when (3.63) holds, one has ([6, eqn. (A.27)])

$$|p_n(e^{i\theta})| \leq C(n+1) \tag{3.68}$$

and the existence of a solution $u_n(e^{i\theta})$ with

$$e^{-in\theta} u_n(e^{i\theta}) \to 1 \tag{3.69}$$

uniformly on $\partial\mathbb{D}$. u_0 is called the Jost function and

$$W(p_{\cdot-1}, u) = u_0 \tag{3.70}$$

We want to use (2.14) where there is a collision of notation, so we let $v(\theta,\varphi)$ be the solution at $E = 2\cos\theta$ and boundary condition φ. Then (2.14) becomes

$$\|A_L(2\cos\theta)\| \leq \max_\varphi \|v(\theta,\varphi)\|_L \min_\varphi \|v(\theta,\varphi)\|_L \tag{3.71}$$

If $u_0(\theta = 0) \neq 0$, we get one solution for θ small, $v(\theta, \varphi_0(\theta))$ which is uniformly bounded in θ and n, and another solution $v(\theta, \varphi_1 = 0)$ $(= p_{\cdot-1})$ bounded by Cn. It follows that

$$\min_\varphi \|v(\theta,\varphi)\|_L \leq CL^{1/2} \qquad \max_\varphi \|v(\theta,\varphi)\| \leq CL^{3/2} \tag{3.72}$$

If $u_0(\theta = 0) = 0$, we start at $n = 1$ since $u_1(\theta = 0) \neq 0$ and construct the bounded and linearly growing solution that way (in essence, the two solutions in this case are $p_{\cdot-1}$ and q, where q is the second kind polynomial), so (3.72) still holds.

We conclude, using Theorem 2.2, that when (3.63) holds, then

$$\sup_{\substack{E,E' \text{ successive zeros of } p_n \\ E,E' \in [2-\varepsilon,2]}} |E - E'| \geq \frac{c}{n^2} \tag{3.73}$$

Define $\varphi(\theta)$ by

$$u_0(e^{i\theta}) = |u_0(e^{i\theta})| e^{i\varphi(\theta)} \tag{3.74}$$

φ can be defined by $\theta \neq 0$ since $u_0(\theta)$ is then nonzero. We can pick θ continuous on $(0, \varepsilon)$. We claim

$$\varphi(0) \equiv \lim_{\theta \downarrow 0} \varphi(\theta) = \begin{cases} 0 & \text{mod } \pi \text{ if 2 is nonresonant} \\ \frac{\pi}{2} & \text{mod } \pi \text{ if 2 is a resonance} \end{cases} \tag{3.75}$$

Postponing the proof of this for now, let us complete the proof of the theorem.

By (3.70) and $W(u, \bar{u}) = z - z^{-1}$ (since $u \sim z^n$), we see that

$$p_n(2\cos\theta) = \frac{\overline{u_0(e^{i\theta})}\, u_n(e^{i\theta}) - u_0(e^{i\theta})\, \overline{u_n(e^{i\theta})}}{2i\sin(\theta)} \tag{3.76}$$

$$= \frac{|u_0(e^{i\theta})|\, |u_n(e^{i\theta})|}{2\sin\theta} \sin(n\theta + \beta_n(\theta)) \tag{3.77}$$

where

$$\beta_n(\theta) \to -\varphi(\theta) \tag{3.78}$$

as n goes to infinity uniformly in a neighborhood of $\theta = 0$. Thus, zeros of $p_n(2\cos\theta)$ are given as solutions of

$$n\theta + \beta_n(\theta) = j\pi \tag{3.79}$$

In the resonant case, since

$$\sup_{|\theta| \leq \frac{c}{n}} \left| \beta_n(\theta) - \frac{\pi}{2} \right|_{\text{mod } \pi} \to 0 \tag{3.80}$$

there is at least one solution asymptotic with

$$n\theta_j^{(n)} \to (j - \tfrac{1}{2})\pi \tag{3.81}$$

If there were multiple solutions for some j, we would have two zeros with $0 < \theta' < \theta$,

$$n(\theta - \theta') \to 0 \qquad \theta < \frac{\pi}{n}$$

for n large (π can be any number strictly larger than $\pi/2$),

$$\begin{aligned}|2\cos\theta' - 2\cos\theta| &\leq |\theta' - \theta|\sin(\theta) \\ &= o\left(\frac{1}{n}\right)O\left(\frac{1}{n}\right)\end{aligned}$$

violating (3.73). Thus, there are unique solutions and (3.64) holds.

In the nonresonant case, (3.79) holds, but instead

$$\sup_{|\theta|\leq\frac{c}{n}} |\beta_n(\theta)|_{\text{mod }\pi} \to 0 \tag{3.82}$$

which proves existence of solutions with

$$n\theta_j^{(n)} \to j\pi \tag{3.83}$$

for $j = 1, 2, \dots$. We must prove uniqueness for $j \geq 1$ and nonexistence for $j = 0$.

The uniqueness argument for $j \geq 1$ is the same as in the resonant case. To show no solution with $j = 0$, we suppose that J has m eigenvalues above $E = 2$ (by Bargmann's bound [17], the number is finite). Let $J(\lambda)$ be the Jacobi matrix with

$$\begin{aligned}a_n(\lambda) &= a_n \\ b_n(\lambda) &= \begin{cases} b_n + \lambda & n \leq m+1 \\ b_n & n > m+1 \end{cases}\end{aligned}$$

It is easy to see that as $\lambda \to \infty$, $J(\lambda)$ has at least $m+1$ eigenvalues. So pick $\lambda_0 = \inf\{\lambda \mid J(\lambda) \text{ has } m+1 \text{ eigenvalues in } (2,\infty)\}$. Then $\lambda_0 > 0$ and $J(\lambda_0)$ has a resonance at 2. By the analysis of the resonant case, $p_n^{(\lambda_0)}(x)$ has m zeros in $(2,\infty)$ and its $(m+1)$st zero asymptotic to $2 - (\frac{1}{2}\frac{\pi}{n})^2$, which means $p_n(x)$, whose zeros are less than those of $p_n^{(\lambda_0)}$, cannot have a zero asymptotic to $\theta^{(n)} \to 0$.

That proves the result subject to (3.75). In the nonresonant case, $u_0(e^{i\theta})$ is continuous and nonvanishing at $\theta = 0$, and $u_0(1)$ is real, $\varphi(0) \equiv 0 \text{ mod } \pi$, so continuity proves the top half of (3.75).

In the resonant case, we note that a Wronskian calculation (see [6, eqn. (A.49)]) shows that

$$\operatorname{Im}(u_1(e^{i\theta})u_0(e^{i\theta})) = \sin\theta \tag{3.84}$$

Since $u_0(1) = 0$, $u_1(1) \neq 0$, and $u_1(1)$ is real, so

$$\lim_{\theta\downarrow 0} \operatorname{Im}\left(\frac{u_0(e^{i\theta})}{\sin\theta}\right) \neq 0 \tag{3.85}$$

On the other hand, $u_0(e^{-i\theta}) = \overline{u_0(e^{i\theta})}$, so $\varphi(\theta) + \varphi(-\theta) \equiv 0 \bmod \pi$, which means that any limit point of $\varphi(\theta)$ is a multiple of $\pi/2$. This is only consistent with (3.85) if the limit is congruent to $\pi/2 \bmod \pi$. $\square$

4. Lower Bounds in the Ergodic Case

Our main goal in this section is to prove the following:

Theorem 4.1 ($\equiv$ Theorem 1.5)**.** *Let $a_n^{(\omega)}, b_n^{(\omega)}$ be ergodic Jacobi parameters. Let E_0 be such that*
(i) *The Lyapunov exponent $\gamma(E_0) = 0$.*
(ii) *The symmetric derivative of ν exists at E_0 and is finite and nonzero.*

Then there exists $C > 0$ so that with probability 1,

$$\liminf_{n\to\infty} n[z_n^{(1)}(E_0) - z_n^{(-1)}(E_0)] > C \tag{4.1}$$

This is particularly interesting because of the connection to Poisson behavior, where:

Definition. We say a probabilistic family of Jacobi matrices has Poisson behavior at E_0 if and only if for some λ (normally $\lambda =$ density of zeros) we have that for any $\alpha_1 < \beta_1 \le \alpha_2 < \beta_2 \le \cdots < \beta_\ell$ and any $k_1, k_2, \dots, k_\ell \in \{0, 1, 2, \dots\}$,

$$\begin{aligned} \text{Prob}\biggl(\#\biggl\{z_n^{(j)}(E_0) \in \biggl[E_0 + \frac{\alpha_m}{n}, E_0 + \frac{\beta_m}{n}\biggr]\biggr\} = k_m \text{ for } m = 1, \dots, \ell\biggr) \\ \to \prod_{m=1}^{\ell} \frac{[\lambda(\beta_m - \alpha_m)]^{k_m}}{k_m!}\, e^{-\lambda(\beta_m - \alpha_m)} \end{aligned} \tag{4.2}$$

Remark. Poisson behavior was proven in certain random Schrödinger operators by Molchanov [33] and for random Jacobi matrices by Minami [32]. See Stoiciu [47, 48] and Davies-Simon [7] for related work on OPUC.

Corollary 4.2. *Let $a_n^{(\omega)}, b_n^{(\omega)}$ be ergodic Jacobi parameters and $E_0 \in \mathbb{R}$ so that the symmetric derivative of ν exists at E_0 and is finite. Suppose there is Poisson behavior at E_0. Then $\gamma(E_0) > 0$.*

Remarks. 1. Basically, (4.1) is a rigid level repulsion inconsistent with Poisson behavior.

2. Ergodicity is critical here. Killip-Stoiciu [27] have examples which are not ergodic for which there is Poisson behavior with $\gamma = 0$.

Proof of Corollary 4.2. Suppose first $\gamma(E_0) = 0$ so Theorem 4.1 applies. For each n, let $f_n(\omega)$ be the characteristic function of $\{\omega \mid |z_n^{(1)}(E_0) - z_n^{(-1)}(E_0)| \leq \frac{1}{2}Cn^{-1}\}$. By (4.1), $f_n(\omega) \to 0$ for a.e. ω, so

$$\int f_n(\omega)\, d\omega \to 0 \tag{4.3}$$

as $n \to \infty$.

Clearly, if there is one $z^{(j)}$ in $[E_0 - \frac{1}{4}Cn^{-1}, E_0]$ and one in $[E_0, E_0 + \frac{1}{4}Cn^{-1}]$, then $f_n(\omega) = 1$. Thus, by the assumption of Poisson behavior,

$$\lim \int f_n(\omega)\, d\omega \geq \left(\frac{\lambda C}{4}\, e^{-\lambda C/4}\right)^2$$

which contradicts (4.3). □

Our proof of Theorem 4.1 will use the complex solutions constructed by Deift-Simon [8] and the estimate Theorem 2.2. It is thus important to be able to estimate A in terms of any pair of solutions with Wronskian 1.

