

Asymptotics of Chebyshev polynomials, I: subsets of \mathbb{R}

Jacob S. Christiansen¹ · Barry Simon² · Maxim Zinchenko³

Received: 18 October 2015 / Accepted: 30 April 2016 / Published online: 19 September 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract We consider Chebyshev polynomials, $T_n(z)$, for infinite, compact sets $\epsilon \subset \mathbb{R}$ (that is, the monic polynomials minimizing the sup-norm, $\|T_n\|_\epsilon$, on ϵ). We resolve a 45+ year old conjecture of Widom that for finite gap subsets of \mathbb{R} , his conjectured asymptotics (which we call Szegő–Widom asymptotics) holds. We also prove the first upper bounds of the form $\|T_n\|_\epsilon \leq QC(\epsilon)^n$ (where $C(\epsilon)$ is the logarithmic capacity of ϵ) for a class of ϵ 's with an infinite number of components, explicitly for those $\epsilon \subset \mathbb{R}$ that obey a Parreau–Widom condition.

Mathematics Subject Classification 41A50 · 30E15 · 30C10

B. Simon's research was supported in part by NSF Grant DMS-1265592 and in part by Israeli BSF Grant No. 2010348. M. Zinchenko's research was supported in part by Simons Foundation Grant CGM-281971.

✉ Barry Simon
bsimon@caltech.edu

Jacob S. Christiansen
stordal@maths.lth.se

Maxim Zinchenko
maxim@math.unm.edu

- ¹ Centre for Mathematical Sciences, Lund University, Box 118, 22100 Lund, Sweden
- ² Departments of Mathematics and Physics, Mathematics 253-37, California Institute of Technology, Pasadena, CA 91125, USA
- ³ Department of Mathematics and Statistics, University of New Mexico, Albuquerque, NM 87131, USA

1 Introduction

This paper is the first of what we hope will be a series studying the asymptotics of Chebyshev polynomials associated to an arbitrary compact subset, $\epsilon \subset \mathbb{C}$, which has an infinite number of points. These are those degree n monic polynomials, T_n , which minimize

$$\|f\|_\epsilon = \sup_{z \in \epsilon} |f(z)| \quad (1.1)$$

See [36] for background on general Chebyshev polynomials and their applications. It is known (see below for the case $\epsilon \subset \mathbb{R}$) that the minimizer is unique. We will denote this minimizer as T_n in cases where the intended ϵ is clear. If $\operatorname{Re}P$ is the polynomial whose coefficients are the real parts of those of P , we have that $|\operatorname{Re}P(x)| < |P(x)|$ for all but finitely many $x \in \mathbb{R}$ unless $\operatorname{Re}P \equiv P$. Since $\operatorname{Re}P$ is monic if P is, $\operatorname{Re}T_n$ is also a Chebyshev polynomial so by the alternation theorem (see Theorem 1.1 below), $|\operatorname{Re}T_n(x)| = \|T_n\|$ has at least $n + 1$ solutions in \mathbb{R} if $\epsilon \subset \mathbb{R}$. Hence $\operatorname{Im}T_n \equiv 0$, that is, T_n is real.

We begin by recalling a basic result which goes back to Borel [4] and Markov [20] (which according to Akhiezer [1] was based on lectures from 1905 but only published in 1948). It depends on a basic notion that comes from ideas of Chebyshev [6]:

Definition We say that P_n , a real degree n polynomial, has an *alternating set* in $\epsilon \subset \mathbb{R}$ if there exists $\{x_j\}_{j=0}^n \subset \epsilon$ with $x_0 < x_1 < \dots < x_n$ so that

$$P_n(x_j) = (-1)^{n-j} \|P_n\|_\epsilon \quad (1.2)$$

Theorem 1.1 (The Alternation Theorem) *Let $\epsilon \subset \mathbb{R}$ be compact. The Chebyshev polynomial of degree n for ϵ has an alternating set in ϵ . Conversely, any monic polynomial with an alternating set in ϵ is the Chebyshev polynomial for ϵ .*

Proof The proof is simple and not so available in our generality, so we include it—it is essentially what Markov gives in [20] for the case $\epsilon = [a, b]$. If T_n is the Chebyshev polynomial, let $y_0 < y_1 < \dots < y_k$ be the set of all the points in ϵ where it takes the value $\pm \|T_n\|_\epsilon$. If there are fewer than n sign changes among these ordered points, then we can find a degree at most $n - 1$ polynomial, Q , non-vanishing at each y_j and with the same sign as T_n at those points. For ϵ small and positive, $T_n - \epsilon Q$ will be a monic polynomial with smaller $\|\cdot\|_\epsilon$. Thus there must be at least n sign flips and therefore an alternating set.

Conversely, let P_n be a degree n monic polynomial with an alternating set and suppose that $\|T_n\|_\epsilon < \|P_n\|_\epsilon$. Then at each point, x_j , in the alternating set

for P_n , $Q \equiv P_n - T_n$ has the same sign as P_n , so Q has at least n zeros, which is impossible, since it is of degree at most $n - 1$. \square

The alternation theorem implies uniqueness of the Chebyshev polynomial. For, if T_n and S_n are two minimizers, so is $Q \equiv \frac{1}{2}(T_n + S_n)$. At the alternating points for Q , we must have $T_n = S_n$, so they must be equal polynomials since there are $n + 1$ points in the alternating set and their difference has degree at most $n - 1$.

The alternation theorem also implies some simple facts about the zeros of T_n :

- (a) All the zeros of the Chebyshev polynomials of a set $\epsilon \subset \mathbb{R}$ lie in \mathbb{R} and all are simple and lie in $\text{cvh}(\epsilon)$, the convex hull of ϵ . This is because there must be at least one zero between any pair of points in an alternating set and this accounts for all n zeros. The same argument shows that for any $\gamma \in (-\|T_n\|_\epsilon, \|T_n\|_\epsilon)$ all n solutions of $T_n(x) = \gamma$ are simple and lie in $\text{cvh}(\epsilon)$. This plus the open mapping theorem implies that (inverse as a function from \mathbb{C})

$$\epsilon_n \equiv T_n^{-1}([-\|T_n\|_\epsilon, \|T_n\|_\epsilon]) \subset \text{cvh}(\epsilon) \tag{1.3}$$

- (b) By a *gap* of $\epsilon \subset \mathbb{R}$, we mean a bounded connected component of $\mathbb{R} \setminus \epsilon$. If there are only finitely many gaps and no component of ϵ is a single point, we speak of a *finite gap set*. Between any two zeros of T_n , there is a point in the alternating set so each gap of $\epsilon \subset \mathbb{R}$ has at most one zero of T_n .
- (c) Above the top zero (resp. below the bottom zero) of T_n , $|T_n(x)|$ is monotone increasing. It follows that $x_n = \sup_{y \in \epsilon} y$ (resp. $x_0 = \inf_{y \in \epsilon} y$) so at the endpoints of $\text{cvh}(\epsilon) \subset \mathbb{R}$ we have that $|T_n(x)| = \|T_n\|_\epsilon$.

To get ahead of our story, a key understanding in our analysis in this paper is that ϵ_n defined in (1.3) is the spectrum of a periodic Schrödinger operator and up to normalization, T_n is its spectral theory discriminant; see Sect. 2. In the end, we found proofs that didn't emphasize this connection which we use to make this paper more accessible to specialists on orthogonal polynomials. But this intuition played an important role in our research. We mention in passing that our methods don't appear to work for weighted Chebyshev polynomials.

Going back at least to Szegő [41] is the idea that potential theory is essential to the study of Chebyshev polynomials. To settle the notation we use, we recall some of the basic definitions. References for the potential theory that we need include [16, 18, 27, 32, 35, 51]. Given a probability measure, $d\mu$, of compact support on \mathbb{C} , we define its *Coulomb energy*, $\mathcal{E}(\mu)$ by

$$\mathcal{E}(\mu) = \int d\mu(x) d\mu(y) \log |x - y|^{-1} \tag{1.4}$$

and we define the *Robin constant* of a compact set $\epsilon \subset \mathbb{C}$ by

$$R(\epsilon) = \inf\{\mathcal{E}(\mu) \mid \text{supp}(\mu) \subset \epsilon \text{ and } \mu(\epsilon) = 1\} \tag{1.5}$$

If $R(\epsilon) = \infty$, we say ϵ is a *polar set* or has *capacity zero*. If something holds except for a polar set, we say it holds q.e. (for *quasi-everywhere*). The *capacity*, $C(\epsilon)$, of ϵ is defined by

$$C(\epsilon) = \exp(-R(\epsilon)), \quad R(\epsilon) = \log(1/C(\epsilon)) \tag{1.6}$$

If ϵ is not a polar set, it follows from weak lower semicontinuity of $\mathcal{E}(\cdot)$ and weak compactness of the family of probability measures that there is a probability measure whose Coulomb energy is $R(\epsilon)$. Since $\mathcal{E}(\cdot)$ is strictly convex on the probability measures, this minimizer is unique. It is called the *equilibrium measure* or *harmonic measure* of ϵ and denoted $d\rho_\epsilon$. The second name comes from the fact (see Conway [9] or Simon [35]) that if f is a continuous function on ϵ , there is a unique function, u_f , harmonic on $(\mathbb{C} \cup \{\infty\}) \setminus \epsilon$, which approaches $f(x)$ for q.e. $x \in \epsilon$ (i.e., solves the Dirichlet problem) and

$$u_f(\infty) = \int_\epsilon f(x) d\rho_\epsilon(x) \tag{1.7}$$

The function $\Phi_\epsilon(z) = \int_\epsilon d\rho_\epsilon(x) \log|x - z|^{-1}$ is called the *equilibrium potential*. The *Green's function*, $G_\epsilon(z)$, of a compact subset, $\epsilon \subset \mathbb{C}$, is defined by

$$G_\epsilon(z) = R(\epsilon) - \Phi_\epsilon(z) \tag{1.8}$$

It follows from Frostman's theorem that it is the unique function harmonic on $\mathbb{C} \setminus \epsilon$ with q.e. boundary value 0 on ϵ and so that $G_\epsilon(z) - \log|z|$ is harmonic at ∞ . Moreover, $G_\epsilon(z) \geq 0$ everywhere and near ∞

$$G_\epsilon(z) = \log|z| + R(\epsilon) + O(1/|z|) \tag{1.9}$$

equivalently,

$$\exp(G_\epsilon(z)) = \frac{|z|}{C(\epsilon)} + O(1) \tag{1.10}$$

If G_ϵ is zero on ϵ and continuous on all of \mathbb{C} , we say that ϵ is *regular* (for *potential theory*).