Lemma 4.3. *Let $u^{(1)}, u^{(2)}$ be any pair of solutions of* (2.1) *with Wronskian* 1. *Then the kernel K of* (2.7) *has the form*

$$K(n,m,E) = u_n^{(1)} u_m^{(2)} - u_n^{(2)} u_m^{(1)} \tag{4.4}$$

In particular,

$$\|A_L(E)\|_{\text{HS}}^2 \leq \|u^{(1)}\|_L^2 \|u^{(2)}\|_L^2 \tag{4.5}$$

Proof. Noting that $u_n^{(j)} u_m^{(j)} - u_m^{(j)} u_n^{(j)} = 0$, we see that K is invariant under linear changes of the u's of determinant 1. This proves (4.4). (4.5) follows as in Theorem 2.1. □

We need the following result of Deift-Simon [8]:

Theorem 4.4. *Let $a_n^{(\omega)}, b_n^{(\omega)}$ be ergodic Jacobi parameters and let* (i)–(ii) *of Theorem 4.1 hold for E_0. Then for a.e. ω, there exists a complex-valued solution $u(\,\cdot\,,\omega)$ of* (2.1) *so that*

(i) *The Wronskian of u and $\bar u$ is $-2i$.*

(ii) *We have that*

$$\limsup_{n\to\infty} |n|^{-1} \sum_{j=0}^{n-1} |u(j,w)|^2 \leq 2\pi\, \frac{d\nu}{dE} \tag{4.6}$$

Remarks. 1. In fact, $|u(j,w)| = |u(0,T^j w)|$ and $\mathbb{E}(|u(0,w)|^2) < \infty$, so by the Birkhoff ergodic theorem, the lim sup on the left of (4.6) can be replaced by a limit which is a.e. constant with a constant bounded by the right side of (4.6)

2. [8] states results for a.e. E with $\gamma(E) = 0$, but the proof shows that what is needed is (i)–(ii).

Proof of Theorem 4.1. By (4.5), taking $u^{(1)} = u$ and $u^{(2)} = (2i)^{-1}\bar{u}$, we see

$$\limsup n^{-1}\|A_n(E)\| \leq \pi \frac{d\nu}{dE} \tag{4.7}$$

by (4.6). Thus, by Theorem 2.2,

$$\liminf_{n\to\infty} n[z_n^{(1)}(E_0) - z_n^{(-1)}(E_0)] \geq \frac{1}{\pi}\left[\frac{d\nu}{dE}\right]^{-1}$$

which is (4.1). $\square$

5. Upper Bounds via Transfer Matrices

Our goal in this section is to prove:

Theorem 5.1. *Let $\{a_n, b_n\}_{n=1}^\infty$ be a set of Jacobi parameters. For any bounded interval $I \subset \mathbb{R}$, we have*

$$\sup_{\substack{E,E' \text{ successive} \\ \text{zeros of } p_n \text{ in } I}} |E - E'| \leq \frac{[8e \sup_{E\in I} \|T_n(E)\|](\prod_{j=1}^n a_j)^{1/n}}{n} \tag{5.1}$$

Corollary 5.2. *Let $\{a_n, b_n\}_{n=1}^\infty$ be a set of Jacobi parameters and $I = [\alpha, \beta]$ a closed interval. Let*

$$A = \sup_n a_n < \infty \qquad T = \sup_{n;\, E\in I} \|T_n(E)\| < \infty \tag{5.2}$$

Let $c = 8eAT$. Then for any $\delta > 0$, there is N_0 so that if $n > N_0$ and $E \in [\alpha + \delta, \beta - \delta]$ is a zero of p_n, then there are at least two additional zeros in $[E - \frac{c}{n}, E + \frac{c}{n}]$, one above E and one below.

Proof of Corollary 5.2. It is known that any point in $\operatorname{spec}(J)$ is a limit point of zeros, so for large enough N_0, there are zeros in $[\alpha, \alpha + \delta)$ and $(\beta - \delta, \beta]$. Thus, (5.1) implies the result. $\square$

Example 5.3. Let $a_n \equiv 1$, $b_n \equiv 0$ so

$$p_n(2\cos\theta) = \frac{\sin((n+1)\theta)}{\sin\theta} \tag{5.3}$$

and for n odd, $p_n(0) = 0$. The next nearest zero is at $\theta = \frac{\pi}{2} - \frac{\pi}{n+1}$, so at $E \sim \frac{2\pi}{n+1}$ $\left(\frac{d}{d\theta}(2\cos\theta)\big|_{\theta=\frac{\pi}{2}} = -2\right)$. In this case, p_n at $E = 0$ $(\theta = \frac{\pi}{2})$ is $(1, 0, -1, 0, 1, 0, \dots)$ for $n = 0, 1, 2, \dots$ and $q_n = (0, -1, 0, 1, \dots)$, so $T_{n\text{ odd}} = \left(\begin{smallmatrix} 0 & +1 \\ \mp 1 & 0 \end{smallmatrix}\right)$ so $\|T_n(0)\| = 1$ and $\|T_n(E)\| \sim 1$ for E near 0. Thus, the correct answer for the spacing is $2\pi \sim 6.3$ and our upper bound is $8e \sim 21.7$, a factor of about 3.5 too large. $\square$

To get Theorem 5.1, we will use

Theorem 5.4. *Let Q be a polynomial with all its zeros real. Let $Q(E_0)=0$, $Q'(E_1)=0$, with $E_0<E_1$ and Q nonvanishing on (E_0,E_1). Then*

$$|E_1-E_0|\leq e\,\frac{|Q(E_1)|}{|Q'(E_0)|} \tag{5.4}$$

Proof. Since $\frac{d^2}{dx^2}\log(x-x_0)=-\frac{1}{|x-x_0|^2}<0$, we see that $g(E)=\log(|Q(E)|/|E-E_0|)$ is concave on $[E_0,E_1]$. Note that $g(E_0)=\log|Q'(E_0)|$ and that the tangent to g at E_1 is

$$\log\biggl(\frac{|Q(E_1)|}{|E_1-E_0|}\biggr)-\frac{1}{|E_1-E_0|}\,(E-E_1) \tag{5.5}$$

Thus,

$$|Q'(E_0)|\leq\biggl(\frac{|Q(E_1)|}{|E_1-E_0|}\biggr)e$$

which is (5.4). □

The Q we will take to get (5.1) is not P, but the discriminant

$$\Delta_n(E)=\operatorname{Tr}(T_n(E)) \tag{5.6}$$

associated to the periodic set of Jacobi parameters

$$a^{(n)}_{mn+q}=a_q \qquad b^{(n)}_{mn+q}=b_q \qquad q=1,\dots,n;\ m\geq 0 \tag{5.7}$$

We have:

Lemma 5.5. *The zeros of p_{n-1} and Δ_n interlace. Thus, if $E_1<E_2<E_3$ are three successive zeros of Δ_n, then p_{n-1} has two zeros, E and E', with*

$$|E-E'|<|E_3-E_1| \tag{5.8}$$

Proof. We need the analysis of Δ_n as a periodic discriminant [30]. Δ_n has n bands given by $\alpha_1<\beta_1\leq\alpha_2<\beta_2\leq\cdots\leq\alpha_n<\beta_n$ and bands $[\alpha_j,\beta_j]$. $\Delta^{-1}(-2,2)=\cup_{j=1}^n(\alpha_j,\beta_j)$ and, in particular, the zeros of Δ lie one per band. p_{n-1} has one zero in each gap $[\beta_1,\alpha_2],[\beta_2,\alpha_3],\dots,[\beta_{n-1},\alpha_n]$. That gives us the interlacing. (5.8) is an immediate consequence of this interlacing. □

To get a bound on Δ_n' at its zeros, we need a bound on the rotation number for ergodic Schrödinger operators found by Deift-Simon [8]. This rotation number is

$$\alpha(E)=\pi(1-\nu(-\infty,E)) \tag{5.9}$$

where ν is the density of states. Thus, α runs from π to 0 as E runs from $\min\operatorname{spec}(J)$ to $\max\operatorname{spec}(J)$. $\cos(\alpha)$ runs from -1 to 1.

Proposition 5.6. *In the periodic case, on* spec(J),

$$\frac{d\cos(\alpha(E))}{dE} \geq \frac{1}{2}\biggl(\prod_{j=1}^{n} a_j\biggr)^{-1/n} \tag{5.10}$$

Remarks. 1. (1.2) of [8] is an integrated form of (5.10). We can take derivatives since spec(J) is a union of intervals.

2. Deift-Simon assume $a_j \equiv 1$. By using the modification of the Thouless formula for general a_j, it is easy to see their proof yields (5.10).

3. In the free case (a_n constant, $b_n = 0$), one has equality in (5.10).

Proof of Theorem 5.1. $\Delta_n(E)$ is related to $\alpha(E)$ by

$$\Delta_n(E) = 2\cos(n\alpha(E)) \tag{5.11}$$

Thus, at zeros of Δ_n where $\cos(n\alpha(E)) = 0$, we have $|\sin(n\alpha(E))| = 1$. So at such points,

$$\begin{aligned}\left|\frac{d\Delta_n}{dE}\right| &= 2n\,\frac{d\alpha}{dE} \geq 2n\sin(\alpha(E))\,\frac{d\alpha}{dE} \\ &= 2n\,\frac{d\cos(\alpha(E))}{dE} \\ &\geq n\biggl(\prod_{j=1}^{n} a_j\biggr)^{-1/n}\end{aligned}$$

by (5.10).

Since $\Delta_n = \mathrm{Tr}(T_n)$, $|\Delta_n| \leq 2\|T_n\|$, so (5.4) becomes

$$|E_1 - E_0| \leq 2e\,\frac{\|T_n\|(\prod_{j=1}^n a_j)^{1/n}}{n} \tag{5.12}$$

Between the zeros of Δ are two (E_0, E_1)-type intervals and so, between the first and third of three zeros are four such intervals. (5.8) and (5.12) imply (5.1). $\square$

6. Prüfer Angles and Bounds on Zero Spacing

There are various possible Prüfer angles. We will exploit one that is ideal for studying the energy dependence of zeros of p_n.

Proposition 6.1. *Fix Jacobi parameters* $\{a_n, b_n\}_{n=1}^\infty$. *For each* $n = 1, 2, \dots$, *there is a unique continuous function* $\theta_n(E)$ *determined by*

$$\tan(\theta_n(E)) = \frac{p_n(E)}{p_{n-1}(E)} \tag{6.1}$$

$$\lim_{E\to-\infty} \theta_n(E) = -\frac{\pi}{2} \tag{6.2}$$

Moreover,

$$\frac{d\theta_n}{dE} = \frac{\sum_{j=0}^{n-1} p_j(E)^2}{a_n(p_{n-1}(E)^2 + p_n(E)^2)} > 0 \tag{6.3}$$

Proof. $p_n(E)/p_{n-1}(E) \to -\infty$ as $E \to -\infty$ and the ratio is continuous on $\mathbb{R} \cup \{\infty\}$, so existence and uniqueness are immediate, as is (6.2) since $\tan(-\pi/2) = -\infty$. Note next that since $\frac{d}{dy}\arctan(y) = \frac{1}{1+y^2}$, we have

$$\frac{d\theta_n}{dE} = \frac{p_{n-1}p_n' - p_n p_{n-1}'}{p_{n-1}^2(1 + p_n^2/p_{n-1}^2)} \tag{6.4}$$

so that (6.3) follows from the CD formula (A.3). $\square$

Remark. (6.3) is closely related to well-known formulae relating the derivative of projective angles in $\mathbb{SU}(1,1)$ or $\mathbb{SL}(2,\mathbb{R})$; see Theorem 10.4.11 of [40]. These formulae have been used by Furstenberg, Carmona, and others; see, for example, Carmona [5] or Simon [45].

The transfer matrix, $T_n(E)$, is a 2×2 matrix of determinant $1/a_{n+1}$, so

$$\|T_n(E)^{-1}\| = a_{n+1}\|T_n(E)\| \tag{6.5}$$

Thus, since $(p_{j+1}, p_j)^t = T_j(E)(1,0)^t$, we have that

$$a_{j+1}^{-2}\|T_j(E)\|^{-2} \le p_{j+1}(E)^2 + p_j(E)^2 \le \|T_j(E)\|^2 \tag{6.6}$$

Since

$$\tfrac{1}{2}\sum_{j=0}^{n-2}(p_j^2 + p_{j+1}^2) \le \sum_{j=0}^{n-1} p_j^2 \le \sum_{j=0}^{n-2}(p_j^2 + p_{j+1}^2) \tag{6.7}$$

(6.3) immediately implies

Theorem 6.2. *Let*

$$t_n(E) = \sup_{0\le j\le n-1}(1 + a_{j+1}^2)\|T_j(E)\|^2 \tag{6.8}$$

Then

$$\frac{n}{2t_n(E)t_{n-1}(E)} \le a_n\frac{d\theta_n}{dE} \le nt_n(E)t_{n-1}(E) \tag{6.9}$$

This, in turn, implies $1/n$ upper and lower bounds on zero spacings sufficient for what we needed in Section 3:

Theorem 6.3. *If Δ is an interval in $\mathbb{R}$ on which $\tau_\Delta \equiv \sup_{E\in\Delta,n}(1 + a_{n+1}^2)\|T_n(E)\|^2 < \infty$, then*

$$\inf_{\substack{E,E'\in\Delta \\ E,E' \text{ successive zeros of } p_n(E)}} |E - E'| \ge \frac{a_n\pi}{\tau_\Delta^2 n} \tag{6.10}$$

and

$$\sup_{\substack{E,E'\in\Delta \\ E,E' \text{ successive zeros of } p_n(E)}} |E-E'| \leq \frac{2a_n\pi\tau_\Delta^2}{n} \tag{6.11}$$

Moreover, if $\Delta = [\alpha,\beta]$, p_n *has zeros in* $[\beta - \frac{2a_n\pi\tau_\Delta^2}{n}, \beta]$ *and* $[\alpha, \alpha + \frac{2a_n\pi\tau_\Delta^2}{n}]$ *once* $|\beta-\alpha| \geq \frac{2a_n\pi\tau_\Delta^2}{n}$.

Proof. Since θ_n is monotone in E and $p_n(E) = 0$ if and only if $\theta_n(E) = \ell\pi$ for some $\ell \in \mathbb{Z}$, we have at successive zeros, $E < E'$, that

$$\pi = \theta_n(E') - \theta_n(E) = \int_E^{E'} \frac{d\theta_n}{dE}\,dE \tag{6.12}$$

(6.10)/(6.11) are then immediate from (6.9). The final assertion comes from the fact that $\theta(E_1) - \theta(E_0) \geq \pi$ implies that $\tan(\theta(E))$ has a zero in $[E_0, E_1]$. $\square$

7. Relations of the Weight to the Christoffel Function

The previous sections were dominated by the transfer matrix. In this section, we shift to the weight where the CD kernel (see (A.1)) will play a major role. This section is a technical interlude: a detailed result that will be useful in the analysis of examples in later sections. Our main result in this section is

Theorem 7.1. *Suppose that*

$$d\mu = w(x)\,dx + d\mu_{\mathrm{s}} \tag{7.1}$$

with $d\mu_{\mathrm{s}}$ *singular, and that for some* x_0, $a > 0$, *and some* $r > 0$,

$$\int_{x_0-a}^{x_0+a} w(x)^{-r}\,dx < \infty \tag{7.2}$$

Then

$$K_n(x_0,x_0) \leq C_r n^{1+r^{-1}} \tag{7.3}$$

where C_r *only depends on* r, a *and the integral in* (7.2).

This result generalizes one of Geronimus (see [11, Remark 3.3 and Table II]) in two ways. His estimate is on $|\varphi_n|^2$ not K_n and, more importantly, his estimates require global estimates on w in the context of OPUC rather than just our local estimate. One reason we can go beyond Geronimus is that he uses the Szegő function and we just use the Christoffel variational principle. Another reason is that we have a powerful result of Nevai [34]:

Proposition 7.2 ([34]). *For any p in $(0,\infty)$, there is a constant D_p so*

$$\int_{-1}^{1} |\pi_n(x)|^p \, dx \geq D_p n^{-1} |\pi_n(0)|^p \tag{7.4}$$

for any polynomial π_n of degree n.

Proof. Since this is a special case of Nevai's result which depends on several arguments, for the reader's convenience, we extract exactly what is needed for (7.4).

Let

$$d\mu_0(x) = \frac{1}{\pi} (1-x^2)^{-1/2} \, dx \tag{7.5}$$

on $[-1,1]$, so if $x = \cos(\theta)$, then

$$d\mu_0 = \frac{d\theta}{\pi}$$

on $[0,\pi]$ which implies, as is well known, that the OPs for (7.5) are given by

$$p_n(\cos(\theta)) = \begin{cases} 1 & n = 0 \\ \sqrt{2}\,\cos(n\theta) & n \geq 1 \end{cases} \tag{7.6}$$

the Chebyshev polynomial (of the first kind).

It follows that for $x = \cos(\theta) \in [-1,1]$,

$$K_n(x,x;d\mu_0) = 1 + \sum_{j=1}^{n} 2\cos^2(jn) \leq 2n+1 \tag{7.7}$$

Thus, by (A.7),

$$\sup_{|x|\leq 1} |\pi_n(x)|^2 \leq (2n+1) \int_{-1}^{1} |\pi_n(x)|^2 (1-x^2)^{-1/2} \, \frac{dx}{\pi} \tag{7.8}$$

If m is an integer, $(\pi_n)^m$ is a polynomial of degree at most mn, so (7.8) implies

$$\sup_{|x|\leq 1} |\pi_n(x)|^{2m} \leq (2mn+1) \int_{-1}^{1} |\pi_n(x)|^{2m} (1-x^2)^{-1/2} \, \frac{dx}{\pi} \tag{7.9}$$

If $2m-2 < p \leq 2m$, we write

$$|\pi_n|^{2m} \leq |\pi_n|^p \big(\sup_{|x|\leq 1} |\pi_n(x)| \big)^{2m-p}$$

to deduce

$$\sup_{|x|\leq 1} |\pi_n(x)|^p \leq \left\{ 2n\left(\left[\frac{p}{2}\right] + 1 \right) \right\} \int_{-1}^{1} |\pi_n(x)|^p (1-x^2)^{-1/2} \, \frac{dx}{\pi} \tag{7.10}$$

Given p, pick ℓ from $1, 2, \dots$ so $\ell p \geq \frac{1}{2}$ and apply (7.10) to the polynomial $\pi_n(x)(1-x^2)^\ell$ which has degree $n+2\ell$ and get

$$|\pi_n(0)|^p \leq \sup_{|x|\leq 1} |(1-x^2)^\ell \pi_n(x)|^p \leq \left\{2(n+2\ell)\left(\left[\frac{p}{2}\right]+1\right)\right\} \int_{-1}^1 |\pi_n(x)|^p \,\frac{dx}{\pi} \tag{7.11}$$

since $(1-x^2)^{\ell p - \frac{1}{2}} \leq 1$.

Find D_p so for $n \geq 1$,

$$2(n+2\ell)\left(\left[\frac{p}{2}\right]+1\right) \leq D_p^{-1} n$$

and (7.11) implies (7.4). $\square$

Proof of Theorem 7.1. By (A.16), we can suppose $d\mu_{\mathrm{s}} = 0$. By scaling and translation, we can suppose $x_0 = 0$, $a = 1$. By Theorem A.2, we need to get lower bounds on $\int_{-1}^1 |\pi_n(x)|^2 w(x)\,dx$. By Hölder's inequality, for any α, β, $p \in (1,\infty)$, and q dual to p,

$$\begin{aligned}\int_{-1}^1 |\pi_n(x)|^\alpha\,dx &= \int_{-1}^1 |\pi_n(x)|^\alpha w(x)^\beta w(x)^{-\beta}\,dx \\ &\leq \left(\int_{-1}^1 |\pi_n(x)|^{\alpha p} w(x)^{\beta p}\right)^{1/p} \left(\int_{-1}^1 w(x)^{-\beta q}\,dx\right)^{1/q}\end{aligned} \tag{7.12}$$

We want to pick β, q, α so $\beta q = r$, $\alpha p = 2$, $\beta p = 1$, that is,

$$q = 1+r \qquad p = \frac{1+r}{r} \qquad \alpha = \frac{2r}{1+r} \qquad \beta = \frac{r}{1+r} \tag{7.13}$$

The result is that

$$\begin{aligned}\int_{-1}^1 |\pi_n(x)|^2 w(x)\,dx &\geq \left(\int_{-1}^1 w(x)^{-r}\,dx\right)^{-1/r} \left(\int_{-1}^1 |\pi_n(x)|^{2r/(1+r)}\,dx\right)^{(1+r)/r} \\ &\geq C|\pi_n(0)|^2 n^{-1-r^{-1}}\end{aligned} \tag{7.14}$$

Taking the inf over all π_n's with $\pi_n(0) = 1$ and using (A.5), we get

$$K_n(0,0)^{-1} \geq C n^{-1-r^{-1}}$$

which is (7.3). $\square$

Example 7.3. Let $d\mu$ be the measure on $[-1,1]$ given by

$$d\mu(x) = C_{a,b}|x|^a(1-|x|^2)^b\,dx \tag{7.15}$$

where $a \geq 0$, $b \geq -1$, and C is a normalization constant.

This is an even measure so $p_{2n-1}(0) = 0$. Moreover,

$$p_{2n}(x) = q_n(x^2) \tag{7.16}$$

where q_n are the OPs for the measure obtained from an $x \to y = x^2$ change of variables. Since $dx = (dy)/y^{1/2}$, we see q_n are the orthogonal polynomials for the measure

$$\tilde{C}_{a,b}|y|^{a/2-1/2}(1-y)^b\,dy \tag{7.17}$$

Thus, up to a constant,

$$K_n(0,0;d\mu) = \tilde{K}_{n,\alpha.\beta}(1,1)$$

where $\tilde{K}_{n;\alpha,\beta}$ is the CD kernel for the Jacobi polynomial associated to $(1-x)^\alpha(1+x)^\beta\,dx$ with $\alpha = \frac{1}{2}a - \frac{1}{2}$, $\beta = b$. If we call these orthogonal polynomials $j_{\alpha,\beta}$, and $J_{\alpha,\beta}$ the conventional normalization, then [51, 31]

$$\|J_{\alpha,\beta}\|_2 \sim C^{(1)}_{\alpha,\beta} n^{-1}$$
$$J_{\alpha,\beta}(1) \sim C^{(2)}_{\alpha,\beta} n^\alpha$$

so

$$j_{\alpha,\beta}(1) \sim (C^{(1)}_{\alpha,\beta})^{-1/2}(n^{1/2})J_{\alpha,\beta}(1) = C^{(3)}_{\alpha,\beta} n^{\alpha+1/2}$$

and

$$\tilde{K}_{n;\alpha,\beta}(1,1) \sim n^{2\alpha+2}$$

Taking $\alpha = \frac{1}{2}(a-1)$, we get

$$K_n(0,0;d\mu) \sim n^{1+a} \tag{7.18}$$

We can take r in Theorem 7.1 arbitrary with $ra < 1$, so (7.3) cannot be improved. $\square$

The following shows that in some cases the power of n in Theorem 7.1 is optimal:

Theorem 7.4. *Let $d\mu(x) = w(x)\,dx$ where $\operatorname{supp}(d\mu) \subset [-1,1]$ and*

$$|w(x)| \le C|x|^\alpha \tag{7.19}$$

for some $\alpha < 1$. Then

$$|K_n(0,0)| \ge C_1 n^{1+\alpha} \tag{7.20}$$

Remark. For $w(x) = C_\alpha|x|^\alpha$ on $[-1,1]$, (7.2) holds for any $r < 1/\alpha$, so (7.20) says (7.3) cannot hold for any smaller power of n in case $r > 1$.