To put our new results in context, we need to remind the reader of some previous results. Using what is now called the Bernstein–Walsh lemma, Szegő [41] proved for all non-polar compact sets $\epsilon \subset \mathbb{C}$

$$\|T_n\|_\epsilon \geq C(\epsilon)^n \tag{1.11}$$

which was improved when $\epsilon \subset \mathbb{R}$ by Schiefermayr [31] (see Sect. 2) to

$$\|T_n\|_\epsilon \geq 2C(\epsilon)^n \quad (1.12)$$

Szegő [41], using in part prior results of Faber [11] and Fekete [13], proved

Theorem 1.2 (FFS Theorem) *For any compact set $\epsilon \subset \mathbb{C}$, one has that*

$$\lim_{n \rightarrow \infty} \|T_n\|_\epsilon^{1/n} = C(\epsilon) \quad (1.13)$$

Upper bounds on $\|T_n\|_\epsilon$ which complement (1.11) or (1.12) in that they also grow like $C(\epsilon)^n$ are clearly interesting. The following is known

Theorem 1.3 (Totik–Widom Theorem) *For any finite gap set $\epsilon \subset \mathbb{R}$, one has, for a constant $Q \geq 2$ depending on ϵ , that*

$$\|T_n\|_\epsilon \leq QC(\epsilon)^n \quad (1.14)$$

Remarks 1. This result follows from work of Widom [53] that we discuss below on asymptotics of $\|T_n\|_\epsilon$ (see Theorem 1.8). Totik [46] proved an equivalent result but in a different form involving control of $C(\epsilon_n)$ where ϵ_n is the set of (1.3). We'll discuss this further in Sect. 4.

2. Neither approach leads to very explicit control on the constant Q (although Widom explicitly finds $\limsup_{n \rightarrow \infty} \|T_n\|_\epsilon / C(\epsilon)^n$ in terms of the solution to a minimization problem and he does have explicit bounds on this lim sup but not on the sup).
3. Widom proved this bound also for certain sets $\epsilon \subset \mathbb{C}$ that have finitely many components. Recently, Andrievskii [2] and Totik–Varga [49] have increased the family of finite component sets in \mathbb{C} for which (1.14) holds.

One of our two main results in this paper extends this last result to a larger class of sets $\epsilon \subset \mathbb{R}$ with a simple explicit bound on Q in terms of G_ϵ . Recall [22, 54].

Definition A set $\epsilon \subset \mathbb{C}$ is said to be a *Parreau–Widom set* if

$$PW(\epsilon) \equiv \sum_{w \in \mathcal{C}} G_\epsilon(w) < \infty \quad (1.15)$$

where \mathcal{C} is the set of critical points of G_ϵ (i.e., points where $\partial G_\epsilon(w) = 0$).

In this paper, we use Wirtinger calculus:

$$\partial = \frac{1}{2} \left(\frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right), \quad \bar{\partial} = \frac{1}{2} \left(\frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right) \quad (1.16)$$

We recall that $\bar{\partial} f = 0$ are the Cauchy–Riemann equations (so that, for harmonic functions, $u, \partial u$ is analytic) and that for analytic functions, f , we have that $\partial f = f'$, the complex derivative. Moreover, by the Cauchy–Riemann equations

$$f \text{ analytic} \Rightarrow 2 \partial(\operatorname{Re}(f)) = f' \tag{1.17}$$

If $\epsilon \subset \mathbb{R}$, it is easy to see all the critical points lie in \mathbb{R} . If ϵ is also regular, there is exactly one critical point in each gap and so (1.15) is the sum over the maxima of G_ϵ in the gaps. In particular, every finite gap set is a Parreau–Widom set. Our new result, proven in Sect. 4, is:

Theorem 1.4 *If $\epsilon \subset \mathbb{R}$ is a regular Parreau–Widom set, then*

$$\|T_n\|_\epsilon \leq 2 \exp(PW(\epsilon))C(\epsilon)^n \tag{1.18}$$

Remarks 1. For a finite gap set, the sum in (1.15) has finitely many terms, so is finite, and thus this result implies Theorem 1.3 with a fairly explicit Q . We note that homogeneous sets in the sense of Carleson [5], and, in particular, positive measure Cantor sets, are regular Parreau–Widom sets [17].

2. Recent work of Goncharov and Hatinoğlu [15] shows that there are very thin Cantor-type sets for which $\|T_n\|_\epsilon / C(\epsilon)^n$ grows subexponentially.
3. It is interesting to know for which infinite gap sets in \mathbb{R} a bound like (1.14) is true. For example, does the classical 1/3 Cantor set, which is not a Parreau–Widom set, obey (1.14)?
4. We wonder if this result extends to Parreau–Widom sets in \mathbb{C} .
5. For the finite gap case, Widom [53] obtains a bound on the lim sup involving $\exp(PW(\epsilon))$ and our result is compatible with his in this finite gap case.

Our main focus will be on pointwise asymptotics of $T_n(z)$ on $\mathbb{C} \setminus \epsilon$. The earliest results on this subject go back to Faber [11] in 1919. Let ϵ be a Jordan region with analytic boundary, i.e., an analytic Jordan curve together with its interior region. By the maximum principle, the Chebyshev polynomials for ϵ are the same as those for the curve. There is a unique Riemann map, $B_\epsilon(z)$, from $(\mathbb{C} \cup \{\infty\}) \setminus \epsilon$ onto \mathbb{D} which is a bijection with $B_\epsilon(\infty) = 0$ and positive “derivative”, $B'_\epsilon(\infty)$, at ∞ . Then:

Theorem 1.5 (Faber [11]) *If ϵ is a Jordan region with an analytic boundary, then*

$$\lim_{n \rightarrow \infty} T_n(z) B_\epsilon(z)^n B'_\epsilon(\infty)^{-n} = 1 \tag{1.19}$$

uniformly for z in a neighborhood of the closure of $(\mathbb{C} \cup \{\infty\}) \setminus \epsilon$.

Remarks 1. Since the curve is assumed analytic, $B_\epsilon(z)$ has a continuation into a neighborhood of the curve.

2. Since B_ϵ maps the curve to $\partial\mathbb{D}$, and G_ϵ is unique, on $(\mathbb{C} \cup \{\infty\}) \setminus \epsilon$, we have that

$$|B_\epsilon(z)| = \exp(-G_\epsilon(z)) \tag{1.20}$$

so that near ∞ , we have that

$$B_\epsilon(z) = C(\epsilon)z^{-1} + O(|z|^{-2}) \tag{1.21}$$

This implies that $B'_\epsilon(\infty) = C(\epsilon)$; thus, Faber’s result implies that $\lim_{n \rightarrow \infty} \|T_n\|_\epsilon / C(\epsilon)^n = 1$, a strong version of the FFS theorem.

3. We call (1.19) *Szegő asymptotics* after Szegő’s famous result [39, 40] on the asymptotics of OPUC. Since Faber’s paper was earlier than Szegő’s, this naming is perhaps unfair, but the term Szegő asymptotics is so common, we use it in this case also.

In 1969, Widom wrote a 100+ page brilliant, seminal work [53] on the asymptotics of Chebyshev and orthogonal polynomials associated to a set ϵ , where ϵ is the union of a finite number of Jordan regions with C^{2+} boundary and C^{2+} Jordan arcs (i.e., not closed, simple curves). As in Faber’s case, the polynomials are the same whether one takes Jordan regions or their boundaries, as Widom does.

Widom began by looking for the replacement for $B_\epsilon(z)$ in (1.19)/(1.20). Since $G_\epsilon(z)$ is harmonic on $\mathbb{C} \setminus \epsilon$, it has a local harmonic conjugate so one can locally define an analytic function, $B_\epsilon(z)$, on $(\mathbb{C} \cup \{\infty\}) \setminus \epsilon$ obeying (1.20) (∞ is a removable singularity if one sets $B_\epsilon(\infty) = 0$). $B_\epsilon(z)$ is determined by (1.20), up to a phase which we can fix by demanding (1.21) near ∞ .

$B_\epsilon(z)$ can be continued along any curve lying in $(\mathbb{C} \cup \{\infty\}) \setminus \epsilon$ and, by the monodromy theorem, the continuation is the same for homotopic curves. Since G_ϵ is continuous, only the phase can change, i.e., the phase change is associated with a character, χ_ϵ , of the fundamental group of $(\mathbb{C} \cup \{\infty\}) \setminus \epsilon$. The character is non-trivial if ϵ is not connected (up to polar sets)—indeed, if a curve loops once around a subset g of ϵ , the phase change in B_ϵ is $\exp(-2\pi i \rho_\epsilon(g))$; see Theorem 2.7.

There is a language introduced by Sodin–Yuditskii [37] for doing the book-keeping for such functions. It relies on the fact that the universal cover of $(\mathbb{C} \cup \{\infty\}) \setminus \epsilon$ is \mathbb{D} . Using the notation from our presentation of this machinery [8], there is a Fuchsian group, Γ , of Möbius transformations on \mathbb{D} , and a map $\mathbf{x}(z)$ from $\mathbb{D} \rightarrow (\mathbb{C} \cup \{\infty\}) \setminus \epsilon$ which is automorphic (i.e., invariant under Γ).

\mathbf{x} is a covering map, so a local bijection. Its “inverse”, $\mathbf{z}(x)$, is a multivalued analytic function which is not character automorphic—rather its values are an orbit of the group Γ . Γ is such that $\sum_{\gamma \in \Gamma} (1 - |\gamma(0)|) < \infty$ so one can form the Blaschke product $B(z) = \prod_{\gamma \in \Gamma} b(z, \gamma(0))$. In this language, one should use a complex variable, x , on $(\mathbb{C} \cup \{\infty\}) \setminus \epsilon$ in which case one has

that $B_\epsilon(x) = B(\mathbf{z}(x))$ and the object whose asymptotics we should look at is $T_n(\mathbf{x}(z))B(z)^n C(\epsilon)^{-n}$ on \mathbb{D} . While we feel this language should be in the back of one’s mind, we will do our analysis with multivalued functions on $\Omega \equiv (\mathbb{C} \cup \{\infty\}) \setminus \epsilon$, as Widom did.

So Widom looked at $T_n(z)B_\epsilon(z)^n C(\epsilon)^{-n}$. Unlike the simply connected case of Ω studied by Faber, this cannot have a pointwise limit because the character of this character automorphic function is χ_ϵ^n which is not constant! Instead Widom found a good candidate for the asymptotics:

Theorem 1.6 (Widom [53]) *Let ϵ be a finite union of disjoint smooth Jordan regions and arcs. For every character, χ , of the fundamental group of Ω there is a character automorphic function with that character, $F(z, \chi)$, on Ω which minimizes $\|f\|_\Omega$ among all character automorphic functions, f , with that character and which obey $f(\infty, \chi) = 1$. Moreover, this minimizer is unique and it and its $\|\cdot\|_\Omega$ are continuous in χ (the functions in the topology of uniform convergence on compact subsets of the universal cover of Ω).*

- Remarks*
1. We should refer to F as a function on the universal cover of Ω , not Ω . The results of Widom then also say that every minimizer is a finite product of Blaschke products on \mathbb{D} .
 2. Continuity in χ and uniqueness are intimately related.
 3. We will use $F_n(z)$ for the function $F(z, \chi_\epsilon^n)$. Note that this function is almost periodic in n (in fact, quasiperiodic) since the minimizer is continuous in χ .

The *Widom surmise* is the notion that

$$\lim_{n \rightarrow \infty} \left[\frac{T_n(z)B_\epsilon(z)^n}{C(\epsilon)^n} - F_n(z) \right] = 0 \tag{1.22}$$

When it holds uniformly on compact subsets of the universal cover of Ω , we will say that ϵ has *Szegő–Widom asymptotics*.

Widom proved two results about the asymptotics of T_n . The first involves the situation where there are no arcs—but only regions:

Theorem 1.7 (Widom [53]) *Let ϵ be the union of a finite number of disjoint Jordan regions with smooth boundaries. Then ϵ has Szegő–Widom asymptotics, i.e., (1.22) holds uniformly on compact subsets of the universal cover of Ω . Moreover,*

$$\lim_{n \rightarrow \infty} \frac{\|T_n\|_\epsilon}{C(\epsilon)^n \|F_n\|_\Omega} = 1 \tag{1.23}$$

The second concerns finite gap sets in \mathbb{R} :

Theorem 1.8 (Widom [53]) *Let ϵ be a finite gap subset of \mathbb{R} . Then*

$$\lim_{n \rightarrow \infty} \frac{\|T_n\|_\epsilon}{C(\epsilon)^n \|F_n\|_\Omega} = 2 \tag{1.24}$$

Widom also conjectured that one had Szegő–Widom asymptotics in this case. At first sight this seems surprising—(1.22) suggests that one might expect the limit in (1.24) to be 1 as it is in (1.23). Widom was clearly motivated by the example of $\epsilon = [-1, 1]$ where both (1.22) and (1.24) hold! Indeed, in that case $\mathbf{x}(z)$ is one half the Joukowski map, $\mathbf{x}(z) = \frac{1}{2}(z + z^{-1})$, and $B_\epsilon(x) = \mathbf{z}(x)$, the inverse of the one half the Joukowski map, i.e., $B_\epsilon(z) = z - \sqrt{z^2 - 1}$ and $1/B_\epsilon(z) = z + \sqrt{z^2 - 1}$. The familiar formula for the Chebyshev polynomials in this case (a multiple the usual Chebyshev polynomials of the first kind with the multiple chosen to make the polynomials monic) is:

$$T_n(\cos(\theta)) = 2^{-n+1} \cos(n\theta), \quad T_n(z) = 2^{-n}[B_\epsilon^n(z) + B_\epsilon^{-n}(z)] \quad (1.25)$$

This implies that $\|T_n\|_\epsilon = 2^{-n+1} = 2C(\epsilon)^n$ since $C([-1, 1]) = 1/2$. This is consistent with the FSS Theorem and saturates Schiefermayr’s bound (1.12). Since (1.20) holds and G_ϵ is 0 (resp. > 0) on ϵ (resp. off ϵ), we have that $|B_\epsilon|$ is 1 (resp. < 1) on ϵ (resp. off ϵ). Thus off ϵ , only B_ϵ^{-n} contributes to the asymptotics while on ϵ , there are points with $B_\epsilon(z) = 1$ so both terms contribute and the norm is twice as large as one might have expected. This explains where Widom’s conjecture came from. Our second main result here is a proof of this conjecture:

Theorem 1.9 *The Chebyshev polynomials for any finite gap set in \mathbb{R} have Szegő–Widom asymptotics.*

It seems to us reasonable to think that any Parreau–Widom subset of \mathbb{R} with the additional requirement that the direct Cauchy theorem holds (see the discussion in section 3 of Christiansen [7] as well as [52, 55]) has Szegő–Widom asymptotics. This will be investigated in an upcoming paper.

Besides these results, we exploit the connection of Chebyshev polynomials to the spectral theory of periodic Jacobi matrices which we present in Sect. 2. In Sect. 3 we discuss several results about root asymptotics and we prove Theorem 1.4 in Sect. 4 and Theorem 1.9 in Sect. 5.

As an aside we note that the limit 2 in (1.24) is special to the case of $\epsilon \subset \mathbb{R}$ even though Widom had conjectured the limit was 2 as long as there was at least one arc (and not just regions) included among the Jordan arcs and regions. Indeed, for the case where ϵ is a connected subset of the unit circle, the limit has been computed by Thiran–Detaille [42] who find it is always strictly between 1 and 2 if the connected set is a proper, non-empty subset. Moreover, Totik–Yuditskii [50] have shown the lim sup is strictly less than 2 if at least one Jordan region is included among the components of a set ϵ of Widom’s class and Totik [48] has shown the lim inf is strictly bigger than 1 if at least one Jordan arc is included among the components of a set ϵ of Widom’s class. But it still seems to us a reasonable, albeit difficult, conjecture that every set of Widom’s class has Szegő–Widom asymptotics.

2 Periodic sets

In this section, we'll see the important role played by the sets, ϵ_n , of (1.3) and the related sets:

$$\mathring{\epsilon}_n \equiv T_n^{-1}((- \|T_n\|_\epsilon, \|T_n\|_\epsilon)) \tag{2.1}$$

Clearly, by the definition of T_n and ϵ_n , we have that

$$\epsilon \subset \epsilon_n \tag{2.2}$$

We will see that the set ϵ_n determines many properties of T_n . In particular,

$$\|T_n\|_\epsilon = 2C(\epsilon_n)^n \tag{2.3}$$

which, by (2.2), implies Schiefermayr's bound, (1.12). If B_n is short for B_{ϵ_n} , we'll also prove (indeed, we'll use this to prove (2.3))

$$\frac{2T_n(z)}{\|T_n\|_\epsilon} = B_n(z)^n + B_n(z)^{-n} \tag{2.4}$$

Given our discussion of Szegő–Widom asymptotics for $\epsilon = [-1, 1]$, it should not be a surprise that (2.4) is a significant part of our proof of Theorem 1.9.

The equilibrium measure for ϵ_n which we'll denote ρ_n will also play a role. We'll prove that, for any gap K of ϵ , one has that

$$\rho_n(K) \leq 1/n \tag{2.5}$$

which will be the key to our proof of Theorem 1.4.

An interesting further fact concerns the weight that ρ_n gives to components of ϵ_n . We will call a compact set $g \subset \mathbb{R}$ a *period- n set* if and only if each connected component of g has harmonic measure k/n for some $k \in \{1, \dots, n\}$ (which, of course implies that g has at most n components and so is a finite gap set). We will prove that any ϵ_n is a period- n set and that conversely, if ϵ is a period- n set, then it is its own ϵ_n .

The name “period- n set” comes from the fact that these sets are precisely the spectra of two-sided periodic Jacobi matrices. The original proofs we had for some of the results we just described used the theory of such matrices and we have kept some of the terminology. While we will prove these results here using only the alternation theorem and some potential theory, we'll end the section with a brief indication of the approach that relies on the fact that $2T_n/\|T_n\|_\epsilon$ is the discriminant of a periodic Jacobi matrix.

We are not the first ones to note the special properties of polynomials, P , for which $P^{-1}([-A, A]) \subset \mathbb{R}$. Their use is implicit in much of the work

on the theory of periodic Schrödinger operators and Jacobi matrices as we'll explain at the end of this section. In the orthogonal polynomial community, there is an initial work of Geronimo–Van Assche [14] and important follow-up of Peherstorfer [23–26] and Totik [43,44,46,47].

Theorem 2.1 *Let ϵ be an infinite, compact subset of \mathbb{R} , T_n its n th Chebyshev polynomial and let ϵ_n and $\overset{\circ}{\epsilon}_n$ be given respectively by (1.3) and (2.1). Then there exist $\alpha_1 < \beta_1 \leq \alpha_2 < \dots < \beta_j \leq \alpha_{j+1} < \dots < \beta_n$ so that*

$$\overset{\circ}{\epsilon}_n = \bigcup_{j=1}^n (\alpha_j, \beta_j), \quad \epsilon_n = \bigcup_{j=1}^n [\alpha_j, \beta_j] \tag{2.6}$$

Moreover on (α_j, β_j) , we have that $(-1)^{n-j} T'_n(x) > 0$, $\{\alpha_1, \beta_n\} \in \epsilon$ and for each $j = 1, \dots, n - 1$, at least one of β_j and α_{j+1} lie in ϵ .

Proof As we noted in the consequences of the alternation theorem, for any $\gamma \in (-\|T_n\|_\epsilon, \|T_n\|_\epsilon)$ all n solutions of $T_n(x) = \gamma$ are simple and lie in $\text{cvh}(\epsilon)$. This implies the claimed structure for $\overset{\circ}{\epsilon}_n$ and ϵ_n , (2.6) and the derivative condition. The α 's and β 's are all the solution of $T_n(x) = \pm \|T_n\|_\epsilon$ so the remainder of the theorem is a restatement of the alternation theorem. \square

We will call $[\alpha_j, \beta_j] = \epsilon_n^{(j)}$, the j th band of ϵ_n . Define

$$\Delta_n(z) \equiv \frac{2T_n(z)}{\|T_n\|_\epsilon} \tag{2.7}$$

so that ϵ_n is exactly the set where $-2 \leq \Delta_n(x) \leq 2$ and Δ_n takes values in $\mathbb{C} \setminus [-2, 2]$ on $\mathbb{C} \setminus \epsilon_n$. The Joukowski map $z \mapsto z + z^{-1}$ takes \mathbb{D} one-one to $\mathbb{C} \setminus [-2, 2]$ and $\partial\mathbb{D}$ two-one to $[-2, 2]$ so its functional inverse $z \mapsto \frac{z}{2} - \sqrt{\left(\frac{z}{2}\right)^2 - 1}$ maps $(\mathbb{C} \cup \{\infty\}) \setminus [-2, 2]$ to \mathbb{D} . The numerical inverse of this, $z \mapsto \frac{z}{2} + \sqrt{\left(\frac{z}{2}\right)^2 - 1}$, thus maps $(\mathbb{C} \cup \{\infty\}) \setminus [-2, 2]$ to $(\mathbb{C} \cup \{\infty\}) \setminus \overline{\mathbb{D}}$. It follows that

$$\frac{\Delta_n(z)}{2} + \sqrt{\left(\frac{\Delta_n(z)}{2}\right)^2 - 1} \tag{2.8}$$

maps $\Omega_n \equiv (\mathbb{C} \cup \{\infty\}) \setminus \epsilon_n$ to $(\mathbb{C} \cup \{\infty\}) \setminus \overline{\mathbb{D}}$. If we take the log of the absolute value of this nonvanishing analytic function, we get a strictly positive harmonic function on $\mathbb{C} \setminus \epsilon_n$. Since this function approaches 0 as one approaches ϵ_n and is $n \log |z| + O(1)$ near ∞ , we have proven the first assertion in:

Theorem 2.2 *Let ϵ be an infinite compact subset of \mathbb{R} , T_n its n th Chebyshev polynomial, Δ_n given by (2.7) and let ϵ_n be given by (1.3). Then the Green's function, G_n , of ϵ_n is given by:*

$$G_n(z) = \frac{1}{n} \log \left| \frac{\Delta_n(z)}{2} + \sqrt{\left(\frac{\Delta_n(z)}{2}\right)^2 - 1} \right| \tag{2.9}$$

Moreover, we have that:

$$\begin{aligned} B_n(z)^n &= \frac{\Delta_n(z)}{2} - \sqrt{\left(\frac{\Delta_n(z)}{2}\right)^2 - 1} \\ B_n(z)^{-n} &= \frac{\Delta_n(z)}{2} + \sqrt{\left(\frac{\Delta_n(z)}{2}\right)^2 - 1} \end{aligned} \tag{2.10}$$

and (2.4) and (2.3) hold.