Proof. Let π_n be the polynomial of Theorem A.7 where $x_0 = 0$, $a = 1$. On $|x| \in [\frac{j}{n}, \frac{j+1}{n}]$, $j = 0,1,2,\dots,n-1$,

$$|\pi_n(x)| \le \begin{cases} 1 & j = 0 \\ \frac{1}{2n} + \frac{1}{2j} & j = 1,2,\dots \end{cases}$$

so, by (7.19),

$$\begin{aligned}
\int |\pi_n(x)|^2\,d\mu &\le 2c\left\{\frac{1}{n^{1+\alpha}} + \sum_{j=1}^{n-1}\left(\frac{j+1}{n}\right)^{\alpha}\left[\frac{1}{2n^2}+\frac{1}{2j^2}\right]\frac{1}{n}\right\} \\
&\le 2c\left\{\frac{1}{2n^2} + \frac{1}{n^{1+\alpha}}\left(1+\sum_{j=1}^{\infty}\frac{(j+1)^\alpha}{j^2}\right)\right\} \\
&\le C_1^{-1} n^{-(1+\alpha)}
\end{aligned}$$

since $n^2 \ge n^{1+\alpha}$ and $\sum_{j=1}^\infty \frac{(j+1)^\alpha}{j^2} < \infty$ since $\alpha < 1$.

By $\deg(\pi_n) = 2n-2$, and Theorem A.2,

$$K_{2n-2}(0,0)^{-1} \le C_1^{-1} n^{-(1+\alpha)}$$

which is (7.20). $\square$

8. Upper Bounds via OP Methods

Our main purpose in this section is to note that the upper bounds produced by the method of Erdös-Turan [9] provide a universal bound. So long as $d\mu_{\mathrm{s}} = 0$ near $x_0 \in \operatorname{supp}(d\mu)$ and $w(x)$ is continuous and nonvanishing at x_0, the bound is independent of the value of w at x_0! Upper bounds on spacing imply lower bounds on the density of zeros. Deift-Simon [8] obtained universal lower bounds on the density of zeros, so the bounds we find are a kind of microscopic analog of theirs. One key to the Erdös-Turan method is

Lemma 8.1. *Let $x_1 < x_2 < \cdots < x_n$ in $\mathbb{R}$ and let $1 \le j \le n-1$. Then there exists a polynomial π of degree at most $n-1$ so*

$$\pi(x_\ell) = 0 \qquad 1 \le \ell \le n;\ \ell \ne j, j+1 \tag{8.1}$$

$$\pi(x_j) = \pi(x_{j+1}) = 1 \tag{8.2}$$

$$\pi(y) \ge 1 \quad \text{in} \quad [x_j, x_{j+1}] \tag{8.3}$$

Proof. Let

$$\pi_0(x) = \prod_{\ell \ne j, j+1} (x - x_j) \tag{8.4}$$

If

$$\pi_0(x_j) = \pi_0(x_{j+1}) \tag{8.5}$$

take

$$\pi(x) = \frac{\pi_0(x)}{\pi_0(x_j)} \tag{8.6}$$

so (8.1)/(8.2) hold (we will look at (8.3) shortly). If (8.5) fails, for $y \in \mathbb{R}\backslash[x_j, x_{j+1}]$, let

$$\pi_y(x) = (x - y)\pi_0(x) \tag{8.7}$$

As y runs through $(-\infty, x_j]$, $\frac{y-x_j}{y-x_{j+1}}$ runs from 1 down to 0, and as y runs from ∞ to x_{j+1}, the ratio runs from 1 to ∞. Since (8.5) fails and $\pi_0(x_j)$ and $\pi_0(x_{j+1})$ have the same sign, there is a unique y with

$$\pi_y(x_j) = \pi_y(x_{j+1}) \tag{8.8}$$

so take

$$\pi(x) = \frac{\pi_y(x)}{\pi_y(x_j)} \tag{8.9}$$

In any event, π obeys (8.1) and (8.2). By Snell's theorem, π' has a zero between any two zeros of π, and so by counting degrees, exactly 1. It follows that π' has a local maximum in $[x_j, x_{j+1}]$ and no local minimum, so (8.3) holds. $\square$

Let $\{x_j\}_{j=1}^n$ be the zeros of the OP, P_n, associated to a measure $d\mu$. Recall (see Theorem A.4) that there are positive weights $\{\lambda_j\}_{j=1}^n$ so

$$\int \tilde\pi(x)\, d\mu(x) = \sum_{\ell=1}^n \lambda_\ell \tilde\pi(x_\ell) \tag{8.10}$$

for any polynomial $\tilde\pi$ with $\deg \tilde\pi \le 2n-1$.

Theorem 8.2. *For any $j = 1, 2, \dots, n-1$,*

$$\mu([x_j, x_{j+1}]) \le \lambda_j + \lambda_{j+1} \tag{8.11}$$

Proof. Let π be the polynomial of degree $n-1$ or less given by Lemma 8.1. Let $\tilde\pi = \pi^2$, so $\deg\tilde\pi \le 2n-2 \le 2n-1$. Since $\tilde\pi(x_\ell) = 0$, $\ell \neq j, j+1$, and $\tilde\pi(x_j) = \tilde\pi(x_{j+1}) = 1$,

$$\text{RHS of } (8.10) = \lambda_j + \lambda_{j+1} \tag{8.12}$$

Since $\tilde\pi \ge 0$ and $\tilde\pi \ge 1$ on $[x_j, x_{j+1}]$,

$$\text{LHS of } (8.10) \ge \mu([x_j, x_{j+1}]) \tag{8.13}$$

so (8.10) implies (8.11). $\square$

To exploit (8.11), we need upper bounds on λ_j.

Suppose

$$E_\pm = {}^{\sup}_{\inf}\, \operatorname{supp}(d\mu) \tag{8.14}$$

and for $E \in [E_-, E_+]$,

$$d(E) = \max(E_+ - E, E - E_-) \le E_+ - E_- \tag{8.15}$$

Theorem 8.3. *Suppose I is a closed interval on which $d\mu$ is purely a.c. and $\max_{x\in I} w(x) = w_+ < \infty$. Then for each $\delta > 0$ and all weights λ_j associated with $x_j \in I$ and $\operatorname{dist}(x_j, \mathbb{R}\backslash I) \geq \delta$, we have with $m = [\frac{n}{2}]$ and n the number of zeros,*

$$\lambda_j \leq \frac{w_+ \pi \max_I d(E)}{m} + O\biggl(\frac{1}{n^2}\biggr) \tag{8.16}$$

where the $O(\frac{1}{n^2})$ is uniform in all λ's with the given δ (and depends on $\max_I d(E)$).

Proof. Let $a = \max_I d(E)$ and x_0 be the x_j for λ_j, and let $\tilde\pi(x) = \pi_m(x; x_0, a)$ given by Theorem A.7. Since $\tilde\pi(x) = 1$ and $\deg(\tilde\pi^2) \leq 2n-1$, we have

$$\begin{aligned}\lambda_j &\leq \sum \lambda_\ell \tilde\pi(x_\ell)^2 \\ &= \int d\mu(x)\tilde\pi(x)^2 = K_1 + K_2\end{aligned}$$

where K_1 is the integral over $(x_0-\delta, x_0+\delta)$, and K_2 the integral over all other x. By (A.34),

$$K_2 = O\biggl(\frac{1}{n^2}\biggr)$$

with estimates only dependent on a and δ. For $\tilde\pi(x)^2 = O(\frac{1}{n^2})$ on the region of integration and $\mu(\mathbb{R}) = 1$.

If $(x_0-\delta, x_0+\delta) \subset I$, $d\mu \leq w_+\, dx$, so (8.16) follows from (A.35). $\square$

Theorem 8.4. *Suppose I is a closed interval on which $d\mu$ is purely a.c. and*

$$0 < w_- \equiv \min_{x\in I} w(x) \leq \max_{x\in I} w(x) \equiv w_+ < \infty \tag{8.17}$$

Then for any $E \in I^{\text{int}}$,

$$\limsup_{n\to\infty} n[z_n^{(1)}(E) - z_n^{(-1)}(E)] \leq 4\pi\, d(E) \lim_{\delta\downarrow 0} \frac{\max\{w(x) \mid |x-E| < \delta\}}{\min\{w(x) \mid |x-E| < \delta\}} \tag{8.18}$$

In particular, if E is a point of continuity of w,

$$\textit{LHS of } (8.18) \leq 4\pi\, d(E)$$

independently of the value of $w(E)$.

Proof. Clearly,

$$\begin{aligned}&|z_n^{(1)}(E) - z_n^{(-1)}(E)| \\ &\qquad \leq [\min\{w(x) \mid z_n^{(-1)}(E) \leq x \leq z_n^{(1)}(E)\}]^{-1} \mu(z_n^{(-1)}(E), z_n^{(1)}(E))\end{aligned} \tag{8.19}$$

From this, (8.11), (8.16), and $\lim n/m = 2$, we get (8.18) by using the fact that since E is a limit point of an infinity of zeros, we have $\lim_{n\to\infty} |z_n^{\pm 1}(E) - E| = 0$. $\square$

This is the promised universal lower bound on the density of zeros. The method is flexible enough to say something if $w(x)$ has a zero of a fixed order.

Theorem 8.5. *Suppose $d\mu$ is purely absolutely continuous in a neighborhood of E_0, and for some $q > 0$,*

$$0 < \gamma_- = \liminf_{x\to E_0} \frac{w(x)}{|x-E_0|^q} \le \limsup_{x\to E_0} \frac{w(x)}{|x-E_0|^q} = \gamma_+ < \infty \tag{8.20}$$

Then

$$\limsup\, n|z_n^{(1)}(E_0) - z_n^{(-1)}(E_0)| < \infty \tag{8.21}$$

Proof. By (8.20) for any δ, there is N so for $n \ge N$,

$$\mu([z_n^{(-1)}(E_0), z_n^{(1)}(E_0)]) \ge (\gamma_- - \delta) 2^{-q}(q+1)^{-1}(z_n^{(1)}(E_0) - z_n^{(-1)}(E_0))^{q+1} \tag{8.22}$$

By the proof of (8.16), the λ's associated to $z_n^{\pm}(E)$ obey

$$\lambda_j \le \frac{C_1(\gamma_+ + \delta)[\max(C_2 n^{-1}, |z_n^{\pm}(E_0) - E_0|)]^q}{n} \tag{8.23}$$

for constants C_1, C_2. (8.11),(8.22), and (8.23) imply (8.21). $\square$

Given our bounds in Section 7, we can also say something when the singularity of the weight is not as regular as some power. The key is an abstraction of an argument of Nevai [34] (see also Golinskii [14]).

Theorem 8.6. *Let $a = \max(\operatorname{supp}(d\mu)) - \min(\operatorname{supp}(d\mu))$. Fix integers p, q so that*

$$(2p-2)^{2q} \le 2n - 1 \tag{8.24}$$

Then for any successive zeros E, E' of p_n, we have

$$|E - E'| \le \frac{a}{p}\,[K_p(\tfrac{1}{2}(E+E'), \tfrac{1}{2}(E+E'))]^{1/2q} \tag{8.25}$$

Proof. Let $\tilde\pi$ be defined in terms of the π of Theorem A.7 by Theorem A.4,

$$\tilde\pi(x) = [\pi_p(x; \tfrac{1}{2}(E+E'), a)]^q \tag{8.26}$$

By (8.24), $\deg[\tilde\pi]^2 \le 2n-1$ so, by (8.10),

$$\int |\tilde\pi(x)|^2\, d\mu(x) = \sum_{j=1}^n \lambda_j |\tilde\pi(E_j)|^2 \tag{8.27}$$

$$\leq \left(\frac{1}{2p} + \frac{a}{2p|E-E'|} \right)^{2q} \tag{8.28}$$

since $\sum \lambda_j = 1$, $\lambda_j \geq 0$, and $\min|E_j - \frac{1}{2}(E+E')| = \frac{1}{2}|(E-E')|$.