Proof We proved (2.9) above. By that formula, the absolute value of the right-hand side of the second equation in (2.10) is $\exp(nG_\epsilon(z))$. Since this expression is analytic on $\Omega_n \setminus \{\infty\}$ and is $Cz^n + O(z^{n-1})$ with $C > 0$ there, it must be $B_n(z)^{-n}$. The first equation in (2.10) holds since both sides are inverses of the two sides of the second equation, which we have just proven. Adding the two equations in (2.10) and using (2.7), we get (2.4).

By (1.10), $B_n(z)^{-n} = z^n C(\epsilon_n)^{-n} + O(z^{n-1})$; we see that $\Delta_n(z) = z^n C(\epsilon_n)^{-n} + O(z^{n-1})$ also. By (2.7) and the fact that T_n is monic, we obtain (2.3). □

Since $\epsilon \subset \epsilon_n$, we have that $C(\epsilon) \leq C(\epsilon_n)$, so (2.3) immediately implies Schiefermayr’s Theorem, (1.12).

Next, we turn to the form of the equilibrium measure, ρ_n , for ϵ_n . We note that Δ_n runs monotonically from -2 to $+2$ or vice versa. We have that:

Theorem 2.3 *In each band of ϵ_n , define $\theta(x) \in [0, \pi]$ by*

$$\Delta_n(x) = 2 \cos(\theta(x)) \tag{2.11}$$

Then

$$d\rho_n(x) = (\pi n)^{-1} |\theta'(x)| dx \tag{2.12}$$

In particular, each band has ρ_n -measure $1/n$. If $\eta_j \in \epsilon_n^{(j)}$ is the zero of T_n in $\epsilon_n^{(j)}$, then each of $[\alpha_j, \eta_j]$ and $[\eta_j, \beta_j]$ has ρ_n -measure $1/2n$.

We will not give a formal proof of this result. The final sentence is an immediate consequence of (2.12) given that Δ_n runs monotonically from 2 to -2 or from -2 to 2 on a band, so that θ runs monotonically from 0 to π or from π to 0 . (2.12) is well known in the mathematical physics literature obtained from the theory of discriminants. For example, Simon [34] has two

proofs of it—one as Theorem 5.3.8 and one as Theorem 5.4.8. A quick proof is to apply the operator ∂ of (1.16) to (2.9), using (1.17) to get

$$\int \frac{d\rho_n(x)}{x-z} = \frac{1}{n} \frac{\Delta'_n(z)}{\sqrt{\Delta_n(z)^2 - 4}} \quad (2.13)$$

Taking imaginary parts of both sides, one gets (2.12) by noting the boundedness of this imaginary part and computing its boundary value. The square root on the right of (2.13) is pure imaginary on ϵ_n . One needs to track carefully its phase from the square root singularity which is compensated in the ratio by the change of the sign of Δ'_n from band to band. This immediately implies a strong form of (2.5).

Theorem 2.4 *Let K be a gap of ϵ . Then (2.5) holds. If T_n has no zero in K , then $1/n$ can be replaced by $1/2n$. Moreover, $K \cap \epsilon_n$, if non-empty, is a single interval.*

- Remarks*
1. When ϵ is a finite gap set, it is an implicit result of Sodin–Yuditskii [36] and explicit result of Peherstorfer [23, 25] that each gap contains no more than one band.
 2. The interval mentioned in the last sentence may be closed (if the band is entirely in K), half open (if one end of the intersection is an end-point of K), or open (if the intersection is all of K).
 3. From (2.13), we deduce that ρ_n is a.c. with respect to dx and

$$d\rho_n(x) = w_n(x)dx, \quad w_n(x) = \frac{1}{\pi n} \frac{|\Delta'_n(x)|}{\sqrt{4 - \Delta_n(x)^2}} \quad (2.14)$$

for $x \in \epsilon_n$

Proof Suppose that $\epsilon_n^{(j)} \cap K \neq \emptyset$. Since K is connected and at least one of β_j or α_{j+1} lies in ϵ , we conclude that K is disjoint from all the $\epsilon_n^{(k)}$ for $k > j$. Similarly, K is disjoint from all the $\epsilon_n^{(k)}$ for $k < j$. Thus K contains at most one band, so (2.5) follows from Theorem 2.3. If T_n has no zero in K , at most half a band lies in K and we get the improved $1/2n$ result. Since $K \cap \epsilon_n$, if non-empty, is a single band, we get the single interval claim. \square

Theorem 2.3 has another immediate consequence:

Theorem 2.5 *ϵ_n is a period- n set.*

Our penultimate result in this section is a converse to this result. We need two preliminaries:

Theorem 2.6 *Suppose that $[\alpha, \beta]$ is a connected component of a compact set $\epsilon \subset \mathbb{R}$. Then:*

- (a) G_ϵ has an analytic continuation across (α, β) , i.e., there is an analytic function in a neighborhood, N , of (α, β) whose real part agrees with G_ϵ on $\{z \in N \mid \text{Im}(z) > 0\}$. G_ϵ vanishes everywhere on (α, β) and is continuous on $\mathbb{C}_\pm \cup (\alpha, \beta)$.
- (b) If ∂ is given by (1.16), then $h(z) \equiv \sqrt{(z - \alpha)(\beta - z)}\partial G_\epsilon$ has an analytic continuation across $[\alpha, \beta]$, i.e., there is an analytic function in a neighborhood, N_1 , of $[\alpha, \beta]$ which agrees with h on $\{z \in N_1 \mid \text{Im}(z) > 0\}$.
- (c) We have that

$$d\rho_\epsilon \upharpoonright [\alpha, \beta] = \frac{q(x)}{\sqrt{(x - \alpha)(\beta - x)}} dx \tag{2.15}$$

where $q(x) > 0$ on (α, β) and continuous on $[\alpha, \beta]$.

- (d) Suppose that $\epsilon = \bigcup_{k=1}^p [a_k, b_k]$. Let $\{c_k\}_{k=1}^{p-1}$ be the critical points of G_ϵ where $a_k < b_k < c_k < a_{k+1} < b_{k+1}$, $k = 1, \dots, p - 1$. Then $d\rho_\epsilon(x) = w(x)dx$ where, for $x \in \epsilon$,

$$w(x) = \frac{\frac{1}{\pi} \prod_{k=1}^{p-1} |x - c_k|}{\prod_{k=1}^p |(x - a_n)(x - b_k)|^{1/2}} \tag{2.16}$$

Remarks 1. We note the compatibility of (2.14) and (2.16). For the leading coefficient of Δ'_n is n times that of Δ_n canceling the $1/n$ yielding a formula like (2.16) but with the product over all the zeros of Δ'_n in the numerator and over all band edges in the denominator. At a closed gap, $4 - \Delta_n^2$ has a double zero and Δ'_n a single so they cancel and (2.16) results in the special case where ϵ is a period- n set.

- 2. (d) is, as we'll note, equivalent to a product formula for $2\partial G_\epsilon$. This formula can be found, for example, as (5.4.88) in Simon [34].
- 3. (2.18) below is not literally true but is a bit of poetry because $Q(x) = 1$ above $\text{cvh}(\epsilon)$, so the integral in (2.18) diverges. One can use the renormalized version of the Herglotz representation, only look at imaginary parts, or put in a cutoff. For use in proving (b), the cutoff is no problem since the remainder is analytic in a neighborhood of $[\alpha, \beta]$. For (d), if the upper cutoff is R above $\text{cvh}(\epsilon)$, we get a $\log(R - z)$ term which is $\log(R) + o(1)$ so if we absorb $\log(R)$ into redefining C , we get a limit and the argument that we give then works.

Proof These results are well known to experts on potential theory and/or spectral theory of Schrödinger operators. Especially relevant are ideas of Craig [10] given that $G_\epsilon(x) = 0$ on ϵ implies that $-\partial G_\epsilon$ is the Stieltjes transform of a measure reflectionless on ϵ . (We caution the reader that Craig's "Green's function" is not G_ϵ but $-2\partial G_\epsilon$.) So we'll only sketch the details.

Since G_ϵ is a positive harmonic function on the upper half plane, \mathbb{C}_+ , there is a Herglotz function f on \mathbb{C}_+ with $\text{Im}(f) = G_\epsilon$, so we can write

a Herglotz representation for it. Since G_ϵ is locally bounded, the measure in this representation is absolutely continuous. Moreover, since q.e. on (α, β) , $\lim_{\epsilon \downarrow 0} G_\epsilon(x + i\epsilon) = 0$, this measure gives zero weight to $[\alpha, \beta]$ which implies (a).

By differentiating the formula for G_ϵ in terms the potential of $d\rho_\epsilon$, we see, by (1.17), that:

$$F(z) \equiv -2\partial G_\epsilon(z) = \int \frac{d\rho_\epsilon(x)}{x - z} \tag{2.17}$$

(a) implies that F is analytic across (α, β) and $\text{Re}(F(x)) = 0$ there. Thus we can use ideas of Craig [10] to write a Herglotz representation for $\log(F)$:

$$\log(F(z)) = C + \int \frac{Q(x)dx}{x - z} \tag{2.18}$$

where C is a real constant and $Q(x) = \frac{1}{\pi} \lim_{\epsilon \downarrow 0} \text{Arg}(F(x + i\epsilon))$. It follows that $Q(x) = 1/2$ on (α, β) and is either identically 0 or identically 1 just below α and similarly just above β . Exponentiating (2.18) implies (b) which easily leads to (c) with $q(x) \geq 0$. To see that $q(x) > 0$, we note that the Herglotz representation (2.18) implies that $\text{Im}(\log(F(z)) - \frac{\pi}{2})$ goes to zero as (α, β) is approached from the upper half plane. By the strong reflection principle, $\log(F(z))$ has continuous boundary values, so in particular F has no zeros on (α, β) . Since $\text{Re}(F) = 0$ there, we see that $\text{Im}(F)$ is non-vanishing there.

To get (d), we use the representation (2.18) (noting that $Q(x)$ is 1/2 on each (a_j, b_j) , 1 on (b_p, ∞) and on each (b_j, c_j) and 0 on $(-\infty, a_1)$ and each (c_j, a_{j+1})) to get a product formula for F . This is equivalent to the formula for w by the integral representation in (2.17) and the theory of boundary values of Stieltjes transforms. (The constant C in (2.18) is determined by the $-1/z$ asymptotics of F .) □

Theorem 2.7 *If $\epsilon \subset \mathbb{C}$ is compact and γ is any rectifiable curve in $\mathbb{C} \setminus \epsilon$, then, $\Delta_\gamma(B_\epsilon)$, the change in phase of B_ϵ in going around γ , is given by*

$$\Delta_\gamma(B_\epsilon) = \exp\left(-2\pi i \int N(\gamma, x) d\rho_\epsilon(x)\right) \tag{2.19}$$

where $N(\gamma, x)$ is the winding number of γ around x . In particular, if γ winds once around $\mathfrak{g} \subset \epsilon$ and around no other points of ϵ , then the multiplicative change of phase of B_ϵ around γ is $\exp(-2\pi i \rho_\epsilon(\mathfrak{g}))$.