Since $\tilde{\pi}(\frac{1}{2}(E+E')) = 1$, by Theorem A.2,

$$K_n(\tfrac{1}{2}(E+E'), \tfrac{1}{2}(E+E'))^{-1} \leq \int |\tilde{\pi}(x)|^2 \, d\mu(x) \tag{8.29}$$

Since $|E-E'| \leq a$,

$$\frac{1}{2p} + \frac{a}{2p|E-E'|} \leq \frac{a}{p|E-E'|} \tag{8.30}$$

(8.25) is immediate from (8.28), (8.29), and (8.30). □

The following abstracts an argument of Golinskii, who needed to make global hypotheses since he relied on estimates of Geronimus:

Corollary 8.7. *Suppose that for some interval I, $A > 0$ and C, we have*

$$\sup_{E \in I} |K_n(E,E)| \leq C(n+1)^A$$

Then for any $\delta > 0$,

$$\lim_{n\to\infty} \sup_{\substack{E,E' \text{ successive zeros} \\ \text{dist}(E,\mathbb{R}\setminus I)>\delta}} |E-E'| \left[\frac{n}{\log n} \right] < \infty \tag{8.31}$$

Proof. Pick $q = [\log n]$ and p as large as possible so that (8.24) holds. Since

$$(n+1)^A = \exp(A \log(n+1))$$

$[K_p]^{1/2q}$ is bounded and (8.25) implies (8.31). □

Combining this corollary and Theorem 7.1, we obtain a local version of Golinskii's [14] result:

Corollary 8.8. *If* (7.1) *and* (7.2) *hold, then we have* (8.31) *for* $I = (x_0 - a, x_0 + a)$.

We also have the following (a local version of results of Nevai [34] and Golinskii [14]):

Theorem 8.9. *Suppose for some interval I we have that*

$$d\mu = w\, dx + d\mu_{\mathrm{s}} \tag{8.32}$$

where

$$\int_I \log w \, dx > -\infty \tag{8.33}$$

Then for any $\delta > 0$,

$$\lim_{n\to\infty} \sup_{\substack{E,E' \text{ successive zeros} \\ \operatorname{dist}(E,\mathbb{R}\setminus I)>\delta}} |E-E'|n^{1/2} < \infty \tag{8.34}$$

We need the following lemma:

Lemma 8.10. *If* J *has Jacobi parameters obeying*

$$\sum_{n=1}^{\infty} |a_n-1|^2 + |b_n|^2 < \infty \tag{8.35}$$

then for any $\delta > 0$, *there is* C_δ *so*

$$\sup_{E\in[-2+\delta,2-\delta]} \|T_n(E)\| \leq \exp\bigl(C_\delta\sqrt{n+1}\,\bigr) \tag{8.36}$$

Proof. Define $u_n^\pm(E) = e^{in\theta}$ where $2\cos\theta = E$ and $0 < \theta < \pi$. By standard variation of parameters about $u_n^\pm$, one proves

$$\text{LHS of (8.36)} \leq \prod_{j=1}^{n} \{1 + C(|b_j| + |a_j-1|)\}$$

Since $1+x \leq \exp(x)$ and

$$\sum_{j=1}^{n} (|b_j| + |a_j-1|) \leq \biggl[\sum_{j=1}^{n} (|b_j|^2 + |a_j-1|^2)\biggr]^{1/2} [2n]^{1/2}$$

(8.36) is immediate. □

Proof of Theorem 8.9. By scaling, we suppose $I = [-2,2]$. By Corollary A.3,

$$K_n(x,x;d\mu) \leq K_n(x,x;d\nu) \tag{8.37}$$

where

$$d\nu = \chi_{[-2,2]} w\,dx \tag{8.38}$$

Let $d\tilde\nu$ be the normalized $d\nu$. By the theorem of Killip-Simon [25], the Jacobi parameters obey (8.35), so by (8.37),

$$\sup_{x\in[-2+\delta,2-\delta]} K_n(x,x;d\mu) \leq \exp\bigl(C_\delta\sqrt{n+1}\,\bigr)$$

In Theorem 8.6, take $q = \bigl[\sqrt{n}\bigr]$ and p as large as can be so (8.24) holds. Then $p \sim c\sqrt{n}$ and (8.25) implies (8.34). □

9. Lower Bounds via OP Methods

In this section, we will get lower bounds in terms of the CD kernel alone. The basic method is due to Golinskii [14], but when he applied the method, he made global assumptions on the measure, and we want to note that local assumptions suffice. Other OP lower bound methods are due to Erdös-Turan [9] and Nevai [34].

Theorem 9.1. *If E, E' are distinct zeros of $P_n(x)$, if $\bar{E} = \frac{1}{2}(E+E')$ and $\delta > \frac{1}{2}|E-E'|$, then*

$$|E-E'| \geq \frac{[\delta^2 - (\frac{1}{2}|E-E'|)^2]}{3n} \left[\frac{K_n(E,E)}{\sup_{|y-\bar{E}|\leq\delta} K_n(y,y)}\right]^{1/2} \tag{9.1}$$

Remarks. 1. In most applications, δ is fixed and $|E-E'| \to 0$, so $\delta^2 - (\frac{1}{2}|E-E'|)^2 \sim \delta^2 > 0$. In typical cases, the inf and sup of $K_n(y,y)$ for $|y-\bar{E}| < \delta$ are comparable and (9.1) gives an $1/n$ lower bound.

2. This theorem also yields a result with the same asymptotics as (2.22) for $K_n(E,E) \geq 1$, while the sup is bounded exponentially in n.

3. Interestingly enough, the proof here depends on (2.19) written as (9.2).

4. It is interesting to compare Theorems 9.1 and 2.2. (2.17) only depends on information at E_0 while (9.1) has a $\sup K_n(y,y)$ over a neighborhood, but (2.17) requires information on both solutions of (2.1) while (9.1) only on p_n.

Proof. Since $p_n(E) = p_n(E') = 0$ and $E \neq E'$, we have

$$K_n(E,E') = 0 \tag{9.2}$$

by (A.3) (the Christoffel-Darboux formula). Thus, (supposing $E < E'$ for notational convenience),

$$\begin{aligned} K_n(E,E) &= K_n(E,E) - K_n(E,E') && (9.3) \\ &\leq |E-E'| \sup_{|y-\bar{E}|\leq\frac{1}{2}|E-E'|} \left|\frac{\partial}{\partial y} K_n(E,y)\right| && (9.4) \\ &\leq |E-E'| \{\delta^2 - [\tfrac{1}{2}(E-E')]^2\}^{-1} \\ &\qquad \sup_{|y-\bar{E}|\leq\delta} [\delta^2 - (y-\bar{E})^2]^{1/2} \left|\frac{\partial}{\partial y} K_n(E,y)\right| && (9.5) \end{aligned}$$

$K_n(E,y)$ is a polynomial in y of degree n, so by (A.29) and (9.5),

$$K_n(E,E) \leq |E-E'| \{\delta^2 - [\tfrac{1}{2}(E-E')]^2\}^{-1} (3n) \sup_{|y-\bar{E}|<\delta} |K_n(E,y)| \tag{9.6}$$

By the Schwartz inequality,

$$\sup_{|y-\bar{E}|<\delta} |K_n(E,y)| \leq K_n(E,E)^{1/2} \sup_{|y-\bar{E}|<\delta} |K_n(y,y)|^{1/2} \tag{9.7}$$

(9.6) plus (9.7) imply (9.1). □

Corollary 9.2. *Let I be an interval on which*

$$t = \sup_{n,y\in I} \|T_n(y)\| < \infty \tag{9.8}$$

Then, for any $E \in I^{\text{int}}$,

$$\liminf_{n\to\infty} n|z_n^{(+1)}(E) - z_n^{(-1)}(E)| \geq \frac{t^{-2}}{3} \operatorname{dist}(E, \mathbb{R}\backslash I)^2 \tag{9.9}$$

Remark. This should be compared with what follows from (2.15) and (2.17) which implies

$$\text{LHS of (9.9)} \geq t^{-2}$$

Proof. (9.9) follows from (9.1) if one notes that for $y \in I$,

$$(n+1)t^{-2} \leq K_n(y,y) \leq (n+1)t^2$$

□

We can also use Theorem 9.1 to get a lower bound in terms of local bounds on the weights.

Theorem 9.3. *Suppose $d\mu = w\,dx + d\mu_{\text{s}}$ where $d\mu_{\text{s}}(x_0-\delta, x_0+\delta) = 0$ and*

$$0 < \inf_{|y-x_0|\leq\delta} w(x) \leq \sup_{|y-x_0|\leq\delta} w(x) < \infty \tag{9.10}$$

Then for any $\varepsilon < \delta$,

$$\inf_{|y-x_0|<\varepsilon} \liminf_{n\to\infty} n|z_n^{(+1)}(y) - z_n^{(-1)}(y)| > 0 \tag{9.11}$$

Proof. By (9.1), it suffices to prove

$$\sup_{|y-x_0|<\varepsilon} [n^{-1}K_n(y,y)] < \infty \tag{9.12}$$

and

$$\inf_{|y-x_0|<\varepsilon} [n^{-1}K_n(y,y)] > 0 \tag{9.13}$$

By (9.10), for ε fixed, uniformly in y with $|y-x_0| < \varepsilon$, we can find a fixed scaling and some translate of $c(4-x^2)^{1/2}\chi_{[-2,2]}\,dx$ lying below $d\mu(x-y)$, so using (A.14) and the explicit K_n for Chebyshev polynomials of the second kind (i.e., the free K_n), we get (9.12).

On the other hand, by (A.4)/(A.5),

$$K_n(x_0,x_0) \geq \left(\int |\pi_n(x)|^2\,d\mu(x)\right)^{-1} \tag{9.14}$$

for any π_n of degree n with $\pi_n(x_0) = 1$. Using a suitable $\pi_{n/2}$ of the form given by Theorem A.7 and estimates we used earlier in this paper, we get an $O(n)$ lower bound on K_n, that is, (9.13) holds. $\square$

10. Zeros of POPUC

While we have discussed OPRL up to now, virtually all the ideas extend to POPUC. POPUC are defined by taking the first $n-1$ recursion parameters (Verblunsky coefficients), $\alpha_0, \dots, \alpha_{n-2}$ in $\mathbb{D}$, and picking $\beta \in \partial\mathbb{D}$ and letting

$$\tilde{\varphi}_n(z) = z\varphi_{n-1}(z) - \bar{\beta}\varphi_{n-1}^*(z) \tag{10.1}$$

Since φ_{n-1}^* is nonvanishing on $\mathbb{D}$ and $|\varphi_{n-1}| = |\varphi_{n-1}^*|$ on $\partial\mathbb{D}$, we have $|\varphi_{n-1}/\varphi_{n-1}^*| < 1$ on $\mathbb{D}$ by the maximum principle. Thus, $\tilde{\varphi}_n(z)$ is nonvanishing on $\mathbb{D}$ and, by symmetry $(\overline{\tilde{\varphi}_n(1/\bar{z})} = z^{-n}[\varphi_{n-1}^* - \beta z\varphi_{n-1}])$, nonvanishing on $\mathbb{C}\backslash\overline{\mathbb{D}}$. Thus, the zeros of $\tilde{\varphi}_n$ lie on $\partial\mathbb{D}$; indeed, they are the zeros of a finite unitary matrix (see Theorem 8.2.7 of [39]). Zeros of POPUC are discussed extensively in Golinskii [14], Cantero-Moral-Velázquez [3, 4], Simon [46], and Wong [55].