Proof Applying 2∂ to both sides of $\log(|B_\epsilon|) = -G_\epsilon$, using the formula (1.8) for G_ϵ in terms of ρ_ϵ and (1.17), we get that

$$B'_\epsilon(z) B_\epsilon^{-1}(z) = \int \frac{d\rho_\epsilon(x)}{x - z} \tag{2.20}$$

where one needs an easy argument to justify interchanging the derivative and integral. Multiplying by $(2\pi i)^{-1}$ and doing the contour integral, one gets (2.19) after interchanging the integrals and using the formula for $N(\gamma, x)$ as a contour integral. \square

Theorem 2.8 *Let $\epsilon \subset \mathbb{R}$ be a period- n set. Then for $k = 1, 2, \dots, \epsilon$ is the set where its Chebyshev polynomial, T_{kn} , takes its values in $[-\|T_{kn}\|_\epsilon, \|T_{kn}\|_\epsilon]$, i.e., $\epsilon_{kn} = \epsilon$.*

Remark It is easy to see that if S_n is the Chebyshev polynomial for $[-1, 1]$ (which is the classical Chebyshev polynomial of the first kind up to a constant), then for the ϵ 's of this theorem, one has that $T_{kn} = \|T_n\|_\epsilon^k S_k(T_n/\|T_n\|_\epsilon)$.

Proof By Theorem 2.7, the argument of B_ϵ^n changes by an integral multiple of 2π as one goes around any connected component of ϵ so it defines a function analytic in $\mathbb{C} \setminus \epsilon$. Since this function is real on \mathbb{R} near $+\infty$, we have on $\mathbb{C} \setminus \epsilon$ that $B_\epsilon^n(\bar{z}) = \overline{B_\epsilon^n(z)}$. Moreover, by Theorem 2.6, this function is continuous as ϵ is approached from one or the other side of ϵ and has magnitude 1 there. This shows that

$$P_n(z) \equiv C(\epsilon)^n (B_\epsilon^n(z) + B_\epsilon^{-n}(z)) \tag{2.21}$$

is continuous across the interior of ϵ and so analytic there. The end points of the intervals are thus removable singularities since P_n is bounded there by Theorem 2.6. It follows that P_n is an entire function and, by the asymptotics, (1.21), of B_ϵ , it is a monic polynomial of degree n .

Since $|B_\epsilon| = 1$ on ϵ , we have that $\|P_n\|_\epsilon \leq 2C(\epsilon)^n$, so by Schiefermayr's inequality, (1.12), and uniqueness of the minimizer, P_n is T_n . Since $|B_\epsilon| < 1$ on $\mathbb{C} \setminus \epsilon$, we see that $\epsilon_n = \epsilon$ as claimed. This proves the $k = 1$ part of the Theorem. But any period- n set is also a period- kn set. \square

Our final result in this section proves a minimality property of ϵ_n .

Theorem 2.9 *Let $\epsilon \subset \mathbb{R}$. Then for any period- n set, $g \supset \epsilon$, we have that*

$$C(\epsilon_n) \leq C(g) \tag{2.22}$$

with equality if and only if $g = \epsilon_n$.

Proof Let H_n be the n th Chebyshev polynomial for g . Since H_n is monic, we have that

$$2C(\epsilon_n)^n = \|T_n\|_\epsilon \leq \|H_n\|_\epsilon \leq \|H_n\|_g = 2C(g)^n \tag{2.23}$$

proving (2.22). If one has equality in (2.22), then one has equality in the first inequality in (2.23), so $T_n = H_n$ which implies, by Theorem 2.8 and the definition of ϵ_n , that $g = \epsilon_n$. \square

The results of this section can be understood from a spectral theory point of view. We end this section with a description of this connection to periodic Jacobi matrices—one place to find the details of the theory of such matrices is Chapter 5 of Simon [34]. We consider two-sided sequences $\{a_j, b_j\}_{j=-\infty}^{\infty}$ with $a_j > 0$, $b_j \in \mathbb{R}$ and so that for some $p > 0$ and all j in \mathbb{Z}

$$a_{j+p} = a_j, \quad b_{j+p} = b_j \tag{2.24}$$

We define doubly infinite tridiagonal matrices, J , with b_j along the diagonal and a_j on the principle subdiagonals (so that row k has non-zero elements $a_{k-1} b_k a_k$ with b_k in column k).

For $z \in \mathbb{C}$ fixed, we are interested in solutions, $\{u_j\}_{j=-\infty}^{\infty}$, of

$$a_j u_{j+1} + b_j u_j + a_{j-1} u_{j-1} = z u_j \tag{2.25}$$

We study the p -step transfer (aka update) matrix:

$$M_p(z) \begin{pmatrix} u_1 \\ a_0 u_0 \end{pmatrix} = \begin{pmatrix} u_{p+1} \\ a_p u_p \end{pmatrix} \tag{2.26}$$

We put a 's in the bottom component so that the one step matrix $\frac{1}{a_j} \begin{pmatrix} z - b_j & -1 \\ a_j^2 & 0 \end{pmatrix}$ has determinant 1 and thus $\det(M_p(z)) = 1$.

In terms of the first and second kind orthogonal polynomials for Jacobi parameters $\{a_n, b_n\}_{n=1}^{\infty}$, as defined in Section 3.2 of [34],

$$M_p(z) = \begin{pmatrix} p_p(z) & -q_p(z) \\ a_p p_{p-1}(z) & -a_p q_{p-1}(z) \end{pmatrix} \tag{2.27}$$

The *discriminant*, $\Delta(z)$, defined by

$$\Delta(z) = \text{Tr}(M_p(z)) = p_p(z) - a_p q_{p-1}(z) \tag{2.28}$$

is a (real) polynomial of degree exactly p . Given the recursion relations for $p_j(z)$ or the form of the one step transfer matrix, we see that $\Delta(z)$ is a polynomial of degree p with leading coefficient $(a_1 \cdots a_p)^{-1}$.

If $M_p(z)$ has an eigenvalue λ , it is easy to see the difference equation has a (Floquet) solution obeying $u_{j+mp} = \lambda^m u_j$ for all $m \in \mathbb{Z}$. Since $\det(M_p(z)) = 1$, if $\lambda \neq \pm 1$, we get two linearly independent solutions, so if $|\lambda| \neq 1$, all solutions are exponentially growing at ∞ and/or at $-\infty$. On the other hand, if $|\lambda| = 1$, there is a bounded solution. Note that $M_p(z)$ has an eigenvalue with $|\lambda| = 1$ if and only if $\Delta(z) \in [-2, 2]$.

Since it is known that the spectrum of J is the closure of the set of z 's for which there are polynomially bounded solutions (Schnol's Theorem), we conclude that $\text{spec}(J) = \Delta^{-1}([-2, 2])$. Since J is self-adjoint, we have that $\Delta^{-1}([-2, 2]) \subset \mathbb{R}$.

If $f(z)$ is an entire function real on the real axis and $f'(x_0) = 0$ for $x_0 \in \mathbb{R}$, because of the local structure of analytic functions, there will be non-real z 's near x_0 with $f(z)$ a real value near $f(x_0)$. Thus $\Delta^{-1}([-2, 2]) \subset \mathbb{R}$ implies that

$$\Delta(x) \in (-2, 2) \Rightarrow \Delta'(x) \neq 0$$

Therefore, between successive points where $\Delta(x_0) = \pm 2$ and where $\Delta(x_1) = \mp 2$, $\Delta(x)$ is strictly monotone and Δ is a bijection. It follows that Δ has an alternating set in $\epsilon = \Delta^{-1}([-2, 2])$. Therefore, $a_1 \cdots a_p \Delta$ is the Chebyshev polynomial for $\epsilon = \epsilon_p$. It is known that J has purely absolutely continuous spectrum of multiplicity 2, which implies that the half line operator is regular in the sense of Stahl–Totik [38] so one has that $(a_1 \cdots a_p)^{1/p} = C(\epsilon)$ and the density of zeros is $d\rho_\epsilon$. In the spectral theory literature, the density of zeros is called the density of states and it is well known that each "band" of the spectrum has density $1/n$.

From a spectral theory point of view, the fact that every period- n set has a discriminant follows from the fact that such a set is the spectrum of a periodic Jacobi matrix. Indeed, if ℓ is the number of gaps of a period- n set, then one constructs an ℓ dimensional torus of such periodic matrices. For any finite gap set there is an isospectral torus which can be constructed as reflectionless Jacobi matrices (see Remling [28]), or as minimal Herglotz functions (see our paper [8]) or using Hardy spaces of character automorphic functions (see Sodin–Yuditskii [37]). The elements of the isospectral torus are almost periodic with frequencies generated by the harmonic measures of the components of the finite gap set and so periodic with period n if all these measures are of the form k/n .

3 Root asymptotics

In this section, before turning to our two main theorems in the final sections, we make a comment on the FSS theorem and a remark on a Theorem of Saff–Totik concerning root asymptotics of the Chebyshev polynomials of any arbitrary infinite, compact set $\epsilon \subset \mathbb{C}$. The first concerns the following theorem of Totik:

Theorem 3.1 (Totik [43]) *Given any infinite, compact subset $\epsilon \subset \mathbb{R}$, there exist period- n sets $\mathfrak{g}_n \supset \epsilon$ so that*

$$\lim_{n \rightarrow \infty} C(\mathfrak{g}_n) = C(\epsilon) \tag{3.1}$$

Remarks 1. This is useful because one can use polynomial mappings to extend some results from $[-1, 1]$ to period- n sets and then this theorem to extend the result to general sets in \mathbb{R} . Polynomial inequalities have been obtained by Totik using this method (see his review article [47]) and, using this method, Lubinsky's approach [19] to universality for the CD kernel has been extended from $[-1, 1]$ to general compact sets in \mathbb{R} by Simon [33] and Totik [45].

2. A stronger result is known for finite gap sets—namely, if ϵ has ℓ gaps, then for $n \geq \ell$, g_n can be picked to also have exactly ℓ gaps. Indeed, this is how Totik proved Theorem 3.1. This stronger result has been discovered and proven independently by several different authors [3, 21, 25, 29, 43].

Our point here is to note the following:

Theorem 3.1 is equivalent to the FSS Theorem

Totik informed us that he knew this but it seems not to be in the literature. To see that FSS \Rightarrow Theorem 3.1, note that one can take $g_n = \epsilon_n$ and use (2.3) and (1.13) plus $2^{1/n} \rightarrow 1$ to get (3.1). Conversely, given Theorem 3.1 and (2.22), we see that (3.1) holds for $g_n = \epsilon_n$. Then (2.3) and $2^{1/n} \rightarrow 1$ implies (1.13).