As explained in Section 10.8 of [40], there is an OPUC analog of (2.7)–(2.8) (namely, (10.8.3)–(10.8.5) of [40]) which immediately leads to an analog of Theorem 2.1. While (10.8.3)–(10.8.5) are stated for the solutions $\psi_{\cdot} + F\varphi_{\cdot}$, they also hold for $\psi_{\cdot}$ and $\varphi_{\cdot}$. Key to this analog is the orthogonality of $\varphi_{\cdot}(z_1)$ and $\varphi_{\cdot}(z_2)$ for two zeros of $\tilde{\varphi}_n(z)$. This follows from the CD formula for OPUC (see Theorem 2.2.7 of [39]) for $\tilde{\varphi}_n(z_1) = \tilde{\varphi}_n(z_2) = 0$ implies $z_j\varphi_n(z_j) = \bar{\beta}\varphi_n^*(z_j)$, so

$$\overline{\varphi_n^*(z_1)}\,\varphi_n^*(z_2) - \bar{z}_1 z_2\,\overline{\varphi_n(z_1)}\,\varphi_n(z_2) = 0$$

and thus, by (2.2.42) of [39],

$$\sum_{j=0}^{n} \overline{\varphi_j(z_1)}\,\varphi_j(z_2) = 0 \tag{10.2}$$

Combined with the techniques of Section 12.1 of [40] and our proof of Theorem 3.4, we get

Theorem 10.1. *Let $\{\alpha_n\}_{n=0}^\infty$ be a set of Verblunsky coefficients that obeys*

$$\sum_{n=0}^{\infty} |\alpha_{n+1} - \alpha_n| < \infty \tag{10.3}$$

and

$$|\alpha_n| \to 0 \tag{10.4}$$

Then the zeros of the POPUC, $\tilde{\varphi}_n(z)$, for any choice of β have uniform clock behavior on any compact subset of $\partial\mathbb{D}\backslash\{1\}$.

Remarks. 1. An interesting example is $\alpha_n = (n+2)^{-\beta}$ for any $\beta > 0$. This is related to a conjecture of [41], albeit the conjecture there is for OPUC, not POPUC.

2. The density of zeros in this case is $d\theta/2\pi$ on $\partial\mathbb{D}$; see Theorem 8.2.7 and Example 8.2.8 of [39].

3. If (10.3)/(10.4) are replaced by

$$\sum_{n=0}^{\infty} |\alpha_n| < \infty \tag{10.5}$$

then $\partial\mathbb{D}\backslash\{1\}$ can be replaced by $\partial\mathbb{D}$. This is a result of [41]. Because we have global control in this case, one does not need a priori $1/n$ bounds on zero spacing.

There are also analogs of the bounds of Sections 5, 6, 8, and 9:

1. One has that

$$\frac{d^2}{d\theta^2} \log|e^{i\varphi} - e^{i\theta}| = -\frac{1}{|e^{i\varphi} - e^{i\theta}|^2} \tag{10.6}$$

so there is a bound like (5.4) for POPUC (all of whose zeros lie on $\partial\mathbb{D}$), and thus, there is an analog of Theorem 5.1.

2. If one defines $\eta_n(z)$ by

$$e^{i\eta_n(\theta)} = \frac{e^{i\theta}\varphi_{n-1}(e^{i\theta})}{\varphi_{n-1}^*(e^{i\theta})} \tag{10.7}$$

then

$$\frac{d\eta_n}{d\theta} = \frac{[\sum_{j=0}^{n-1} |\varphi_j(e^{i\theta})|^2]}{|\varphi_{n-1}(e^{i\theta})|^2} \tag{10.8}$$

This follows from (2.2.71) of [39] which implies

$$\frac{\partial}{\partial r} \log|\varphi_{n+1}(re^{i\theta})|^2 \bigg|_{r=1} = (n+1) + |\varphi_{n+1}(e^{i\theta})|^{-2} \sum_{j=0}^{n} |\varphi_j(e^{i\theta})|^2 \tag{10.9}$$

By the Cauchy-Riemann equations

$$\text{LHS of (10.9)} = 2\,\frac{\partial}{\partial\theta} \arg[\varphi_{n+1}(e^{i\theta})] \tag{10.10}$$

Since (10.7) implies

$$\eta_n = \theta - (n-1)\theta + 2\arg[\varphi_{n-1}(e^{i\theta})] \tag{10.11}$$

we obtain (10.8).

(10.8) implies $d\eta_n/d\theta > 0$ and, given that zeros of $\tilde{\varphi}_n$ occur when $\eta_n = \arg\bar{\beta} \pmod{2\pi}$, bounds like those of Theorem 6.3 on zero spacing for POPUCs.

3. Since all the techniques of the Appendix extend to OPUC, the estimates of Sections 8 and 9 extend to POPUC; indeed, somewhat weaker variants occur already in Golinskii [14].

11. Examples, Counterexamples, Conjectures, and Questions

(a) *Clock behavior based only on local behavior.* Suppose (7.1) holds on for some $(c,d) \subset \mathbb{R}$, we have $d\mu_{\rm s}([c,d]) = 0$, $w > 0$ on (c,d) and w is C^∞ there. We have proven $O(1/n)$ upper and lower bounds in this case. This leads to the natural question:

Open Question 11.1. Under the above hypothesis, does one have clock behavior on (c,d)?

This is a very subtle question because clock behavior involves the density of states, and it is not even clear that exists on (c,d) only under the above hypothesis. What is clear is that if the density of states exists, it is a global quantity and not just dependent on w on (c,d). We want to demonstrate this by example. We will need the following:

Proposition 11.2. *Let $d\mu_0$ be given by*

$$d\mu_0 = (2\pi)^{-1}\sqrt{4-x^2}\,\chi_{[-2,2]}\,dx \tag{11.1}$$

Suppose f is a C^2 function on $[-2,2]$ with $f'(2) = f'(-2) = 0$, $f''(-2) = f''(2)$, and

$$f \geq \alpha > 0 \qquad \int f\,d\mu_0 = 1 \tag{11.2}$$

where $\alpha > 0$ is a positive real. Let

$$d\mu(x) = f(x)\,d\mu_0(x) \tag{11.3}$$

Then the density of states exists for $d\mu$ and is given by (1.8) *and there is clock behavior uniformly on each interval $[-2+\varepsilon, 2-\varepsilon]$.*

Proof. By Theorem 13.2.1 of [40], there is a map Sz_2 of real measures, $d\rho$, on $\partial\mathbb{D}$ (i.e., those measures with real Verblunsky coefficients) to those measures $d\mu$ on $[-2,2]$ which are of the form $f\,d\mu_0$ where

$$\int_{-2}^{2} f(x)(4-x^2)^{-1/2}\,dx < \infty \tag{11.4}$$

and under this map, $d\rho(\theta) = w(\theta)\,d\theta$ where

$$w(\theta) = c\,f(2\cos\theta) \tag{11.5}$$

Thus, w is C^2 (including at $\theta = 0, \pi$) and so $\sum_{n=-\infty}^{\infty} n^2 |\widehat{w}_n|^2 < \infty$, which implies $\sum_{n=-\infty}^{\infty} |\widehat{w}_n| < \infty$. Thus, by Baxter's theorem (Theorem 5.2.1 of [39]), the Verblunsky coefficients are in ℓ^1. By (13.2.20)/(13.2.21) of [40], (1.15) holds, which implies the claimed result. $\square$

Example 11.3. Let $d\tilde\mu_0$ be $d\mu_0$ scaled to $[-1,1]$. We pick f_1 obeying the hypothesis of Proposition 11.2 and f_2 scaled to $[-1,1]$, and so that $d\mu_1 = f_1\,d\mu_0$ and $d\mu_2 = f_2\,d\tilde\mu_0$ obey $d\mu_1 \equiv d\mu_2$ on $[-\frac{1}{2}, \frac{1}{2}]$. Both have clock behavior on $[-\frac{1}{2}, \frac{1}{2}]$ but with different density of states, namely (1.8) and (1.8) scaled. $\square$

(b) *Pointwise upper bounds.* We obtained lower bounds on $z_n^{(1)}(E_0) - z_n^{(-1)}(E_0)$ if $T_n(E_0)$ is bounded, but our upper bounds required control of $T_n(E)$ for E in a neighborhood of E_0.

Open Question 11.4. Are there upper bounds on spacing if we only know that $T_n(E_0)$ is bounded?

(c) *Improved spacing estimates.*

Open Question 11.5. Can $\sqrt{n}$ in (2.25) be improved?

(d) *More on spacing and $\gamma(E)$.* We saw that $\gamma(E_0) = 0$ plus some regularity of ν at E_0 implies an $O(1/n)$ lower bound. Does it imply clock spacing? In particular,

Open Question 11.6. Is there local clock behavior for a.e. E_0 with $\gamma(E_0) = 0$ in the case of almost periodic Jacobi parameters?

Example 11.7. [27] has proven, for $\alpha < \frac{1}{2}$, the OPUC analog of Poisson behavior for $a_n \equiv 1, b_n$ independent random variable of the form $b_n = Cn^{-\alpha} w_n$ where w_n is uniformly distributed in $[-1,1]$. We assume their result is true in the Jacobi case. Of course, $\gamma(E) = 0$ in this case. We do not have a contradiction with Corollary 4.2 since this model is not ergodic. The example does show though that ergodicity is a necessary hypothesis. $\square$

Example 11.8. Corollary 4.2 shows that ergodicity along with Poisson behavior imply positive Lyapunov exponent. This raises the natural question: Does ergodicity along with a positive Lyapunov exponent imply Poisson behavior? The answer is negative, as can be shown by the following example: Consider the Jacobi matrix with $a_n = 1$, $b_n = \lambda\cos(2\pi\alpha n+\theta)$, where $|\lambda| > 2$ and α is a (Liouville) irrational for which there is a sequence of rationals $\{p_n/q_n\}_{n=1}^\infty$ such that $|\alpha-p_n/q_n| < n^{-q_n}$. This is an ergodic Jacobi matrix and it is well known (see, e.g., [1]) that its Lyapunov exponent $\gamma(E)$ is positive for any E. By using the results of Avron-van Mouche-Simon [2] and considering scales of the form mq_n, where $m > 2$ is an integer, one can show that, for each θ, there would be clusters of $m-1$ zeros each of which is contained in an interval whose length is of order $(2/|\lambda|)^{q_n/2}$. As θ is varied, these clusters will move over regions whose size is roughly of order $1/q_n$. This behavior contradicts Poisson behavior. More precisely, it is possible to show that Poisson behavior does not occur for Lebesgue a.e. E in the spectrum.

(e) *Zero spacing and the Szegő condition.*

Open Question 11.9. Does one have $O(n^{-1})$ bounds (upper and lower) when a Szegő or quasi-Szegő condition holds?

(f) *Spacing at zeros of* $w(x)$.

Open Question 11.10. What can one say at zero spacing at points x_0 where $w(x)$ has a "regular" zero, that is, $w(x) \sim |x-x_0|^\alpha$ for some $\alpha > 0$?

(g) *Edge zeros when* $a_n = 1 - n^{-\gamma}$. The following illuminates Theorem 3.8.

Example 11.11. Let $b_n \equiv 0$, $a_n = 1 - n^{-\gamma}$ for $\gamma > 0$. Then Theorem 3.4 applies and there is clock behavior away from -2 and 2. If $\gamma > 2$, Theorem 3.8 applies and the largest E_j has $E_j = 2 - Cn^{-2} + o(n^{-2})$. We claim for γ in general

$$2 - C_2 n^{-\gamma} + O(n^{-1}) \le E_j^{\max} \le 2 - C_1 n^{-\gamma} \tag{11.6}$$

capturing the leading behavior for $\gamma \in (0,1)$. The upper bound in (11.6) comes from monotonicity of $E_j^{\max}$ in the a's and the fact that for $J_{n;F}$ $\max_{1\le j\le n}|a_j| \sim 1 - n^{-\gamma}$. The lower bound comes with a trial vector that lives in $[n/2, 2]$ with maximum at $3n/4$ and constant slope in between. $\square$

(h) *OPUC.* This paper has a fairly complete analysis of OPRL and POPUC. Many questions remain for general OPUC.

Appendix A. Tools of the OP Trade

As explained in the introduction, this paper is intended for two audiences, so we include here a summary of tools well known to the OP community but not so well to the Schrödinger operator community. Because the tools, while powerful, are simple, we can even give complete proofs. We will discuss the Christoffel variational principle, Gauss-Jacobi quadrature, Bernstein's inequality, and Dirichlet-Fejér trial polynomials.