Our other result on root asymptotics is:

Theorem 3.2 For any compact non-polar set, $\epsilon \subset \mathbb{C}$,

$$|T_n(z)|^{1/n} \rightarrow C(\epsilon) \exp(G_\epsilon(z)) = \exp(-\Phi_\epsilon(z)) \quad (3.2)$$

uniformly on compact subsets of $\mathbb{C} \setminus \text{cvh}(\epsilon)$.

Remarks 1. This theorem is not new. It appears in Saff–Totik [30, Chapter 3.3] as Theorem 3.9 in the more general context of weighted Chebyshev polynomials.

2. Our proof is different. They first control the density of zeros and use that to prove this result; shortly, we'll go in the other direction.

Proof Recall the Bernstein–Walsh Lemma, which says that, for any polynomial, P , of degree n , and any compact set $\epsilon \subset \mathbb{C}$, and so for T_n , one has that for all $z \in \mathbb{C}$

$$|T_n(z)| \leq \|T_n\|_\epsilon \exp(nG_\epsilon(z)) \quad (3.3)$$

Taking n th roots and using $C(\epsilon) = \exp(-R(\epsilon))$, we get that

$$|T_n(z)|^{1/n} \leq Y(n) \exp(-\Phi_\epsilon(z)), \quad Y(n) \equiv \|T_n\|_\epsilon^{1/n} / C(\epsilon) \quad (3.4)$$

Fejér’s theorem [12] says that all the zeros of T_n lie in $\text{cvh}(\epsilon)$ so on $\tilde{\Omega} \equiv (\mathbb{C} \cup \{\infty\}) \setminus \text{cvh}(\epsilon)$, we have that

$$h_n(z) \equiv \log(Y(n)) - \Phi_\epsilon(z) - \frac{1}{n} \log(|T_n(z)|) \tag{3.5}$$

are non-negative harmonic functions with $h_n(\infty) = \log(Y(n))$. By the FFS Theorem, $Y(n) \rightarrow 1$, so by Harnack’s inequality, h_n goes to zero, uniformly on compact subsets of $\tilde{\Omega}$. \square

Standard methods of going from root asymptotics to control on the density of zeros (see, for example [32, 38]) imply the following result (which is a special case of Theorem 4.7 of [30]):

Theorem 3.3 *Let $\epsilon \subset \mathbb{R}$ be compact and non-polar. Then the density of zeros measures for T_n converge to the equilibrium measure for ϵ .*

4 Totik–Widom bounds

Proof of Theorem 1.4 Consider the function

$$h(z) \equiv G_\epsilon(z) - G_{\epsilon_n}(z) \tag{4.1}$$

This function is harmonic on $(\mathbb{C} \cup \{\infty\}) \setminus \epsilon_n$. h is harmonic at ∞ since the $\log(|z|)$ terms cancel and one can use the removable singularities theorem. One has that

$$h(\infty) = R(\epsilon) - R(\epsilon_n) = \log \left[\frac{C(\epsilon_n)}{C(\epsilon)} \right] \tag{4.2}$$

Since $d\rho_n$ is harmonic measure, (1.7) holds for h . Since $h(x) = G_\epsilon(x)$ on ϵ_n , if $\{K_j\}_{j=1}^M$ are the gaps for ϵ , then, using (2.5),

$$h(\infty) \leq \sum_{j=1}^M \rho_n(K_j) \max_{x \in K_j} (G_\epsilon(x)) \leq \frac{1}{n} \sum_{j=1}^M G_\epsilon(w_j) \tag{4.3}$$

Since regularity of ϵ implies that G_ϵ vanishes at the ends of each gap, the maximum is taken at a critical point, w_j , and the sum is precisely the Parreau–Widom sum. Exponentiating and using $\|T_n\|_\epsilon = 2C(\epsilon_n)^n$ and (4.2), we get the result, (1.18). \square

Remarks 1. Because of the final assertion in Theorem 2.4, one can replace $PW(\epsilon)$ in (1.18) by $\frac{1}{2}PW(\epsilon) + \frac{1}{2}S_n$ where S_n is the sum of the n largest values among the $G_\epsilon(w_j)$.

2. Just as the FFS Theorem is equivalent to information about $C(\epsilon_n)$, so (1.14) is equivalent to

$$C(\epsilon_n) \leq C(\epsilon) \left(1 + \frac{q}{n}\right) \tag{4.4}$$

for some q , which is the form that Totik proved in the finite gap case. To see that (4.4) \Rightarrow (1.14), we note that it is well known that $(1 + \frac{q}{n})^n$ is monotone increasing in n to $\exp(q)$, so given (2.3), (4.4) implies (1.14) with $Q = 2 \exp(q)$. In the other direction, since

$$e^x - 1 = \int_0^x e^y dy \leq x e^x \quad \text{for } x \geq 0 \tag{4.5}$$

we have for $n \geq 1$ and $\tilde{Q} \geq 1$ that

$$\tilde{Q}^{1/n} = e^{\log(\tilde{Q})/n} \leq 1 + \frac{\log(\tilde{Q})}{n} e^{\log(\tilde{Q})/n} \leq 1 + \frac{q}{n} \tag{4.6}$$

where $q = \tilde{Q} \log(\tilde{Q})$. This shows that given (2.3), (1.14) implies (4.4) with $\tilde{Q} = Q/2$ and q as just given.

3. This seems to be the first example of an upper bound of the form (1.14) for an ϵ with an infinite number of components although there have been a number of papers, as we noted, for fairly general finite component sets in \mathbb{C} . Since T_n is defined variationally, in principle, upper bounds shouldn't be hard—one need only guess a clever trial polynomial and indeed, that's what the earlier work does. Our approach uses potential theory and doesn't seem to involve a variational guess although, it might be argued that underlying (2.3) is using the discriminant of ϵ as a trial polynomial. But that's of course T_n which is not merely a trial polynomial!

5 Szegő–Widom asymptotics

In this section, we will prove Theorem 1.9 settling Widom's 1969 conjecture. We begin with some notation and some preliminaries. $\epsilon \subset \mathbb{R}$ will be a finite gap set (although some results like Theorem 5.1 hold more generally). As before,

$$\Omega \equiv (\mathbb{C} \cup \{\infty\}) \setminus \epsilon, \quad G \equiv G_\epsilon, \quad B \equiv B_\epsilon \tag{5.1}$$

ϵ_n is given by (1.3), G_n is its Green's function, $B_n \equiv B_{\epsilon_n}$, and $\Omega_n \equiv (\mathbb{C} \cup \{\infty\}) \setminus \epsilon_n$. We'll let

$$\tilde{\Omega} \equiv (\mathbb{C} \cup \{\infty\}) \setminus \text{cvh}(\epsilon), \quad \tilde{G} \equiv G_{\text{cvh}(\epsilon)}, \quad \tilde{B} \equiv B_{\text{cvh}(\epsilon)} \tag{5.2}$$

Since Green’s functions of regular sets decrease as the set increases (by an application of the maximum principle), we have that for all $z \in \mathbb{C}$

$$\begin{aligned} \epsilon \subset \epsilon_n \subset \text{cvh}(\epsilon) &\Rightarrow \tilde{G}(z) \leq G_n(z) \leq G(z) \\ &\Rightarrow |B(z)| \leq |B_n(z)| \leq |\tilde{B}(z)| \end{aligned} \tag{5.3}$$

Define

$$L_n(z) \equiv \frac{T_n(z)B(z)^n}{C(\epsilon)^n} \tag{5.4}$$

so that (1.22) says that $\lim_{n \rightarrow \infty} (L_n(z) - F_n(z)) = 0$ uniformly on compact subsets of the universal cover of Ω . By the Bernstein–Walsh lemma, (3.3), and (1.14), for any n and z ,

$$|L_n(z)| \leq \|T_n\|_\epsilon C(\epsilon)^{-n} \leq Q \tag{5.5}$$

Thus, by the Vitali convergence theorem, it suffices to prove that $\lim_{n \rightarrow \infty} (L_n(z) - F_n(z)) = 0$ uniformly on compact subsets of $\tilde{\Omega}$.

By (2.4) and (2.3), we have that

$$\begin{aligned} L_n(z) &= \frac{1}{2} \|T_n\|_\epsilon (B_n(z)^{-n} + B_n(z)^n) B(z)^n C(\epsilon)^{-n} \\ &= (1 + B_n(z)^{2n}) \frac{C(\epsilon_n)^n B(z)^n}{C(\epsilon)^n B_n(z)^n} \\ &= (1 + B_n(z)^{2n}) H_n(z) \end{aligned} \tag{5.6}$$

where

$$H_n(z) \equiv \frac{C(\epsilon_n)^n B(z)^n}{C(\epsilon)^n B_n(z)^n} \tag{5.7}$$

Since, by (5.3), $|B_n(z)| \leq |\tilde{B}(z)|$ and $|\tilde{B}(z)| < 1$ on $\tilde{\Omega}$, we conclude that to prove (1.22), it suffices to prove that uniformly on compact subsets of $\tilde{\Omega}$,

$$\lim_{n \rightarrow \infty} (H_n(z) - F_n(z)) = 0 \tag{5.8}$$

Lest it go by too fast, we want to note that, in the above, we canceled two factors of 2—namely those in (2.3) and (2.4)—and it is this cancelation that enables the proof to work. For an additional use of one of the factors of 2 allows us to rewrite (1.24) as

$$\lim_{n \rightarrow \infty} \frac{C(\epsilon_n)^n}{C(\epsilon)^n \|F_n\|_\Omega} = 1 \tag{5.9}$$

Since the first inequality in (5.3) implies that

$$\|H_n\|_{\tilde{\Omega}} \leq \frac{C(\epsilon_n)^n}{C(\epsilon)^n} \leq \frac{Q}{2} \tag{5.10}$$

we see that (1.24) and (2.3) imply that

$$\limsup_{n \rightarrow \infty} \frac{\|H_n\|_{\tilde{\Omega}}}{\|F_n\|_{\Omega}} \leq 1 \tag{5.11}$$

Our ability to settle Widom’s conjecture relies on a careful analysis of the sets ϵ_n . We also make use of the norm asymptotics obtained by Widom.

Since there is a uniform upper bound on H_n and F_n , Montel’s theorem implies compactness, so we need only show convergence of enough subsequences and we can pick them so that $H_{n(j)}$ has a limit which we’ll show is a trial function for Widom’s variational problem—(5.11) will then imply the limit must be the limit of the $F_{n(j)}$. For this to work, we need to consider H_n on a larger region than $\tilde{\Omega}$. H_n is, of course, defined as a multivalued function on Ω_n so we’ll need the following to control ϵ_n for n large. Recall (see point (b) after the alternation theorem) that each gap, K , of ϵ has at most one zero, $\zeta_n^{(K)}$, of T_n .