(a) Christoffel variational principle. We will define OPRL for arbitrary positive measures (with finite moments) even if $\mu(\mathbb{R}) \neq 1$. The monic polynomials, P_n, are independent of normalization, but the orthonormal polynomials, p_n, are not. For example, $p_0(x, d\mu) = d\mu(\mathbb{R})^{-1/2}$. The *Christoffel-Darboux* (a.k.a. CD) *kernel* or *reproducing kernel* is defined by

$$K_n(x,y) = \sum_{j=0}^{n} \overline{p_j(y)}\, p_j(x) \tag{A.1}$$

We will use $K_n(x,y;d\mu)$ if the measure needs to be more explicit. The name reproducing kernel comes from

$$(P_n f)(x) = \int K_n(x,y) f(y)\, d\mu(y) \tag{A.2}$$

where P_n is the projection in $L^2(\mathbb{R}, d\mu)$ onto the space of polynomials of degree n.

We need the following in Section 6:

Theorem A.1 (CD formula)**.** *We have*

$$K_n(x,y) = a_{n+1} \left[\frac{\overline{p_{n+1}(y)}\, p_n(x) - \overline{p_n(y)}\, p_{n+1}(x)}{\bar{y} - x} \right] \tag{A.3}$$

Proof. This is a discrete version of integrating a Wronskian. Take the equation (1.3) at $\bar{y}$, multiply by $p_n(x)$ and subtract (1.3) at x, multiplied by $\overline{p_n(y)}$, and obtain

$$Q_{n+1}(x,y) = \overline{p_n(y)}\, p_n(x) + Q_n(x,y)$$

where Q_{n+1} is the right side of (A.3). Since $p_{-1}(x) = 0$, $Q_0 \equiv 0$, so (A.3) follows by iteration. $\square$

Remark. For x, y real, $p_n(x), p_n(y)$, are real so the bars are not needed. Indeed, one can drop all the bars for complex x, y, but given the OPUC analogs, it is natural to use the bars.

Theorem A.2 (Christoffel variational principle)**.** *Let*

$$\lambda_n(x_0; d\mu) = \inf\biggl(\int |\pi_n(x)|^2\, d\mu(x) \biggm| \deg \pi_n \le n;\ \pi_n(x_0) = 1\biggr) \tag{A.4}$$

Then

$$\lambda_n(x_0; d\mu) = K_n(x_0, x_0; d\mu)^{-1} \tag{A.5}$$

Remarks. 1. λ_n are called *Christoffel numbers.* More generally, we have p-Christoffel numbers defined, for $0 < p < \infty$, by

$$\lambda_n(x_0, p; d\mu) = \inf\biggl(\int |\pi_n(x)|^p\, d\mu(x) \biggm| \deg \pi_n \le n;\ \pi_n(x_0) = 1\biggr) \tag{A.6}$$

2. Another way of writing (A.4) is that λ_n is the optimal constant in

$$|\pi_n(x_0)|^2 \le \lambda_n(x_0; d\mu)^{-1} \int |\pi_n(x)|^2\, d\mu(x) \tag{A.7}$$

or

$$|\pi_n(x_0)|^p \le \lambda_n(x_0, p; d\mu)^{-1} \int |\pi_n(x)|^p\, d\mu(x) \tag{A.8}$$

3. Our proof shows the inf in (A.4) is a min and the minimizing π is given by

$$\pi(x) = K(x, x_0) \tag{A.9}$$

Proof. Expand π_n in terms of the orthonormal basis $\{p_j\}_{j=0}^n$:

$$\pi_n(x) = \sum_{j=0}^n a_j p_j(x) \tag{A.10}$$

$\pi_n(x_0) = 1$ is equivalent to

$$\sum_{j=0}^n a_j p_j(x_0) = 1 \tag{A.11}$$

By the Schwartz inequality

$$1 \le K(x_0, x_0) \sum_{j=0}^n a_j^2 \tag{A.12}$$

$$= K(x_0, x_0) \int \pi_n(x)^2\, d\mu(x) \tag{A.13}$$

where equality occurs in (A.12) if $a_j = \overline{p_j(x_0)}/K(x_0, x_0)$, that is, if π_n is given by (A.10). (A.5) is immediate from this case of equality and (A.13). $\square$

Christoffel numbers have been a critical tool in OP theory for over a century, with important uses by Erdös-Turán [9] and turned to high art by Freud and Nevai; see Nevai [35]. They can be used for lower bounds on K, that is, upper bounds for λ_n, by using any convenient trial polynomial for π_n (see (d) below). One gets upper bounds on K, that is, lower bounds for λ_n, by the immediate

Corollary A.3. *If $d\mu \geq d\nu$, then*

$$K_n(x_0, x_0; d\mu) \leq K_n(x_0, x_0; d\nu) \tag{A.14}$$

Remarks. 1. This shows the true power of Theorem A.2 and the need to allow $\mu(\mathbb{R}) \neq 1$.

2. In particular, if

$$d\mu = f(x)\,dx + d\mu_{\rm s} \tag{A.15}$$

then

$$K_n(x_0, x_0; d\mu) \leq K_n(x_0, x_0; f\,dx) \tag{A.16}$$

(b) Gauss-Jacobi quadrature. The main result here is

Theorem A.4 (Gauss-Jacobi quadrature)**.** *Let μ be an arbitrary positive nontrivial measure on $\mathbb{R}$ with finite moments. Fix n and define $d\mu_n$ to be the point measure with weights only at the zeros $\{x_j^{(n)}\}_{j=1}^n$ of $p_n(x)$ and weights*

$$d\mu_n(\{x_j^{(n)}\}) = \lambda_n(x_j^{(n)}; d\mu) \tag{A.17}$$

the Christoffel numbers of $d\mu$. Then, if π is a polynomial of degree $2n-1$ or less, we have

$$\int \pi(x)\,d\mu(x) = \int \pi(x)\,d\mu_n(x) \tag{A.18}$$

Remark. In our applications, we will care much more that the masses are at the zeros than the variational formulae for the weights.

Sketch of Proof. Here is a proof intended for Schrödinger operator experts. (For the more usual OP proof, see Freud's book [10].) Let $J_{n;F}$ be an $n \times n$ matrix in the upper corner of the Jacobi matrix, (1.4), associated to $d\mu$. Then the recursion (1.3) implies that if

$$u_j(z) = p_{j-1}(z) \qquad j = 1, 2, \dots, n \tag{A.19}$$

then

$$[(J_{n;F} - z)u]_j = -a_{n+1}\delta_{jn}p_n(z) \tag{A.20}$$

Thus the eigenvalues of $J_{n;F}$ are the zeros of p_n and the normalized eigenvectors are $p_{j-1}(z)/K(z,z)^{1/2}$. It follows that $d\mu_n$ is the spectral measure of $J_{n;F}$ with eigenvector $(1\,0\dots0)^t$.

For any measure $d\eta$ and associated Jacobi matrix, J, since $d\eta$ is the spectral measure for $\delta \equiv (1\,0\ldots)^t$, we have

$$\int x^\ell \, d\eta = \langle \delta, J^\ell \delta \rangle \tag{A.21}$$

so $\int x^{\ell+k}\, d\eta = \langle J^k\delta, J^\ell \delta\rangle$ for $k,\ell = 0,\ldots,n-1$ depends only on $\{J^k\delta\}_{k=0}^{n-1}$ and so on $\{a_k,b_k\}_{k=0}^{n-1}$. Moreover,

$$\int x^{2n-1}\, d\eta = \langle J^{n-1}\delta, JJ^{n-1}\delta\rangle \tag{A.22}$$

only depends on $\{a_k,b_k\}_{k=0}^{n-1} \cup \{b_n\}$. Thus, J and $J_{n;F}$, which have the same set of these parameters, have the same moments of order up to $2n-1$, that is,

$$\int x^k\, d\mu = \int x^k \, d\mu_{n;F} \qquad 0 \leq k \leq 2n-1 \tag{A.23}$$

which is (A.18). $\square$

(c) Bernstein's inequality. These inequalities control π_n' in terms of n and π_n for polynomials π_n of degree at most n.

Theorem A.5. *Let $\partial\mathbb{D}$ be the unit circle in $\mathbb{C}$, $\partial\mathbb{D} = \{z \mid |z| = 1\}$. Let π_n be a polynomial of degree n. Then*

$$\sup_{z\in\partial\mathbb{D}} |\pi_n'(z)| \leq n \sup_{z\in\partial\mathbb{D}} |\pi_n(z)| \tag{A.24}$$

Remark. If $\pi_n(z) = z^n$, one has equality in (A.24).

Proof. (Szegő [50]) Since $\pi_n(e^{i\theta}) = \sum_{j=0}^n a_j e^{ij\theta}$, we have

$$\pi_n(e^{i\theta}) = \int_0^{2\pi} \sum_{j=0}^n e^{ij(\theta-\varphi)} \pi_n(e^{i\varphi})\, \frac{d\varphi}{2\pi} \tag{A.25}$$

so

$$\begin{aligned} -i\pi_n'(e^{i\theta}) &= \int_0^{2\pi} \sum_{j=1}^n j e^{ij(\theta-\varphi)} \pi_n(e^{i\theta})\, \frac{d\rho}{2\pi} \\ &= \int F_n(\theta-\varphi) e^{in(\theta-\varphi)} \pi_n(e^{i\varphi})\, \frac{d\rho}{2\pi} \end{aligned} \tag{A.26}$$

where

$$F_n(\theta) = \sum_{j=-n+1}^{n-1} (n-|j|) e^{ij\theta} \tag{A.27}$$

(for the $j > 0$ terms in (A.26), integrate to zero).

By cancellation,

$$(1-\cos\theta)F_n(\theta) = 1-\cos(n\theta) \tag{A.28}$$

so $F_n(\theta)\geq 0$ and, by (A.27), $\int F_n(\theta)\frac{d\theta}{2\pi}=n$. Thus,

$$|\pi_n'(e^{i\theta})| \leq \|\pi_n\|_\infty \int |F_n(\theta-\varphi)|\,\frac{d\varphi}{2\pi} = n\|\pi_n\|_\infty \qquad \square$$

Theorem A.6. *Let π_n be an arbitrary polynomial of degree n. Then*

$$\sup_{x\in[-a,a]} \bigl[|\pi_n'(x)|(a^2-x^2)^{1/2}\bigr] \leq 3n \sup_{x\in[-a,a]} |\pi_n(x)| \tag{A.29}$$

Proof. By scaling, we need only check the case $a=2$. Define

$$\tilde\pi_n(z) = z^n\pi_n\biggl(z+\frac{1}{z}\biggr) \tag{A.30}$$

$\tilde\pi_n$ is a polynomial of degree $2n$ so, by (A.24),

$$\sup_\theta |\tilde\pi_n'(e^{i\theta})| \leq 2n \sup_{x\in[-2,2]} |\pi_n(x)| \tag{A.31}$$

since $e^{i\theta}\to e^{i\theta}+e^{-i\theta}=2\cos\theta$ maps $\partial\mathbb{D}$ to $[-2,2]$.

By (A.30),

$$\begin{aligned}\pi_n'\biggl(z+\frac{1}{z}\biggr)(1-z^{-2}) &= \frac{d}{dz}\,z^{-n}\tilde\pi_n(z)\\ &= -nz^{-n-1}\tilde\pi_n(z)+z^{-n}\tilde\pi_n'(z)\end{aligned}$$

so, by (A.31),

$$\sup_{e^{i\theta}\in\partial\mathbb{D}} |\pi_n'(2\cos\theta)2\sin\theta| \leq 3n\|\tilde\pi_n\|_\infty$$

which is (A.29) for $a=2$. $\square$

(d) Dirichlet trial polynomials. For use in both (A.4) and (A.18), we want a rich set of trial polynomials, $\pi_n(x)$. In particular, we want π_n's concentrated near $x=x_0$ and otherwise small in some interval $[x_0-a,x_0+a]$. By scaling, we may as well consider $x_0=0$, $a=1$. An analyst might try $(1-x^2)^n$, but that has width $n^{-1/2}$ — and we will see that one can do better. We will get width n^{-1}. One can't do better than this, by Bernstein's inequality, if $\pi_n(\theta)=1$ and $\|\pi_n\|_\infty=1$, then $\pi_n(x)\geq\frac{1}{2}$ for $|x|\leq\frac{1}{2n}-O(\frac{1}{n^2})$.