Theorem 5.1 *Let $K = (r, s)$ be a gap of ϵ .*

- (a) *If, for some subsequence, $\{n(j)\}_{j=1}^\infty$, $\zeta_{n(j)}^{(K)}$ has a limit $\zeta_\infty^{(K)} \in K$, then for large j , $\epsilon_{n(j)} \cap K$ is a closed interval containing $\zeta_{n(j)}^{(K)}$ of size bounded by $e^{-Dn(j)}$ for some $D > 0$.*
- (b) *If, for some subsequence, $\{n(j)\}_{j=1}^\infty$, $K \setminus \epsilon_{n(j)}$ is connected, then for some $C > 0$,*

$$|K \cap \epsilon_{n(j)}| \leq Cn(j)^{-2} \tag{5.12}$$

- (c) *If, for some subsequence, $\{n(j)\}_{j=1}^\infty$, $\zeta_{n(j)}^{(K)}$ has a limit which is r or s and if $K \setminus \epsilon_{n(j)}$ is not connected, then for some $C > 0$,*

$$|K \cap \epsilon_{n(j)}| \leq Cn(j)^{-1} \tag{5.13}$$

and that intersection approaches r or s .

Remarks 1. For our application here, all we need is that in the first case, the band shrinks to a point and, in the last two cases, the band moves to the edges. But the quantitative estimates are not hard, are interesting and we think optimal (as to order) in the first two cases and perhaps also in the third.

- The second case can occur if the band of ϵ_n that intersects K has a piece in ϵ (which is always the case if there are no zeros of T_n in K) or if the band is entirely in $[r, s]$ but one of the touching gaps is closed, i.e., one edge is r or s . It is also the case if $K \cap \epsilon_n = \emptyset$.

Lemma 5.2 *If, for some subsequence, $\{n(j)\}_{j=1}^\infty, \zeta_{n(j)}^{(K)}$ has a limit $\zeta_\infty^{(K)} \in K$, then for some $\delta > 0$ and all large j , we have that for some $D > 0$ and all $x \in [\zeta_\infty^{(K)} - \delta, \zeta_\infty^{(K)} + \delta]$,*

$$\frac{T_{n(j)}(x)}{\|T_{n(j)}\|_\epsilon} = \left(x - \zeta_{n(j)}^{(K)}\right) Q_j(x), \quad |Q_j(x)| > e^{Dn(j)} \tag{5.14}$$

Proof The first part of (5.14) holds where Q_j is the product of $x - x_j^{(n)}$ over all zeros other than $\zeta_{n(j)}^{(K)}$ divided by $\|T_{n(j)}\|_\epsilon$. By Theorems 1.2 and 3.3, for δ small, $\lim_{j \rightarrow \infty} n(j)^{-1} \log |Q_j(x)| = G(x)$ uniformly on $[\zeta_\infty^{(K)} - \delta, \zeta_\infty^{(K)} + \delta]$. Since G is bounded away from 0 on this interval, the second part of (5.14) is valid. \square

Recall that we use Δ_n for $2T_n / \|T_n\|_\epsilon$.

Lemma 5.3 *Let $v \in K$ be the unique critical point of G in K . Suppose that for some subsequence, $\{n(j)\}_{j=1}^\infty$, and $\delta > 0$, we have that $\epsilon_{n(j)} \cap (v - \delta, v + \delta) = \emptyset$. Then for j large, there is a single zero, $c_{n(j)}$, of $\Delta'_{n(j)}$ in $(v - \delta, v + \delta)$ and*

$$\lim_{j \rightarrow \infty} c_{n(j)} = v \tag{5.15}$$

Proof We have that uniformly in $[v - \delta/2, v + \delta/2]$, $\partial G_{n(j)} \rightarrow \partial G$. By (2.9), zeros of $\Delta'_{n(j)}$ are precisely the critical points of $G_{n(j)}$. This implies uniqueness of the zero while the uniform convergence implies existence and convergence. \square

It will be useful to have notation for the connected components of

$$\epsilon = \bigcup_{k=1}^p [a_k, b_k], \quad a_k < b_k < a_{k+1} < b_{k+1}, \quad k = 1, \dots, p - 1 \tag{5.16}$$

It will also be convenient to have a notation for connected components of ϵ_n . We will denote these by

$$[a_{n,k}, b_{n,k}], \quad k = 1, \dots, s_n, \quad 1 \leq s_n < 2p \tag{5.17}$$

since there is at most one extra band or partial band in each gap. Then $\epsilon_n = \bigcup_{k=1}^{s_n} [a_{n,k}, b_{n,k}]$ is a disjoint union. We let $\{c_{n,k}\}_{k=1}^{s_n-1}$ be the zeros of $\Delta'_n(x)$

not in ϵ_n , labeled so that $a_{n,k} < b_{n,k} < c_{n,k} < a_{n,k+1} < b_{n,k+1}$. By (2.16), we have that

$$w_n(x) = \frac{1}{\pi n} \frac{|\Delta'_n(x)|}{\sqrt{4 - \Delta_n(x)^2}} = \frac{\frac{1}{\pi} \prod_{k=1}^{s_n-1} |x - c_{n,k}|}{\prod_{k=1}^{s_n} |(x - a_{n,k})(x - b_{n,k})|^{1/2}} \tag{5.18}$$

Lemma 5.4 For $k = 2, \dots, s_n - 1$ and $x \in [a_{n,k}, b_{n,k}]$,

$$w_n(x) \geq \frac{1}{\pi} \frac{|x - c_{n,k}| |x - c_{n,k-1}|}{|x - b_p| |x - a_1|} \frac{1}{|(x - a_{n,k})(x - b_{n,k})|^{1/2}} \tag{5.19}$$

When $x \in [a_{n,1}, b_{n,1}]$,

$$w_n(x) \geq \frac{1}{\pi} \frac{|x - c_{n,1}|}{|x - b_p|} \frac{1}{|(x - a_{n,1})(x - b_{n,1})|^{1/2}} \tag{5.20}$$

When $x \in [a_{n,s_n}, b_{n,s_n}]$,

$$w_n(x) \geq \frac{1}{\pi} \frac{|x - c_{n,s_n}|}{|x - a_1|} \frac{1}{|(x - a_{n,s_n})(x - b_{n,s_n})|^{1/2}} \tag{5.21}$$

Proof Suppose first that $k = 2, \dots, s_n - 1$. Since $a_{n,j} < b_{n,j} < c_{n,j} < a_{n,j+1} < b_{n,j+1}$, $j = 1, \dots, s_n - 1$, the following estimates hold,

$$\frac{|x - c_{n,j}|}{|(x - a_{n,j})(x - b_{n,j})|^{1/2}} \geq 1 \quad \text{for all } x < a_{n,j} \tag{5.22}$$

$$\frac{|x - c_{n,j-1}|}{|(x - a_{n,j})(x - b_{n,j})|^{1/2}} \geq 1 \quad \text{for all } x > b_{n,j} \tag{5.23}$$

Thus, we get a lower bound if we drop the first zero and second band, second zero and third band, etc. and similarly drop the last zero and next to last band, etc. We can then use, since $x > b_1$, that $|x - b_1| < |x - a_1|$ and similarly on the other end to get the lower bound (5.19).

For the cases $k = 1$ or s_n , we only need to do the zero shielding on one side and we get (5.20) and (5.21). □

Proof of Theorem 5.1 (a) By Lemma 5.2, $T_{n(j)}(x) = \pm \|T_{n(j)}\|_\epsilon$ has solutions within $e^{-Dn(j)}$ of $\zeta_{n(j)}^{(K)}$, so there is a band of $\epsilon_{n(j)}$ of size less than $2e^{-Dn(j)}$ containing that zero. Since $K \cap \epsilon_{n(j)}$ contains at most one band, we have the claimed result.

(b), (c) (common part) We begin by showing that when the zeros don't have a limit point in (r, s) , then any part of $K \cap \epsilon_{n(j)}$ approaches the edges of K . If there is a point in $(r + 2\epsilon, s - 2\epsilon) \cap \epsilon_{n(j)}$ and a zero in $(r, r + \epsilon)$, then since $\epsilon_{n(j)} \cap K$

is connected, we have $[r + \epsilon, r + 2\epsilon] \equiv I_- \subset \epsilon_{n(j)} \cap K$ and similarly with $I_+ \equiv [s - 2\epsilon, s - \epsilon]$ if the zero is within ϵ of s . If there is no zero in K , then the band extends past one of the edges of K and again either I_- or I_+ is in $\epsilon_{n(j)} \cap K$. Thus, if it is not eventually true that $(r + 2\epsilon, s - 2\epsilon) \cap \epsilon_{n(j)} = \emptyset$, then infinitely often, either $C(\epsilon_n) \geq C(\epsilon \cup I_-)$ or the same for I_+ . This is inconsistent with Theorems 2.9 and 3.1 which imply that $\lim_{n \rightarrow \infty} C(\epsilon_n) = C(\epsilon)$ proving, by contradiction, the desired result that the bands in ϵ approach the edges.

(b) We consider the case where the band spreads into the left side of K and $k = 2, \dots, s_n - 1$. The argument is similar for $k = 1$ or $k = s_n$ or at the other ends of the gaps. We let $[a, b]$ be the connected component of ϵ with $b = r$. We thus suppose that $K \cap \epsilon_{n(j)} = [\alpha_j, \beta_j]$ with $\alpha_j = b$ and $\beta_j \rightarrow b$ and that $b_{n(j),k(j)} = \beta_j$ and $a_{n(j),k(j)} \leq a$. We assume that $c_{n(j)}$ and v are as in Lemma 5.3 and that (5.15) holds. Thus, we can suppose that j is so large that for $x < \beta_j$ we have that $|x - c_{n(j)}| > \frac{v-b}{2}$. (5.19) then becomes

$$w_{n(j)}(x) \geq \frac{v - b}{2\pi} \frac{|b - a|^{1/2}}{|b_p - a_1|^2} \frac{1}{|x - \beta_j|^{1/2}} \tag{5.24}$$

Integrating from α_j to β_j and using Theorem 2.4, we get that

$$\frac{1}{n(j)} \geq \int_{\alpha_j}^{\beta_j} w_{n(j)}(x) dx \geq \frac{v - b}{\pi} \frac{|b - a|^{1/2}}{|b_p - a_1|^2} |\beta_j - \alpha_j|^{1/2} \tag{5.25}$$

proving (5.12).

(c) The argument is similar to (b) except that now we don't have $c_{n(j),k(j)-1} < a < b = \alpha_j$ but only $b < c_{n(j),k(j)-1} < \alpha_j = a_{n(j),k(j)}$ so (5.24) is replaced by

$$w_{n(j)}(x) \geq \frac{v - b}{2\pi} \frac{|x - \alpha_j|^{1/2}}{|b_p - a_1|^2} \frac{1}{|x - \beta_j|^{1/2}} \tag{5.26}$$

which leads to (5.13) using $\int_0^1 \sqrt{\frac{x}{1-x}} < \infty$ and the analog of (5.25). □

Proof of Theorem 1.9 As noted, by boundedness and Montel's theorem, it suffices to prove that any subsequence has a subsubsequence for which (5.8) holds. We can choose this subsubsequence, $\{n(j)\}_{j=1}^\infty$, so that:

- (1) The characters $\chi_{n(j)} \rightarrow \chi_\infty$ for some character χ_∞ . This implies, by Theorem 1.6, that $F_{n(j)} \rightarrow F_\infty$ uniformly on compact subsets and $\|F_{n(j)}\|_\Omega \rightarrow \|F_\infty\|_\Omega$.
- (2) In each gap, K_ℓ , of ϵ , either $T_{n(j)}$ has a zero for j large and the limit of the zeros is $x_\ell \in K_\ell$ or any zero in the gap, K_ℓ , approaches ϵ in the limit or there is no zero in that gap.

- (3) On $\tilde{\Omega}$, the subsequence $H_{n(j)}(z)$ has a limit $H_\infty(z)$ by Montel's theorem. In this case, it is easy to see that $\|H_\infty\|_{\tilde{\Omega}} \leq \liminf \|H_{n(j)}\|_{\tilde{\Omega}}$.

Let \mathcal{L} be the set of gaps, K_ℓ , with a limit point of zeros. The $H_{n(j)}$ can be continued along any curve in $\Omega_{n(j)}$. Since all the harmonic measures of sets in e_n are multiples of $1/n$, B_n^n is analytic on Ω_n . It follows that $H_{n(j)}$ are defined and character automorphic with character $\chi_{n(j)}$ on sets which converge to the universal cover of Ω with the points that lie over the set $\{x_\ell\}_{\{K_\ell \in \mathcal{L}\}}$ removed. Thus, by (5.10) and Vitali's theorem, H_∞ has a continuation to that set. By (5.10) again, x_ℓ are removable singularities.

Since near ∞ , for any g , $B_g(z) = z/C(g) + O(1)$, we see that for each n , $H_n(\infty) = 1 \Rightarrow H_\infty(\infty) = 1$. Thus, H_∞ is a trial function for the problem where F_∞ is the minimizer. By (5.11) and continuity of $\|F_n\|_\Omega$, we see that $\|H_\infty\|_\Omega \leq \|F_\infty\|_\Omega$. Thus, the uniqueness of the minimizer implies that $H_\infty = F_\infty$, proving the desired convergence on $\tilde{\Omega}$. \square

Acknowledgements We would like to thank V. Totik and P. Yuditskii for useful communications. J. S. C. and M. Z. would like to thank D. Ramakrishnan and T. Soifer for the hospitality of Caltech where much of this work was done.

References

1. Akhiezer, N.I.: Function theory according to Chebyshev in Mathematics of the 19th century. In: Kolmogorov, A.N., Yushkevich, A.P. (eds.) Function Theory According to Chebyshev, Ordinary Differential Equations, Calculus of Variations, Theory of Finite Differences. Translated from the: Russian original by Roger Cooke, p. 1998. Birkhäuser, Basel (1987)
2. Andrievskii, V.V.: Chebyshev polynomials on a system of continua. *Constr. Approx.* **43**, 217–229 (2016)
3. Bogatyřev, A.B.: On the efficient computation of Chebyshev polynomials for several intervals, *Sb. Math.* **190** (1999), 1571–1605; Russian original in. *Mat. Sb.* **190**(11), 15–50 (1999)
4. Borel, É.: Leçons sur les fonctions de variables réelles et les développements en séries de polynômes. Gauthier-Villars, Paris (1905)
5. Carleson, L.: On H^∞ in multiply connected domains. In: Conference on Harmonic Analysis in Honor of Antoni Zygmund, vols. I, II, Chicago, IL, 1981. Wadsworth, Belmont, pp. 349–372 (1983)
6. Chebyshev, P.L.: Sur les fonctions qui diffèrent le moins possible de zéro. *J. Math. Pures Appl.* **19**, 319–346 (1874)
7. Christiansen, J.S: Dynamics in the Szegő class and polynomial asymptotics. *J. Anal. Math.* (to appear). [arXiv:1602.04033](https://arxiv.org/abs/1602.04033)
8. Christiansen, J.S., Simon, B., Zinchenko, M.: Finite gap Jacobi matrices. I. The isospectral torus. *Constr. Approx.* **32**, 1–65 (2010)
9. Conway, J.B.: Functions of One Complex Variable. II, Graduate Texts in Mathematics, vol. 159. Springer, New York (1995)
10. Craig, W.: The trace formula for Schrödinger operators on the line. *Commun. Math. Phys.* **126**, 379–407 (1989)
11. Faber, G.: Über Tschebyscheffsche Polynome. *J. Reine Angew. Math.* **150**, 79–106 (1919)

12. Fejér, L.: Über die Lage der Nullstellen von Polynomen, die aus Minimumforderungen gewisser Art entspringen. *Math. Ann.* **85**, 41–48 (1922)
13. Fekete, M.: Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten. *Math. Z.* **17**, 228–249 (1923)
14. Geronimo, J.S., Van Assche, W.: Orthogonal polynomials on several intervals via a polynomial mapping. *Trans. Am. Math. Soc.* **308**, 559–581 (1988)
15. Goncharov, A., Hatinoğlu, B.: Widom factors. *Potential Anal.* **42**, 671–680 (2015)
16. Helms, L.L.: *Introduction to Potential Theory*, Pure and Applied Mathematics, vol. 22. Wiley-Interscience, New York (1969)
17. Jones, P.W., Marshall, D.: Critical points of Green's function, harmonic measure, and the corona problem. *Ark. Mat.* **23**, 281–314 (1985)
18. Landkof, N.S.: *Foundations of Modern Potential Theory*. Springer, Berlin (1972)
19. Lubinsky, D.: A new approach to universality limits involving orthogonal polynomials. *Ann. Math.* **170**, 915–939 (2009)
20. Markov, A.: *Selected Papers on Continued Fractions and the Theory of Functions Deviating Least from Zero*. OGIZ, Moscow (1948)
21. McKean, H.P., van Moerbeke, P.: The spectrum of Hills equation. *Invent. Math.* **30**, 217–274 (1975)
22. Parreau, M.: Théorème de Fatou et problème de Dirichlet pour les lignes de Green de certaines surfaces de Riemann. *Ann. Acad. Sci. Fenn. Ser. A. I* **250/25** (1958)
23. Peherstorfer, F.: On the asymptotic behaviour of functions of the second kind and Stieltjes polynomials and on the Gauss-Kronrod quadrature formulas. *J. Approx. Theory* **70**, 156–190 (1992)
24. Peherstorfer, F.: Orthogonal and extremal polynomials on several intervals. *J. Comput. Appl. Math.* **48**, 187–205 (1993)
25. Peherstorfer, F.: Deformation of minimal polynomials and approximation of several intervals by an inverse polynomial mapping. *J. Approx. Theory* **111**, 180–195 (2001)
26. Peherstorfer, F.: Inverse images of polynomial mappings and polynomials orthogonal on them. *J. Comput. Appl. Math.* **153**, 371–385 (2003)
27. Ransford, T.: *Potential Theory in the Complex Plane*. Cambridge University Press, New York (1995)
28. Remling, C.: The absolutely continuous spectrum of Jacobi matrices. *Ann. Math.* **174**, 125–171 (2011)
29. Robinson, R.M.: Conjugate algebraic integers in real point sets. *Math. Z.* **84**, 415–427 (1964)
30. Saff, E., Totik, V.: *Logarithmic Potentials with External Fields*, Grundlehren der Mathematischen Wissenschaften, vol. 316. Springer, Berlin (1997)
31. Schiefermayr, K.: A lower bound for the minimum deviation of the Chebyshev polynomial on a compact real set. *East J. Approx.* **14**, 223–233 (2008)
32. Simon, B.: Equilibrium measures and capacities in spectral theory. *Inverse Probl. Imaging* **1**, 713–772 (2007)
33. Simon, B.: Two extensions of Lubinsky's universality theorem. *J. Anal. Math.* **105**, 345–362 (2008)
34. Simon, B.: *Szegő's Theorem and Its Descendants: Spectral Theory for L^2 Perturbations of Orthogonal Polynomials*. Princeton University Press, Princeton (2011)
35. Simon, B.: *A Comprehensive Course in Analysis, Part 3. Harmonic Analysis*. American Mathematical Society, Providence (2015)
36. Sodin, M.L., Yuditskii, P.: Functions deviating least from zero on closed subsets of the real axis. *St. Petersburg Math. J.* **4**, 201–249 (1993)
37. Sodin, M., Yuditskii, P.: Almost periodic Jacobi matrices with homogeneous spectrum, infinite-dimensional Jacobi inversion, and Hardy spaces of character-automorphic functions. *J. Geom. Anal.* **7**, 387–435 (1997)

38. Stahl, H., Totik, V.: *General Orthogonal Polynomials*. Cambridge University Press, Cambridge (1992)
39. Szegő, G.: Beiträge zur Theorie der Toeplitzschen Formen, I. *Math. Z.* **6**, 167–202 (1920)
40. Szegő, G.: Beiträge zur Theorie der Toeplitzschen Formen, II. *Math. Z.* **9**, 167–190 (1921)
41. Szegő, G.: Bemerkungen zu einer Arbeit von Herrn M. Fekete: Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten. *Math. Z.* **21**, 203–208 (1924)
42. Thiran, J.P., Demaille, C.: *Chebyshev Polynomials on Circular Arcs in the Complex Plane*, *Progress in Approximation Theory*, vols. 771–786. Academic Press, Boston (1991)
43. Totik, V.: Asymptotics for Christoffel functions for general measures on the real line. *J. Anal. Math.* **81**, 283–303 (2000)
44. Totik, V.: Polynomial inverse images and polynomial inequalities. *Acta Math.* **187**, 139–160 (2001)
45. Totik, V.: Universality and fine zero spacing on general sets. *Ark. Mat.* **47**, 361–391 (2009)
46. Totik, V.: Chebyshev constants and the inheritance problem. *J. Approx. Theory* **160**, 187–201 (2009)
47. Totik, V.: The polynomial inverse image method in *Approximation theory XIII: San Antonio 2010*, pp. 345–365, *Springer Proc. Math.*, vol. 13. Springer, New York (2012)
48. Totik, V.: Chebyshev polynomials on compact sets. *Potential Anal.* **40**, 511–524 (2014)
49. Totik, V., Varga, T.: Chebyshev and fast decreasing polynomials. *Proc. L.M.S.* (2015)
50. Totik, V., Yuditskii, P.: On a conjecture of Widom. *J. Approx. Theory* **190**, 50–61 (2015)
51. Tsuji, M.: *Potential Theory in Modern Function Theory*, reprinting of the 1959 original. Chelsea, New York (1975)
52. Volberg, A., Yuditskii, P.: Kotani-Last problem and Hardy spaces on surfaces of Widom type. *Invent. Math.* **197**, 683–740 (2014)
53. Widom, H.: Extremal polynomials associated with a system of curves in the complex plane. *Adv. Math.* **3**, 127–232 (1969)
54. Widom, H.: \mathcal{H}_p sections of vector bundles over Riemann surfaces. *Ann. Math.* **94**, 304–324 (1971)
55. Yuditskii, P.: On the L^1 extremal problem for entire functions. *J. Approx. Theory* **179**, 63–93 (2014)