Our choice is related to Dirichlet and Fejér kernels and is, in fact, essentially the minimizer for the Christoffel problem with $x_0=0$ and $d\mu=\chi_{[-1,1]}(1-x^2)^{-1/2}\,dx$.

Theorem A.7. *For any $x_0\in\mathbb{R}$ and $a>0$, there exist, for each n, polynomials $\pi_n(x;x_0,a)$ so that*

(i)
$$\deg \pi_n = 2n - 2 \tag{A.32}$$
(ii)
$$\pi_n(x_0) = 1 \tag{A.33}$$
(iii)
$$|\pi_n(x)| \le \min\biggl(1, \frac{1}{2n} + \frac{a}{2n|x-x_0|}\biggr) \qquad \textit{if } |x-x_0| \le a \tag{A.34}$$
(iv) *For any* $\delta \le a$*,*
$$\int_{x_0-\delta}^{x_0+\delta} |\pi_n(x)|^2\, dx = \frac{\pi a}{n} + O\biggl(\frac{1}{n^2}\biggr) \tag{A.35}$$

Remark. (A.34) implies $|\pi_n(x)| \le C_\delta/n$ if $|x-x_0| > \delta$, and for any ε, $|\pi_n(x)| < \varepsilon$ if $|x-x_0| \ge C_\varepsilon/n$.

Proof. By scaling, we can suppose that $x_0 = 0$, $a = 1$, in which case we will call the polynomials D_n, that is,
$$\pi_n(x; x_0, a) = D_n\biggl(\frac{x-x_0}{a}\biggr) \tag{A.36}$$
Recall there are polynomials $T_n(x)$ (Chebyshev of the first kind) with
$$\deg T_n = n \tag{A.37}$$
so that
$$T_n(\cos\theta) = \cos(n\theta) \tag{A.38}$$
Define D_n by
$$D_n(x) = \frac{1}{n} \sum_{j=0}^{n-1} (-1)^j T_{2j}(x) \tag{A.39}$$

By (A.37), (A.32) holds for D_n. By (A.38) (and $\cos\theta = 0 \Leftrightarrow \theta = \frac{\pi}{2}$ mod π), $T_{2j}(0) = (-1)^j$, so D_n obeys (A.33) for $x_0 = 0$.

By (A.38), $|T_n(x)| \le 1$ on $[-1,1]$, so
$$|D_n(x)| \le 1 \qquad \text{on } [-1,1]$$
which is half of (A.34). For the other half, sum the geometric series to see that
$$D_n(\cos\theta) = \frac{1}{2n} + \frac{(-1)^{n-1}}{2n} \, \frac{\cos((2n-1)\theta)}{\cos\theta} \tag{A.40}$$
which implies the other half of (A.34).

Since
$$\int_{-\pi}^{\pi} \cos(2k\theta)\cos(2j\theta)\, \frac{d\theta}{2\pi} = \begin{cases} 1 & \text{if } k=j=0 \\ \frac{1}{2} & \text{if } k=j\ne 0 \\ 0 & \text{if } k\ne j \end{cases}$$

we have

$$\begin{aligned}\int_{-\pi}^{\pi} D_n^2(\cos\theta)\,\frac{d\theta}{2\pi} &= \frac{1}{n^2}\left[1+\frac{1}{2}\,(n-1)\right]\\ &= \frac{1}{2n}+O\biggl(\frac{1}{n^2}\biggr)\end{aligned}$$

Since $d\theta = (1+O(x^2))\,dx$ near $x=0$ and $D_n^2 = O(\frac{1}{n^2})$ away from $x=0$, we obtain (A.35) when $a=1$. $\square$

References

[1] J. Avron and B. Simon, *Almost periodic Schrödinger operators, II. The integrated density of states*, Duke Math. J. **50** (1983), 369–391.

[2] J. Avron, P. M. H. van Mouche, and B. Simon, *On the measure of the spectrum for the almost Mathieu operator*, Comm. Math. Phys. **132** (1990), 103–118.

[3] M. J. Cantero, L. Moral, and L. Velázquez, *Measures and para-orthogonal polynomials on the unit circle*, East J. Approx. **8** (2002), 447–464.

[4] M. J. Cantero, L. Moral, and L. Velázquez, *Measures on the unit circle and unitary truncations of unitary operators*, J. Approx. Theory **139** (2006), 430–468.

[5] R. Carmona, *Exponential localization in one-dimensional disordered systems*, Duke Math. J. **49** (1982), 191–213.

[6] D. Damanik and B. Simon, *Jost functions and Jost solutions for Jacobi matrices, II. Decay and analyticity*, to appear in Int. Math. Res. Not.

[7] E. B. Davies and B. Simon, *Eigenvalue estimates for non-normal matrices and the zeros of random orthogonal polynomials on the unit circle*, to appear in J. Approx. Theory.

[8] P. A. Deift and B. Simon, *Almost periodic Schrödinger operators, III. The absolutely continuous spectrum in one dimension*, Comm. Math. Phys. **90** (1983), 389–411.

[9] P. Erdös and P. Turán, *On interpolation. III. Interpolatory theory of polynomials*, Ann. of Math. (2) **41** (1940), 510–553.

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

[11] 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.

[12] D. J. Gilbert, *On subordinacy and analysis of the spectrum of Schrödinger operators with two singular endpoints*, Proc. Roy. Soc. Edinburgh Sect. A **112** (1989), 213–229.

[13] D. J. Gilbert and D. B. Pearson, *On subordinacy and analysis of the spectrum of one-dimensional Schrödinger operators*, J. Math. Anal. **128** (1987), 30–56.

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

[15] E. Hairer and G. Wanner, *Analysis by Its History*, Undergraduate Texts in Mathematics. Readings in Mathematics. Springer-Verlag, New York, 1996.

[16] E. Harrell and B. Simon *The mathematical theory of resonances whose widths are exponentially small*, Duke Math. J. **47** (1980), 845–902.

[17] D. Hundertmark and B. Simon, *Lieb-Thirring inequalities for Jacobi matrices*, J. Approx. Theory **118** (2002), 106–130.

[18] S. Jitomirskaya and Y. Last, *Dimensional Hausdorff properties of singular continuous spectra*, Phys. Rev. Lett. **76** (1996), 1765–1769.

[19] S. Jitomirskaya and Y. Last, *Power-law subordinacy and singular spectra, I. Half-line operators*, Acta Math. **183** (1999), 171–189.

[20] S. Jitomirskaya and Y. Last, *Power law subordinacy and singular spectra, II. Line operators*, Comm. Math. Phys. **211** (2000), 643–658.

[21] S. Jitomirskaya, H. Schulz-Baldes, and G. Stolz, *Delocalization in random polymer models*, Comm. Math. Phys. **233** (2003), 27–48.

[22] R. Johnson and J. Moser, *The rotation number for almost periodic potentials*, Comm. Math. Phys. **84** (1982), 403–438.

[23] S. Kahn and D. B. Pearson, *Subordinacy and spectral theory for infinite matrices*, Helv. Phys. Acta **65** (1992), 505–527.

[24] R. Killip, A. Kiselev, and Y. Last, *Dynamical upper bounds on wavepacket spreading*, Am. J. Math. **125** (2003), 1165–1198.

[25] R. Killip and B. Simon, *Sum rules for Jacobi matrices and their applications to spectral theory*, Ann. of Math. (2) **158** (2003), 253–321.

[26] R. Killip and B. Simon, *Sum rules and spectral measures of Schrödinger operators with L^2 potentials*, preprint.

[27] R. Killip and M. Stoiciu, in preparation.

[28] W. Kirsch and F. Martinelli, *On the density of states of Schrödinger operators with a random potential*, J. Phys. A **15** (1982), 2139–2156.

[29] R. J. Kooman, *Asymptotic behaviour of solutions of linear recurrences and sequences of Möbius-transformations*, J. Approx. Theory **93** (1998), 1–58.

[30] Y. Last, *On the measure of gaps and spectra for discrete 1D Schrödinger operators*, Comm. Math. Phys. **149** (1992), 347–360.

[31] MathWorld on Jacobi Polynomials,
`http://mathworld.wolfram.com/JacobiPolynomial.html`

[32] N. Minami, *Local fluctuation of the spectrum of a multidimensional Anderson tight binding model*, Comm. Math. Phys. **177** (1996), 709–725.

[33] S. A. Molchanov, *The local structure of the spectrum of the one-dimensional Schrödinger operator*, Comm. Math. Phys. **78** (1980/81), 429–446.

[34] P. Nevai, *Orthogonal polynomials*, Mem. Amer. Math. Soc. **18** (1979), no. 213, 185 pp.

[35] P. Nevai, *Géza Freud, orthogonal polynomials and Christoffel functions. A case study*, J. Approx. Theory **48** (1986), 167 pp.

[36] R. G. Newton, *Scattering Theory of Waves and Particles*, second edition, Springer, New York, 1982.

[37] K. Pan, *Asymptotics for orthogonal polynomials beyond the analytic boundary*, Rocky Mountain J. Math. **26** (1996), 269–279.

[38] L. A. Pastur, *Spectra of random selfadjoint operators*, Uspekhi Mat. Nauk **28** (1973), 3–64.

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

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

[41] B. Simon, *Fine structure of the zeros of orthogonal polynomials, I. A tale of two pictures*, to appear in Proc. Constructive Functions Tech-04.

[42] B. Simon, *Fine structure of the zeros of orthogonal polynomials, II. OPUC with competing exponential decay*, J. Approx. Theory **135** (2005), 125–139.

[43] B. Simon, *Fine structure of the zeros of orthogonal polynomials, III. Periodic recursion coefficients*, Comm. Pure Appl. Math. **59** (2005) 1042–1062.

[44] B. Simon, *Meromorphic Szegő functions and asymptotic series for Verblunsky coefficients*, to appear in Acta Math.

[45] B. Simon, *Orthogonal polynomials with exponentially decaying recursion coefficients*, preprint

[46] B. Simon, *Rank one perturbations and the zeros of paraorthogonal polynomials on the unit circle*, preprint

[47] M. Stoiciu, *The statistical distribution of the zeros of random paraorthogonal polynomials on the unit circle*, J. Approx. Theory **139** (2006), 29–64.

[48] M. Stoiciu, *Zeros of Random Orthogonal Polynomials on the Unit Circle*, Ph.D. dissertation, 2005.
`http://etd.caltech.edu/etd/available/etd-05272005-110242/`

[49] J. Szabados and P. Vértesi, *Interpolation of Functions*, World Scientific Publishing, Teaneck, NJ, 1990.

[50] G. Szegő, *Über einen Satz des Herrn Serge Bernstein*, Schriften Königsberg **5** (1928), 59–70.

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

[52] P. Vértesi, *On the zeros of Jacobi polynomials*, Studia Sci. Math. Hungar. **25** (1990), 401–405.

[53] P. Vértesi, *On the zeros of generalized Jacobi polynomials. The heritage of P. L. Chebyshev: a Festschrift in honor of the 70th birthday of T. J. Rivlin*, Ann. Numer. Math. **4** (1997), 561–577.

[54] P. Vértesi, *Uniform asymptotics of derivatives of orthogonal polynomials based on generalized Jacobi weights*, Acta Math. Hungar. **85** (1999), 97–130.

[55] M.-W. L. Wong, *First and second kind paraorthogonal polynomials and their zeros*, in preparation.